Astronomy & Astrophysics manuscript no. aanda
February 13, 2024
©ESO 2024

Assembly bias in eBOSS

R. Paviot[1]⋆, A. Rocher[2,3], S. Codis[1], A. de Mattia[2], E. Jullo[4], and S. de la Torre[4]

[1] Université Paris Cité, Université Paris-Saclay, CEA, CNRS, AIM, F-91191, Gif-sur-Yvette, France
[2] IRFU Université Paris-Saclay, Université Paris Cité, CEA, DPhP, 91191, Gif-sur-Yvette, France
[3] Laboratory of Astrophysics, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland
[4] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

Received XXX; accepted XXX

arXiv:2402.07715v1 [astro-ph.CO] 12 Feb 2024

ABSTRACT

Context. Analytical models of galaxy-halo connection such as the Halo Occupation Distribution (HOD) model have been widely used over the past decades as a means to intensively test perturbative models on quasi-linear scales. However, these models fail to reproduce the galaxy-galaxy lensing signal on non-linear scales, over-predicting the observed signal up to 40%.
Aims. With ongoing Stage-IV galaxy surveys such as DESI and EUCLID, it is now crucial to accurately model the galaxy-halo connection up to intra-halo scales to accurately estimate theoretical uncertainties of perturbative models.
Methods. This paper compares the standard HOD model (based on halo mass only) to an extended HOD framework that incorporates as additional features galaxy assembly bias and local environmental dependencies on halo occupation. These models have been calibrated against the observed clustering and galaxy-galaxy lensing signal of eBOSS Luminous Red Galaxies (LRG) and Emission Lines Galaxies (ELG) in the range $0.6 < z < 1.1$. A combined clustering-lensing cosmological analysis is then performed on the simulated galaxy samples of both standard and extended HOD frameworks to quantify the systematic budget of perturbative models.
Results. By considering not only the mass of the dark matter halos but also these secondary properties, the extended HOD model offers a more comprehensive understanding of the connection between galaxies and their surroundings. In particular, we found that the LRGs preferentially occupy denser and more anisotropic environments. Our results highlight the importance of considering environmental factors in galaxy formation models, with an extended HOD framework that reproduces the observed signal within 20% on scales below 10 h^{-1}Mpc. Our cosmological analysis reveals that our perturbative model yields similar constraints regardless of the galaxy population, with a better goodness of fit for the extended HOD. These results suggest that the extended HOD framework should be used to quantify modeling systematics. This extended framework should also prove useful for forward modeling techniques.

Key words. Galaxies: halos – Galaxies: statistics (Cosmology:) large-scale structure of Universe

1. Introduction

In the standard model of cosmology, galaxies are predicted to form and evolve within dark matter halos. To extract cosmological information from observed galaxy clustering, it is mandatory to understand galaxy formation and thus the connection between galaxies and their host dark matter halos beyond their linear relationship on large scales (Kaiser 1984). One of the most popular and efficient models of the galaxy-halo connection for large-scale cosmological studies is the Halo Occupation Distribution model (HOD, Berlind & Weinberg (2002); Zheng et al. (2005, 2007)). In the standard HOD framework, it is assumed that all galaxies live inside dark matter halos and that galaxy occupation, which corresponds to the probability for a halo of hosting a galaxy, is solely determined by halo mass. This assumption has emerged as halo mass is the halo property that most strongly correlates with halo abundances and clustering as well as galaxy intrinsic properties (White & Rees 1978). However, a large number of studies has shown that the mass-only HOD does not properly model galaxy clustering on small scales (below $10h^{-1}$Mpc), and that halo occupation depends on secondary halo properties beyond halo mass. (Artale et al. 2018; Zehavi et al. 2018; Bose et al. 2019; Hadzhiyska et al. 2020, 2023a,b). In particular, it has been shown that galaxy occupation is correlated with halo formation time, a phenomenon referred to as *galaxy assembly bias*,(Dalal et al. 2008) as well as environmental properties, such as local density and local anisotropies, as one can predict from first principles for instance using excursion-set theory (Musso et al. 2018). In the previous study of Yuan et al. (2021, 2022), it has been shown that the standard framework fails to model accurately the full two-dimensional redshift space clustering of the Sloan Digital Sky Survey BOSS CMASS sample (Dawson et al. 2013), and that an extended HOD framework which incorporates assembly bias and environmental dependence of halo occupation agrees better with the data. In particular, the observed discrepancy of 30 − 40% between the observed galaxy-galaxy lensing (GGL) signal of CMASS galaxies and the standard HOD framework (Leauthaud et al. 2017), a tension referred to as 'lensing is low', can be partially resolved with the extended model (Contreras et al. 2023) so as to account for the impact of galaxy physics (Gouin et al. 2019; Chaves-Montero et al. 2023). However, no study has shown if these small scales discrepancy might bias or not the extraction of cosmological information. Here, we investigate this matter with the galaxies observed with the extended BOSS survey (Dawson et al. 2016), by directly fitting the GGL signal from the standard and extended HOD framework.
The paper is organized as follows: In Section 2, we briefly

⋆ e-mail: romainpaviot@gmail.com

present the observations. In Section 3, we describe the implementation of the standard and extended HOD framework as well as our fitting procedure. Finally, we present the results of our analysis in Section 4 and conclude in Section 5. We also provide additional information in the Appendix A and B.

2. Data

2.1. eBOSS and DES catalogs

The extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. (2016)) measured thousands of spectra of galaxies and quasars in the range $0.6 < z < 2.4$ in order to probe the geometry and growth rate of the Universe over cosmic time. In this work, we study the properties of the Emission Line Galaxies (ELGs) and eBOSS Luminous Red Galaxies (LRGs) combined with the BOSS CMASS sample[1]. Description of these samples can be found in Ross et al. (2020) and their cosmological analyses in Tamone et al. (2020); de Mattia et al. (2021); Bautista et al. (2021); Gil-Marín et al. (2020). We focus our attention on the South Galactic Cap (SGC), where 121 717 LRGs and 89 967 ELGs were observed in the range $0.6 < z < 1.0$ and $0.6 < z < 1.1$. In addition to the clustering properties of these galaxies, we aim to extract information from the GGL signal. This is achieved with the shape measurements of the Dark Energy Survey (DES) Year 3(Dark Energy Survey Collaboration et al. 2016). We used the main Y3 DES catalog (referred as 'Gold') [2]; Information of the photometry and shape measurements of this sample can be found in Sevilla-Noarbe et al. (2021); Gatti et al. (2021) and the corresponding cosmological 3x2pt analysis is presented in Abbott et al. (2022). The radial distribution and the footprint of each survey are presented in Fig 1. eBOSS and DES overlap in a small portion, around 754 deg^2, of the SGC. In the overlapping area, 14 413 244 sources can be used to infer the GGL signal of 44 456 LRG lenses, while all of the observed ELGs are within the DES footprint. [3]

2.2. The UCHUU simulation

In order to model the observed signals, we relied on the N-Body simulation UCHUU (Ishiyama et al. 2021). This simulation was run with GreeM, a massively parallel TreePM code (Ishiyama et al. 2009, 2012), with initial conditions produced by second-order Lagrangian perturbation theory starting at a redshift $z = 127$. The cosmological flat ΛCDM parameters of the simulation are compatible with the ones measured by Planck Collaboration et al. (2016), $(\Omega_{\rm m}, \Omega_{\rm b}, h, n_s, \sigma_8) = (0.3089, 0.0486, 0.6774, 0.9667, 0.8159)$. UCHUU is one of the most precise N-Body simulations to date, with $12\,800^3$ dark matter particles with mass $3.27 \times 10^8 h^{-1}\ M_\odot$ in a box of side-length $L = 2h^{-1}$Gpc. This large volume enables us to study a large number of dark matter halos up to a very high mass of $M \approx 10^{14.8} h^{-1}\ M_\odot$. These halos were identified with the RockStar halo finder (Behroozi et al. 2013). We extracted from the simulation [4] the halo and sub-halo catalogs of mass $M > 10^{11} h^{-1}$ M$_\odot$ and the particle distribution (sub-sampled by a factor 0.005), at redshifts $z = 0.86$ and $z = 0.7$, corresponding to the effective redshift of the ELG and LRG samples respectively.

[1] We will refer to it as a LRG sample, for simplicity. Catalogs can be found at https://data.sdss.org/sas/dr17/eboss/lss/catalogs/DR16/

[2] https://des.ncsa.illinois.edu/releases/y3a2

[3] These values were computed with Healpix (Górski et al. 2005) with NSIDE = 512

[4] https://skiesanduniverses.iaa.es/Simulations/Uchuu/

3. Halo Occupation model

3.1. Standard HOD parametrisation

In order to simulate realistic galaxy distributions from the UCHUU simulation, we populate halos using the standard halo occupation distribution (HOD) description. For LRG, we consider the HOD framework first described in Zheng et al. (2005, 2007) and widely used in previous BOSS and eBOSS studies (White et al. 2011; Zhai et al. 2017). For emission line galaxies, we use the Gaussian HOD model (GHOD) Avila et al. (2020); Rocher et al. (2023a) for central occupation. Hence, let us first define the mean central halo occupations as [5]

$$\left\langle N_{\rm cen}^{\rm LRG}(M)\right\rangle = \frac{f_{\rm ic}}{2}\left[1 + {\rm erf}\left(\frac{\log M - \log M_{\rm min}^{\rm LRG}}{\sigma_{\rm M,LRG}}\right)\right], \quad (1)$$

$$\left\langle N_{\rm cen}^{\rm ELG}(M)\right\rangle = \frac{A_c}{\sqrt{2\pi}\sigma_{\rm M,ELG}} \cdot \exp - \frac{\left(\log M - \log M_{\rm min}^{\rm ELG}\right)^2}{2\sigma_{\rm M,ELG}^2}, \quad (2)$$

where the $M_{\rm min}$ are the halo mass for which half of the halos have a central galaxy, the $\sigma_{\rm M}$ correspond to the scatter of central occupations, and $f_{\rm ic}$ and A_c correspond to completeness factors [6]. The satellite occupation is defined as

$$\langle N_{\rm sat}(M)\rangle = f(M)\left[\frac{M - \kappa M_{\rm min}}{M_1}\right]^\alpha, \quad (3)$$

$$\text{with } f(M) = \begin{cases} \left\langle N_{\rm cen}^{\rm LRG}(M)\right\rangle \text{ for LRG,} \\ A_s = \text{cst for ELG,} \end{cases} \quad (4)$$

with M_1 the halo mass for which a halo has one satellite galaxy, α the power-law index and κ a parameter that allows the cutoff of satellite occupation to vary with halo mass. For the ELG HOD we assume the relations given in Table 2 of Avila et al. (2020),

$$\kappa = 0.8 \text{ and } \log M_1 = \log M_{\rm min} + 0.3. \quad (5)$$

From these mean occupations, the number of central galaxies per halo of mass M then follows a Bernoulli distribution with mean probability equals to $N_{\rm cen}$. These galaxies are positioned at the center of their host halos. The number of satellite galaxies per halo follows a Poisson distribution with mean $\langle N_{\rm sat}\rangle$, with positions sampled from a Navarro-Frank-White profile (Navarro et al. 1996). Satellite velocities are normally distributed around the mean halo velocity, with a dispersion being equal to the one of the dark-matter halo particles,

$$\boldsymbol{v}_{\rm sat} \sim \mathcal{N}\left(\boldsymbol{v}_h, \cdot\sigma_{v_h}\right). \quad (6)$$

Eventually, the final number density of mock galaxies can be computed as

$$\bar{n}_{\rm gal} = \int \frac{{\rm d}n(M)}{{\rm d}M}\left[\langle N_{\rm cent}(M)\rangle + \langle N_{\rm sat}(M)\rangle\right]{\rm d}M \quad (7)$$

and the satellite fraction is given by

$$f_{\rm sat} = \frac{1}{\bar{n}_{\rm gal}}\int \frac{{\rm d}n(M)}{{\rm d}M}\langle N_{\rm sat}(M)\rangle\,{\rm d}M. \quad (8)$$

Obviously, the mean number density of galaxies is fixed which gives a constraint on the free parameters of the HOD. For ELGs, we follow the methodology described in Rocher et al. (2023a).

[5] Throughout this paper, the logarithm notation refers to the decimal logarithm unless specified otherwise.

[6] We will drop the subscripts for simplicity in the following.

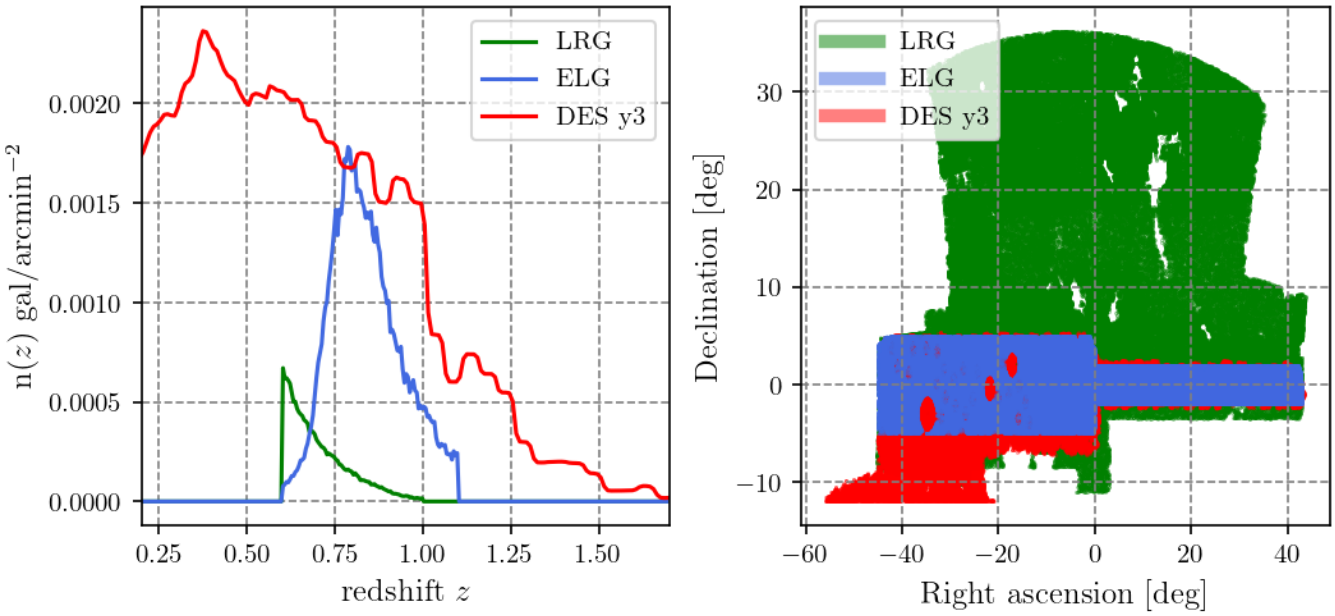

Fig. 1. Galaxy number density (left panel) and footprint (right panel) of the LRG sample (green), the ELG sample (blue), and the DES y3 GOLD catalog (red). The n(z) of DES corresponds to the concatenation of the n(z) on each tomographic bin (Myles et al. 2021). It has been sub-sampled by a factor 20 for clarity. On the right panel, we present the DES footprint with a cut below a declination of - 12.

