Galaxy alignments from multiple angles

Fortuna, M.C.

3 KiDS-1000: Constraints on the intrinsic alignment of luminous red galaxies

M.C Fortuna, *H. Hoekstra, H. Johnston, M. Vakili, A. Kannawadi, C. Georgiou, B. Joachimi, A. H. Wright, M. Asgari, M. Bilicki, C. Heymans, H. Hildebrandt, K. Kuijken, M. Von Wietersheim-Kramsta* *2021, A&A, 654, A76*

We constrain the luminosity and redshift dependence of the intrinsic alignment (IA) of a nearly volume-limited sample of luminous red galaxies selected from the fourth public data release of the Kilo-Degree Survey (KiDS-1000). To measure the shapes of the galaxies, we used two complementary algorithms, finding consistent IA measurements for the overlapping galaxy sample. The global significance of IA detection across our two independent luminous red galaxy samples, with our favoured method of shape estimation, is $\sim 10.7\sigma$. We find no significant dependence with redshift of the IA signal in the range $0.2 < z < 0.8$, nor a dependence with luminosity below $L_r \lesssim 2.9 \times 10^{10} h^{-2} L_{r,\odot}$. Above this luminosity, however, we find that the IA signal increases as a power law, although our results are also compatible with linear growth within the current uncertainties. This behaviour motivates the use of a broken power law model when accounting for the luminosity dependence of IA contamination in cosmic shear studies.

3.1 Introduction

Galaxies that form close to a matter over-density are affected by the tide induced by the quadrupole of the surrounding gravitational field, and the distribution of stars will adjust accordingly. This process, which starts during the initial stages of galaxy formation (Catelan et al. 2001), can persist over their entire lifetime, as galaxies have continuous gravitational interactions with the surrounding matter (e.g. Bhowmick et al. 2020), and leads to the intrinsic alignment (IA) of galaxies.

This tendency of neighbouring galaxy pairs to have a similar orientation of their intrinsic shapes is an important contaminant for weak gravitational lensing measurements (e.g. Joachimi et al. 2015). The matter distribution along the line-of-sight distorts the images of background galaxies, resulting in apparent correlations in their shapes. Intrinsic alignment contributes to the observed correlations, complicating the interpretation. To infer unbiased cosmological parameter estimates it is therefore crucial to account for the IA contribution. This is particularly important in the light of future surveys, such as *Euclid*[1] (Laureijs et al. 2011) and the Large Synoptic Survey Telescope (LSST)[2] at the Vera C. Rubin Observatory (Abell et al. 2009), which aim to constrain the cosmological parameters with sub-percent accuracy (for a forecast of the IA impact on current and upcoming surveys see Kirk et al. 2010; Krause et al. 2016, among others). Some recent results on current weak lensing studies are available in, for example, Aihara et al. (2018); Asgari et al. (2021); DES Collaboration et al. (2021).

To provide informative priors to lensing studies, it is essential to learn as much as possible from direct observations of IA. It is, however, also important that such results can be related to the properties of galaxies that give rise to the alignment signal in cosmic shear surveys (Fortuna et al. 2021a). Intrinsic alignment studies are typically limited to relatively bright galaxies, which often sit at the centre of their own group or cluster, and it is thus possible to connect their alignment to the underlying dark matter halo alignment via analytic models (Hirata & Seljak 2004). The picture becomes more complicated when considering samples that contain a significant fraction of satellite galaxies: The alignment of satellites arises as a result of the continuous torque exercised by the intra-halo tidal fields while the satellite orbits inside the halo (Pereira et al. 2008; Pereira & Bryan 2010). This leads to a radial alignment, which also depends on the galaxy

[1] https://www.euclid-ec.org
[2] https://www.lsst.org

distance from the centre of the halo (Georgiou et al. 2019b). At the same time, satellites fall into halos through the filaments of the large-scale structure, and this persists as an anisotropic distribution within the halo, which has been detected both in simulations (Knebe et al. 2004; Zentner et al. 2005) and observations (West & Blakeslee 2000; Bailin et al. 2008; Huang et al. 2016; Johnston et al. 2019; Georgiou et al. 2019b). The combination of these two effects complicates the picture. At small scales, where the satellite contribution is expected to be important, their signal may be described using a halo model formalism (Schneider & Bridle 2010; Fortuna et al. 2021a), but their contribution to IA on large scales remains poorly constrained (Johnston et al. 2019); although it is expected that they are not aligned, they do affect the inferred amplitude because they contribute to the overall mix of galaxies. This prevents a straightforward interpretation of any secondary sample dependence of the IA signal sourced by the central galaxy population, such as the dependence on luminosity or colour, in mixed samples where the fraction of satellites is relevant.

Observational studies have found discordant results regarding the presence of a luminosity dependence of the IA signal, with the bright end being well described by a steep power law with index ~ 1.2 (Hirata et al. 2007; Joachimi et al. 2011; Singh et al. 2015), while less luminous galaxies do not show any significant dependence of the IA signal with luminosity (Johnston et al. 2019). A recent investigation using hydrodynamic simulations by Samuroff et al. (2020a) supports a flatter slope, in agreement with Johnston et al. (2019) and Fortuna et al. (2021a) at low luminosities but in tension with previous studies that probe more luminous galaxies. The interpretation of these results is also affected by the presence of satellites, whose fraction varies with luminosity and depends on the specific selection function of the data. At low redshift, a cosmic shear survey is dominated by faint galaxies, and improving our understanding of the IA signal at low luminosities is one of the most urgent questions for IA studies.

Another relevant aspect that is often neglected is the dependence of IA on the shape measurement method (Singh & Mandelbaum 2016). The tendency to align in the direction of the surrounding tidal field is a function of galaxy scale (Georgiou et al. 2019b), with the outermost parts – which are more weakly gravitationally locked to the galaxy – showing a more severe twist. It increases the IA signal associated with shapes measured via algorithms that assign more importance to the galaxy outskirts. In contrast, lensing studies typically prefer shape methods that give more weight to the inner part of a galaxy. Accounting for this discrepancy is potentially

relevant for future cosmic shear studies.

In this work we focus on investigating the luminosity dependence of the IA signal in the least constrained regime, $M_r \gtrsim -22$. We employ two different samples, which differ in mean luminosity and number density. We limit the analysis to the large-scale alignment, for which a theoretical framework is already available and where the luminosity dependence is known to play a crucial role (Fortuna et al. 2021a). We also provide estimates of the satellite fractions present in our samples in order to guide future work on the modelling of satellite alignment at large scales. We also explore the dependence of our signal on the shape measurement algorithm used to create the shape catalogue. We compare the signal as measured by two complementary algorithms: DEIMOS (DEconvolution In MOment Space; Melchior et al. 2011), which has been widely used in IA studies (Georgiou et al. 2019a; Johnston et al. 2019; Georgiou et al. 2019b), and *lens*fit (Miller et al. 2007, 2013) which has been used for the cosmological analysis of the Canada-France-Hawaii Telescope Lensing Survey (Heymans et al. 2013, CFHTLenS;) and the Kilo-Degree Survey (KiDS; see Asgari et al. 2021, and references therein).

One of the main limitations for measuring IA is the necessity of simultaneously relying on high-quality images and precise redshifts to properly identify physically close pairs of galaxies that share the same gravitational tidal shear. Wide field image surveys provide high-quality images, but the uncertainty in the photometric redshifts is too large for useful IA measurements. Fortunately, using a specific selection in colours, it is possible to obtain a sub-sample of galaxies with more precise photometric redshifts: the luminous red galaxies (LRGs). At any given redshift, LRGs populate a well-defined region in the colour-magnitude diagram, known as the red-sequence ridgeline. Using this unique property, it is possible to design a specific algorithm to select LRGs in photometric surveys, which results in both precise and accurate redshifts (Rozo et al. 2016; Vakili et al. 2019, 2020). Luminous red galaxies have also been shown to be strongly affected by the surrounding tidal fields, making them an extremely suitable sample for exploring the behaviour of IA at different redshifts and as a function of secondary galaxy properties, such as luminosity and type (central or satellites).

Joachimi et al. (2011) first studied the IA signal of an LRG sample with photometric redshifts. In this paper we follow their main approach but use a catalogue of LRGs selected by Vakili et al. (2020) using the KiDS fourth public data release (KiDS-1000 Kuijken et al. 2019).

The paper is structured as follows. In Sect. 3.2 we describe our data and the characteristics of our two main samples. In Sect. 3.3 we introduce the two shape measurement methods employed in the analysis and present the strategy adopted to calibrate the bias in the measured shapes. Section 3.4 presents the estimators we use to extract the signal from the data, while Sect. 3.5 illustrates the theoretical framework we rely on when modelling the signal: the way the model accounts for the use of photometric redshifts as well as the way we account for astrophysical contaminants. Finally, we present our main results in Sect. 5.6 and conclude in Sect. 4.8.

Throughout the paper, we assume a flat Λ cold dark matter cosmology with $h = 0.7, \Omega_\mathrm{m} = 0.25, \Omega_\mathrm{b} = 0.044, \sigma_8 = 0.8$, and $n_\mathrm{s} = 0.96$.

3.2 KiDS

The Kilo-Degree Survey is a multi-band imaging survey designed for weak lensing studies, currently at its fourth data release (KiDS-1000; Kuijken et al. 2019). The data are obtained with the OmegaCAM instrument (Kuijken 2011) on the VLT Survey Telescope (VST; Capaccioli et al. 2012). This combination of telescope and camera was designed specifically to produce high-quality images in the *ugri* filters, with best seeing-conditions in the r–band, and a mean magnitude limit of ~ 25 (5σ in a $2''$ aperture). These measurements are combined with results from the VISTA Kilo-degree INfrared Galaxy survey (VIKING; Edge et al. 2013), which surveyed the same area in five infrared bands ($ZYJHK_\mathrm{s}$). This resulted in high-quality photometry in nine bands across approximately $1000\ \mathrm{deg}^2$ imaged by the fourth data release[3]. The VIKING data are important for the LRG selection at high redshift (Vakili et al. 2020): the Z band is included in the red-sequence template and improves the constraints on the redshift of the high-redshift galaxies, while the K_s band allows for a clean separation between galaxies and stars in the $(r - K_\mathrm{s}) - (r - z)$ colour-colour space.

3.2.1 The LRG sample

Red-sequence galaxies are characterised by a tight colour-redshift relation, so that at any given redshift they follow a narrow ridgeline in the colour-magnitude space. This relation can be exploited to select red galaxies from photometric data and obtain precise photometric redshifts. Here we use the catalogue of LRGs presented in Vakili et al. (2020). It uses a variation of

[3]The survey was recently completed, imaging a final total of 1350 deg^2.

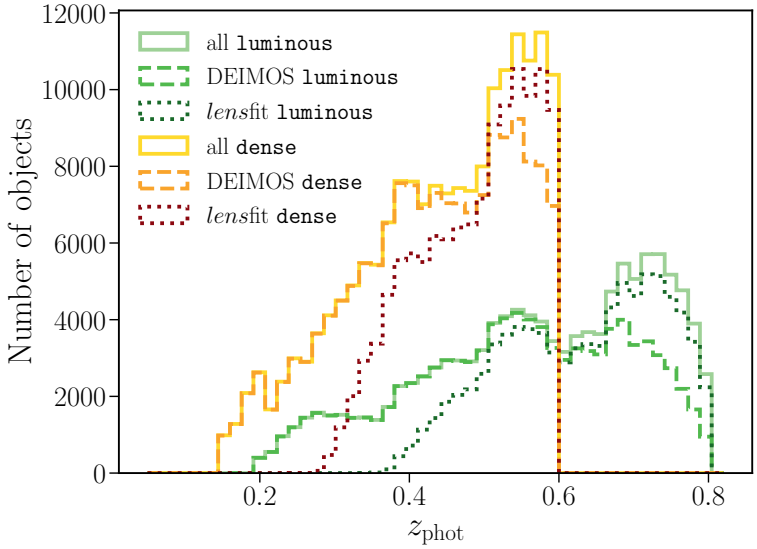

Figure 3.1: Photometric redshift distributions for our density (all) and shape catalogues (*lens*fit and DEIMOS; see text for details). The orange histograms show the distribution for the dense samples, which is limited to $z_{\rm phot} < 0.6$, whereas the luminous sample (green) is restricted to $z_{\rm phot} < 0.8$.

the redMagiC algorithm (Rykoff et al. 2014) to select LRGs from the KiDS-1000 data. As detailed in Vakili et al. (2019) and Vakili et al. (2020), the red-sequence template is calibrated using the regions of KiDS that overlap with a number of spectroscopic surveys: SDSS DR13 (Albareti et al. 2017), 2dFLenS (Blake et al. 2016), GAMA (Driver et al. 2011), together with the GAMA G10 region, which overlaps with COSMOS (Davies et al. 2015).

The algorithm is designed to return a sample of LRGs with a constant comoving number density. It achieves this by imposing a redshift-dependent magnitude cut that depends on $m_r^{\rm pivot}(z)$, the characteristic r-band magnitude of the Schechter (1976) function, assuming a faint-end slope $\alpha = 1$ (for more details, see Vakili et al. 2019, sect. 3.1). We use this to define two samples that differ from each other in terms of their minimum luminosity relative to the luminosity $L_{\rm pivot}(z)$. We refer to them as our `luminous` sample (high luminosity, low number density, $L_{\rm min}/L_{\rm pivot}(z) = 1$) and `dense` sample (lower luminosity, higher number density, $L_{\rm min}/L_{\rm pivot}(z) = 0.5$). To ensure that the two samples are separate, we removed the galaxies in the `dense` sample that also belong to the `luminous` one. However, this does not mean they do not overlap in their physical properties. In particular,

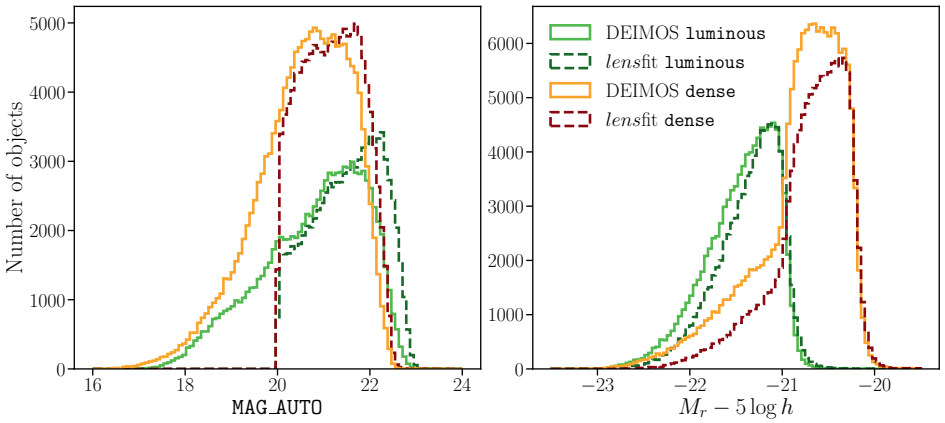

Figure 3.2: The magnitude distributions of the samples used in the analysis. *Left panel:* Histograms of the apparent magnitude, `MAG_AUTO` in the r-band for the galaxies in the `dense` (orange lines) and `luminous` (green lines) samples with shapes measured by *lens*fit (darker colours) and DEIMOS (lighter colours). *Right panel:* Histograms of the absolute magnitudes in the r-band ($K+e$ corrected) for the same samples.

they overlap partially in luminosity, a feature that we will exploit later in the paper.