As the mock clustering is governed by the ratio A_s/A_c, A_c is a not a free parameter but is instead fixed to provide the desired number density of objects, in our case $\bar{n}_{\rm gal}^{\rm ELG} = 2.3 \times 10^{-4} h^3{\rm Mpc}^{-3}$ by matching $n_{\rm gal}^{\rm HOD}$ to Eq. 7. Similarly, for LRG, the HOD is parametrised by (M_1,$\sigma_{\rm M}, \alpha, \kappa, f_{\rm ic}$), $M_{\rm min}$ is fixed in order to obtain twice the mean observed density of the LRG sample (to reduce sample noise), where $\bar{n}_{\rm gal}^{\rm LRG} = 7.8 \times 10^{-5} h^3{\rm Mpc}^{-3}$ as in Zhai et al. (2017). This is done by matching $n_{\rm gal}^{\rm HOD}$ to $\bar{n}_{\rm gal}^{\rm target} = 15.6 \times 10^{-5} h^3{\rm Mpc}^{-3}$ in Eq. 7. We therefore have 3 free parameters in our ELG HOD model, $\theta \subset (A_s, M_{\rm min}, \alpha)$, and 5 parameters $\theta \subset (M_1, \sigma_{\rm M}, \alpha, \kappa, f_{\rm ic})$ with the LRG HOD model.

In recent works (Avila et al. 2020; Rocher et al. 2023a,b; Yuan et al. 2022), it is common to enhance this description with extra parameters, that describe deviation from satellite NFW profile and from the velocity assignment of Eq. 6 in order to properly model the anisotropic correlation function. Since we are only interested in describing projected quantities (see Section 3.3), we do not introduce these additional parameters. Indeed, the GGL signal, being a real space quantity, is insensitive to galaxy velocities, while redshift space distortion has a subdominant impact on projected clustering (van den Bosch et al. 2013; Rocher et al. 2023a).

Eventually, the HOD model described in this section is then calibrated against data to provide realistic mocks with similar clustering properties as the observations.

3.2. Extended HOD including assembly bias

Since we aim to provide realistic galaxy mocks that are in agreement with observations, we extend our HOD framework with additional parameters to take into account secondary halo properties. This is motivated by studies such as Leauthaud et al. (2017) which found strong differences between the observed galaxy-galaxy lensing signal of CMASS galaxies, and the one measured in the mock catalogs using the standard HOD model described above. Here, we follow the formalism described in Xu et al. (2021); Yuan et al. (2022). The standard HOD framework is extended with four additional parameters, two depending on halo properties, specifically the concentration and shape of the halo, and two external properties namely the local density and local shear (local density anisotropies). Let us now define in practice how we estimate those additional parameters.

First, the halo concentration c and shape e are defined as

$$c = \frac{r_{\rm vir}}{r_s}, \quad e = \frac{1 - b/a}{1 + b/a}, \tag{9}$$

with $r_{\rm vir}$ the virial radius, r_s the scale radius of the density profile, and b/a the ratio between the major and semi-major axis of the dark matter halo ellipsoid. These halo properties are directly extracted from the halo catalogs.

To specify halos environment within the simulation, we employ two different techniques. The first one consists of finding for each halo all neighboring halos (including subhalos) beyond its virial radius but within $r_{\rm max} = 5h^{-1}$Mpc of the halo center. This is done with the Python package scipy-Kdtree [7] which keeps track of pairs of halo per each object. The $r_{\rm max}$ scale was chosen in Yuan et al. (2021) as values of $r_{\rm max}$ in the range $\approx 4 - 6h^{-1}$Mpc were found to accurately reproduce the clustering of CMASS galaxies. After summing up the mass of all these neighboring halos, the environmental parameter $f_{\rm env}$, can then be computed as

$$f_{\rm env} = M_{\rm env}/\bar{M}_{\rm env}(M) - 1, \tag{10}$$

where $\bar{M}_{\rm env}(M)$ corresponds to the mean environment mass within a given halo mass bin dM. This environmental parameter hence corresponds to the halo over-density in an annulus around each halo.

The second method consists in painting the dark-matter particle field onto a mesh in order to evaluate the over-density field $\delta(\mathbf{x})$ Paranjape et al. (2018); Delgado et al. (2022); Hadzhiyska

[7] https://scipy.org/

et al. (2023b). The field δ is then smoothed with a Gaussian kernel of radius R and interpolated at the position of each halo. This is done with the Python package NBODYKIT[8]. From the smoothed density field, we further define the environmental shear, which quantifies the amount of local density anisotropy. First, we compute the tidal tensor field as

$$T_{ij}(\mathbf{x}) = \partial_i \partial_j \psi_R(\mathbf{x}), \tag{11}$$

where the normalized gravitational potential $\psi_R(\mathbf{x})$ obeys Poisson equation

$$\nabla^2 \psi_R(\mathbf{x}) = \delta_R(\mathbf{x}). \tag{12}$$

This is easily done in Fourier space, the expression for the tidal tensor being equal to

$$T_{ij}(\mathbf{x}) = \mathrm{FT}\left\{\left(k_i k_j / k^2\right) \delta_R(\mathbf{k})\right\}, \tag{13}$$

from which we determine the tidal shear q_R^2

$$q_R^2 \equiv \frac{1}{2}\left|(\lambda_2 - \lambda_1)^2 + (\lambda_3 - \lambda_1)^2 + (\lambda_3 - \lambda_2)^2\right|. \tag{14}$$

where the λ correspond to the eigenvalues of T_{ij}. Different definitions of the smoothing radius have been used in the literature. Delgado et al. (2022) used a Top Hat filter of fixed size of $5h^{-1}$Mpc [9]. Hadzhiyska et al. (2023b) estimated the density field with a Gaussian filter at different smoothing scales, and selected for each halo the closest smoothing scale to the virial radius of the halo rounding up. In Paranjape et al. (2018), they found that the correlation between large-scale linear bias and local shear was maximum at scales of 4-6 times the virial radii. We therefore adopt two definitions for our smoothing scale. The first one is a constant smoothing of size $\mathrm{R} = 5/\sqrt{5} \approx 2.25h^{-1}$Mpc from which we determine q_R^2 for each halo. Our second definition uses a smoothing scale of $2.25r_{\rm vir}$ for each halo, from which we determine local density and local shear. The density field is smoothed from 1 to 5 h^{-1}Mpc with a step $\mathrm{d}x = 0.25h^{-1}$Mpc and interpolated at the position of each halo to determine $\delta_R(\mathbf{x})$ for $R = 2.25r_{\rm vir}$.

These secondary halo properties are used to modify the two mass parameters of the HOD model, $M_{\rm min}$ and M_1 as

$$\log M_{\rm min}^{\rm mod} = \log M_{\rm min} + A_{\rm cent} f_A + B_{\rm cent} f_B \tag{15}$$

$$\log M_1^{\rm mod} = \log M_1 + A_{\rm sat} f_A + B_{\rm sat} f_B \tag{16}$$

where f_A and f_B correspond to the distributions of internal (either concentration or shape) and environmental (either shear or density) properties of the halos. This type of effective description which modifies central and satellite occupation based on secondary properties was also proposed in Hadzhiyska et al. (2023a). The authors found that the simulation-predicted quantities $N_{\rm cent}(M, f_A, f_B)$ and $N_{\rm sat}(M, f_A, f_B)$, for each mass bin are well approximated by hyper-plane surfaces which are nearly constant over mass. These distributions are normalized in the range $[-1, 1]$ within narrow halo mass bins of dex $\mathrm{d}M = 0.1$. This is done by first taking the natural logarithm of each quantity ($1+\delta$ for the density) before scaling and normalizing it around zero. Taking the logarithm of each quantity first has the advantage of keeping the shape of their distributions intact, as we are much less sensitive to outliers that squeeze the distribution after normalization. We present these distributions in Appendix A. The high volume of UCHUU enables us to probe assembly bias and environmental dependencies on halo occupation up to very high halo masses, $M = 14.8h^{-1}$Mpc, with more than 100 halos in the mass bin $[14.7, 14.8]\ h^{-1}$Mpc. Extras parameters in Eq. 15 are fixed to zero above this threshold as higher masses bins are not populated enough for a robust statistical description. Our extended HOD framework has therefore 7 parameters for an ELG HOD $\theta \subset (A_s, M_{\rm min}, \alpha, A_{\rm cent}, B_{\rm cent}, A_{\rm sat}, B_{\rm sat})$, and 9 parameters $\theta \subset (M_1, \sigma_{\rm M}, \alpha, \kappa, f_{\rm ic}, A_{\rm cent}, B_{\rm cent}, A_{\rm sat}, B_{\rm sat})$ for a LRG HOD.
In the following, we will test different combinations in Eqs. 15-16, between internal halo properties f_A (concentration and shape) and local environment f_B (density and shear).

3.3. Fitting Procedure

Our fitting method uses the HOD model described above to populate dark matter halos with galaxies for which clustering and lensing statistics are measured and compared to the ones of the data. Here, we are interested in the projected correlation function, which is almost insensitive to the effect of redshift space distortions at small scales. We compute the galaxy anisotropic two-point function $\xi\left(r_p, \pi\right)$, as a function of galaxy pair separation along (π) and transverse to the line-of-sight (los)(r_p). Integrating over the line of sight yields

$$w_p\left(r_p\right) = 2\int_0^{\pi_{\rm max}} \xi\left(r_p, \pi\right) \mathrm{d}\pi\ . \tag{17}$$

We use the DESI wrapper, PYCORR[10] around the CORRFUNC package Sinha & Garrison (2020) to compute $\xi\left(r_p, \pi\right)$ with the natural estimator, which normalizes galaxy pairs counts to analytical estimation of random pairs in a periodic box. For the data, we use the minimum variance Landy & Szalay (1993) estimator

$$\xi(r_p, \pi) = \frac{DD(r_p, \pi) - 2DR(r_p, \pi) + RR(r_p, \pi)}{RR(r_p, \pi)}, \tag{18}$$

with R a random catalog with the same angular and redshift distribution as the data. We apply for each galaxy and random catalogs the total weight $w = w_{\rm noz} w_{\rm cp} w_{\rm syst} w_{\rm FKP}$, see Ross et al. (2020) for an in-depth definition of each individual weight. Briefly, $w_{\rm noz}$ corrects for redshift failures, $w_{\rm syst}$ corrects for spurious fluctuation of target density due to observational systematic effects, FKP weights Feldman et al. (1994) are optimal weights used to reduce variance in clustering measurements and close pair weights $w_{\rm cp}$ correct for fiber collisions. Correcting for fiber collision is an important matter in our case, as we want to extract information from the one-halo regime, to constrain the satellite distribution in our HOD model. Traditionally, these weights are applied by up-weighting observed galaxies in a collision group.
While this method provides a good correction scheme at large scales, which is enough for the study of galaxy anisotropic clustering, it fails on smaller scales, and we therefore need another method to correct for fiber collisions. In particular, we use here the pairwise-inverse probability (PIP) and angular upweighting (ANG) weights described in Mohammad et al. (2020). These weights were shown to properly account for fiber collisions at small separations. However, PIP weights were only produced for eBOSS galaxies, and we therefore only apply ANG weights to the combined sample (including the cp weights), as they are providing sufficient correction below 1 h^{-1}Mpc. For $w_p(r_p)$, we use

[8] https://nbodykit.readthedocs.io/en/latest/
[9] The equivalent between of a Top Hat filter of size R, is a Gaussian filter of size $R/\sqrt{5}$, see Paranjape et al. (2018).
[10] https://py2pcf.readthedocs.io/en/latest/index.html

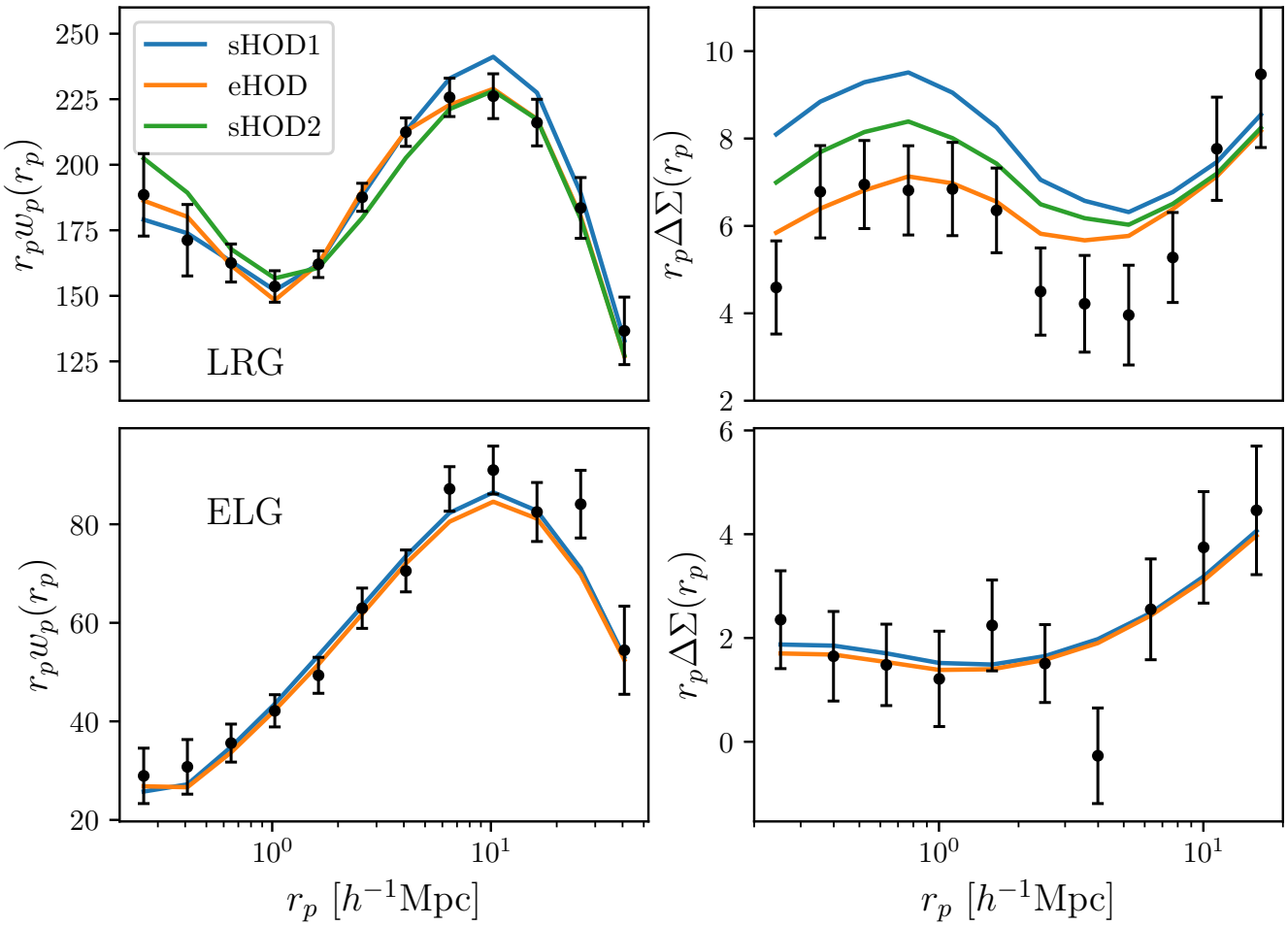

Fig. 2. Best fit HOD models for LRG (top) and ELG (bottom) panels. The subscript s/e denotes the standard and extended HOD model. The points with errors bar correspond to the measured w_p (left) and $\Delta\Sigma$ (right) signal. The blue lines (sHOD1) correspond to a fit to the projected correlation function w_p only, while the green (SHOD2) and orange lines (eHOD) correspond to standard and extended HOD fits including the measurement of $\Delta\Sigma$ in the fitting pipeline. The extended HOD model provides a good description of both the clustering and lensing signal.

13 logarithmic bins in r_p between 0.2 and 50 h^{-1}Mpc, with $\pi_{\max}$ = 80 h^{-1}Mpc. In addition to the clustering, we also fit the GGL signal, similarly to the work of Contreras et al. (2023), observed with eBOSS lenses around DES sources in order to further constrain the assembly bias parameters. The measured GGL differential excess surface density is defined as

$$\Delta\Sigma_{\rm gm}(R) = \bar{\Sigma}_{\rm gm}(R) - \Sigma_{\rm gm}(R), \tag{19}$$

where the mean projected surface density is defined as

$$\bar{\Sigma}_{\rm gm}(R) = \frac{2}{R^2}\int_0^R \Sigma_{\rm gm}(r) r {\rm d}r. \tag{20}$$

The projected surface density $\Sigma_{\rm gm}(r)$ can be written as a function of the galaxy-matter cross-correlation as (Guzik & Seljak 2001; de la Torre et al. 2017; Jullo et al. 2019)

$$\Sigma_{\rm gm}(R) = \Omega_{\rm m}\rho_{\rm c}\int_{-\infty}^{\infty} \xi_{\rm gm}\left(\sqrt{R^2+\chi^2}\right){\rm d}\chi, \tag{21}$$

with the mean matter density $\rho_{\rm m} = \Omega_{\rm m}\rho_{\rm c} = 3\Omega_{\rm m0}H_0^2/8\pi G$ constant in redshift in comoving coordinates. The galaxy-matter correlation function is measured in the range $[5\times10^{-3}, 150]h^{-1}$Mpc with 41 logarithm bins, by cross-correlating galaxy mock catalogs with the dark-matter particles of the simulation, subsampled by a factor of 2000. We then estimate $\Sigma_{\rm gm}(r)$ by integrating up to 100 h^{-1}Mpc. The estimation of $\Delta\Sigma$ is then compared to the one observed in the data, which is computed as (Jullo et al. 2019)

$$\Delta\Sigma_{\rm gm}\left(r_{\rm p}\right) = \frac{\sum_{\rm l,s}^{N_{\rm ls}} \Sigma_{\rm cr}\left(z_{\rm l}, z_{\rm s}\right) w_{\rm l,s}\epsilon_+\left(r_{\rm p}\right)}{\sum_{\rm l,s}^{N_{\rm l,s}} w_{\rm l,s}} - \frac{\sum_{\rm r,s}^{N_{\rm r,s}} \Sigma_{\rm cr}\left(z_{\rm r}, z_{\rm s}\right) w_{\rm r,s}\epsilon_+\left(r_{\rm p}\right)}{\sum_{\rm r,s}^{N_{\rm r,s}} w_{\rm r,s}}, \tag{22}$$

where $z_{\rm s}$ and $z_{\rm l}$ correspond to the source and lens redshift respectively. The subtraction around random lenses reduces the contribution of lensing systematics and leads to a reduced variance on small scales (Singh et al. 2017). The critical density $\Sigma_{\rm cr}$ is defined as

$$\Sigma_{\rm cr}\left(z_{\rm l}, z_{\rm s}\right) = \frac{c^2}{4\pi G\left(1+z_{\rm s}\right)^2}\frac{D_{\rm S}}{D_{\rm LS}D_{\rm L}}, \tag{23}$$

where $D_{\rm S}, D_{\rm LS}, D_{\rm L}$ are the observer-source, lens-source, and observer-lens angular diameter distances. We measured $\Delta\Sigma(r_p)$

Table 1. Best fit parameters for all fit configurations. The baseline eHOD configuration in this work corresponds to the case where the concentration and an annular definition of δ are used as additional parameters.

Method	sHOD1	sHOD2	eHOD $\delta_{\rm vir}$,c	eHOD $T^{ij}_{\rm vir}$,c	eHOD $T^{ij}_{\rm vir}$, b/a	eHOD $\delta_{\rm annulus}$,c	eHOD $\delta_{\rm annulus}$,b/a	eHOD $T^{ij}_{\rm 5Mpc}$,c
$M_{\rm min}$	13.24	12.81	12.95	13.23	12.93	13.03	13.11	13.26
$M_{\rm cut}$	14.18	13.84	13.91	13.80	13.72	13.76	13.76	13.87
σ	0.54	0.11	0.44	0.74	0.53	0.58	0.67	0.72
α	1.20	1.36	0.91	0.93	1.05	0.93	0.90	0.97
κ	0.39	0.36	0.64	0.48	0.39	0.56	0.51	0.46
$f_{\rm ic}$	0.64	0.28	0.32	0.38	0.25	0.30	0.32	0.43
$A_{\rm cent}$			0.11	- 0.01	0.10	0.18	0.20	0.15
$B_{\rm cent}$			- 0.27	-0.27	-0.24	-0.25	-0.25	-0.18
$A_{\rm sat}$			0.01	-0.15	-0.05	-0.06	-0.08	-0.02
$B_{\rm sat}$			0.00	-0.04	-0.10	0.01	0.02	0.01
$f_{\rm sat}$%	8.0	14.1	13.3	12.8	16.6	14.8	14.6	12.0
χ^2_r	1.05	2.45	1.43	1.11	1.06	1.08	1.09	1.62

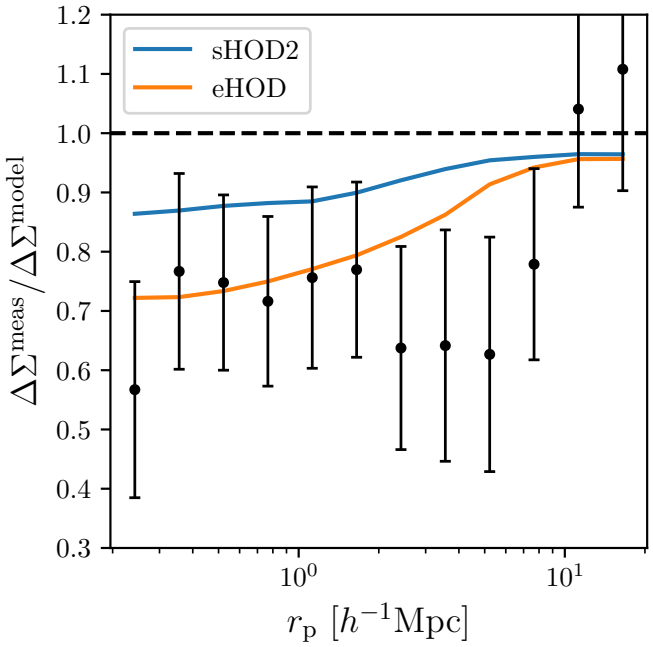

Fig. 3. Difference between the observed galaxy-galaxy lensing signal (points with error bars) in eBOSS LRG with the prediction of a standard HOD (sHOD) determined by fitting the projected correlation function $w_{\rm p}$ only. The continuous line in blue represents an sHOD fit to $w_{\rm p} + \Delta\Sigma$ while the orange line corresponds to the extended HOD model.

with 13 (11 for the ELGs) logarithmic bins in the range between 0.2 and 20 h^{-1}Mpc with the python package DSIGMA [11] (Lange & Huang 2022). We used METACALIBRATION shape measurements and corrected for potential systematic effects induced by this method. We used for photometric redshifts of source galaxies the ones estimated with the Directional Neighbourhood Fitting (DNF) algorithm (De Vicente et al. 2016). Errors in photo-z measurements can lead to bias in the estimation of $\Delta\Sigma$. We matched DES galaxies with galaxies observed spectroscopically with BOSS and eBOSS. While these samples are not representative of DES galaxies, they can however be used to obtain a rough estimate of the bias introduced by the photo-z. We found that photo-z leads to an underestimation of the observed signal by approximately 3 − 5% at all scales. Since this effect is subdominant compared to the signal-to-noise (S/N) of our measurements, we do not include it. We also do not include any Boost factor correction Miyatake et al. (2015) as we have found that it has nearly no impact on our measurements, even at the smallest scale.

We then compare the clustering and lensing signal to the observed ones in eBOSS by mean of a likelihood analysis

$$-2\ln\mathcal{L}(\theta) = \sum_{i,j}^{N_p} \mathbf{\Delta}_i(\theta)\hat{\Psi}_{ij}\mathbf{\Delta}_j(\theta), \tag{24}$$

where θ is the set of HOD parameters, Δ is the data-model difference vector, N_p is the total number of data points and $\hat{\Psi}_{ij}$ is the precision matrix, the inverse of the covariance matrix. The covariance matrix is determined by jackknife (JK) subsampling with $N = 100$ patches. Since all ELGs are within the DES footprint, we can use the same JK regions for the clustering and lensing measurements in order to estimate the cross-correlation between $w_{\rm p}$ and $\Delta\Sigma$. For the LRGs, only about a third of the sample is within DES footprint. We decided to use the whole sample to estimate $w_{\rm p}$, to get a sufficient S/N measurement. Since $\Delta\Sigma$ is only measured within the overlapping footprint, the cross-covariance cannot be estimated through the jackknife method, and we assume it to be zero. We have checked with the ELG sample that this assumption does not change significantly the results: the jackknife cross-covariance is mainly dominated by noise.

Our fitting procedure corresponds to the one described in (Rocher et al. 2023a,b). The parameter space is first sampled with a quasi-uniform grid, which serves as a training sample for a Gaussian process regressor. Here we use the one implemented in SKLEARN[12]. For each sampled point in the parameter space, the clustering and lensing signals are averaged over ten realizations to reduce stochastic noise from halo occupation. We then repeat iteratively the following procedure: we use an MCMC sampler, here `emcee`[13] (Foreman-Mackey & al. 2013), to sample our parameter space given our surrogate model generated by the Gaussian process regressor. We then randomly pick 3 points within our final MCMC chains from which we compute the likelihood. We then add these points to the initial training sample and re-train our Gaussian process. In Rocher et al. (2023a), they

[11] https://dsigma.readthedocs.io/en/stable/index.html

[12] https://scikit-learn.org/stable/

[13] https://emcee.readthedocs.io/

found that iterating up to $N = 800$ from a training sample of $N_{\rm ini} \approx 1000$ was enough to get robust contours. These findings are highly dependent on the parameter space we want to explore and the dynamic range of the likelihood. While we do find the same conclusion for an ELG-HOD, the degree of degeneracy of a LRG-HOD, calls for a higher number of iterations. We therefore use $N_{\rm ini} = 1000$ (resp. 2000) and $N = 1000$ (resp. 2000) for ELG (resp. LRG) fit respectively.

4. Results

4.1. Luminous Red Galaxies

As a baseline configuration for the extended HOD (eHOD) model, we use the concentration as a proxy of assembly bias and the local density, described as an annular definition of the density field δ, following the formalism of Yuan et al. (2022). We present in Fig. 2 the best-fit results obtained with the standard (s) and extended (e) HOD models. The standard model sHOD2 and the extended model both fit the combined w_p, $\Delta\Sigma$ data vector, while the sHOD1 model only fits w_p. Best-fit values are given in Table 1 while full posterior distributions and prior information can be found in Fig. 4, Fig. 6 and in Appendix A. For the sHOD1, we can see that the model over-predicts the signal at scales around 10 h^{-1}Mpc while obtaining a satellite fraction of 8%, slightly lower than those obtained in Alam et al. (2020); Yuan et al. (2022) and 50% below the best-fit value of Zhai et al. (2017). We found that this difference was due to noise in the jackknife covariance estimate of $w_{\rm p}$. When we performed a fit only considering the diagonal elements, we obtained a satellite fraction of 11%, the model matches well the peak at 10 h^{-1}Mpc but in turn provides a poor fit at scales $1-3h^{-1}$Mpc, similarly to the observed fit for the sHOD2 case, an issue in the modeling also reported in Zhai et al. (2017).

The introduction of secondary HOD parameters enables us to provide a good description of the small and intermediate scales of $w_{\rm p}$, while also yielding a more accurate description of $\Delta\Sigma$, decreasing the reduced chi-square from $\chi^2_r = 2.45$ down to $\chi^2_r = 1.08$, an improvement mainly due to a better description of $\Delta\Sigma$ on the scales below 3 h^{-1} Mpc. This is illustrated in Fig. 3, which displays the ratio of $\Delta\Sigma$ measurements in the data and as measured in the mocks generated by sHOD2 and eHOD models (calibrated on both w_p and $\Delta\Sigma$ measurements), compared to the $\Delta\Sigma$ measurement as measured in the mock generated by the sHOD1 (calibrated only against the projected clustering w_p). The standard model dashed line) over-predicts the GGL signal by $30-40\%$ at scales below $10h^{-1}$ Mpc, while a combined fit (blue solid line) poorly alleviates this issue, with a slight improvement of only $\approx 5-15\%$. By considering additional parameters describing internal and external properties of the halos (orange solid line), the modeling of the clustering and lensing signals at small scales is significantly improved compared to using the standard HOD model, with the main improvements coming from the dependence on the local density environment (Yuan et al. 2021; Hadzhiyska et al. 2023a). On scales between $2-10h^{-1}$ Mpc, we observe an improvement of the extended model of 20% while it matches the data at lower scales, providing a better goodness of fit to the observed lensing signal. Thus, adding secondary halo parameters to the HOD models helps to reproduce the lensing signal at small scales, and gives a partial answer to the 'lensing is low' problem. Indeed, baryonic feedback, not considered in this work, smoothes out the matter distribution, also contributing to a decrease of $\Delta\Sigma$ (Amodeo et al. 2021). This better description

Table 2. Posterior and best-fit results for the ELG HOD. We show here the 16%, 50%, and 84% percentiles of the posterior distribution. The first column presents the prior ranges for each parameter. The two other columns show the 16%, 50% and 84% percentiles of the posterior distributions, with the best-fit values in parenthesis defined as the minimum of the chi^2 of the MCMC chains, for both sHOD and eHOD models. The initial value of A_c is fixed to 0.01 meaning that A_s is allowed to vary in the range $[0.2,20]A_c$. Other columns present the rescaled values of A_c and A_s in order to match the ELG number density. For the sHOD, we show in parenthesis the reduced chi-square including $\Delta\Sigma$.