As shown in Fig. 3.1, the two samples also span different redshift ranges. The `luminous` sample extends from $z = 0.2$ to $z = 0.8$. After applying a conservative mask to select only objects with a high probability to be red-sequence galaxies (corresponding to objects with a clear separation from the star sequence in the colour-colour diagram), we are left with $117\,001$ galaxies, which comprise our density sample. By density sample—not to be confused with the `dense` sample described above—we refer to the sample used to trace galaxy positions, as opposed to the shape sample, which is the sample used for the measurement of galaxy orientations and is composed by the galaxies of the corresponding density sample for which a given shape measurement algorithm is able to measure the galaxy shape. The density and shape samples used in this analysis are visible in Fig. 3.1, where the density samples of the `luminous` and `dense` samples are referred to as 'all' galaxies. The `dense` sample is obtained with the same strategy, but we further impose $z < 0.6$ to ensure the completeness and purity of the sample (see Fig. 4 in Vakili et al. (2020)). This leads to a final sample of $173\,445$ galaxies. As shown in Vakili et al. (2020), the redshift errors are well described by a Student's t–distribution. The width of the distribution increases slightly with redshift, with typical values around $\sigma_z \sim 0.014 - 0.019$.

For further details on the sample selection and redshift estimation, we refer the interested reader to Vakili et al. (2020).

We infer galaxy absolute magnitudes using Lephare[4] (Arnouts & Ilbert 2011), assuming the dust extinction law from (Calzetti et al. 1994) and the stellar population synthesis model from Bruzual & Charlot (2003). We correct our magnitudes to $z = 0$; the K-correction is provided by Lephare and the correction for the evolution of the stellar populations (e–correction) is computed with the python package EzGal[5] (Mancone & Gonzalez 2012), assuming Salpeter initial mass function (Chabrier 2003) and a single star formation burst at $z = 3$. These corrections are based on the magnitudes used to define the colours (`MAG_GAAP`), which are measured using Gaussian apertures (Kuijken et al. 2019). Although ideal for colour estimates, these underestimate the flux and should not be used to compute the luminosity. For that purpose we correct[6] them using the Kron-like `MAG_AUTO` measured from the r-band images by SExtractor (Bertin & Arnouts 1996).

The left panel of Fig. 3.2 shows the distribution in apparent magnitude `MAG_AUTO` for galaxies in the `dense` and `luminous` samples for which shapes were determined by *lens*fit or DEIMOS. In Sect. 3.3 we describe the two shape measurement methods and explain the difference in their number counts. We note that the LRGs are much brighter than the limiting magnitude of KiDS in the r-band. The corresponding distributions in absolute magnitude in the rest-frame r filter, K+e corrected to $z = 0$, are presented in the right panel of Fig. 3.2. This shows that the `dense` sample overlaps somewhat with the `luminous` sample in terms of luminosity, as a consequence of the photometric redshift uncertainty[7].

[4] `https://www.cfht.hawaii.edu/~arnouts/LEPHARE/lephare.html`

[5] `http://www.baryons.org/ezgal`

[6] The total flux in the x filter can be computed using m_x = `MAG_AUTO`$_r$ + (`MAG_GAAP`$_x$ − `MAG_GAAP`$_r$), which implicitly assumes that colour gradients are negligible.

[7] The selection through the redshift-dependent apparent magnitude cut results in an overlap in apparent magnitudes of the `dense` and `luminous` samples. Because the cut is redshift-dependent, this implies a threshold in luminosity: In the case of perfect redshifts, this would result in a disjoint sample, because we removed the galaxies from the `dense` sample that overlap with the `luminous` one. The photometric redshift uncertainty, however, assigns to galaxies with the same apparent magnitude different luminosities, and thus a portion of the `dense` sample extends above the luminosity threshold of the `luminous` sample.

3.2.2 Satellite galaxy fraction estimation

Observations suggest that satellite galaxies are only weakly aligned (see e.g. Georgiou et al. 2019b, for recent constraints) and thus suppress the IA signal at large scales. We do not take this into account in our analysis but provide here an estimate of the fraction of satellites we expect in our samples. Such information will be useful for future modelling studies.

We used the publicly available `G3GGal` and `G3GFoFGroup` catalogues (Robotham et al. 2011) from the GAMA survey (Driver et al. 2009, 2011; Liske et al. 2015). Since KiDS overlaps with GAMA, these catalogues provide group information for a subset of our galaxies, obtained with a Friends-of-Friends algorithm. We cross-matched our LRG samples with the `G3GGal` catalogue and selected galaxies with $z < 0.21$ ($z < 0.32$), which provide a roughly volume-complete match to the `dense` (`luminous`) sample. With the information in both group catalogues, we identify both the brightest group galaxies and ungrouped galaxies as centrals, and the rest as satellites. With this strategy, we obtain $f_{\rm sat} = 0.34$ for our `dense`×GAMA sample and $f_{\rm sat} = 0.23$ for the `luminous`×GAMA[8]. Since our samples are selected to resemble the same galaxy populations at different redshifts, these estimates should be fairly representative beyond the redshift range probed by our direct comparison.

3.3 Shape measurements

In addition to precise redshifts, a successful IA measurement requires accurate shape measurements. In this work, we compare two different algorithms, DEIMOS and *lens*fit both in terms of their ability to recover reliable ellipticity measurements and the resulting IA signal. Exploring the dependence of the IA signal on the shape measurement algorithm is important if one aims to provide informative priors to lensing studies (Singh & Mandelbaum 2016). Both algorithms have been used to analyse KiDS data: DEIMOS to provide the shape catalogue (Georgiou et al. 2019a) for a number of IA studies, while *lens*fit was used for cosmic shear analyses (see Giblin et al. 2021, for the most recent shape measurements).

[8]These estimates refer to the full samples, but should be representative for the shape samples as well.

3.3.1 DEIMOS

DEIMOS (Melchior et al. 2011) is a moment-based shape measurement algorithm designed to measure the moments of the surface brightness distribution from an image, which are subsequently used to estimate the ellipticity. The main features of DEIMOS are its rigorous treatment of the PSF moments to arbitrary order, the lack of model assumptions and the flexibility in changing the size of the weight function so that it is possible to assign more importance to different parts of a galaxy while performing the shape measurement (bulge or outskirts).

The unweighted moments of the surface brightness $G(\vec{x})$ are defined as

$$Q_{ij} \equiv \{G\}_{ij} = \int G(\vec{x})\, x^i y^j \,\mathrm{d}x\,\mathrm{d}y\,, \tag{3.1}$$

where (x, y) are the Cartesian coordinates with origin at the galaxy's centroid. The complex ellipticity is then defined in terms of the second-order moments as

$$\epsilon \equiv \epsilon_1 + \mathrm{i}\epsilon_2 = \frac{\mathrm{Q}_{20} - \mathrm{Q}_{02} + 2\mathrm{i}\,\mathrm{Q}_{11}}{\mathrm{Q}_{20} + \mathrm{Q}_{02} + 2\sqrt{\mathrm{Q}_{20}\,\mathrm{Q}_{02} - \mathrm{Q}_{11}^2}}\,. \tag{3.2}$$

In practice, unweighted moments cannot be used because of noise in the images, and weighted moments have to be employed instead. We will return to this issue later. Moreover, the galaxy images are smeared and distorted by the atmospheric blurring and the telescope optics, so that the observed image, G^*, is convolved with the PSF kernel $P(\vec{x})$,

$$G^*(\vec{x}) = \int G(\vec{x}')\, P(\vec{x} - \vec{x}')\,\mathrm{d}\vec{x}'\,. \tag{3.3}$$

The DEIMOS algorithm estimates the unweighted moments by correcting the observed weighted moments of the galaxy surface brightness for the convolution by the PSF. The underlying mathematical framework is a deconvolution in moment space. In order to measure the moments in Eq. (3.1) we then need to deconvolve them. This can easily be achieved in Fourier space, where the convolution becomes a product. Using the Cauchy product, we can write (Melchior et al. 2011):

$$\{G^*\}_{ij} = \sum_k^i \sum_l^j \binom{i}{k} \binom{j}{l} \{G\}_{kl} \{P\}_{i-k,j-l}\,, \tag{3.4}$$

which shows that the $(i + j)$-order convolved moments are determined by the same- or lower-order moments of the galaxy and the PSF kernel. The deconvolution procedure to estimate the galaxy moments is to invert the above hierarchical system of equations, starting from the zeroth order.

As mentioned above, it is necessary to introduce a weight function to avoid noise dominating the second-order moments outside the galaxy light profile. In this work, we adopt an elliptical Gaussian weight function with size $r_{\rm wf} = r_{\rm iso}$, where $r_{\rm iso}$ is the isophotal radius, defined as $r_{\rm iso} = \sqrt{A_{\rm iso}/\pi}$, following Georgiou et al. (2019a). The area $A_{\rm iso}$ of the galaxy's isophote is computed using the `ISOAREA_IMAGE` by SExtractor (Bertin & Arnouts 1996). The shape measurement procedure is the same as described in Georgiou et al. (2019a) and we point the interested reader to their Section 2 for a detailed description of the algorithm. In Appendix A1 we report our analysis of the measured shape bias for different setups, which led to our final choice reported above.

Using DEIMOS, we successfully measured the shapes of 96 863 galaxies from the `luminous` sample, $\sim 83\%$ of the corresponding density sample, and 152 832 shapes from the `dense` sample, roughly $\sim 88\%$ of its density sample. The shape measurements mainly fail[9] for the faintest galaxies in the sample.

3.3.2 *lens*fit

The second shape catalogue is obtained using the self-calibrating version of *lens*fit (Miller et al. 2013), described in more detail in Fenech Conti et al. (2017). It is a likelihood-based model-fitting method that fits a PSF-convolved two-component bulge and disk galaxy model. This is applied simultaneously to the multiple exposures in the KiDS-1000 r-band imaging, to get an ellipticity estimate for each galaxy.

*lens*fit provides shapes for 84 785 galaxies from the `luminous` sample (72% of the density sample), and for 121 500 galaxies from the `dense` sample (70% of the density sample). The lower completeness with respect to DEIMOS is largely explained by the fact that *lens*fit has been optimised for cosmic shear studies, where the signal is maximised for high-redshift galaxies, which are typically small and faint. Whilst *lens*fit could determine ellipticity measurements for the large bright galaxies with `MAG_AUTO` < 20, this model-fitting algorithm becomes prohibitively slow given the large number

[9] We only considered shapes with `flag_DEIMOS==0000`, corresponding to measurements that do not raise any flag (see Georgiou et al. 2019a).

of pixels that these bright galaxies span. Therefore, the *lens*fit catalogue only contains galaxies fainter than `MAG_AUTO` > 20 (hence the sharp cut-off in apparent magnitude in Fig. 3.2). It performs better than DEIMOS for relatively faint and low signal-to-noise (S/N) galaxies. As these are preferentially found at higher redshifts, this also explains the different redshift distributions, as illustrated in Fig. 3.1.

3.3.3 Image simulations

We want to measure the shapes of galaxies from images that are corrupted by noise and blurred by the atmosphere and telescope optics. These bias the inferred shapes and thus need to be carefully corrected for. Although both DEIMOS and *lens*fit are designed to do so, residual biases remain. These can be expressed as (Heymans et al. 2006)

$$\epsilon_i^{\rm obs} = (1 + m_i)\epsilon_i^{\rm true} + c_i \ , \tag{3.5}$$

with $i \in \{1, 2\}$ the ellipticity components introduced in 3.2. Here $\epsilon_i^{\rm true}$ is the true ellipticity, while $\epsilon_i^{\rm obs}$ is the output of the shape measurement algorithm; m_i is the multiplicative bias and c_i is the additive bias. Differently from what is done in lensing studies (e.g. Kannawadi et al. 2019), here we calibrated the ellipticity rather than the shear. Our aim is to determine the biases in our shape measurements using realistic image simulations, with a precision that is better than the statistical error on our IA signal.

We stress that although it is important to start with an algorithm that does not lead to a large bias in the first place, what matters the most is to calibrate the residual bias on realistic image simulations in order to properly account for galaxy blending and the different observing conditions (Hoekstra et al. 2017; Kannawadi et al. 2019; Samuroff et al. 2018; MacCrann et al. 2020). We use dedicated image simulations generated with the COllege pipeline (COSMOS-like lensing emulation of ground experiments; Kannawadi et al. 2019). These simulations reproduce the observations from the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007), for which we have both KiDS imaging (KiDS-COSMOS) and deeper images from the *Hubble* Space Telescope (HST). We use the HST observations to generate our input catalogue and simulate the KiDS observations by varying the observation conditions. Under the assumption that COSMOS is representative of our galaxy sample (in practice we only require that it covers the S/N and size parameter space, while we do not need the galaxy distributions to match) we study the m–bias properties of the LRGs in our KiDS-COSMOS

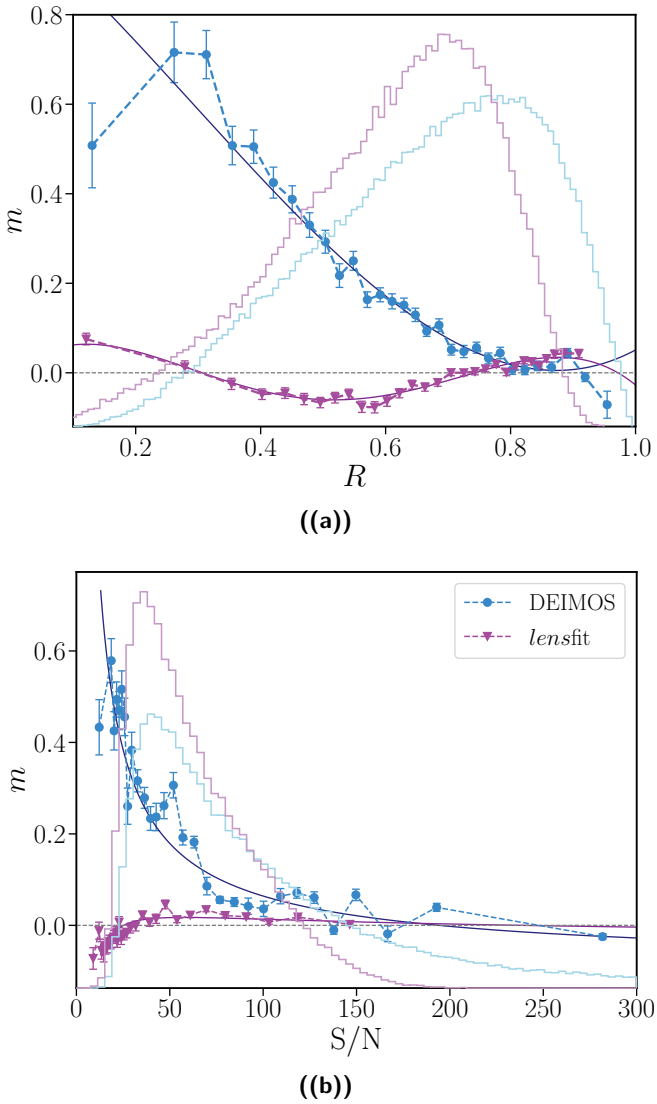

Figure 3.3: Average multiplicative bias, $m = (m_{\epsilon_1} + m_{\epsilon_2})/2$, as a function of (a) the galaxy resolution, R, and (b) the signal-to-noise ratio, S/N. Each point is measured on the same number of simulated galaxies and the error bars are estimated using bootstraps. For a comparison we also display in the background the weighted distribution of the two definitions of R and the S/N in the real data for the `dense` shape samples (pink: *lens*fit; blue: DEIMOS). The solid lines show the polynomial fit to $m(R)$ and $m(\mathrm{S/N})$, which guided the construction of the two-dimensional bias surface.

field and use the bias model obtained from this set of galaxies to calibrate our full sample.