Method	sHOD	eHOD
A_s [0.002,0.02]	$0.017^{+0.006}_{-0.005}$(0.015)	$0.017^{+0.006}_{-0.005}$(0.015)
$M_{\rm min}$ [11.5,12.5]	$12.19^{+0.10}_{-0.12}$ (12.28)	$12.11^{+0.17}_{-0.19}$(12.21)
σ [0.1,1.2]	$0.52^{+0.23}_{-0.23}$(0.41)	$0.42^{+0.22}_{-0.23}$(0.28)
$A_{\rm cent}$ [-0.5,0.5]		$-0.01^{+0.19}_{-0.22}$ (-0.08)
$B_{\rm cent}$ [-0.5,0.5]		$-0.09^{+0.11}_{-0.19}$ (-0.04)
$A_{\rm sat}$ [-0.5,0.5]		$-0.04^{+0.19}_{-0.18}$ (0.02)
$B_{\rm cent}$ [-0.5,0.5]		$-0.04^{+0.20}_{-0.12}$(-0.08)
A_c [0.01]	$0.034^{+0.012}_{-0.011}$ (0.047)	$0.028^{+0.017}_{-0.012}$ (0.038)
$f_{\rm sat}$%	10.4	13.3
χ^2_r	1.13(1.22)	1.40

can be achieved in the model by lowering the mass threshold above which central and satellite galaxies occupy dark matter halos. To maintain a similar galaxy bias, galaxies in lower mass halos should therefore be located in higher density peaks. This is what we obtain in the posterior distributions of Fig. 6. A negative value of $B_{\rm cent}$ implies that halos located in denser environments are more likely to be populated, as a decrease of $M^{\rm mod}_{\rm min}$ in Eq. 15 increases the probability of populating such halos, a result also found for CMASS galaxies (Yuan et al. 2022) and in hydrodynamical simulations (Artale et al. 2018).

Our results also suggest that central LRGs preferentially populate halos of low concentration, as we obtained a positive best-fit value for $A_{\rm cent} = 0.18$ (see Table 1), while Yuan et al. (2022) find an opposite trend for the CMASS sample when fitting $\xi(r_\perp, r_\parallel)$. In recent studies, halo occupation was found to be strongly dependent on formation time (Zehavi et al. 2018; Artale et al. 2018), and concentration (Bose et al. 2019; Delgado et al. 2022), as both are tightly linked to one another Wechsler et al. (2002). These results show that, at low mass, central galaxies are more likely to be hosted by more concentrated, early-formed halos, this trend reverses at higher mass due to the contributions of satellite galaxies. For the satellite occupation, there seems to be a slight correlation with concentration, with satellite galaxies most likely to be hosted in more concentrated halos. This is in disagreement with the results of Bose et al. (2019): at all halos mass, concentrated halos have fewer satellite galaxies, as they had more time to merge with the central. However, given our measurements, the only parameter that is well constrained is $B_{\rm cent}$, being non-zero at the 3σ level, while all the other parameters are consistent with zero at the 1σ level. Similarly to Yuan et al. (2022), we also find that including secondary assembly bias parameters (or just fitting $\Delta\Sigma$) does increase the satellite fraction. This can be explained by the fact that central-particle correlation is dominant on small scales (Leauthaud et al. 2017), which causes the fit to naturally provide higher $f_{\rm sat}$ values. For completeness, we present the full posterior distribution of the extended HOD in Appendix B.

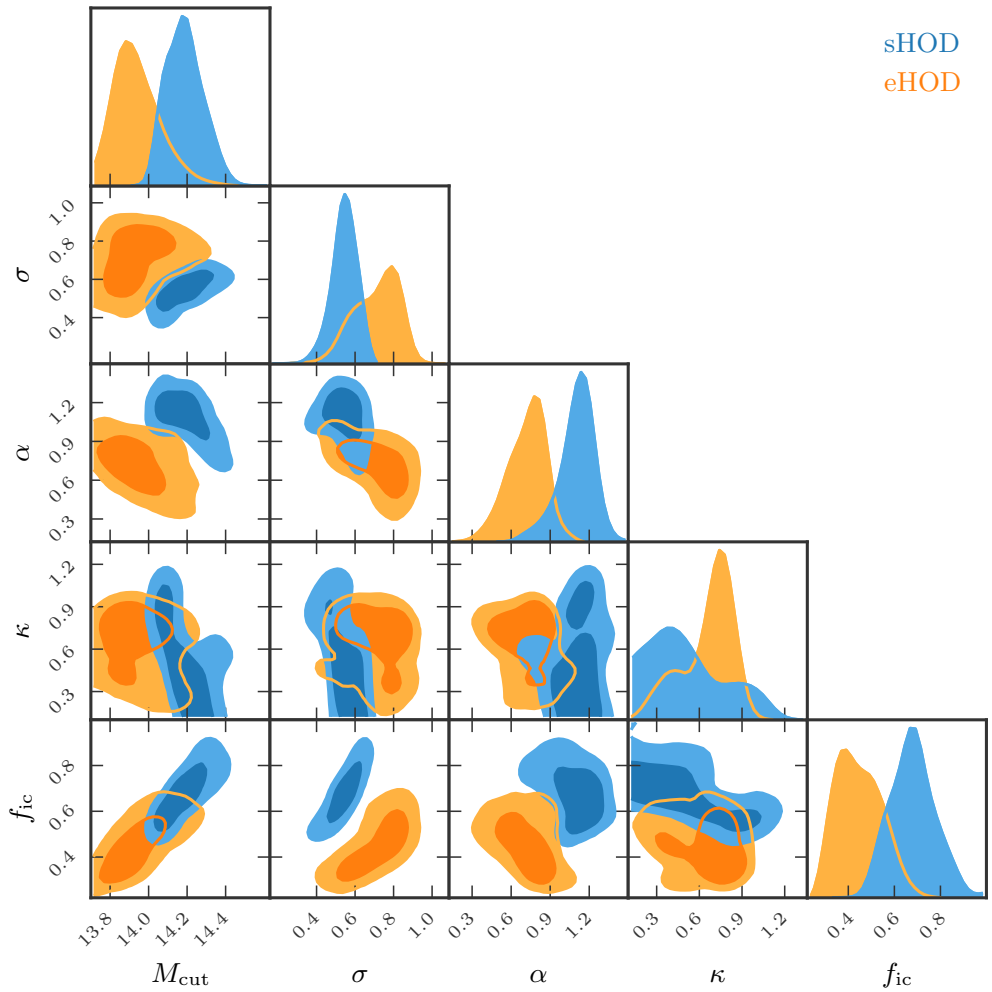

Fig. 4. Full posterior distribution of the sHOD (in blue) and eHOD (in orange) in our baseline configuration (including concentration and annular density as additional parameters) for the LRG sample. We can observe that the extra parameters contained in the eHOD model have a strong impact on the posterior distributions

4.2. Emission Line Galaxies

The best-fit results for ELGs are represented on the bottom panel of Fig. 2 while the posterior contours are shown in Fig. 5 with the corresponding constraints in Table 2. We do not observe significant differences between the standard and the extended HOD framework, when either considering the best-fit model or the full posterior distribution. Both models provide similar goodness of fit and reproduce the observations, with a $\chi^2 = 23.1(22.0)$ for the sHOD and eHOD models respectively, resulting in a worst reduced χ^2 for the eHOD, as the four additional free parameters do not significantly improve the fit. From the posterior distributions in Fig. 6, we can see that all the secondary parameters are consistent with zero at the 1σ level. In the literature, it was shown that ELGs are preferentially located in higher dense regions than average (Artale et al. 2018; Delgado et al. 2022) and in more concentrated halos (Zehavi et al. 2018; Artale et al. 2018; Rocher et al. 2023b). This is in agreement with our negative best-fit values for $A_{\rm cent}$ and $B_{\rm cent}$, even though we cannot derive more decisive conclusions from our analysis.

In addition, we find a satellite fraction of 10.4% for the sHOD, significantly lower than that reported in Avila et al. (2020), $f_{\rm sat} = 22\%$. However, for ELGs the mean satellite occupation is independent of that of the centrals, which means that an ELG satellite can orbit around a non-ELG central. Thus, it only contributes to the two-halo term of the ELG correlation function, affecting the amplitude of the linear bias of the galaxy sample. The lower satellite fraction (compared to Avila et al. (2020)) can be explained by the fact that our mean value for $M_{\rm min} = 12.19$ is larger than that reported in Avila et al. (2020), $M_{\rm min} = 11.71$, and that we allow α to vary (resulting in a different satellite occupation over halo mass). In Rocher et al. (2023b), when they include strict conformity, i.e. satellite galaxies can only populate halo in the presence of a central ELG, they reported a very small fraction of satellite of $\approx 2-3\%$. For DESI ELGs, strict conformity is needed in order to provide a positive value for α, which is predicted by semi-analytical models Gonzalez-Perez et al. (2018, 2020).

4.3. Dependence on parameter choices

We present in Fig. 7 multiple eHOD best-fit models for different configurations. We tested as environment proxy either the local shear or the local density, combined with internal properties, i.e. concentration or shape of the dark matter halos. As a reminder, the baseline configuration corresponds to the case where we apply to Eq. 15, f_A = concentration and $f_B = \delta_{\rm annulus}$, with $\delta_{\rm annulus}$ calculated by counting neighboring halos. Estimat-

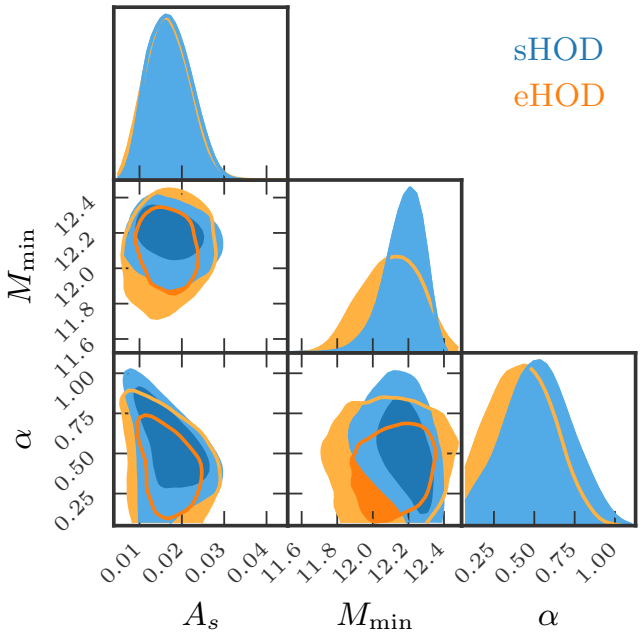

Fig. 5. Full posterior distribution for the sHOD (in blue) and eHOD (in orange) models. Taking into account assembly bias parameters has little impact on halo occupation.

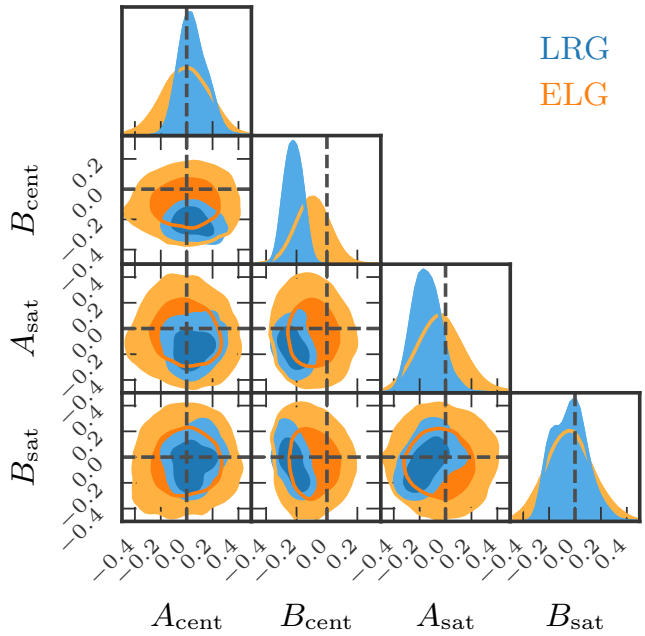

Fig. 6. Full posteriors distributions of the assembly bias and environmental parameters in our baseline configuration (see Section 4.1. For the LRGs, the negative value of $B_{\rm cent}$, detected at 3σ, indicates that LRG centrals most likely occupy a dense environment. A negative value of $A_{\rm sat}$ at the 1σ level, indicates that satellite galaxies could be preferentially located in more concentrated halos than average. For the ELGs, all parameters are consistent with zero at the 1σ level, suggesting that ELGs do not exhibit significant assembly bias or environmental dependencies.

ing the local density field with this definition provides a better description of the observed LRG signal, with a $\chi^2_{\rm r}$ improved by $\Delta\chi^2_{\rm r}$ = -1.39 compared to the case where the density field is estimated on a mesh with a Gaussian smoothing kernel of size = $2.25 r_{\rm vir}$, with $\Delta\chi^2_{\rm r}$ = -0.98. These results are in agreement with those of Yuan et al. (2022) who found that estimating the density field with a Gaussian kernel of a fixed size of 3 h^{-1} Mpc yields no significant improvement compared to the sHOD. Using instead of the concentration, the shape of the dark-matter halo as a proxy for assembly bias provides nearly indistinguishable best-fit model results. The resulting constraints, presented in Table 1 and in Appendix B, suggest that red galaxies preferentially occupy more ellipsoidal halos than average, as the best-fitting value $A_{\rm cent}$ is positive when either considering shear or density as external property. The two extra configurations including the local density are in agreement with the baseline configuration, in favor of a negative value for $B_{\rm cent}$, such that galaxies preferentially occupy more dense environment.

Using the local shear instead of the local density also provides good $\chi^2_{\rm r}$ values when the tidal tensor is computed with a Gaussian kernel of size 2.25 $r_{\rm vir}$. The best $\chi^2_{\rm r}$ value is obtained when combining the local shear with the halo shape, $\chi^2_{\rm r}$ = 1.06. It better reproduces the small scales of $\Delta\Sigma$ but gives a slightly worse fit around the peak of $w_{\rm p}$ observed at $10h^{-1}$ Mpc. Surprisingly, using the local shear with a fixed Gaussian smoothing of size 2.25 h^{-1}Mpc provides the worst $\chi_{\rm r^2}$ value out of all configurations. This demonstrates that local density anisotropy should be considered at halo scales to retrieve the maximum amount of information, as explained in Paranjape et al. (2018). Our results suggest that galaxies preferentially occupy more anisotropic environments, in agreement with Delgado et al. (2022) that found this correlation for low-mass halos, and with DESI ELG Rocher et al. (2023b).