The image simulations used in this work differ slightly from those presented in Kannawadi et al. (2019) because we required a larger number of simulated LRGs for our calibration. To achieve this, we adopted the ZEST catalogue (Zurich Estimator of Structural Type; Scarlata et al. 2007; Sargent et al. 2007) for the input galaxy parameters. We generated 52 KiDS-like images by varying the observing conditions and rotating the galaxies. We used 13 different PSF sets and four rotations per each image. Since our underlying galaxy selection is identical for both the *lens*fit and DEIMOS shape catalogues, we employed the same suite of simulations for both calibrations.

The shape measurement bias depends on the size, S/N, radial surface brightness profile and ellipticity of the galaxy, as well as the observing conditions. Of these, the size and S/N are the most relevant, and we use these to capture the dependence of the bias for our set of simulated galaxies. Rather than the intrinsic size of the galaxy, we use a proxy for how well it is resolved: R quantifies the relative size of the PSF compared to the size of the galaxy. Here, we adopt two slightly different definitions, depending on the shape algorithm employed. For DEIMOS we use

$$R^{\rm DEIMOS} = 1 - \frac{T^{\rm PSF}}{T^{\rm gal}}\,, \tag{3.6}$$

where $T^{\rm PSF} = Q_{20}^{\rm PSF} + Q_{02}^{\rm PSF}$ and $T^{\rm gal} = Q_{20}^{*\rm gal} + Q_{02}^{*\rm gal}$, where $Q_{ij}^{*\rm gal}$ are the unweighted moments of the PSF-convolved surface brightness profile (see Eqs. 3.4 and 3.1). In the case of *lens*fit we use

$$R^{lens\rm fit} = 1 - \frac{r_{\rm PSF}^2}{\left(r_{\rm ab}^2 + r_{\rm PSF}^2\right)}\,, \tag{3.7}$$

where $r_{\rm PSF}^2 = \sqrt{P_{11}P_{22} - P_{12}^2}$ and $r_{\rm ab} = r_{\rm e}\sqrt{q}$. Here, P_{ij} are the *lens*fit PSF weighted quadrupole moments (see Eq. (2) in Giblin et al. 2021), measured with a circular Gaussian function of size 2.5 pixels; $r_{\rm e}$ is the half-light radius measured along the major axis of the best-fit elliptical profile by *lens*fit, which is an estimate of the true galaxy size before PSF-convolution, while q is the axis ratio, such that $r_{\rm ab}$ is the azimuthally averaged size of the galaxy. As we can see, R can in practice only assume values between 0 and 1, where 1 corresponds to galaxies with sizes that are much larger than the PSF.

We evaluate the multiplicative bias m in bins of S/N and R that contain an equal number of galaxies and the error bars are computed using 500 bootstrap realisations. The resulting biases are presented in Fig. 3.3 for both *lens*fit and DEIMOS. We find that the two components $\epsilon_{1,2}$ show similar dependencies, and, therefore, we calibrate the bias for the two components jointly. The additive bias for both components is consistent with zero, and thus we do not consider it further in our calibration.

For both $m(\mathrm{S/N})$ and $m(R)$, we find that *lens*fit has a small bias and thus also our correction is small; in general, it performs better than DEIMOS for poorly resolved galaxies and low S/N. It is, however, prohibitively slow when measuring shapes for large galaxies, limiting the *lens*fit sample to galaxies with $m_r > 20$. In contrast, DEIMOS shows a large bias for low values of R: the galaxy size correlates with its ellipticity, and we find that removing the highly elliptical galaxies significantly reduces the bias. However, once we calibrate the shapes of those galaxies, we recover a very similar signal for the full shape sample and the one cut in ellipticity. Similarly, we have also tested that adding inverse-variance weights to account for these noisy galaxies does not significantly improve our signal. This motivates our choice to keep all galaxies in our sample and not to introduce additional weighting; we assume that the measurements are dominated by shape noise only.

We can see that $m(R)$ for both DEIMOS and *lens*fit is well described by a polynomial curve, which we truncate at degree 3 and 4, respectively, while $m(\mathrm{S/N})$ is well described by the expansion: $d(\mathrm{S/N}) = d_1/\sqrt{\mathrm{S/N}} + d_2/(\mathrm{S/N})$. We combine the two individual bias dependencies into a single bias surface as detailed in Appendix A1. The specific functional forms for the two shape methods differ to better adapt the surface to our observed bias. We use these empirical relations to infer the m-bias associated with each galaxy, given its S/N and R.

To ensure that our empirical correction performs well on our sample, we selected sets of galaxies from the image simulations that resemble our LRG samples by reproducing the observed distributions in S/N and R. We measured the residual biases for these samples, defined as the difference in the estimated m-bias (inferred using our model for the bias) and the bias measured directly from the simulations for the given set of galaxies. For the DEIMOS shape method, we find an average residual of -0.002 ± 0.007 for the `dense`-like sample, while this is -0.002 ± 0.008 for the `luminous`-like sample. Similarly, in the case of *lens*fit the residuals for the `luminous`-like and `dense`-like galaxies are, respectively, -0.0014 ± 0.0013 and -0.0019 ± 0.0020.

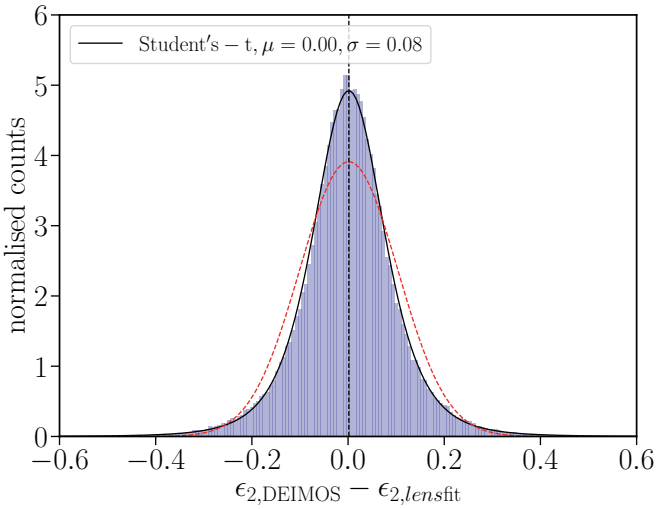

Figure 3.4: Histogram of the difference of the ϵ_1 component of the ellipticity measured by the two shape measurement algorithms, *lens*fit and DEIMOS, on a common sub-sample of galaxies, after applying the $m-$bias correction as described in the text. The ϵ_2 component shows the same behaviour. The distribution is more peaked than a Gaussian (red dashed line) and is best described by a Student's $t-$distribution with $\nu = 4.3$, and a width $\sigma = 0.08$ with zero mean (black solid line).

As we will see later, this is much smaller than the uncertainty in the IA measurements: the average bias introduced by the shape measurement process is subdominant and does not affect our best estimate of the IA amplitude.

The LRGs are relatively bright and we thus expect the shape measurements to be shape noise-dominated. This also implies that the DEIMOS and *lens*fit measurements are correlated. To quantify this, we show the distribution of the difference between the $m-$corrected ellipticities measured by the two algorithms in Fig. 3.4. The distribution is more peaked than a Gaussian, and well described by a Student's $t-$distribution centred on zero, with $\nu = 4.30$ (degrees of freedom) and with scale parameter $\sigma = 0.08$. This is to be compared to the intrinsic ellipticity of galaxies, which is about $\epsilon_{\mathrm{rms}} = 0.12$ based on DEIMOS measurements for galaxies with apparent magnitude $m_r < 20$. It is interesting to note that our sample is considerably rounder than a typical cosmic shear sample, as expected for an LRG sample (see for example van Uitert et al. 2012); this implies that it might be affected differently by a weighting scheme in a lensing analysis. The differences between the DEIMOS and *lens*fit measurements are caused by differences in how each method deals with noise in the images.

3.4 Correlation function measurements

We measured the IA signal using the two-points statistic w_{g+}, defined as the projection along the line-of-sight of the cross-correlation between galaxy positions and galaxy shapes. It measures the tendency of galaxies to point in the direction of another galaxy as a function of their comoving transverse separation, r_p, and comoving line-of-sight separation, Π. To quantify the alignment signal in our data, we employed the estimator presented in Mandelbaum et al. (2006)[10],

$$\hat{\xi}_{g+}(r_p, \Pi) = \frac{S_+D - S_+R_D}{R_S R_D}, \tag{3.8}$$

where R_D and R_S are catalogues of random points designed to reproduce the galaxy distribution of the density and shape samples, respectively. We indicate with D the density sample that provides the galaxy positions, while S_+ is the shape sample, such that the quantity

$$S_+D = \sum_{i \neq j} \gamma_+(i|j), \tag{3.9}$$

gives us the tangential shear component of the galaxy pair (i, j), $\gamma_+(i|j)$, where i is extracted from the shape sample and j from the density sample. γ_+, in turn, is defined as

$$\gamma_+(i|j) = \frac{1}{\mathcal{R}} \Re \left[\epsilon_i \exp(-2i\phi_{ij}) \right], \tag{3.10}$$

where $\Re$ denotes the real part; ϵ_i is the complex ellipticity associated with the galaxy i, $\epsilon_i = \epsilon_{1,i} + i\epsilon_{2,i}$, whose components 1,2 are measured by the shape measurement algorithms presented in Sect. 3.3; ϕ_{ij} is the polar angle of the vector that connects the galaxy pair; $\mathcal{R} = \partial\epsilon/\partial\gamma$ is the shear responsivity and it quantifies by how much the ellipticity changes when a shear is applied: for an ensemble of sources, $\mathcal{R} = 1 - \epsilon^2_{\rm rms}$.

The galaxy clustering signal is computed with the standard estimator (Landy & Szalay 1993),

$$\hat{\xi}_{gg}(r_p, \Pi) = \frac{DD - 2DR_D - R_DR_D}{R_DR_D}. \tag{3.11}$$

[10] Instead of normalising by $R_S R_D$, we actually normalise by the density - randoms vs. shapes pair count, $R_D D_S$. This significantly speeds up the computation and has been tested to have negligible impact (Johnston et al. 2019).

To measure our clustering and IA signals, we used uniform random samples that reproduce the KiDS footprint, accounting for the masked regions; to these we assigned redshifts randomly extracted from the galaxy unconditional photometric redshift distributions. For each sample, we constructed the random sample to match their redshift distribution.

To account for the spatial variation in the survey systematics, we applied weights to the galaxies when computing the signal, as discussed in Vakili et al. (2020). These weights are designed to remove the systematic-induced variation in the galaxy number density across the survey footprint. For a detailed discussion of how the weights are generated and tested, we refer to Sect. 4 in Vakili et al. (2020). To capture the variation in the survey systematics along the line-of-sight, we split each sample into three redshift bins and assigned the weights to those sub-samples. We tested that this procedure does not induce a correlation between the galaxy weights and the redshifts themselves. We also verified that the impact of the weights is very small and can be neglected when considering the split in luminosity of the samples (see Sect. 3.6.1). We applied such weights to both the density and shape samples.

In this work, we measured the clustering and IA signals using an updated version of the pipeline presented in Johnston et al. (2019), which makes use of the publicly available software Treecorr (Jarvis et al. 2004)[11] for clustering correlations. ξ_{g+} and ξ_{gg} are then projected by integrating over the line-of-sight component of the comoving separation, Π,

$$\hat{w}_{gi}(r_p) = \int_{-\Pi_{max}}^{\Pi_{max}} d\Pi \, \hat{\xi}_{gi}(r_p, \Pi) \quad i = \{+, g\} \, . \tag{3.12}$$

The largest scales probed in this analysis are limited by the effective survey area ($\sim$ 777 deg^2). We set a maximum transverse separation of $60\,h^{-1}\mathrm{Mpc}$ and measure the signal in 10 logarithmically spaced bins, from $r_{p,min} = 0.2\,h^{-1}\mathrm{Mpc}$.

We performed the measurements for three different setups: we adopt $\Pi_{max} = 120\,h^{-1}\mathrm{Mpc}$ as the fiducial case, but repeated the analysis for $\Pi_{max} = 90\,h^{-1}\mathrm{Mpc}$ and $\Pi_{max} = 180\,h^{-1}\mathrm{Mpc}$ (see Appendix A5). We always bin our galaxies in equally spaced bins with $\Delta\Pi = 10\,h^{-1}\mathrm{Mpc}$. We observe an extended signal to $\Pi > 180\,h^{-1}\mathrm{Mpc}$, but the signal is comparable to the noise at those distances.

[11] https://github.com/rmjarvis/TreeCorr

Our choice of $\Pi_{\rm max}$ is conservative since the uncertainties in the photometric redshifts are $\sigma_z < 0.02(1+z)$ for both the `dense`and `luminous` samples (Vakili et al. 2020), and if we choose $\Pi_{\rm max}$ based on the 1σ uncertainty in the photometric redshifts (Joachimi et al. 2011), we could potentially reduce $\Pi_{\rm max}$ to $70\,h^{-1}{\rm Mpc}$. However, this might be too optimistic given that the error on σ_z increases with redshift. The choice of $\Pi_{\rm max}$ is motivated by two opposite necessities: to maximise the S/N, we want to minimise the amount of signal that we discard, whilst we also want to avoid adding uncorrelated pairs that would increase the noise. To find the best balance, we calculate the S/N of our signal as a function of $(r_{\rm p}, \Pi)$ by dividing the measured $w_{\rm gg}(r_{\rm p}, \Pi)$ by the root-diagonal of the jackknife covariance. We truncate at $\Pi_{\rm max}$ based on the 10 σ detection, which roughly corresponds to $\Pi_{\rm max} = 120\,h^{-1}{\rm Mpc}$. In addition to these considerations, there is a further motivation to limit the integral to modest line-of-sight separations: as discussed in Appendix A3, the contamination from galaxy-galaxy lensing has a shallower dependence on the line-of-sight separation; as we move along the Π direction, we see an increase in the contamination with a mild increase in the IA signal, until lensing dominates.

The error bars are computed via a delete-one jackknife re-sampling of the observed volume. The covariance matrix is constructed as

$$\mathrm{Cov}_{\rm jack.} = \frac{N-1}{N}\sum_{\alpha=1}^{N}(w^{\alpha}-\bar{w})(w^{\alpha}-\bar{w})^{\top}, \tag{3.13}$$

where w^{α} is the signal measured from jackknife sample α, while $\bar{w}$ is the average over N samples; $\top$ denotes the transpose of the vector.