4.4. Cosmological fit

In current analysis, perturbative and semi-analytical models that aim at extracting cosmological information from the observations, are tested intensively against galaxy mock catalogs. In eBOSS analysis, different HOD models were used in order to quantify modeling systematics of redshift-space distortions models (Rossi et al. 2021; Alam et al. 2021). The authors found consistent results amongst a large variety of HOD models which validated the accuracy of the theoretical models. In the work of de la Torre et al. (2017); Jullo et al. (2019), the measured anisotropic clustering $\xi(r_\perp, r_\parallel)$ at an effective redshift z, is combined at the likelihood level with lensing measurements of $\Delta\Sigma$ in order to provide direct constraints on $f(z)$, the growth rate of structure, $\sigma_8(z)$ the amplitude of matter fluctuation and $\Omega_{\rm m}$, the present time matter density. These studies extract information at scales below 3 h^{-1}Mpc where our HOD models, as illustrated in Fig. 2, show significant differences which might introduce biases when quantifying modeling systematics of $\Delta\Sigma$.

Here, we perform the cosmological analysis of galaxy catalogs populated with the standard and extended HOD framework in order to test the accuracy of a combined analysis of galaxy clustering and galaxy-galaxy lensing. We consider as a redshift space model the Taruya et al. (2010) model extended to non-linearly biased tracers, hereafter referred to as TNS, described and robustly tested in detail in Bautista et al. (2021). The expression for the redshift space power spectrum of biased tracers is

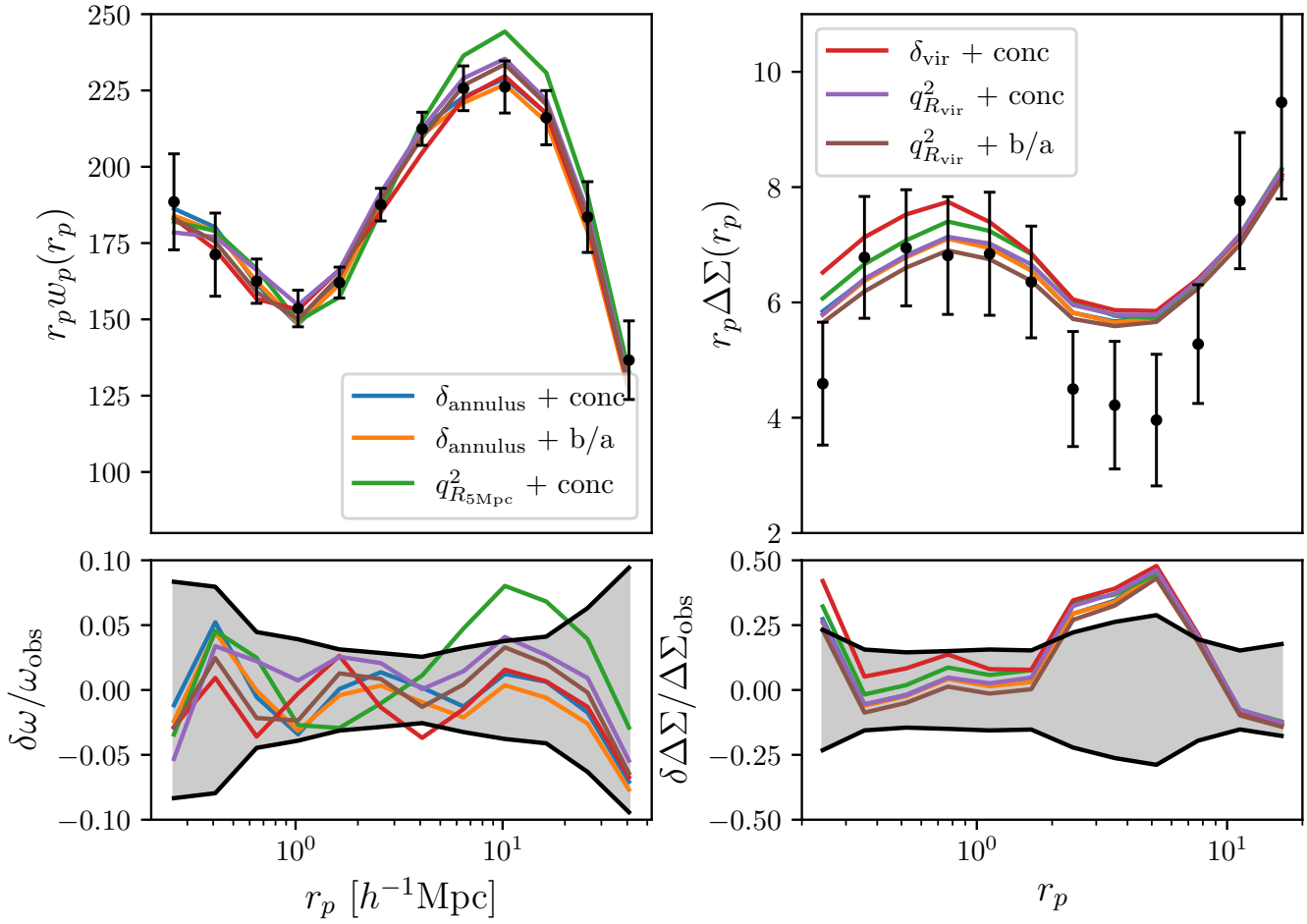

Fig. 7. Result for different combinations of assembly bias (concentration and shape) and local environments (density and shear) parameters. Top panels present the measurements while the bottom panels show the relative differences. All the models provide similar best-fitting clustering predictions, lensing predictions and goodness of fit results, except the shear measured with a Gaussian Kernel of size 2.25 h^{-1}Mpc which has a slightly worse goodness of fit.

given by

$$P^s(k,\nu) = D(k\nu\sigma_v)[P_{\rm gg}(k) + 2\nu^2 f P_{\rm g\theta}(k) + \nu^4 f^2 P_{\theta\theta}(k) + C_A(k,\nu,f,b_1) + C_B(k,\nu,f,b_1)], \quad (25)$$

where k is the wave-vector norm, ν is the cosine angle between the wave-vector and the line of sight, θ is the divergence of the velocity field $\mathbf{v}$ defined as $\theta = -\nabla\cdot\mathbf{v}/(aHf)$, and f is the linear growth rate parameter. $C_A(k,\nu,f)$ and $C_B(k,\nu,f)$ are the two correction terms given in Taruya et al. (2010), which reduce to one-dimensional integrals of the linear matter power spectrum. The damping function $D(k\nu\sigma_v)$ describes the Finger-of-God effect induced by random motions in virialized systems, inducing a damping effect on the power spectra. We adopted a Lorentzian form, $D(k,\nu,\sigma_v) = (1 + k^2\nu^2\sigma_v^2/2)^{-2}$, where σ_v represents an effective pairwise velocity dispersion treated as an extra nuisance parameter. The expressions of the real space quantities $P_{\rm gg}$, $P_{\rm g\theta}$ and $P_{\rm gm}$ are given following the bias expansion formalism of Assassi et al. (2014)

$$\begin{aligned}
P_{\rm gg}(k) =& b_1^2 P_{\delta\delta}(k) + b_2 b_1 I_{\delta^2}(k) + 2 b_1 b_{\mathcal{G}_2} I_{\mathcal{G}_2}(k) \\
&+ 2\left(b_1 b_{\mathcal{G}_2} + \frac{2}{5} b_1 b_{\Gamma_3}\right) F_{\mathcal{G}_2}(k) + \frac{1}{4} b_2^2 I_{\delta^2\delta^2}(k) \\
&+ b_{\mathcal{G}_2}^2 I_{\mathcal{G}_2\mathcal{G}_2}(k) \frac{1}{2} b_2 b_{\mathcal{G}_2} I_{\delta_2\mathcal{G}_2}(k), \\
P_{\rm g\theta}(k) =& b_1 P_{\delta\theta}(k) + \frac{b_2}{2} I_{\delta^2\theta}(k) + b_{\mathcal{G}_2} I_{\mathcal{G}_2\theta}(k) \\
&+ \left(b_{\mathcal{G}_2} + \frac{2}{5} b_{\Gamma_3}\right) F_{\mathcal{G}_2\theta}(k), \\
P_{\rm gm}(k) =& b_1 P_{\delta\delta}(k) + \frac{b_2}{4} I_{\delta^2}(k) + b_{\mathcal{G}_2} I_{\mathcal{G}_2}(k) \\
&+ \left(b_{\mathcal{G}_2} + \frac{2}{5} b_{\Gamma_3}\right) F_{\mathcal{G}_2}(k),
\end{aligned} \quad (26)$$

where each function of k is a 1-loop integral of the linear matter power spectrum $P_{\rm lin}$, which expression can be found in Simonović et al. (2018). The expressions of $I_{\delta^2\theta}(k)$, $I_{\mathcal{G}_2\theta}(k)$ and $F_{\mathcal{G}_2\theta}(k)$ are the same as $I_{\delta^2}(k)$, $I_{\mathcal{G}_2}(k)$ and $F_{\mathcal{G}_2}(k)$ except than the G_2 kernel replaces the F_2 kernel in the integrals. To estimate the velocity divergence power spectra $P_{\theta\theta}$ and $P_{\delta\theta}$, we used Bel et al.

(2019) fitting function, that depends solely on σ_8,

$$P_{\theta\theta}(k) = P_L(k)e^{-k(a_1+a_2k+a_3k^2)},$$
$$P_{\delta\theta}(k) = (P_{\delta\delta}(k)P_{\rm lin}(k))^{\frac{1}{2}}\, e^{-\frac{k}{k_\delta}-bk^6}, \tag{27}$$

with the expressions of each parameter given in Bel et al. (2019); Bautista et al. (2021). The linear matter power spectrum is estimated with CAMB (Lewis & Challinor 2011) [14] while the non-linear power spectrum $P_{\delta\delta}$ is estimated with the latest version of HMCODE (Mead et al. 2021). This is different from Bautista et al. (2021); Paviot et al. (2022) where we used the response function formalism RESPRESSO described in Nishimichi et al. (2017). Indeed the latest version of HMCODE provides a nearly indistinguishable (at least in configuration space) baryon acoustic oscillations damping due to non-linear effects, while also matching at higher k the power spectrum of RESPRESSO.

The TNS correlation function multipole moments are finally determined by performing the Hankel transform of the model power spectrum multipole moments,

$$\xi_\ell^{\rm TNS}(s) = i^\ell \frac{2\ell+1}{2}\int {\rm d}k\frac{k^2}{2\pi^2}j_\ell(ks)\int_{-1}^{1}{\rm d}\nu\, P^s(k,\nu)L_\ell(\nu), \tag{28}$$

where j_ℓ and L_ℓ denote the spherical Bessel function and Legendre polynomials of order ℓ. The Hankel transform, i.e. the outer integral in the above equation, is performed rapidly using FFTLOG algorithm (Hamilton 2000).

Under the limber approximation Limber (1953), the expression of the galaxy-galaxy lensing signal is given by Marian et al. (2015)

$$\Delta\Sigma_{\rm gm}(r_{\rm p}) = \rho_{\rm m}\int\frac{{\rm d}^2k_\perp}{(2\pi)^2}P_{\rm gm}(k_\perp, z_l)\, J_2(k_\perp R), \tag{29}$$

with J_2 the second order Bessel function of first kind and $\rho_{\rm m}$ the comoving matter density. In order to remove the contribution of non-linear modes, we computed the annular differential excess surface density (ASAD) estimator given by (Baldauf et al. 2010)

$$\Upsilon\left(r_{\rm p}, R_0\right) = \Delta\Sigma_{\rm gm}\left(r_{\rm p}\right) - \left(\frac{R_0}{r_{\rm p}}\right)^2\Delta\Sigma_{\rm gm}(R_0), \tag{30}$$

where R_0 corresponds to a cut-off radius, below which scale the signal is damped.

In addition, we parametrize the Alcock-Paczyński (AP) distortions (Alcock & Paczynski 1979) induced by the assumed fiducial cosmology in the measurements via two dilation parameters that scale transverse, $\alpha_\perp$, and radial, $\alpha_\parallel$, separations. These quantities are related to the comoving angular diameter distance, $D_M = (1+z)D_A(z)$, and Hubble distance, $D_H = c/H(z)$, respectively, as

$$\alpha_\perp = \frac{D_M(z_{\rm eff})}{D_M^{\rm fid}(z_{\rm eff})}, \tag{31}$$

$$\alpha_\parallel = \frac{D_H(z_{\rm eff})}{D_H^{\rm fid}(z_{\rm eff})}, \tag{32}$$

where c is the speed of light in the vacuum, and $z_{\rm eff}$ is the effective redshift of the sample. We apply these dilation parameters to the theoretical TNS power spectrum $P^s(k,\nu)$ in Eq. 28, so that $P^s(k,\nu) \to P^s(k',\nu')$ where

$$k' = \frac{k}{\alpha_\perp}\left[1+\nu^2\left(\frac{1}{F_{\rm AP}^2}-1\right)\right]^{1/2}, \tag{33}$$

$$\nu' = \frac{\nu}{F_{\rm AP}}\left[1+\nu^2\left(\frac{1}{F_{\rm AP}^2}-1\right)\right]^{-1/2}, \tag{34}$$

and $F_{\rm AP} = \alpha_\parallel/\alpha_\perp$. $\Delta\Sigma$ and Υ, being projected quantities, are only distorted by the perpendicular distortion factor

$$\Delta\Sigma(R') = \Delta\Sigma(\alpha_\perp R),$$
$$\Upsilon(R') = \Upsilon(\alpha_\perp R). \tag{35}$$

Given a theoretical model for ξ_ℓ and $\Delta\Sigma$ (Υ), we can perform a cosmological analysis assuming a Gaussian likelihood as defined in Eq. 24. Here, our data vector is the combined multipoles galaxy-galaxy lensing signal, with the free parameters of the model $\theta \subset (b_1, b_2, b_{\Gamma_3}, f, \sigma_8, \alpha_\perp, \alpha_\parallel, \sigma_v)$, $b_{\mathcal{G}_2}$ being fixed to Lagrangian prescription as $b_{\mathcal{G}_2} = -2/7(b_1-1)$ (Saito et al. 2014). It is worth noting that, as the TNS corrections and the galaxy biased expansion are functions of the linear power spectrum only, these can be rescaled at each iteration by the normalization σ_8. The non-linear power spectrum evaluated with HMCODE is first sampled into a broad range of σ_8 and then interpolated at any value of σ_8 within our prior, such that the only quantities that we re-computed at each iteration are the real space quantities $P_{\theta\theta}$ and $P_{\delta\theta}$. We will assume as fiducial cosmology the cosmology of the UCHUU simulation, such that AP distortions should be equal to unity.