The number of regions N is ultimately set by the size of the survey and the scales we aim to probe. A maximum value of $r_{\rm p} = 60\,h^{-1}{\rm Mpc}$ corresponds to an angular separation of ~ 8 degrees (`dense` sample) and ~ 6 degrees (`luminous` sample) at the lowest redshifts probed in the analysis. However, to increase the number of jackknife regions, we decided to set the minimum angular scale to 5 degrees, which strictly satisfies our requirement only for $z \gtrsim 0.2$. This is motivated by the fact that the majority of our galaxies are at high redshift and hence only $\lesssim 5\%$ of our galaxies have unreliable error estimates in the last $r_{\rm p}$–bin. The total number of jackknife regions that we are able to obtain for our samples is $N = 37$. We corrected our inverse covariance matrices, which enter into our likelihood estimations, as recommended in Hartlap et al. (2007): because of the presence of noise, the inverse of a covariance matrix obtained from a finite number of jackknife (or bootstrap) realisations is a biased estimator of the true inverse

covariance matrix.

3.5 Modelling

The linear alignment model (Catelan et al. 2001; Hirata & Seljak 2004) predicts a linear relation between the contribution to the shear induced by IA and the quadrupole of the gravitational field responsible of the tidal effect. This can be expressed as

$$\gamma^{\rm I} = (\gamma^{\rm I}_+, \gamma^{\rm I}_\times) = -\frac{C_1}{4\pi G}(\partial_x^2 + \partial_y^2, \partial_x\partial_y)\Phi_{\rm p}\,, \tag{3.14}$$

where the partial derivatives are with respect to comoving coordinates and provide the tangential and cross components of the shear with respect to the x-axis; $\Phi_{\rm p}$ is the gravitational potential at the moment of galaxy formation, assumed to take place during the matter-dominated era (Catelan et al. 2001); C_1 is a normalisation constant and G is the gravitational constant.

Using Eq. (3.14), by correlating the intrinsic shear with itself or with the matter density field δ, we can construct the relevant equations for the IA correlation functions (Hirata & Seljak 2004). In Fourier space, the matter density-shear power spectrum (δI) becomes

$$P^{\rm LA}_{\delta{\rm I}}(k,z) = A_{\rm IA}C_1\rho_{\rm c}\frac{\Omega_m}{D(z)}P^{\rm lin}_{\delta\delta}(k,z)\,. \tag{3.15}$$

Here, $D(z)$ is the linear growth factor, normalised to unity at $z = 0$, $\rho_{\rm c}$ is the critical density of the Universe today, and $P^{\rm lin}_{\delta\delta}$ is the linear matter power spectrum. We set $C_1 = 5\times10^{-14}h^{-2}M_\odot^{-1}{\rm Mpc}^3$ based on the IA amplitude measured at low redshifts using SuperCOSMOS (Hartlap, Simon & Schneider Bro), which is the standard normalisation for IA power spectra.

Galaxies are biased tracers of the matter density field, and at large scales this relation is linear, $\delta_{\rm g} \sim b_{\rm g}\delta$. We can thus relate the galaxy position–intrinsic shear power spectrum to the matter density–intrinsic shear power spectrum via the galaxy bias $b_{\rm g}$:

$$P^{\rm LA}_{\rm gI}(k,z) = b_{\rm g}P^{\rm LA}_{\delta{\rm I}}(k,z)\,, \tag{3.16}$$

which is the power spectrum of interest for our analysis.

A successful modification of the LA model replaces the linear matter power spectrum in Eq. (3.15) with the non-linear one, to account for the non-linearities arising at intermediate scales (Bridle & King 2007). This

so-called NLA model was succesfully employed in a number of studies (e.g. Blazek et al. 2011; Joachimi et al. 2011) and here we follow the same approach to model our signal. More sophisticated treatments of the IA signal, which include the modelling of the mildly or fully non-linear scales, have been developed in the last decade (Schneider & Bridle 2010; Blazek et al. 2019; Fortuna et al. 2021a), but given the scales probed in our analysis (see Sect. 3.5.3) and the homogeneous characteristics of the galaxy population studied, the NLA model provides a sufficient description for this work. Unless stated otherwise, in the following we always assume the NLA model as our reference choice. To generate the linear matter power spectrum we use CAMB[12] (Lewis et al. 2000; Lewis & Bridle 2002), while the non-linear modifications are computed using Halofit (Smith et al. 2002) with the implementation presented in Takahashi et al. (2012). In the rest of the paper, we simply refer to the non-linear matter power spectrum as $P_{\delta\delta}(k, z)$.

3.5.1 Incorporating the photometric redshift uncertainty into the model

The use of photometric redshifts results in an uncertainty in the estimated distance of the galaxies, which has to be included in the model. In particular, if we express the correlation function $\xi_{\rm gI}$ in terms of the two components of the galaxy separation vector $\mathbf{r}$, $(r_{\rm p}, \Pi)$, we can map the redshift probability distribution into the probability that the true values of $r_{\rm p}$ and Π correspond to their photometric estimates. Here, we follow the approach derived in Joachimi et al. (2011) and use their approximated expression,

$$\xi^{\rm ph}_{\rm gI}(\bar{r}_{\rm p}, \bar{\Pi}, \bar{z}_{\rm m}) = \int \frac{{\rm d}\ell\ell}{2\pi} J_2\left(\ell\theta(\bar{r}_{\rm p}, \bar{z}_{\rm m})\right) C_{\rm gI}\left(\ell; \bar{z}_1(\bar{z}_{\rm m}, \bar{\Pi}), \bar{z}_2(\bar{z}_{\rm m}, \bar{\Pi})\right) \ . \quad (3.17)$$

The observables are: $\bar{z}_1$ and $\bar{z}_2$, the photometric redshift estimates of the pair of galaxies for which we are measuring the correlation, and their angular separation θ. These can be related to $(\bar{r}_{\rm p}, \bar{\Pi}, \bar{z}_{\rm m})$, through the approximate relations

$$z_{\rm m} = \frac{1}{2}(z_1 + z_2) \ , \quad (3.18)$$

$$r_{\rm p} \approx \theta\chi(z_{\rm m}) \ , \quad (3.19)$$

$$\Pi \approx \frac{c}{H(z_{\rm m})}(z_2 - z_1) \ , \quad (3.20)$$

[12] https://camb.info

where $\chi(z_\mathrm{m})$ and $H(z_\mathrm{m})$ are, respectively, the comoving distance and the Hubble parameter at redshift z_m, and c is the speed of light.

The conditional redshift probability distributions are incorporated into the angular power spectrum C_gI, which can be expressed in terms of the three-dimensional power spectrum $P_\mathrm{gI}(k, z)$,

$$C_\mathrm{gI}(\ell, \bar{z}_1, \bar{z}_2) = \int_0^{\chi_\mathrm{hor}} \mathrm{d}\chi' \frac{p_n(\chi'|\chi(\bar{z}_1))p_\epsilon(\chi'|\chi(\bar{z}_2))}{\chi'^2} \times P_\mathrm{gI}\left(\frac{\ell+1/2}{\chi'}, z(\chi')\right) \quad (3.21)$$

where we have implicitly assumed the flat-sky and Limber approximations, and n and ϵ indicate the density and shape sample respectively. $p(\chi'|\chi)$ are the conditional comoving distance probability distributions, which are related to the redshift distributions via $p(\chi'|\chi)\mathrm{d}\chi = p(z|\bar{z})\mathrm{d}z$. When computing our predictions, we bin our photometric data and compute the corresponding $p(z|\bar{z}) \equiv p(z_\mathrm{spec}|z_\mathrm{phot})$ per each bin; z_1 and z_2 in Eq. (3.18) corresponds to the mean values of the probability distribution with z_1 being the mean of the i-th bin and z_2 of the j-th bin. In Appendix A2 we show the redshift distributions entering our analysis. We refer the interested reader to appendices A.2 and A.3 in Joachimi et al. (2011) for the full derivation of equation 3.21. The exact same formalism can then be applied to the clustering signal, where $C_\mathrm{gI} \to C_\mathrm{gg}$, $J_2 \to J_0$ and the redshift distributions are those corresponding to the density sample.

The projected correlation functions $w_\mathrm{g+}$ and w_{gg} can then be obtained as:

$$w_\mathrm{g+}(r_\mathrm{p}) = \int \mathrm{d}\bar{\Pi} \int \mathrm{d}z_\mathrm{m} \mathcal{W}(\bar{z}_\mathrm{m}) \xi^\mathrm{ph}_\mathrm{gI}(\bar{r}_\mathrm{p}, \bar{\Pi}, \bar{z}_\mathrm{m}) \quad (3.22)$$

and

$$w_\mathrm{gg}(r_\mathrm{p}) = \int \mathrm{d}\bar{\Pi} \int \mathrm{d}z_\mathrm{m} \mathcal{W}(\bar{z}_\mathrm{m}) \xi^\mathrm{ph}_\mathrm{gg}(\bar{r}_\mathrm{p}, \bar{\Pi}, \bar{z}_\mathrm{m}) \,, \quad (3.23)$$

where the redshift window function $\mathcal{W}(z)$ is defined as (Mandelbaum et al. 2011):

$$\mathcal{W}(z) = \frac{p_i(z)p_j(z)}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z} \left[\int \mathrm{d}z \frac{p_i(z)p_j(z)}{\chi^2(z)\mathrm{d}\chi/\mathrm{d}z}\right]^{-1} , \quad (3.24)$$

where $p_{i,j}(z)$ with $i, j \in S, D$ are now the unconditional redshift distributions for the shape and density samples, and $\chi(z)$ is the comoving distance to redshift z.

3.5.2 Contamination to the signal

All possible two-point correlations between galaxy shapes and positions contribute to the estimator in Eq. (3.12). Following the notation in Joachimi & Bridle (2010), here we consider: the correlation between the intrinsic shear and the galaxy position (g+), which is the quantity we aim to constrain; but also the correlation between gravitational shear and galaxy position, sourced by the galaxy lensing of a background galaxy by a foreground galaxy (gG); and the apparent modification of the galaxy number counts due to the effect of lensing magnification, which affects both the correlations with the intrinsic shear and the gravitational shear (mI and mG).

Among these effects, galaxy-galaxy lensing is the main contaminant to our signal. While IA requires physically close galaxies, galaxy-galaxy lensing occurs between galaxies at different redshifts. This implies that the level of contamination depends on our ability to select close pairs of galaxies, which ultimately depends on the photometric redshift precision. For this reason, the width and the tails of the redshift distributions play an important role in the amount of contamination. Since our $p(z^{\rm spec}|z^{\rm phot})$ are quite narrow (see Appendix A2) we do not expect this to be a major effect in our data. Nevertheless, we fully model both lensing and magnification effects, and account for them when interpreting the signal. We note that the sign of the gI and gG terms are opposite, such that adding the lensing to the model allows us to remove its suppressing contribution and capture the true IA signal.

It is convenient to write the various correlations in terms of the projected angular power spectra: indicating with n the density sample (that provides the galaxy positions) and with ϵ the shape sample, we have

$$C^{(ij)}_{n\epsilon}(\ell) = C^{(ij)}_{\rm gI}(\ell) + C^{(ij)}_{\rm gG}(\ell) + C^{(ij)}_{\rm mI}(\ell) + C^{(ij)}_{\rm mG}(\ell)\,, \tag{3.25}$$

where, in a flat cosmology, these read

$$C^{(ij)}_{\rm gG}(\ell) = b_g \int_0^{\chi_{\rm hor}} {\rm d}\chi \frac{p_n^{(i)}(\chi) q_\epsilon^{(j)}(\chi)}{\chi^2} P_{\delta\delta}\left(\frac{\ell+1/2}{\chi}, \chi\right)\,, \tag{3.26}$$

$$C^{(ij)}_{\rm mI}(\ell) = 2(\alpha^{(i)} - 1) C^{(ij)}_{\rm IG}(\ell), \tag{3.27}$$

and

$$C^{(ij)}_{\rm mG}(\ell) = 2(\alpha^{(i)} - 1) C^{(ij)}_{\rm GG}(\ell)\,. \tag{3.28}$$

Here $\alpha^{(i)}$ is the slope of the faint-end logarithmic luminosity function[13]. The lensing weight function, q_X, $X \in \{n, \epsilon\}$ is defined as

$$q_X(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{\chi}{a(\chi)}\int_0^{\chi_{\rm hor}} \mathrm{d}\chi' p_X(\chi')\frac{\chi'-\chi}{\chi'} \,. \tag{3.29}$$

$C_{\rm IG}^{(ij)}$ is the intrinsic-shear power spectrum. It models the correlation between the shearing of source galaxies by a foreground matter overdensity and the simultaneous IA of galaxies located near that overdensity:

$$C_{\rm IG}^{(ij)}(\ell) = \int_0^{\chi_{\rm hor}} \mathrm{d}\chi \frac{p_n^{(i)}(\chi) q_\epsilon^{(j)}(\chi)}{\chi^2} P_{\delta \rm I}\left(\frac{\ell+1/2}{\chi}, \chi\right) \,; \tag{3.30}$$

$C_{\rm GG}^{(ij)}$ is instead defined as:

$$C_{\rm GG}^{(ij)}(\ell) = \int_0^{\chi_{\rm hor}} \mathrm{d}\chi \frac{q_n^{(i)}(\chi) q_\epsilon^{(j)}(\chi)}{\chi^2} P_{\delta\delta}\left(\frac{\ell+1/2}{\chi}, \chi\right) \,. \tag{3.31}$$

We note that with respect to the usual shear power spectrum, we require here that one of the samples refers to the density sample, n.

To account for these sources of contamination in the fit, we replace $\xi_{\rm gI}$ with $\xi_{{\rm n}\epsilon}$, which can be obtained from Eq. (3.25). The prediction for $\xi^{\rm obs}$ is then used to constrain the measured signal $\hat{w}_{\rm g+}$. In Appendix A3 we expand further on the impact of lensing on our measurements, while in Appendix A4 we describe our strategy to measure the values of $\alpha^{(i)}$ in our data.

3.5.3 Likelihoods

We perform the fits to the data using a Markov Chain Monte Carlo (MCMC) that samples the multi-dimensional parameter posterior distributions and finds the set of parameters that maximise the likelihood. We assume a Gaussian likelihood of the form $\mathcal{L} \propto \exp(-\chi^2/2)$, where

$$\chi^2 = \chi^2_{w_{\rm gg}} + \chi^2_{w_{\rm g+}} \tag{3.32}$$

and we simultaneously fit for the galaxy bias, $b_{\rm g}$ and the IA amplitude, $A_{\rm IA}$.

[13]Formally, the magnification of the *lens*fit sample is also affected by the slope of the luminosity function at the bright end of $m_r = 20$. We ignore such complexity: we find magnification to be a subdominant effect for the faint distant galaxies, thus the contribution of low-redshift galaxies is expected to be negligible for our analysis.

To correct for the effects of a partial-sky survey window, we also introduce an integral constraint, IC, when modelling the clustering, signal,

$$w_{\rm gg} \rightarrow w_{\rm gg} + {\rm IC} \, . \tag{3.33}$$

This term, which becomes important only on large scales, has the function of capturing the bias that arises from a mis-estimation of the global mean density (Roche & Eales 1999). We treat this term as a nuisance parameter, such that our parameter vector reads

$$\lambda = \{b_g, A_{\rm IA}; {\rm IC}\} \, . \tag{3.34}$$

We limit our fits to the quasi-linear regime, $r_{\rm p} > 6\,h^{-1}{\rm Mpc}$, to ensure that the linear bias approximation is satisfied and the IA signal is well described by the NLA model. To perform our fits, we make use of the Emcee (Foreman-Mackey et al. 2013) package as implemented in the cosmology software CosmoSIS[14] (Zuntz et al. 2015). When analysing the chains, we exclude the first 30% of samples for a burn-in phase.

3.6 Results

The left panels in Fig. 3.5 show the measurements of the projected position-shape correlation function $w_{\rm g+}$ for the `luminous` (top panel) and `dense` (bottom panel) samples. We present results for both the *lens*fit (dark green triangles) and DEIMOS (light green squares) shape catalogues. As described in Sect. 3.5.3, we simultaneously fit the IA and the clustering signals. We show the resulting best-fit models to measurements with $r_{\rm p} > 6\,h^{-1}{\rm Mpc}$ of $w_{\rm g+}$ and $w_{\rm gg}$ as solid lines in the figures. The estimates from the two shape measurement algorithms are fit independently, but given that the corresponding clustering signal is the same, here we only show the best-fit curve for the DEIMOS fit. The clustering measurements use the full density samples, and thus do not rely on a successful shape measurement.

We observe similar signals for the DEIMOS and *lens*fit samples, with the *lens*fit measurements having a lower S/N, because of the lack of shape measurements for galaxies with $m_r < 20$. We note that we do not necessarily expect to observe the same signal, because DEIMOS contains more bright, low-redshift galaxies, whereas the *lens*fit sample includes fainter, distant galaxies (see Figs. 3.1 and 3.2). If the alignment signal depends

[14] `http://bitbucket.org/joezuntz/cosmosis/wiki/Home`

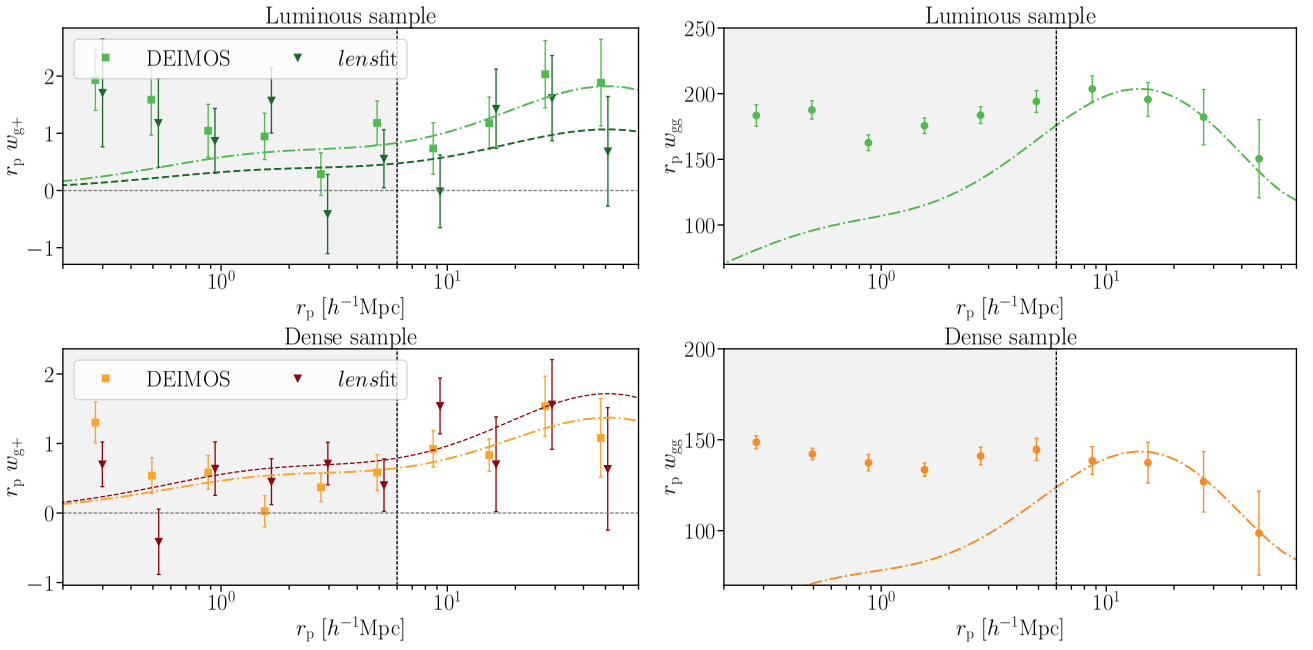

Figure 3.5: Projected correlation functions (IA and clustering signal) measured in this work and the best-fit curve predicted by our model. *Left:* Projected position-shape correlation function, w_{g+}, measured for our `luminous` (top panel) and `dense` (bottom panel) samples. We show results for shapes measured with DEIMOS (light squares) and *lens*fit (dark triangles). The best-fit models to the data with $r_p > 6\,h^{-1}$Mpc (indicated by the vertical dashed line), are shown as well, with the same colour scheme (DEIMOS: dash-dotted lines, *lens*fit: dashed lines). For clarity, the *lens*fit results have been slightly offset horizontally. *Right:* Projected clustering signal, w_{gg}, of the `dense` and `luminous` samples. The dot-dashed lines corresponds to the best-fit models. As we do not include a scale-dependent bias in our model, the mismatch between data and prediction at small scales is expected.

Table 3.1: Properties of the individual galaxy samples used in our analysis and the corresponding best-fit galaxy bias (b_g) and IA amplitude ($A_{\rm IA}$) as constrained by our model.

Samples	$\langle z \rangle$	$N_{\rm D}$	$N_{\rm S}$	$[L_{\rm min}, L_{\rm max}]$	$\langle L \rangle / L_0$	$b_{\rm g}$	$A_{\rm IA}$	$\chi^2_{\rm red}$
DEIMOS								
`dense`	0.44	173 445	152 832		0.38	$1.59^{+0.04}_{-0.04}$	$3.69^{+0.66}_{-0.65}$	0.78
`luminous`	0.54	117 001	96 863		0.64	$2.06^{+0.04}_{-0.04}$	$4.03^{+0.81}_{-0.79}$	1.19
`D1`	0.41	173 445	39 108	$[0.09, 1.13]$	0.21	$1.60^{+0.04}_{-0.04}$	$3.02^{+1.53}_{-1.48}$	1.00
`D2`	0.42	173 445	39 322	[1.13, 1.43]	0.27	$1.60^{+0.04}_{-0.04}$	$1.21^{+1.63}_{-1.64}$	0.91
`D3`	0.43	173 445	39 229	[1.43, 1.92]	0.35	$1.59^{+0.04}_{-0.04}$	$4.11^{+1.48}_{-1.48}$	1.05
`D4`	0.45	173 445	19 333	[1.92, 2.81]	0.49	$1.59^{+0.04}_{-0.04}$	$3.02^{+2.37}_{-2.33}$	1.52
`D5`	0.45	173 445	19 235	$\geqslant 2.81$	0.89	$1.59^{+0.04}_{-0.04}$	$8.39^{+1.04}_{-1.30}$	0.47
`L1`	0.53	117 001	48 588	$[0.29, 2.66]$	0.46	$2.06^{+0.04}_{-0.04}$	$1.80^{+0.96}_{-0.95}$	1.17
`L2`	0.55	117 001	24 208	[2.66, 3.51]	0.65	$2.06^{+0.04}_{-0.04}$	$4.95^{+1.24}_{-1.21}$	1.19
`L3`	0.56	117 001	24 067	$\geqslant 3.51$	1.00	$2.06^{+0.04}_{-0.04}$	$5.71^{+1.57}_{-1.60}$	2.03
*lens*fit								
`dense`	0.49	173 445	121 500		0.33	$1.60^{+0.04}_{-0.04}$	$4.94^{+1.24}_{-1.22}$	1.52
`luminous`	0.63	117 001	84 785		0.59	$2.06^{+0.04}_{-0.04}$	$2.95^{+1.49}_{-1.42}$	1.54
DEIMOS + *lens*fit								
Z1 ($z \leqslant 0.585$)	0.44	56 754	56 754		0.63	$2.01^{+0.06}_{-0.06}$	$3.84^{+1.10}_{-1.06}$	0.22
Z2 ($z > 0.585$)	0.70	57 613	57 613		0.61	$2.39^{+0.08}_{-0.08}$	$3.97^{+2.02}_{-2.04}$	2.43

The galaxy properties are summarised by: the mean redshift, $\langle z \rangle$; the number of galaxies in the density (shape) sample, $N_{\rm D}$ ($N_{\rm S}$); the mean luminosity in terms of a pivot luminosity $L_0 = 4.6 \times 10^{10} h^{-2} L_\odot$; the bias, $b_{\rm g}$. To compute the ratio $\langle L \rangle / L_0$, we only consider the galaxies in the corresponding shape sample. For our $L-$cuts sub-samples, we also provide the range in luminosity they probe, $[L_{\rm min}, L_{\rm max}]$, in units of $10^{10} h^{-2} L_\odot$. Similarly, we provide in brackets the cut adopted to split our sample in two redshift bins. When cross-correlating different samples, N_D refers to the density sample used in the correlation and the bias is the best-fit bias of the density tracer as obtained for that given measurement. All measurements are performed assuming $\Pi_{\rm max} = 120\, h^{-1}{\rm Mpc}$. Since the best-fit parameters and the medians of the marginal posterior distributions are in agreement, we quote the marginal values, while the χ^2 refers to the maximum likelihood. In all cases, the degrees of freedom are 5; the $p-$values are all above 0.03, with the majority of them being in the range 0.3-0.7.

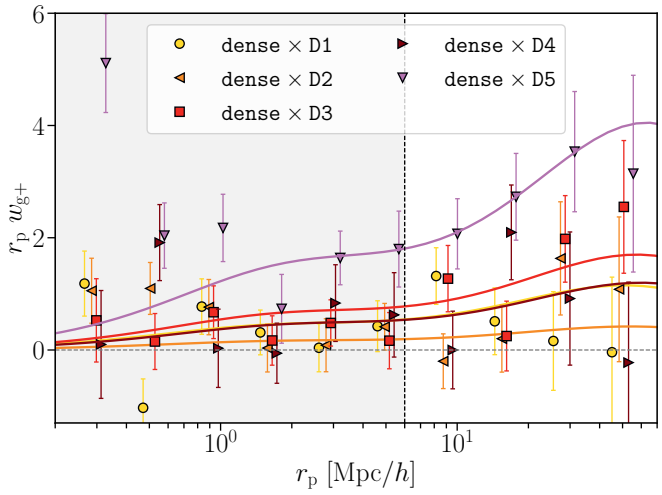

Figure 3.6: Projected correlation function, $w_{\mathrm{g+}}$, measured for the different cuts in luminosity of the DEIMOS `dense` sample. The best-fit curves are plotted on top of the data points, and the fits are performed for $r_\mathrm{p} > 6\,h^{-1}\mathrm{Mpc}$. All but the yellow points have been slightly offset horizontally; to better visualise the goodness of fit, the corresponding best-fit curves have been displaced accordingly.

on luminosity or redshift, the two shape samples would give different signals. In Appendix A6 we restrict the comparison to the sample of galaxies with shape measurements from both methods, and find that the average difference $\langle r_\mathrm{p} \Delta w_{\mathrm{g+}} \rangle = 0.003 \pm 0.13$ is negligible, especially compared to the amplitude of the IA signal quantified as $\langle r_\mathrm{p} w_{\mathrm{g+}} \rangle = 0.90 \pm 0.17$ (DEIMOS shapes; see Appendix A6 for details).

We also show the models that provide the best-fit to the combined w_{gg} and $w_{\mathrm{g+}}$ measurements in Fig. 3.5, and report the values for the bias b_g and IA amplitude A_IA in Table 3.1. The results for DEIMOS and *lens*fit are consistent.

Our constraints on the galaxy bias of the `dense` and `luminous` samples are in broad agreement with the values presented in Vakili et al. (2020): We find a larger bias for the `luminous` sample than for the `dense` one, as expected by its higher luminosity and the higher redshift baseline.

3.6.1 Luminosity dependence

Previous studies of LRGs (Joachimi et al. 2011; Singh et al. 2015) have found a significant dependence of their IA signal with luminosity, with

more luminous galaxies showing stronger alignments. On average our LRG sample probes somewhat lower luminosities than those earlier studies, but the overlap with these earlier works also enables a direct comparison. Thanks to the large range in luminosity it covers, the `dense` sample is particularly suited to explore the dependence with luminosity. To do so, we use the DEIMOS shape catalogue[15] and split the `dense` LRG galaxies in five sub-samples: `D1`, `D2`, and `D3`, correspond to the lowest three quartiles in luminosity; the remaining two, `D4` and `D5`, are obtained by splitting the highest luminosity quartile into two equally sized samples. The motivation to split the quartile with the highest luminosities is that it encompasses a very large range in luminosity, which complicates the interpretation if the signal depends on luminosity (see below). Relevant details for the sub-samples are listed in Table 3.1. We keep the `dense` and `luminous` samples separate, in order to better isolate the effect of the luminosity dependence from any redshift evolution of the sample itself. For instance, as listed in Table 3.1, the mean redshift of the sub-samples increases somewhat from `D1` to `D5`.

We cross-correlate the DEIMOS shape catalogues for the individual sub-samples with the positions of galaxies in the full `dense` sample. In this way, we can disentangle the luminosity dependence of the IA signal from the luminosity dependence of the density tracer (brighter galaxies are typically found in denser environments). The measurements and the best-fit models are presented in Fig. 3.6. In Table 3.1 we list the best-fit values for the galaxy bias $b_{\rm g}$ and IA amplitude $A_{\rm IA}$, as well as the reduced χ^2, as before, using the measurements for $r_{\rm p} > 6\,h^{-1}{\rm Mpc}$. We also show the measurements in Fig. 3.7 as orange stars as a function of L/L_0, where $L_0 = 4.6 \times 10^{10} h^{-2} L_\odot$.

We repeat the same analysis for the `luminous` sample, which we divide in three bins, with a similar bin refining approach as for the `dense` sample (in this case `L1` contains half of the `luminous` galaxies, while `L2` and `L3` the remaining quarters). The best-fit amplitudes for these samples are reported in Table 3.1, and presented as green stars in Fig. 3.7. In the luminosity range where the `luminous` and `dense` samples overlap, we find the results between the two samples to be compatible. The `luminous` sample seems to show a more pronounced luminosity dependence compared to the `dense` sample, which can either be an effect of being brighter overall (from `L1` to `L3`, $L/L_0 = 0.46, 0.64, 1.01$) or due to the satellite fraction being lower (see Sect. 3.2.2), or a combination of the two. We note that the measure-

[15] The internal cut at $m_r < 20$ in *lens*fit makes it less suitable for this analysis, as we have fewer galaxies at high luminosities.