The covariance matrix is estimated with analytical prescriptions. The Gaussian covariance for the multipoles of the correlation function is given by (Grieb et al. 2016)

$$C^{\xi}_{\ell_1\ell_2}\left(s_i, s_j\right) = \frac{i^{\ell_1+\ell_2}}{2\pi^2}\int_0^\infty k^2\sigma^2_{\ell_1\ell_2}(k)\bar{j}_{\ell_1}(ks_i)\,\bar{j}_{\ell_2}\left(ks_j\right){\rm d}k, \tag{36}$$

where the $\bar{j}_\ell$ correspond to bin-averaged spherical bessel functions,

$$\bar{j}_\ell(ks_i) \equiv \frac{4\pi}{V_{s_i}}\int_{s_i-\Delta s/2}^{s_i+\Delta s/2} s^2 j_\ell(ks){\rm d}s, \tag{37}$$

with V_{si} the volume element of the shell. The expression for $\sigma^2_{\ell_1\ell_2}(k)$ is given by

$$\sigma^2_{\ell_1\ell_2}(k) \equiv \frac{(2\ell_1+1)(2\ell_2+1)}{V_{\rm s}} \times\int_{-1}^{1}\left[P(k,\mu)+\frac{1}{\bar{n}}\right]^2\mathcal{L}_{\ell_1}(\mu)\mathcal{L}_{\ell_2}(\mu){\rm d}\mu, \tag{38}$$

with V_s the volume of the periodic box. The expression for the Gaussian covariance of $\Delta\Sigma$ is given by Marian et al. (2015)

$$\begin{aligned}{\rm Cov}[\widehat{\Delta\Sigma}]\left(R_i, R_j\right) &= \frac{1}{A_s}\left(\rho_m^0\right)^2\int\frac{{\rm d}^2k_\perp}{(2\pi)^2}\bar{J}_2(k_\perp R_i)\,\bar{J}_2\left(k_\perp R_j\right)\\ &\times\left\{\left[P_{\delta\delta}(k_\perp)+\frac{1}{\bar{n}_p}\right]\left[P_{\rm gg}(k_\perp)+\frac{1}{\bar{n}_g}\right]+P^2_{\rm gm}(k_\perp)\right\},\end{aligned} \tag{39}$$

where A_s is the surface area of the box. The $\bar{J}_2$ correspond to bin-averaged Bessel functions, computed in the same way as in Eq. 37 by integrating over the surface element instead. In Eqs 36-39, we assume linear theory for the power spectra, such that $P_{\delta\delta} = P_{\rm lin}$, $P_{\rm gm} = b_1P_{\delta\delta}$, $P_{\rm gg} = b_1^2P_{\delta\delta}$ and

[14] wrapped around cosmoprimo https://cosmoprimo.readthedocs.io/en/latest/index.html

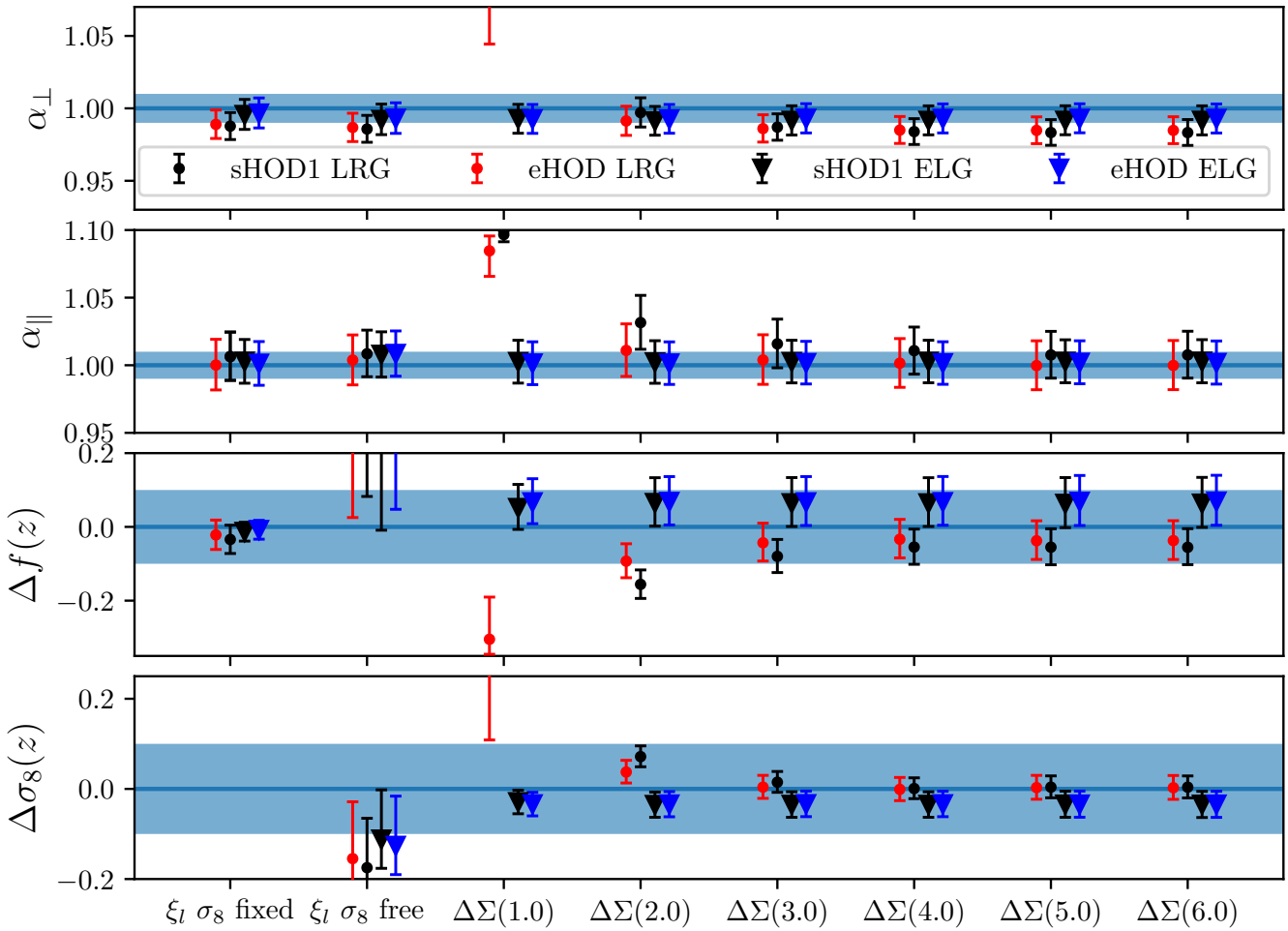

Fig. 8. Results from the cosmological fit. Circle points with error bars correspond to the fit of LRG catalog at redshift $z = 0.70$ populated with the standard and extended HOD framework while the triangles correspond to the fit of ELGs at redshift $z = 0.86$. The value of f and σ_8 for LRG(ELG) are 0.814(0.849) and 0.570(0.529) at the redshift $z = 0.70(0.86)$. Values in parenthesis correspond to the minimum scale used in the fit in h^{-1}Mpc. The shaded area corresponds to a 10% variation around the fiducial value of each parameter.

Table 3. Result of the cosmological fit from the optimal configurations determined from Fig. 8 and from Fig. 9 for the extended HOD model.

Parameter	$\Delta\alpha_\perp$	$\Delta\alpha_\parallel$	Δf	$\Delta\sigma_8$
LRG ξ_ℓ σ_8 fixed	-0.012 ± 0.010	0.006 ± 0.018	-0.035 ± 0.039	
LRG ξ_ℓ σ_8 free	-0.013 ± 0.010	0.003 ± 0.018	0.325 ± 0.277	-0.155 ± 0.099
LRG $\xi_\ell + \Delta\Sigma$ $r_{\rm min} = 3h^{-1}$Mpc	-0.013 ± 0.009	0.004 ± 0.017	-0.048 ± 0.051	0.007 ± 0.026
LRG $\xi_\ell + \Upsilon$ $r_0 = 1.2h^{-1}$Mpc	-0.014 ± 0.009	-0.002 ± 0.017	-0.014 ± 0.054	-0.003 ± 0.028
ELG ξ_ℓ σ_8 fixed	-0.004 ± 0.010	0.002 ± 0.016	-0.012 ± 0.026	
ELG ξ_ℓ σ_8 free	-0.007 ± 0.010	0.008 ± 0.017	0.341 ± 0.257	-0.128 ± 0.077
ELG $\xi_\ell + \Delta\Sigma$ $r_{\rm min} = 1h^{-1}$Mpc	-0.007 ± 0.010	0.001 ± 0.015	0.062 ± 0.057	-0.033 ± 0.025
ELG $\xi_\ell + \Upsilon$ $r_0 = 1h^{-1}$Mpc	-0.008 ± 0.010	0.001 ± 0.016	0.072 ± 0.062	-0.035 ± 0.027

$P(k,\mu) = b_1^2 P_{\rm gg}(k)\left(1+\beta^2\mu^2\right)^2$ (Kaiser 1987), with $\beta = f/b$. We infer the value of f from the UCHUU cosmology at the effective redshift of the LRG and ELG sample, respectively $z = 0.7$ and $z = 0.86$, while the value of b_1 are taken from Bautista et al. (2021); Tamone et al. (2020). The shot-noise terms $\bar{n}_p$ and $\bar{n}_g$ correspond to the number density of dark-matter particles and galaxies in the simulation box. No analytical cross-covariance exists between $\Delta\Sigma$ and the ξ_ℓ such that we approximate it to be zero. To estimate the covariance of Υ, we generate a sample of a thousand $\Delta\Sigma$ from a multivariate normal distribution, given the $\Delta\Sigma$ covariance estimate. We then apply Eq. 30 to each fake sample by considering its mean value for $\Delta\Sigma(R_0)$. this has the effect of reducing noise in the final covariance estimation.

We measured the multipoles of the correlation function with bins of size 5 h^{-1} Mpc in the range between 0 and 150 h^{-1} Mpc. $\Delta\Sigma$ is measured with 22(16) logarithm bins in the range between 10^{-1} and 10^2 h^{-1} Mpc for LRGs and ELGs respectively. This binning was chosen for each sample given the S/N of the observations. We perform fit of the multipoles on scales between 25 and 140 h^{-1} Mpc, with a minimum scale determined from the analysis of Bautista et al. (2021), as it

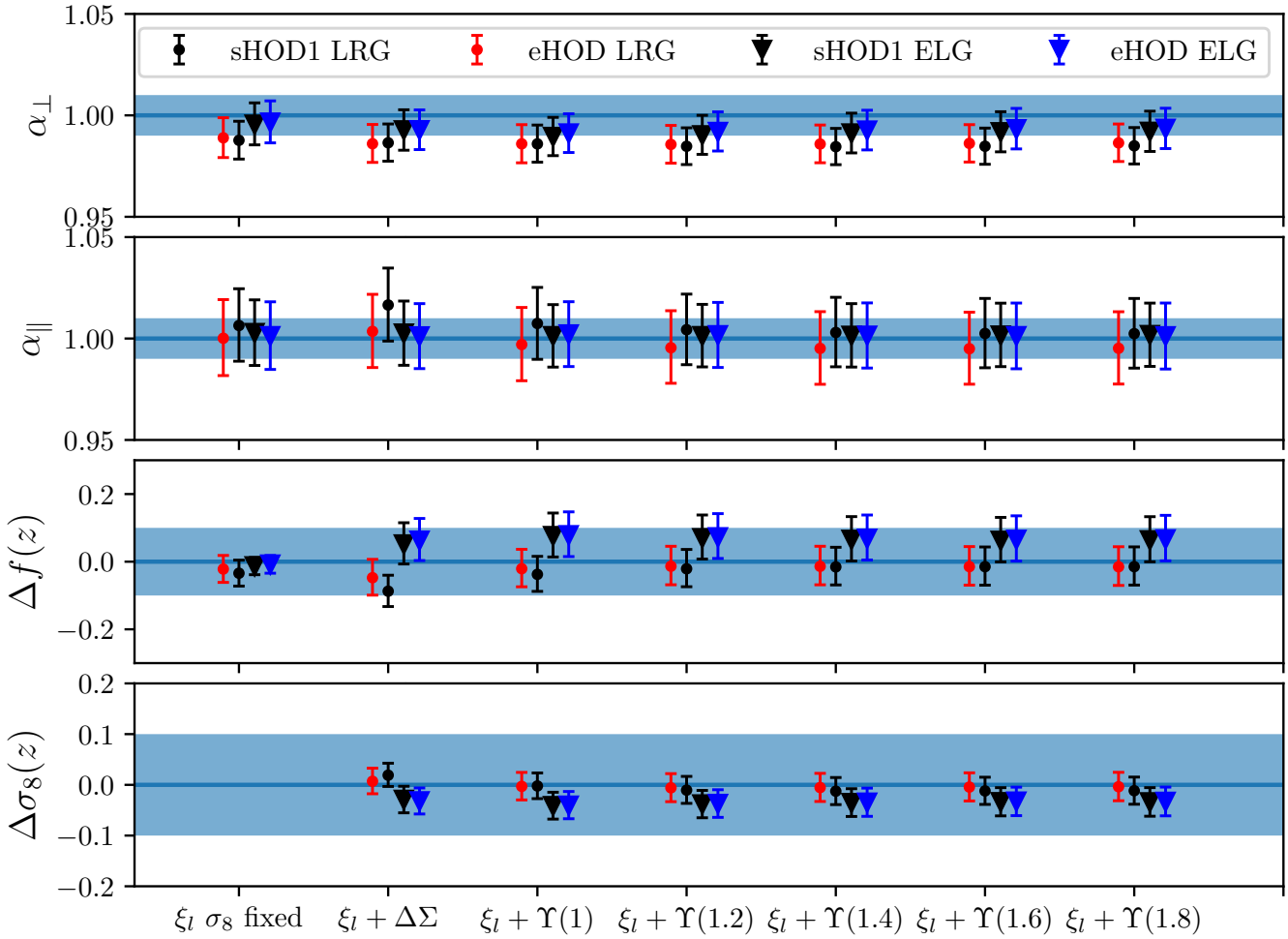

Fig. 9. Same as Fig. 8, but this time considering the annular excess surface density Υ instead.

provides unbiased constraints on cosmological parameters.