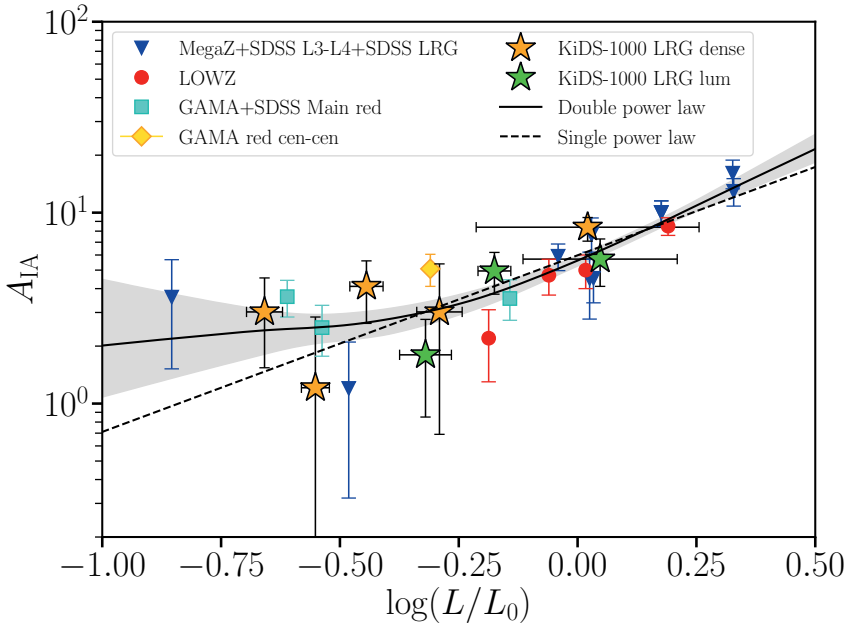

Figure 3.7: Luminosity dependence of the IA amplitude as measured by different observational studies (Joachimi et al. 2011; Singh et al. 2015; Johnston et al. 2019; Fortuna et al. 2021a); our new measurements on the LRG samples are shown as star markers. We provide horizontal error bars to indicate that the measurement is performed on a bin in luminosity, here plotted as the weighted standard deviation of the luminosity distribution of each sample, with the marker placed at the weighted mean. The solid (dashed) black line shows the median of the distribution of the MCMC sample associated with the double (single) power law; the shaded area corresponds to the 68% confidence region.

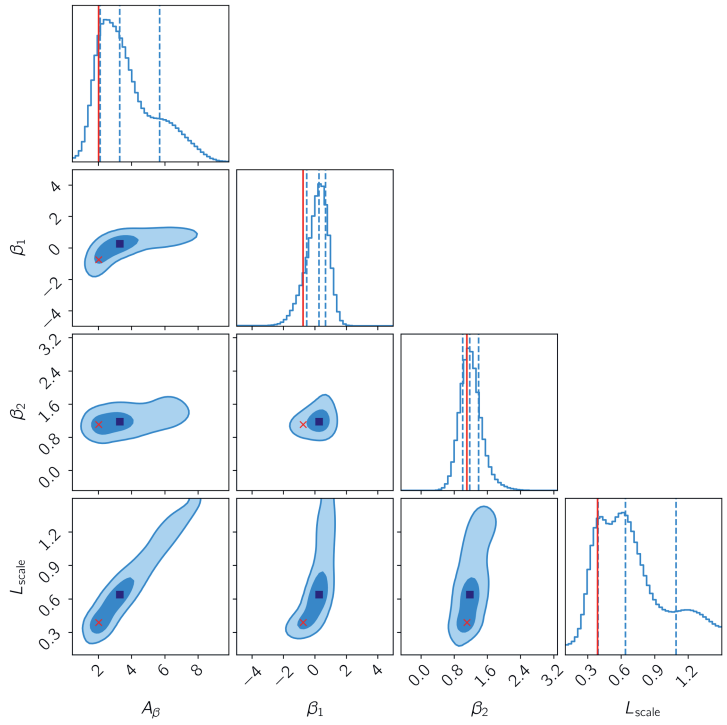

Figure 3.8: Constraints on the double power law parameters described in equation 3.35 by jointly fitting all the measurements in Fig. 3.7. The red crosses indicate the value of the parameters that maximise the likelihood, while the blue squares correspond to the medians.

Table 3.2: Best-fit parameters of the single and double power law fit on the measurements in Fig. 3.7.

Model	A_β	β_1	β_2	$L_{\rm break}$	χ^2/dof	dof
Double power law	$3.28^{+2.41}_{-1.17}$ (2.01)	$0.26^{+0.42}_{-0.77}$ (-0.75)	$1.17^{+0.21}_{-0.17}$ (1.11)	$0.64^{+0.45}_{-0.24}L_0$ ($0.39L_0$)	1.36 (1.33)	22
Single power law	$5.98^{+0.27}_{-0.27}$ (6.0)	$0.93^{+0.11}_{-0.10}$ (0.92)	-	L_0	1.61 (1.61)	24

The listed values correspond to the medians of the marginal posterior distributions, and the associated errors correspond to the 16th and 84th percentiles, while in brackets we report the parameters that maximise the likelihood. The same scheme is adopted for the corresponding reduced χ^2. $L_{\rm break}$ is the pivot luminosity that enters in the denominator of the power law argument. For convenience, the slope of the single power law model is here reported as β_1.

ments of the L3 sample appear to scatter more than the covariance predicts, which results in higher χ^2. A similar issue is present in the D4 sample and it is visible in Fig. 3.6.

The horizontal error bars in Fig. 3.7 indicate the weighted standard deviation of the luminosity distribution within the bin for each sample, with the measurement placed at the luminosity-weighted mean of the bin. If the range is too large, and the IA signal varies within the bin, the resulting amplitude is difficult to interpret, and may even appear discrepant. For instance, when we combine the D4 and D5 samples we obtain $A_{\rm IA} = 6.70^{+1.15}_{-1.14}$. We note, however, that the luminosity range probed by this combined bin is particularly extended, and the high signal measured is mainly driven by the galaxies in the high luminosity tail of the bin (D5, $A_{\rm IA} = 8.39^{+1.04}_{-1.30}$). The other half of the bin has a relatively low signal with very large uncertainties (D4, $A_{\rm IA} = 3.02^{+2.37}_{-2.33}$). This is relevant because it suggests that the alignment of galaxies with luminosities below $L/L_0 \sim 0.60 - 0.70$ hardly depends on luminosity, and thus with a similar amplitude to D1 and D3, the smaller sample is less constraining. As soon as we exceed this approximate threshold, the signal increases significantly, suggesting a luminosity dependence. This overall picture is enhanced when we also consider previous results for LRGs (Joachimi et al. 2011; Singh et al. 2015; Johnston et al. 2019; Fortuna et al. 2021a)[16]. These are also shown in Fig. 3.7. We investigate how well the current measurements support the picture of a single or double power law by fitting the data points in Fig. 3.7, assuming them to be uncorrelated. For each data point, we only use the quoted L/L_0 as we do not have the underlying luminosity distribution for most of the measurements. We propose a double power law with knee at $L_{\rm break}$, amplitude A_β and slopes $\beta_{1,2}$:

$$A(L) = A_\beta \left(\frac{L}{L_{\rm break}}\right)^\beta \text{ with } \begin{cases} \beta = \beta_1 & \text{for } L < L_{\rm break} \\ \beta = \beta_2 & \text{for } L > L_{\rm break} \end{cases} \tag{3.35}$$

and fit for

$$\lambda = \left\{A_\beta, \beta_1, \beta_2, L_{\rm scale}\right\} \ , \tag{3.36}$$

where $L_{\rm scale} = L_{\rm break}/L_0$. We explore the parameter space using a MCMC and assuming a Gaussian likelihood. Figure 3.8 shows our parameter constraints, while the model prediction is shown in Fig 3.7 as a solid black line.

[16]The GAMA points (Johnston et al. 2019) have been adjusted to homogenise the units convention, as discussed in Fortuna et al. (2021a).

Our best-fit parameters are reported in Table 3.2[17]. We repeat the same analysis assuming a single power law, as parametrised in Joachimi et al. (2011). The best-fit parameters are also reported in Table 3.2. The larger χ^2/dof of the single power law compared to the double power law suggests that the latter is a better description of our current data, although the scatter between the points at low L is still too large to draw definitive conclusions and the data are also mildly inconsistent in that regime. The degeneracy between the parameters, and in particular between A_β and $L_{\rm scale}$, shows that the data can weakly constrain the model. Nevertheless, the emerging picture seems to support more the broken power law scenario presented in Fortuna et al. (2021a), but with a transition luminosity around $0.4 - 0.6L_0$, also in line with the results from simulations by Samuroff et al. (2020a). The double power law is also supported by the fact that the alignment of redMaPPer clusters (van Uitert & Joachimi 2017; Piras et al. 2018), not included in this analysis, forms a smooth extension towards higher mass of the alignment observed for the high luminosity LRGs. This result is hard to reconcile with a single shallow power law, but finds a natural framework in the double power law scenario, where the slope of the relation at high luminosities recovers the trend in Joachimi et al. (2011); Singh et al. (2015).

We caution that this analysis does not aim to be fully comprehensive, but rather to provide a sense of the current trends. A proper analysis should jointly fit all of the measurements incorporating the full luminosity distributions of each sample, as well as accounting for the presence of satellites, which might suppress the signal at low luminosities.

3.6.2 Redshift dependence

Having assessed that the two shape measurements produce compatible IA signals and that their calibrations are robust, we merge the two shape catalogues to span the largest possible range in redshift. This allows us to extend the sample from the low-z, high S/N galaxies, where only DEIMOS provides shapes, to the high-z, low S/N galaxies, where we preferentially measure the shapes via *lens*fit. In the case of overlap between DEIMOS and *lens*fit, we select the DEIMOS shapes. We only focus on the `luminous` sample as we are interested in a long redshift baseline with the same luminosity cut. In this way, we can probe the redshift evolution of the sample, without confusing the results with any luminosity dependence.

[17]We note that the parameters that maximise the likelihood differ from the medians of the posterior distributions as a consequence of the degeneracies between the parameters. This is particularly evident for β_1, which has negative slope, $\beta_1 = -0.75$.

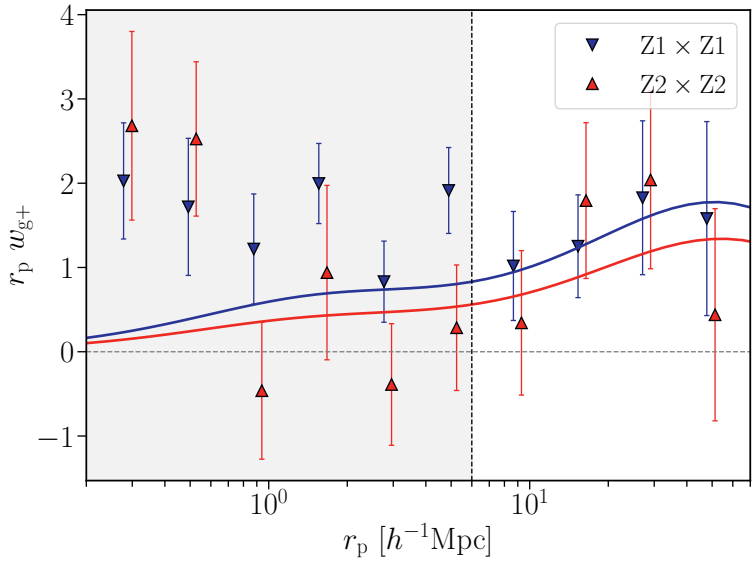

Figure 3.9: Projected correlation function, $w_{\mathrm{g+}}$, measured on our different cuts in redshift of the `luminous` sample. The best-fit curves are plotted on top of the data points, and the fits are performed for $r_\mathrm{p} > 6\,h^{-1}\mathrm{Mpc}$. The red points are slightly displaced for clarity and the corresponding best-fit curve has been displaced accordingly.

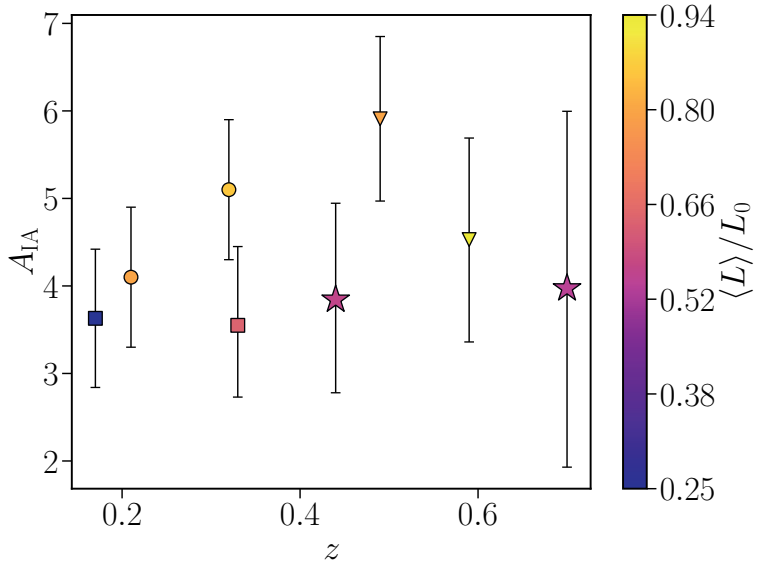

Figure 3.10: Intrinsic alignment amplitude, $A_{\rm IA}$, as a function of redshift and luminosity for different best-fit values in the literature. Different markers refer to different studies and are colour-coded based on their luminosity: MegaZ (Joachimi et al. 2011) is shown as triangles, LOWZ (Singh et al. 2015) as circles, GAMA (Johnston et al. 2019) as squares and the LRG `luminous` sample Z1 and Z2 as stars.

Our final catalogue contains 115 322 galaxies that we split at $z = 0.585$, which roughly provides two equally populated bins. We call these two samples Z1 and Z2. The measurements for $w_{\rm g+}$ are presented in Fig. 3.9. The best-fit values for the two redshift bins are listed in Table 3.1 and agree within their error bars, despite their mean redshift being $\langle z \rangle = 0.44$ and $\langle z \rangle = 0.70$, respectively.

We note that the χ^2 of our Z2 sample is quite high: This is driven by the poor fit of the clustering signal. We attribute this to our photo$-z$, which at high redshift are less reliable. We note, however, that the uncertainty in the IA amplitude is large enough to absorb the inaccuracies in $p(z_{\rm spec}|z_{\rm phot})$, such that modifying the redshift distributions has little impact on the recovered IA amplitude.

Figure 3.10 compares our results with the best-fit amplitudes at various redshifts found by previous studies (Joachimi et al. 2011; Singh et al. 2015; Johnston et al. 2019). The colour of the data points reflects the lu-

minosity of the sample used to measure the signal[18]. As previously discussed, galaxies with different luminosities may manifest different levels of IA, and hence even with a lack of redshift dependence, we should still expect points at different amplitudes: the bottom part of the plot should be mainly populated by darker points and the upper part by brighter points. Figure 3.10 confirms this scenario: overall, the points exhibit a similar alignment and the scatter between the different points is consistent with the extra luminosity dependence. We can conclude that there is little evidence for a strong redshift dependence of the IA signal.

3.7 Conclusions

We have constrained the IA signal of a sample of LRGs selected by Vakili et al. (2020) from KiDS-1000, which images ~ 1000 deg^2. These data allowed us to investigate the luminosity dependence and the redshift evolution of the signal. To do so, we measured the shapes of the LRGs with two different algorithms, DEIMOS and *lens*fit. We used custom image simulations to calibrate and correct the residual biases that arise from measurements of noisy images.

We used the calibrated ellipticities to compute the projected position-shape correlation function $w_{\mathrm{g+}}$ and analyse the signals obtained by the two different algorithms independently, thus exploring the dependence of IA on the specific shape method employed. We found *lens*fit measurements to be overall noisier than the DEIMOS ones and we attributed this to the prevalence of faint galaxies in the sample, due to the internal magnitude cut in the *lens*fit algorithm. Because bright galaxies typically carry more alignment signal, this cut, which removes galaxies with $m_r < 20$, can potentially reduce the IA contamination in KiDS cosmic shear analyses, which employ *lens*fit as the shape method. For a sub-sample of galaxies, where both shape methods return successful measurements of the shapes, we find a remarkable agreement in the measured $w_{\mathrm{g+}}$, with a difference in the signal of 0.003 ± 0.13 (amplitude of a fitted power law).

We explored the luminosity dependence and the redshift evolution independently, selecting our galaxies in such a way that ensures the two do not mix. Within the luminosity range probed by the measurements our re-

[18]The colour of the marker corresponds to the bin centre, which may not be sufficient if the range in luminosity is large, as it is typically the case for these samples. The information provided by the colour has therefore only qualitative meaning and should be considered as such.

sults agree with previous studies (Joachimi et al. 2011; Singh et al. 2015; Johnston et al. 2019). However, a single power law fit, as was used in Joachimi et al. (2011) and Singh et al. (2015) does not describe the measurements well. Instead, our results suggest a more complex dependence with luminosity: for $L_r \lesssim 2.9 \times 10^{10} h^{-2} L_{r,\odot}$ the IA amplitude does not vary significantly, whereas the signal rises rapidly at higher luminosity. This also has implications for the width of the luminosity binning, as the use of broad bins may complicate the interpretation of the measurements. Analyses that aim to combine these measurements to model the luminosity dependence should incorporate the underlying luminosity distributions to properly link the signal to the galaxy luminosity. Nevertheless, we provide a preliminary fit on the current measurements available in the literature and found that the data are best described by a broken power law. This result can already be used by cosmic shear analyses to improved their modelling of the IA carried by the red galaxy population. We remind the reader that this sample is not representative of the galaxy population. Different galaxy samples carry different alignment signals and should thus be individually modelled as described in Fortuna et al. (2021a).

To probe the redshift dependence of the IA signal with the largest baseline to date, we merged the DEIMOS and *lens*fit catalogues. We find no evidence for redshift evolution of the IA signal. This result is in line with previous studies of LRG samples (Joachimi et al. 2011; Singh et al. 2015), and it is consistent with the current paradigm that IA is set at the moment of galaxy formation. However, it is also possible that galaxy mergers counteract the evolution of the tidal alignment, such that the net signal does not change. Further improvements in the measurements are needed to distinguish between scenarios.

Acknowledgements

We thank Sandra Unruh for providing useful comments to the manuscript. MCF, AK, MV and HH acknowledge support from Vici grant 639.043.512, financed by the Netherlands Organisation for Scientific Research (NWO). HH also acknowledges funding from the EU Horizon 2020 research and innovation programme under grant agreement 776247. HJ acknowledges support from the Delta ITP consortium, a program of the Netherlands Organisation for Scientific Research (NWO) that is funded by the Dutch Ministry of Education, Culture and Science (OCW). We also acknowledge support from: the European Research Council under grant agreement No.

770935 (AHW, HHi) and No. 647112 (CH and MA); the Polish Ministry of Science and Higher Education through grant DIR/WK/2018/12 and the Polish National Science Center through grants no. 2018/30/E/ST9/00698 and 2018/31/G/ST9/03388 (MB); the Max Planck Society and the Alexander von Humboldt Foundation in the framework of the Max Planck-Humboldt Research Award endowed by the Federal Ministry of Education and Research (CH); the Heisenberg grant of the Deutsche Forschungsgemeinschaft Hi 1495/5-1 (Hi); the Royal Society and Imperial College (KK) and from the Science and Technology Facilities Council (MvWK).

The MICE simulations have been developed at the MareNostrum supercomputer (BSC-CNS) thanks to grants AECT-2006-2-0011 through AECT-2015-1-0013. Data products have been stored at the Port d'Informació Científica (PIC), and distributed through the CosmoHub webportal (cosmohub.pic.es). Funding for this project was partially provided by the Spanish Ministerio de Ciencia e Innovacion (MICINN), projects 200850I176, AYA2009-13936, AYA2012-39620, AYA2013-44327, ESP2013-48274, ESP2014-58384 , Consolider-Ingenio CSD2007- 00060, research project 2009-SGR-1398 from Generalitat de Catalunya, and the Ramon y Cajal MICINN program.

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under programme IDs 177.A- 3016, 177.A-3017 and 177.A-3018, and on data products produced by Tar- get/ OmegaCEN, INAF-OACN, INAF-OAPD and the KiDS production team, on behalf of the KiDS consortium. OmegaCEN and the KiDS production team acknowledge support by NOVA and NWO-M grants. Members of INAF-OAPD and INAF-OACN also acknowledge the support from the Department of Physics & Astronomy of the University of Padova, and of the Department of Physics of Univ. Federico II (Naples).

Author contributions: All authors contributed to the development and writing of this paper. The authorship list is given in three groups: the lead authors (MCF, HH, HJ, MV, AK, CG) followed by two alphabetical groups. The first alphabetical group includes those who are key contributors to both the scientific analysis and the data products. The second group covers those who have either made a significant contribution to the data products, or to the scientific analysis.

A1 m-bias calibration

In this Appendix, we detail our procedure to calibrate the m-bias in our shape measurements. We follow the same procedure for both DEIMOS and *lens*fit, but we present the results separately.

A1.1 DEIMOS

One of the key features of DEIMOS that was exploited by Georgiou et al. (2019b) is that the weight function that is used to measure the moments of the surface brightness distribution can be adjusted. As explained in Sect. 3.3, we follow Georgiou et al. (2019a) and adopt a Gaussian weight function with a width $r_{\rm iso}$. However, not only the radial profile can be changed, but one can also choose between a circular or an elliptical weight function. Hence, before proceeding with the shape calibration, we investigate which choice of weight function would suit our data best.

In both cases, the weight function is centred on the centroid of the galaxy, with the size and ellipticity iteratively matched to those measured for the galaxy (see Georgiou et al. 2019a, for details). While an elliptical weight function matches the shape of an elliptical galaxy better, a circular one generally performs better on small and faint objects.

The circular weight function performs similar to the elliptical weight function for low-to-intermediate S/N (S/N< 60), but with an overall constant bias of ~ 0.2 as the S/N increases. Hence, the elliptical weight function performs significantly better for more than half of the (real) galaxy sample, which motivates our choice to adopt an elliptical weight function in our analysis.

DEIMOS measured the shapes of 13 301 simulated LRGs from our image simulations, and we use these to calibrate our ellipticity estimates. To do so, we first explore the dependence of the m-bias on the individual galaxy parameters S/N and R, as discussed in Sect. 3.3. Figure 3.3 indicates that $m(R)$ is well described by a polynomial curve, which we truncate at degree 3, $p(R) = p_1 R + p_2 R^2 + p_3 R^3$, while $c(\mathrm{S/N})$ is well described by: $d(\mathrm{S/N}) = d_1/\sqrt{\mathrm{S/N}} + d_2/(\mathrm{S/N})$.

We have tested different combinations of the two functions $m(\mathrm{S/N})$ and $m(R)$, and explored if higher-order polynomials are needed: while the fit to $m(R)$ is indeed better described by a polynomial of degree 5, we stress that we are not interested to reproduce all of the noisy features in the data, but rather to capture the trend in the two components. We therefore keep

Table 3: Best-fit parameters for the empirical correction of the two-dimensional multiplicative bias surface (Sect. 3.3.3).

Parameter	DEIMOS	*lens*fit
b_0	-0.895	0.1794
d_1	5.238	-5.081
d_2	-0.006	1.292
p_1	-1.900	-0.972
p_2	5.147	0.669
p_3	-3.148	0.783
p_4	–	-0.698

the number of the parameters as low as possible. This is also motivated by the fact that the image simulations suffer from galaxy repetitions.

The final expression for our empirical correction for the DEIMOS measurements is then:

$$m(\mathrm{S/N}, R) = b_0 + \frac{1 + d(\mathrm{S/N})}{1 + p(R)} \,. \tag{37}$$

To find the best-fit parameters in 37, we re-computed the value of the $m-$bias by binning the data in 64 regions using the $k-$means algorithm[19]. We then measured the bias for the two components $\epsilon_{1,2}$ in each region, identifying the bin coordinate in S/N and R as their mean value within the bin. We then fit the average of the two components $(m_1 + m_2)/2$ with equation 37. Some of the galaxies have very small shape measurement errors, and to avoid them dominating the fit, we also added an intrinsic scatter σ_{int} to our error-bars. This accounts for the fact that the number of unique galaxies in our simulations is limited and mitigates the importance of the highly resolved ones. The intrinsic scatter σ_{int} is chosen such that the reduced χ^2 is ~ 1. The best-fit parameters are reported in Table 3. We stress here that since we are only correlating shapes with positions, we are not interested in a perfect calibration of the bias per galaxy but rather want to ensure that the mean ellipticity of an ensemble is unbiased.

[19] `https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html`

A1.2 *lens*fit

In the case of *lens*fit we follow a very similar procedure to calibrate the residual m-bias. *lens*fit successfully measured the shapes of 17 573 simulated galaxies, which are used for the calibration. The dependence of the m-bias with S/N can be described by the same parametrisation that we used for the DEIMOS sample, $d(\mathrm{S/N})$, while $m(R)$ is better described by a polynomial of degree four, $p(R) = p_1 R + p_2 R^2 + p_3 R^3 + p_4 R^4$.

The combination that best reproduces our measurements of the m-bias in k-means cells of the two-dimensional space (S/N,R) is

$$m(\mathrm{S/N}, R) = b_0 + d(\mathrm{S/N}) + p(R) \,, \tag{38}$$

with the specific values of the parameters reported in Table 3. We note that compared to DEIMOS, the *lens*fit-bias is small, and hence so is our correction.

A2 Redshift distributions

We describe here the redshift distributions, $p(z_{\rm spec}|z_{\rm phot})$, employed in our analysis as reported in Sect. 3.5 and which are used in the computation of the angular power spectra in Eq. (3.25). We bin the galaxies for which we have spectroscopic redshifts in bins of $\Delta z^{\texttt{dense}}_{\rm phot} = 0.0146(1+z)$ and $\Delta z^{\texttt{luminous}}_{\rm phot} = 0.0139(1+z)$ with an iterative procedure; this constructs unequal binning whose size increases with z. The last bin is adjusted to avoid spurious results: If the maximum redshift found with the iterative procedure exceed the maximum redshift of the sample, we remove the last bin and extend the second-to-last up to $z_{\rm max}$. In the case of the `luminous` sample we further increase the scatter at high redshift to account for the increasing uncertainty of our photometric redshifts: for $z > 0.7$ we increase the bin width to $\Delta z^{\texttt{luminous}}_{\rm phot} = 0.027$. We adopt the same approach for the Z1 and Z2 samples, for which we use, respectively, $\sigma_z = 0.0133$ and 0.0190. We use the resulting spec-z histograms in our analysis. We employ the same conditional redshift distributions for both our density and shape samples; while this is a very good approximation for DEIMOS, *lens*fit lacks bright galaxies that would populate our spec-z, and thus this approximation might partially be responsible for the worse fit of the model.

We tested that our IA constraints are only marginally dependent on the width of the bins adopted, and the changes in the best-fit amplitude are subdominant to the statistical uncertainty.

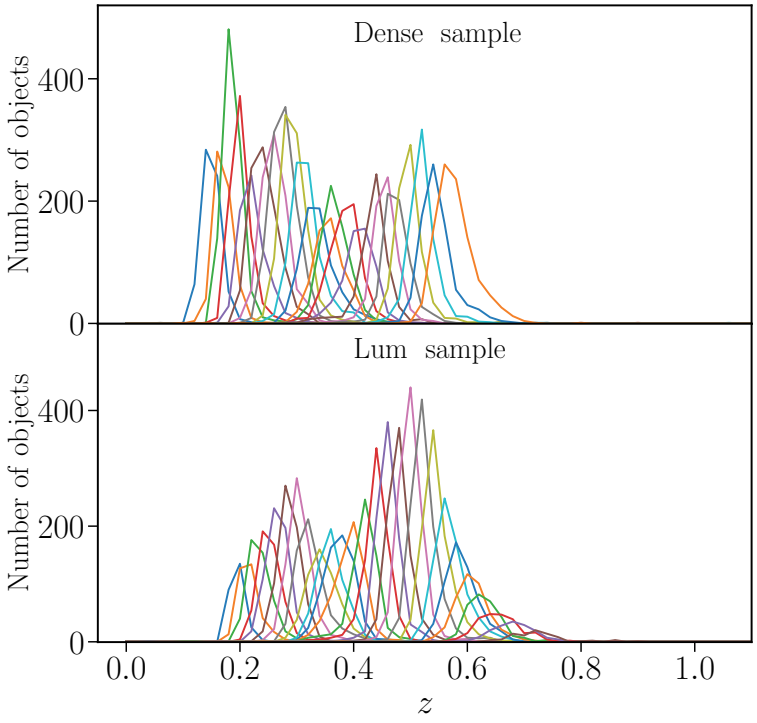

Figure 11: The $p(z_{\rm spec}|z_{\rm phot})$ of our `dense` and `luminous` samples.

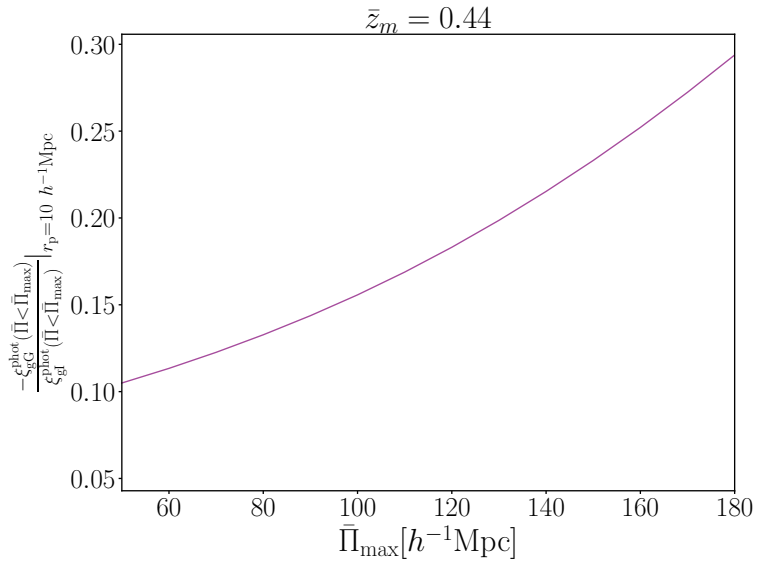

Figure 12: Ratio of the cumulative galaxy-galaxy lensing signal over the cumulative IA signal as a function of $\Pi_{\rm max}$ at the mean redshift of the `dense` sample, $z = 0.44$.

A3 Contamination from galaxy-galaxy lensing

As discussed in Sect. 3.5.2, galaxy-galaxy lensing is the main astrophysical contaminant to our signal. Here, we focus on its dependence on the line-of-sight integration range. The lensing and the IA signals scale differently with distance: this can be used to maximise the signal and avoid an excess of contamination. In this Appendix we therefore explore in more detail the modelling of the galaxy-galaxy lensing and how this has guided our choice for the value of $\Pi_{\rm max}$.