Here, we test the accuracy of the TNS model as a function of different minimum scales used when fitting either $\Delta\Sigma$ or Υ. The maximum scale that we used here is 80 h^{-1} Mpc as Limber approximation breaks down on larger scales. We present in Fig. 8 the results of the EMCEE MCMC chains of the combined $\xi_\ell + \Delta\Sigma$ fit and in Table 3 the corresponding constraints. As expected, we obtain unbiased constraints when fitting only the clustering while fixing the value of $\sigma_8(z)$ to the fiducial value of the simulation. However, in the scenario where σ_8 is allowed to vary, we can observe a strong degeneracy between f and σ_8 that is not captured properly by the model given the amount of information present in the multipoles. It is only when combining anisotropic clustering with galaxy-galaxy lensing measurements that one can break this degeneracy. For the LRG mocks, our model provides unbiased constraints on each cosmological parameter, regardless of the HOD, on scales above $3h^{-1}$ Mpc for $\Delta\Sigma$, with the constraints on the eHOD which seems to converge faster toward consistent results. This is surprising, given the observed difference in Fig. 2. The main difference comes from the goodness of fit, with the TNS model providing a better description of the $\Delta\Sigma$ measured with the eHOD galaxy population, $\Delta\chi^2 \approx 20$ compared to the sHOD. The fact that the TNS model provides a good fit, $\chi^2_r = 0.8$, of the eHOD measurements suggests that our model is also capable of yielding a good description of the observations down to scale of $\approx 3\ h^{-1}$Mpc, in agreement with the result of de la Torre et al. (2017), even if baryonic feedback that affects the shape of $\Delta\Sigma$ up to these scales, are not considered in the modeling.

For ELGs, we can see that both HOD models provide indistinguishable constraints from one another, as both methods provide very similar clustering and lensing measurements. Surprisingly, we find that the model provides unbiased and consistent constraints over all the considered range down to $r_{\rm min} = 1h^{-1}$Mpc. This means that our estimator of $P_{\rm gm}$, which is built upon a one-loop biased expansion of the non-linear matter power spectrum estimated with HMCODE, is valid up to non-linear scales to describe ELGs, at least at an effective redshift of $z = 0.86$. This might be related to their spatial distribution, indeed Gonzalez-Perez et al. (2020) found that half of DESI ELGs reside in filaments and a third in sheets, where density fluctuations are small compared to the nodes where LRGs usually lie. This is consistent with the results of Hadzhiyska et al. (2021) who found that it was twice more likely for ELG to reside in sheets compared to a mass-selected sample. In addition, ELGs occupy lower mass halos, $\log M \approx 12.2h^{-1}M_\odot$, compared to LRGs, with $\log M$ in the range $12.8 - 13.2h^{-1}M_\odot$, such that the 1-halo term contribution to $\Delta\Sigma$ is excepted to be smaller for ELGs at around $1h^{-1}$Mpc. This explains why our bias perturbative expansion provides a good description of the galaxy-galaxy lensing signal of the ELGs down to small scales. As baseline

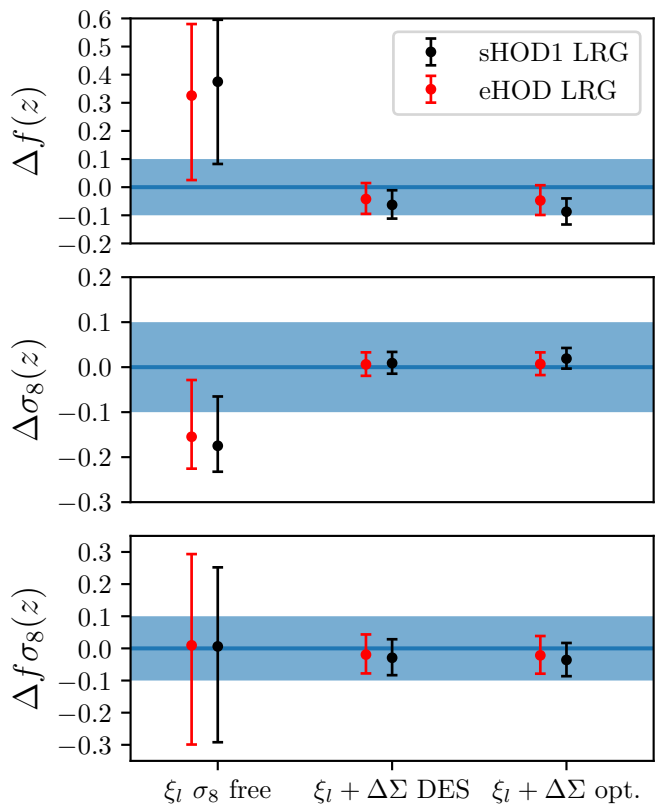

Fig. 10. Comparison between a fit to the multipoles ξ_ℓ only (left) and a fit to the combined $\xi_\ell + \Delta\Sigma$ data vector. We present constraints for the scale cuts used for γ_t in the DES y3 analysis (Krause et al. 2021), $r_{\rm min}$ = 6 h^{-1}Mpc and for the optimal configuration determined in this work, $r_{\rm min}$ = 3 h^{-1}Mpc.

configuration, we adopt $r_{\rm min} = 1(3)h^{-1}$Mpc for ELGs and LRGs respectively.

Our analysis demonstrates the constraining power of $\Delta\Sigma$, in combination with anisotropic clustering measurements, as illustrated in Fig. 10. This methodology provides unbiased constraints on the cosmological parameters $\alpha_\perp, \alpha_\parallel, f$ and σ_8, improving the precision compared to a ξ_ℓ fit only by a factor of about 6 and 3 for f and σ_8 respectively. However, let us emphasize that this conclusion is obviously optimistic, as the variance in observation is dominated by the shape noise contribution Gatti et al. (2021) which is not considered in this work. We compare the optimal LRG configuration determined in this work, $r_{\rm min} = 3h^{-1}$Mpc to the one determined in DES analysis (Krause et al. 2021), $r_{\rm min} = 6h^{-1}$Mpc. We observe similar constraints for the two configurations, with a small 3% improvement in precision for σ_8. As we move down to highly non-linear scales, the two-point function captures less and less information, such that the gain in precision becomes smaller.

We present in Fig. 9 the results of the fits when combining the ξ_ℓ with Υ, compared to the baseline configuration determined from Fig. 8. Since the ELG fits for $\Delta\Sigma$ provide consistent results up to $1h^{-1}$Mpc, the constraints when fitting Υ should be the same, regardless of the value of R_0, which is exactly what we observe on this figure. The trend is similar for LRGs, with consistent results when fitting Υ with a value of R_0 above $1.2h^{-1}$Mpc. As baseline configuration we therefore adopt $R_0 = 1(1.2)h^{-1}$Mpc for ELGs and LRGs respectively. We present in Table 3 the corresponding constraints with our baseline configuration for $\Delta\Sigma$ and Υ. Both methods provide similar constraints with consistent results at the 1 σ level. The difference in χ^2 is smaller between the extended and standard framework, $\Delta\chi^2 = 5$, as observed differences in $\Delta\Sigma$ are smoothed out with the annular excess surface density estimator. It seems that some information is lost when using Υ as a proxy for galaxy-galaxy lensing since the constraints on f and σ_8 are larger by 5 to 8%. However, this effect is marginal, below 1%, compared to the statistical precision. We therefore conclude that both statistics, either $\Delta\Sigma$ or Υ, can be used interchangeably when doing a joint cosmological analysis of galaxy clustering and lensing.

5. Conclusion

In this paper, we revisit the extended HOD framework introduced in Yuan et al. (2022), built on top of the state-of-art UCHUU simulation. This HOD model includes internal (halo concentration and shape) and environmental (local density and local shear) dependencies of central and satellite occupation. We apply this model to two eBOSS tracers: the emission lines galaxies (ELGs) in the range $0.6 < z < 1.1$, and the luminous red galaxies (LRGs) in the range $0.6 < z < 1.0$, by modelling the projected correlation w_p and the galaxy-galaxy lensing signal $\Delta\Sigma$ from highly non-linear ($\approx$ Mpc scale) up to linear scales.

We find that the extended HOD model provides a better description of both the clustering and lensing signal of LRGs, reducing the observed discrepancy on $\Delta\Sigma$ from 30-40 % to 0-20% compared to the standard HOD model based on halo mass only. Our results demonstrate the necessity to include in the HOD framework environment-based secondary parameters with a strong detection, at the 3σ level, that LRGs preferentially occupy anisotropic and denser environments. As internal halo property, we have tested either the concentration and shape of the dark-matter halo, and we have found that both properties can be used to improve the HOD model. For ELGs, we find given our statistical precision, that either the standard or extended HOD framework provides consistent and robust modeling for both the clustering and lensing signal, with a slight preference for ELGs living in denser regions, in agreement with previous results Artale et al. (2018); Delgado et al. (2022).

We then conducted a cosmological analysis of the redshift-space correlation function and galaxy-galaxy lensing signal for both HOD frameworks using a perturbative model. For LRGs, we find that our model provides unbiased and consistent constraints on the linear growth rate f, on the amplitude of fluctuations σ_8 and on the Alcock-Paczynski parameters on scales above $3h^{-1}$Mpc and $1.2h^{-1}$Mpc when fitting the galaxy-galaxy lensing signal $\Delta\Sigma$ and the annular excess surface density Υ respectively. In addition, we find that our model provides a better description of the extended HOD framework, suggesting that the observed eBOSS galaxy-galaxy lensing signal can be accurately modeled down to scales below $5h^{-1}$Mpc, where significant discrepancies can be observed between the HODs and the data. Moreover, we find for ELGs that we can derive robust constraints from the galaxy-galaxy lensing signal down to very small scales, $r_{\rm min} = 1.0h^{-1}$Mpc, regardless of the estimator of the galaxy-galaxy lensing signal. This analysis will serve as a cornerstone for a companion paper that will extract cosmological information from eBOSS galaxy-galaxy lensing signal and pave the way for future cosmological analysis of joint clustering and galaxy-galaxy lensing measurement for future cosmological experiments such as DESI, EUCLID or VRO/LSST.

References

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2022, Phys. Rev. D, 105, 023520
Alam, S., de Mattia, A., Tamone, A., et al. 2021, MNRAS, 504, 4667
Alam, S., Peacock, J. A., Kraljic, K., Ross, A. J., & Comparat, J. 2020, MNRAS, 497, 581
Alcock, C. & Paczynski, B. 1979, Nature, 281, 358
Amodeo, S., Battaglia, N., Schaan, E., et al. 2021, Phys. Rev. D, 103, 063514
Artale, M. C., Zehavi, I., Contreras, S., & Norberg, P. 2018, MNRAS, 480, 3978
Assassi, V., Baumann, D., Green, D., & Zaldarriaga, M. 2014, J. Cosmology Astropart. Phys., 2014, 056
Avila, S., Gonzalez-Perez, V., Mohammad, F. G., et al. 2020, MNRAS, 499, 5486
Baldauf, T., Smith, R. E., Seljak, U. c. v., & Mandelbaum, R. 2010, Phys. Rev. D, 81, 063531
Bautista, J. E., Paviot, R., Vargas Magaña, M., et al. 2021, MNRAS, 500, 736
Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109
Bel, J., Pezzotta, A., Carbone, C., Sefusatti, E., & Guzzo, L. 2019, A&A, 622, A109
Berlind, A. A. & Weinberg, D. H. 2002, ApJ, 575, 587
Bose, S., Eisenstein, D. J., Hernquist, L., et al. 2019, MNRAS, 490, 5693
Chaves-Montero, J., Angulo, R. E., & Contreras, S. 2023, MNRAS, 521, 937
Contreras, S., Chaves-Montero, J., & Angulo, R. E. 2023, MNRAS, 525, 3149
Dalal, N., White, M., Bond, J. R., & Shirokov, A. 2008, ApJ, 687, 12
Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270
Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
de la Torre, S., Jullo, E., Giocoli, C., et al. 2017, A&A, 608, A44
de Mattia, A., Ruhlmann-Kleider, V., Raichoor, A., et al. 2021, MNRAS, 501, 5616
De Vicente, J., Sánchez, E., & Sevilla-Noarbe, I. 2016, MNRAS, 459, 3078
Delgado, A. M., Wadekar, D., Hadzhiyska, B., et al. 2022, MNRAS, 515, 2733
Feldman, H. A., Kaiser, N., & Peacock, J. A. 1994, ApJ, 426, 23
Foreman-Mackey, D. & al. 2013, Publications of the Astronomical Society of the Pacific, 125, 306
Gatti, M., Sheldon, E., Amon, A., et al. 2021, MNRAS, 504, 4312
Gil-Marín, H., Bautista, J. E., Paviot, R., et al. 2020, MNRAS, 498, 2492
Gonzalez-Perez, V., Comparat, J., Norberg, P., et al. 2018, MNRAS, 474, 4024
Gonzalez-Perez, V., Cui, W., Contreras, S., et al. 2020, MNRAS, 498, 1852
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Gouin, C., Gavazzi, R., Pichon, C., et al. 2019, A&A, 626, A72
Grieb, J. N., Sánchez, A. G., Salazar-Albornoz, S., & Dalla Vecchia, C. 2016, MNRAS, 457, 1577
Guzik, J. & Seljak, U. 2001, MNRAS, 321, 439
Hadzhiyska, B., Bose, S., Eisenstein, D., Hernquist, L., & Spergel, D. N. 2020, MNRAS, 493, 5506
Hadzhiyska, B., Eisenstein, D., Hernquist, L., et al. 2023a, MNRAS, 524, 2507
Hadzhiyska, B., Hernquist, L., Eisenstein, D., et al. 2023b, MNRAS, 524, 2524
Hadzhiyska, B., Tacchella, S., Bose, S., & Eisenstein, D. J. 2021, MNRAS, 502, 3599
Hamilton, A. J. S. 2000, MNRAS, 312, 257
Ishiyama, T., Fukushige, T., & Makino, J. 2009, PASJ, 61, 1319
Ishiyama, T., Nitadori, K., & Makino, J. 2012, arXiv e-prints, arXiv:1211.4406
Ishiyama, T., Prada, F., Klypin, A. A., et al. 2021, MNRAS, 506, 4210
Jullo, E., de la Torre, S., Cousinou, M. C., et al. 2019, A&A, 627, A137
Kaiser, N. 1984, ApJ, 284, L9
Kaiser, N. 1987, MNRAS, 227, 1
Krause, E., Fang, X., Pandey, S., et al. 2021, arXiv e-prints, arXiv:2105.13548
Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64
Lange, J. & Huang, S. 2022, dsigma: Galaxy-galaxy lensing Python package, Astrophysics Source Code Library, record ascl:2204.006
Leauthaud, A., Saito, S., Hilbert, S., et al. 2017, MNRAS, 467, 3024
Lewis, A. & Challinor, A. 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library, record ascl:1102.026
Limber, D. N. 1953, ApJ, 117, 134
Marian, L., Smith, R. E., & Angulo, R. E. 2015, MNRAS, 451, 1418
Mead, A. J., Brieden, S., Tröster, T., & Heymans, C. 2021, MNRAS, 502, 1401
Miyatake, H., More, S., Mandelbaum, R., et al. 2015, ApJ, 806, 1
Mohammad, F. G., Percival, W. J., Seo, H.-J., et al. 2020, MNRAS, 498, 128
Musso, M., Cadiou, C., Pichon, C., et al. 2018, MNRAS, 476, 4877
Myles, J., Alarcon, A., Amon, A., et al. 2021, MNRAS, 505, 4249
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Nishimichi, T., Bernardeau, F., & Taruya, A. 2017, Phys. Rev. D, 96, 123515
Paranjape, A., Hahn, O., & Sheth, R. K. 2018, MNRAS, 476, 3631
Paviot, R., de la Torre, S., de Mattia, A., et al. 2022, MNRAS, 512, 1341
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Rocher, A., Ruhlmann-Kleider, V., Burtin, E., & de Mattia, A. 2023a, J. Cosmology Astropart. Phys., 2023, 033
Rocher, A., Ruhlmann-Kleider, V., Burtin, E., et al. 2023b, arXiv e-prints, arXiv:2306.06319
Ross, A. J., Bautista, J., Tojeiro, R., et al. 2020, MNRAS, 498, 2354
Rossi, G., Choi, P. D., Moon, J., et al. 2021, MNRAS, 505, 377
Saito, S., Baldauf, T., Vlah, Z., et al. 2014, Phys. Rev. D, 90, 123522
Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. 2021, ApJS, 254, 24
Simonović, M., Baldauf, T., Zaldarriaga, M., Carrasco, J. J., & Kollmeier, J. A. 2018, J. Cosmology Astropart. Phys., 2018, 030
Singh, S., Mandelbaum, R., Seljak, U., Slosar, A., & Vazquez Gonzalez, J. 2017, MNRAS, 471, 3827
Sinha, M. & Garrison, L. H. 2020, Monthly Notices of the Royal Astronomical Society, 491, 3022
Tamone, A., Raichoor, A., Zhao, C., et al. 2020, MNRAS, 499, 5527
Taruya, A., Nishimichi, T., & Saito, S. 2010, Phys. Rev. D, 82, 063522
van den Bosch, F. C., More, S., Cacciato, M., Mo, H., & Yang, X. 2013, MNRAS, 430, 725
Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, ApJ, 568, 52
White, M., Blanton, M., Bolton, A., et al. 2011, ApJ, 728, 126
White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341
Xu, X., Zehavi, I., & Contreras, S. 2021, MNRAS, 502, 3242
Yuan, S., Garrison, L. H., Hadzhiyska, B., Bose, S., & Eisenstein, D. J. 2022, MNRAS, 510, 3301
Yuan, S., Hadzhiyska, B., Bose, S., Eisenstein, D. J., & Guo, H. 2021, MNRAS, 502, 3582
Zehavi, I., Contreras, S., Padilla, N., et al. 2018, ApJ, 853, 84
Zhai, Z., Tinker, J. L., Hahn, C., et al. 2017, ApJ, 848, 76
Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791
Zheng, Z., Coil, A. L., & Zehavi, I. 2007, ApJ, 667, 760

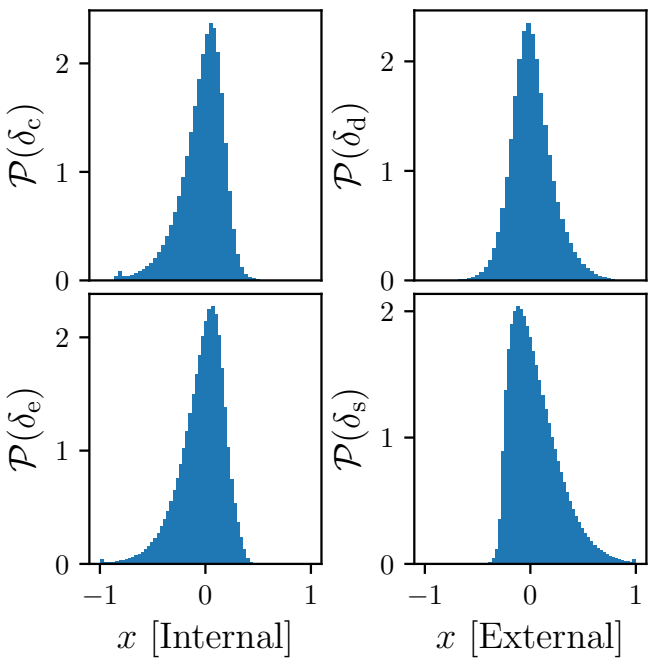

Fig. A.1. Distribution of internal and external halo properties as defined in Sec 3.2. The left panels present the distribution of concentration (c) and ellipticity (e) while the right panels present the local density (d) and local shear (s).

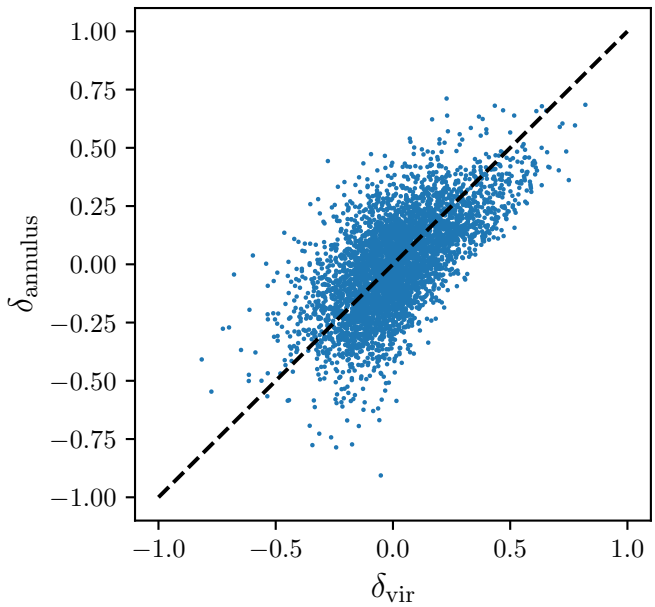

Fig. A.2. Comparison between the local density estimated on a mesh with a Gaussian filter of size $R = 2.25\ r_{\rm vir}$ and the local density estimated by counting halo and sub-halo around $r_{\rm vir}$ and 5 h^{-1}Mpc.

Appendix A: Additional parameters

We present in Fig. A.1 the normalized distribution in the range [-1,1] of internal (concentration and ellipticity) and external (local density and local shear) halo properties on the left and right panel respectively. We present in Fig. A.2 the relation between the two different local density estimators described in this work.

Appendix B: Posterior of the LRG HOD

We present here the full posterior distributions for the different configurations presented in Section 4.3. We only present $M_{\rm min}$ for the posterior distributions presented in Fig. 4 to save computational time, as $M_{\rm min}$ is not a free parameter and is computed for each point of the chain given Eq. 7. In addition, for completeness, we also present the full posterior contours.

Acknowledgements. RP is supported by a DIM-ACAV+ fellowship from *Region Ile-de-France*. SC acknowledges financial support from Fondation MERAC and by the SPHERES grant ANR-18-CE31-0009 of the French *Agence Nationale de la Recherche*.

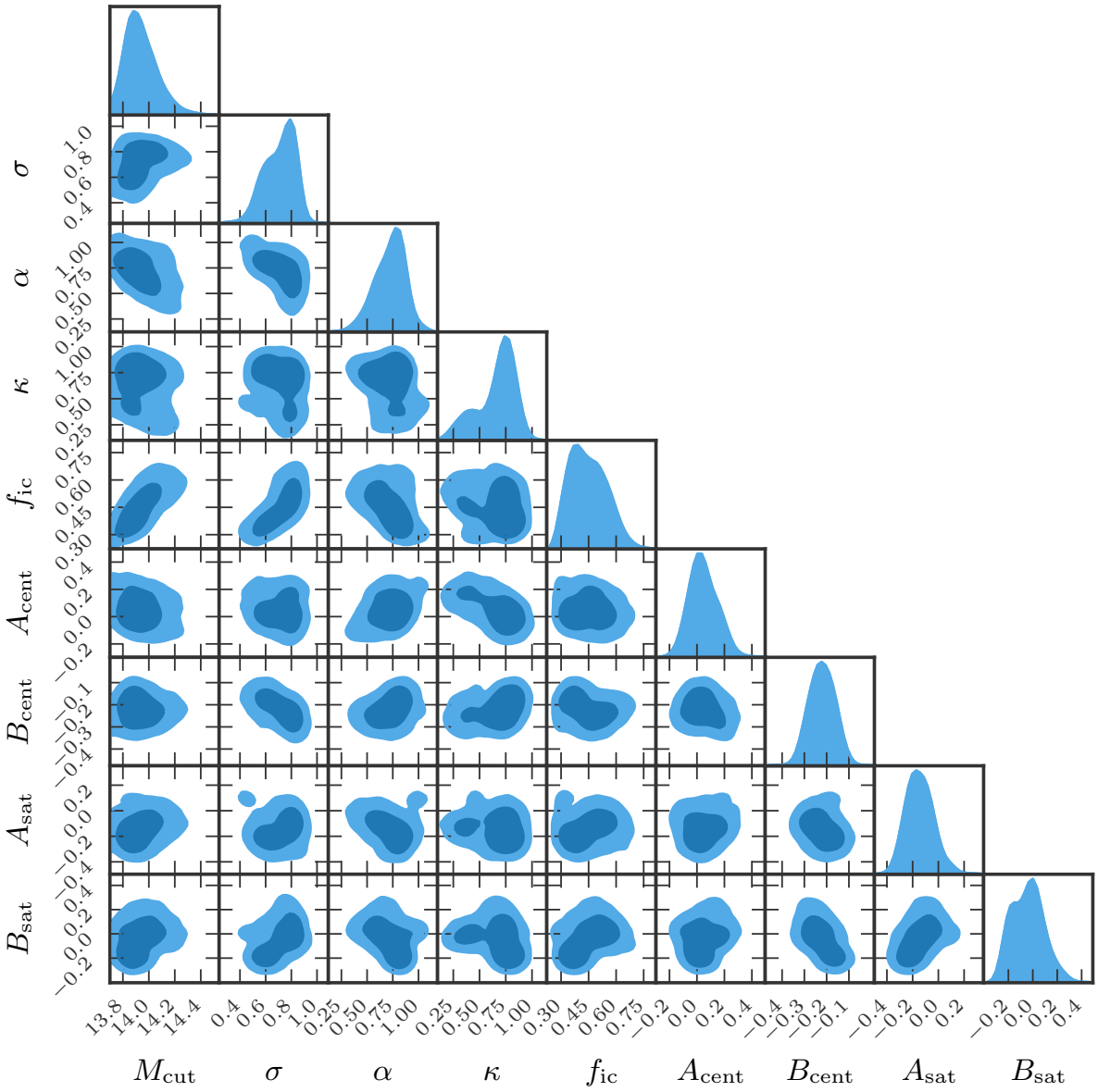

Fig. A.3. Full posteriors distribution of the LRG extended HOD, including the assembly bias parameters.

Table A.1. Result of the Gaussian process emulating the likelihood function after 4000 iterations for the eBOSS LRG sample. We present here the 16%, 50% and 84% percentiles of the posterior distributions. The prior are presented in bracket and correspond to uniform priors within the quoted ranges.

Method	sHOD1	sHOD2	eHOD $\delta_{\rm vir}$,c	eHOD $T^{ij}_{\rm vir}$,c	eHOD $T^{ij}_{\rm vir}$,b/a	eHOD $\delta_{\rm annulus}$,c	eHOD $\delta_{\rm annulus}$,b/a	eHOD $T^{ij}_{\rm 5Mpc}$,c
$M_{\rm cut}$ [13.7,15.0]	$14.17^{+0.13}_{-0.12}$	$14.02^{+0.07}_{-0.06}$	$14.02^{+0.12}_{-0.09}$	$13.91^{+0.15}_{-0.14}$	$13.86^{+0.10}_{-0.09}$	$13.92^{+0.14}_{-0.11}$	$13.91^{+0.12}_{-0.11}$	$13.94^{+0.13}_{-0.11}$
σ [0.1,1.2]	$0.53^{+0.11}_{-0.12}$	$0.40^{+0.11}_{-0.13}$	$0.50^{+0.11}_{-0.13}$	$0.77^{+0.10}_{-0.12}$	$0.65^{+0.13}_{-0.13}$	$0.74^{+0.09}_{-0.14}$	$0.79^{+0.11}_{-0.10}$	$0.79^{+0.10}_{-0.09}$
α [0.1,1.5]	$1.11^{+0.14}_{-0.15}$	$1.12^{+0.16}_{-0.10}$	$0.83^{+0.20}_{-0.20}$	$0.83^{+0.15}_{-0.14}$	$0.92^{+0.13}_{-0.19}$	$0.73^{+0.14}_{-0.18}$	$0.82^{+0.11}_{-0.11}$	$1.02^{+0.09}_{-0.16}$
κ [0.1,1.5]	$0.59^{+0.30}_{-0.31}$	$0.72^{+0.21}_{-0.22}$	$0.79^{+0.13}_{-0.16}$	$0.29^{+0.29}_{-0.15}$	$0.73^{+0.20}_{-0.32}$	$0.69^{+0.14}_{-0.31}$	$0.58^{+0.26}_{-0.32}$	$0.45^{+0.16}_{-0.12}$
$f_{\rm ic}$ [0.1,1.0]	$0.64^{+0.20}_{-0.15}$	$0.45^{+0.10}_{-0.07}$	$0.41^{+0.08}_{-0.07}$	$0.42^{+0.15}_{-0.18}$	$0.34^{+0.08}_{-0.07}$	$0.45^{+0.11}_{-0.10}$	$0.43^{+0.10}_{-0.09}$	$0.45^{+0.09}_{-0.07}$
$A_{\rm cent}$ [-0.5,0.5]			$-0.01^{+0.15}_{-0.12}$	$0.24^{+0.23}_{-0.31}$	$0.01^{+0.19}_{-0.13}$	$0.04^{+0.14}_{-0.10}$	$0.12^{+0.24}_{-0.16}$	$0.26^{+0.15}_{-0.38}$
$B_{\rm cent}$ [-0.5,0.5]			$-0.24^{+0.09}_{-0.08}$	$-0.27^{+0.10}_{-0.09}$	$-0.20^{+0.07}_{-0.06}$	$-0.23^{+0.06}_{-0.06}$	$-0.26^{+0.07}_{-0.08}$	$-0.20^{+0.06}_{-0.06}$
$A_{\rm sat}$ [-0.5,0.5]			$0.13^{+0.20}_{-0.31}$	$-0.21^{+0.36}_{-0.18}$	$-0.02^{+0.24}_{-0.14}$	$-0.14^{+0.12}_{-0.11}$	$0.01^{+0.16}_{-0.32}$	$-0.10^{+0.12}_{-0.24}$
$B_{\rm sat}$ [-0.5,0.5]			$-0.08^{+0.11}_{-0.09}$	$0.20^{+0.14}_{-0.27}$	$-0.08^{+0.21}_{-0.10}$	$-0.02^{+0.12}_{-0.15}$	$0.04^{+0.21}_{-0.22}$	$0.04^{+0.08}_{-0.10}$
$M_{\rm min}$	$13.28^{+0.09}_{-0.09}$					$13.27^{+0.14}_{-0.17}$		