Figure 12 shows the amount of lensing contamination as a function of the maximum Π used in the integral along the line-of-sight. We illustrate it by plotting the cumulative contribution of the galaxy-galaxy lensing over the one of IA for different values of the truncation, $\Pi_{\rm max}$. To generate the signal, we used the $p(z_{\rm spec}|z_{\rm phot})$ associated with the `dense` sample and evaluated the correlation functions at the mean redshift of the sample, assuming the fiducial bias and IA amplitude reported in Table 3.2. The ratio is almost constant in $r_{\rm p}$, thus we plot it for fixed $r_{\rm p} = 10\,h^{-1}{\rm Mpc}$. We also note that the lensing signal has negligible impact for negative Π because the source is in front of the lens in that case.

In principle, if one had perfect knowledge of the galaxy-galaxy lensing contribution, extending the integration up to very large line-of-sight separations would allow us recover the full IA signal from the measurements, without discarding any information. In practice, even though we fully model the galaxy-galaxy lensing contribution, we are limited by the accuracy of the lensing modelling we rely on, and thus it is safer to truncate the integral to values of Π that are not severely affected by it.

We use Fig. 12 to choose the fiducial $\Pi_{\rm max}$ that enters in Eq. (3.12): although the specific values of the ratio depend on the input parameters ($b_{\rm g}$, $A_{\rm IA}$), it provides a realistic estimate of the amount of contamination for our LRG samples. We chose as our fiducial setup a conservative value of $\Pi_{\rm max} = 120\,h^{-1}{\rm Mpc}$, which ensures that the mean contamination is below $\sim 20\%$ of the signal.

A4 Contamination from magnification

The changes in the galaxy number counts determined by lensing magnification arise as a result of two competing effects: on one hand, the lensing locally stretches the sky, diluting the observed number density; on the other hand, it enlarges the apparent sizes of the galaxies without modifying the surface brightness: at the faint end, this allows the detection of galaxies that are intrinsically fainter than the magnitude limit, enhancing the observed number density.

The theory of magnification for flux-limited surveys is well established and allows us to relate the changes in the number density to the differential galaxy count $n(m)$ over a given band magnitude range from m to $m + {\rm d}m$ (Bartelmann & Schneider 2001; Joachimi & Bridle 2010):

$$\alpha(m) = 2.5\frac{{\rm d}\log[n(m)]}{{\rm d}m}\,. \tag{39}$$

The case of a non-flux-limited sample, such as our LRG sample, is more complicated and we lack a proper theoretical framework for the interpretation of α. Here, we follow von Wietersheim-Kramsta et al. (2021) and calibrate α using dedicated mocks, which we present in Appendix A7. We remind the reader that our samples are selected by imposing a luminosity threshold, which implies a redshift-dependent magnitude selection.

The calibration works as follows: the mocks provide the reference relation between the convergence κ and the slope α, which we can measure

as the difference in the number density of a 'magnified' sample and a 'non-magnified' one,

$$\frac{n(<m) - n_0(<m)}{n_0(<m)} \approx 2(\alpha - 1)\kappa \,. \tag{40}$$

Here, $n(<m)$ is the local number density of magnified sources with magnitudes below m, while $n_0(<m)$ is the underlying true number density without the enhancement due to the flux magnification and the simultaneous lensing dilution.

We used our mocks to measure α in Eq. (40), obtained as the mean value of κ on sufficiently small patches of the sky. To partition the sky we used the public available python module Healpy[20] (Zonca et al. 2019), based on the HEALPix pixellization of the sphere[21] (Górski et al. 2005). We used this value of α to calibrate the magnitude range over which the observable α in equation 39 best agrees with the *true* one obtained from equation 40. If the mocks reproduce the data selection function to good accuracy, this provides the optimal magnitude range to use to measure α via observable quantities (Eq. (39)) in the data.

To evaluate Eq. (39) we used the $r-$band magnitude and we ensured that the magnitude distribution of the mocks and the data agree to high accuracy. We find that, when applied to the data, the method results in values of α that depend somewhat on the binning scheme employed along the redshift baseline. While the values of α are robust against changing the bins at intermediate and high redshifts, the very low-z bins are poorly constrained by the method. However, at such low redshifts magnification is negligible, and our samples contain only a few galaxies, so it is reasonable to expect the same value of α to hold for the entire sample. Moreover, the LRG selection ensures a constant comoving number density, which reduces the sensitivity to magnification even further.

We find $\alpha \sim 1.5$ for both our `dense` and `luminous` sample. In Appendix A5 we show that the effect of including magnification is subdominant in our analysis.

Table 4: Reduced χ^2 statistics to assess the significance of our signals w_{g+} and $w_{g\times}$ against the null hypothesis.

Sample	Shapes	Signal	$\chi^2_{\nu,\mathrm{null}}$	p–value
`dense`	DEIMOS	w_{g+}	8.01 (7.56)	4.88×10^{-13} (4.35×10^{-6})
		$w_{g\times}$	0.59 (0.36)	0.83 (0.83)
	*lens*fit	w_{g+}	3.52 (4.99)	0.0001 (0.0005)
		$w_{g\times}$	0.66 (0.64)	0.76 (0.63)
`luminous`	DEIMOS	w_{g+}	9.46 (5.85)	6.66×10^{-16} (0.0001)
		$w_{g\times}$	0.37 (0.33)	0.96 (0.85)
	*lens*fit	w_{g+}	2.48 (1.49)	0.006 (0.20)
		$w_{g\times}$	0.40 (0.40)	0.95 (0.81)

A detection of $w_{g\times}$ would hint at the presence of unaccounted systematics in the measurements. The numbers in brackets refer to the signal for $r_\mathrm{p} > 6\ h^{-1}\mathrm{Mpc}$.

A5 Systematic tests and significance of the detection

To ensure the robustness of our analysis, we performed a number of tests for residual systematics. We present the results of these in this Appendix. Many of these are commonly used to test weak gravitational lensing signals.

In one of the most basic tests, the galaxy shapes are rotated by $45\,\mathrm{deg}$ and the correlation between $\epsilon_\times$ and galaxy position, $w_{g\times}$ is measured. This correlation is expected to vanish, and any detection of a non-vanishing signal is therefore an indication of residual systematics. Table 4 reports the reduced $\chi^2_{\nu,\mathrm{null}}$, which we used to assess the significance of the signal against the null hypothesis for both w_{g+} and $w_{g\times}$. We choose a significance level of 5%: for p–values below 0.05 we discard the null hypothesis. We can see that all of our $w_{g\times}$ measurements have a p–value above 0.05 and thus support the null hypothesis. In contrast, we observe a significant detection for all of our w_{g+} measurements, for both DEIMOS and *lens*fit shapes.

As a further look into possible systematics in the data, we measured the signal for a very large value of the line-of-sight truncation, $\Pi_\mathrm{max} = 1000\,h^{-1}\mathrm{Mpc}$, using our `dense` sample. Extending the value of Π_max to very large separations introduces uncorrelated pairs into the estimator, and thus

[20] `https://healpy.readthedocs.io/en/latest/`
[21] `http://healpix.sourceforge.net`

we expect the IA signal to vanish, while the galaxy-galaxy lensing can potentially arise. We find a signal consistent with a null detection, with $\chi^2_{\nu,\mathrm{null}} = 0.35$ and p–value of 0.96.

We also investigated the impact of specific choices for the setup of our modelling, with a particular focus on how our results depend on the value of the Π_{max} adopted in the analysis. To do so, we repeat our analysis of the `dense` sample using two different values of Π_{max}: 90 and 180 $h^{-1}\mathrm{Mpc}$. Table 5 reports our results. We find compatible results that also agree with our fiducial value of $\Pi_{\mathrm{max}} = 120\,h^{-1}\mathrm{Mpc}$.

In Table 5 we also report the results when we include magnification in the modelling for the `dense` sample, or ignore it for the `luminous` sample. The resulting parameter estimates agree with the baseline results (also see Sect. 5.6), suggesting magnification is small in our data, as expected from theory (Unruh et al. 2020).

A6 IA dependence on the shape measurement method

Singh & Mandelbaum (2016) compared the IA signal measured with different shape methods and found that the signal depends on the specific algorithm employed. Georgiou et al. (2019a) explored this further, and used DEIMOS to show that the IA signal depends on the width of the weight function. Since different methods use different weight functions, the difference in the IA detection can be linked to the parts of the galaxies they probe.

In this Appendix, we therefore explore how the IA signal depends on the shape measurement methods used in our analysis. To ensure this is done consistently, we only selected galaxies that belong to both our DEIMOS and *lens*fit catalogues, irrespective whether they are part of the `luminous` or `dense` sample. We identify 173 499 galaxies in common between the two shape catalogues.

We measure $w_{\mathrm{g+}}$ for this sub-sample for both shape catalogues, and show the difference in the signal, $\Delta w_{\mathrm{g+}} = w_{\mathrm{g+,DEIMOS}} - w_{\mathrm{g+},\mathit{lens}\mathrm{fit}}$ (indigo squared markers) in Fig. 13. The error bars are computed via bootstrap; we are only interested in the shape noise contribution: We are measuring the difference of signals obtained using the same sample of galaxies and thus the sample variance should vanish. We generated 215 re-samplings with replacement of our input galaxies and provided the same input catalogue to both our DEIMOS and *lens*fit measurement of $w_{\mathrm{g+}}$. The error

Table 5: Tests of the modelling setup.

Sample	b_g	A_{IA}	χ^2_{red}
DEIMOS			
dense (120, baseline)	$1.59^{+0.04}_{-0.04}$	$3.69^{+0.66}_{-0.65}$	0.78
dense (90)	$1.60^{+0.04}_{-0.04}$	$3.99^{+0.73}_{-0.72}$	0.67
dense (180)	$1.58^{+0.05}_{-0.05}$	$3.50^{+0.69}_{-0.70}$	0.71
dense (120, w/o magnification)	$1.59^{+0.04}_{-0.04}$	$3.67^{+0.66}_{-0.64}$	0.78
dense (120, w/o lensing and magnification)	$1.59^{+0.04}_{-0.04}$	$3.47^{+0.67}_{-0.66}$	0.78
luminous (120, baseline)	$2.06^{+0.04}_{-0.04}$	$4.03^{+0.81}_{-0.79}$	1.19
luminous (120, w/o magnification)	$2.06^{+0.04}_{-0.04}$	$4.01^{+0.82}_{-0.81}$	1.19
luminous (120, w/o lensing and magnification)	$2.06^{+0.04}_{-0.04}$	$3.84^{+0.80}_{-0.80}$	1.19

The value of Π_{max} adopted for the measurement is reported in brackets. The same value is assumed in the model. The tests are always performed using the DEIMOS shape catalogue.

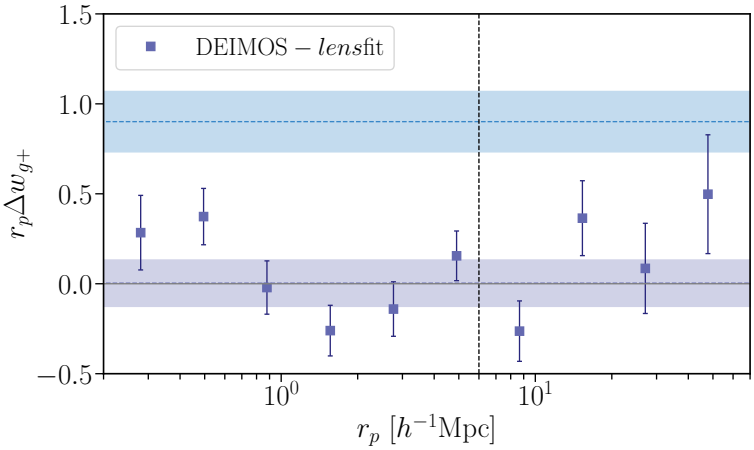

Figure 13: Difference in the w_{g+} measurements as measured by DEIMOS and *lens*fit. The indigo dashed line shows the best-fit amplitude of the difference, here parametrised as A/r_p. Similarly, the light blue dashed line illustrates the best-fit amplitude for the DEIMOS sample, both performed for $r_p > 6\,h^{-1}\mathrm{Mpc}$. The shaded areas delimit the 1σ contour of the fit.

Table 6: Parameters of the Student's t-distributions that best-fit the residuals $(z^{\mathrm{phot}} - z^{\mathrm{spec}})/\sigma_z$ of our samples.

Sample	ν	μ	s
`dense`	3.79	0.06	0.90
`luminous`	3.99	-5.43×10^{-6}	0.86

bars are then computed as the standard deviation of the difference in the measured signal for this ensemble.

To quantify the amplitude of the signal to the potential differences in measurement method, we fit both $w_{g+,\mathrm{DEIMOS}}$ and Δw_{g+} with a curve of the form $f(r_p) = A/r_p$, for $r_p > 6\,h^{-1}\mathrm{Mpc}$. The best-fit amplitudes are, respectively, 0.90 ± 0.17 and 0.003 ± 0.13, which means that we detect a signal that is more than six sigma above the uncertainty due to the choice in the shape measurement algorithm adopted.

A7 Mock catalogues

To investigate the impact of magnification bias on the interpretation of our measurements, we generated two mock catalogues that resemble our LRG samples. Our simulated catalogues are obtained from the KiDS photomet-

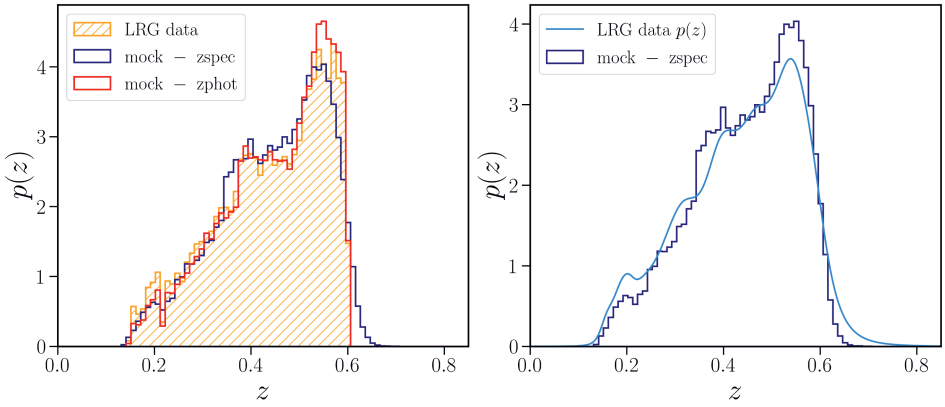

Figure 14: Comparison between the redshift distributions of the data and those reproduced by the mocks. *Left:* The photometric redshift distribution of our data is shown as an orange hatched histogram, while the solid red line shows the distribution of the photometric redshifts of the mock, obtained from the true ('spectroscopic') redshifts (blue solid line) as detailed in the text. *Right:* Comparison of the mock spectroscopic redshift distribution (solid blue line) and the estimated spectroscopic distribution of our data (light blue line).

ric mock catalogue presented in van den Busch et al. (2020), which is based on the MICEv2 simulation[22] (Fosalba et al. 2015b; Crocce et al. 2015; Fosalba et al. 2015a; Carretero et al. 2017; Hoffmann et al. 2015) and is specifically designed to reproduce the KiDS photometry. We did not run the LRG selection algorithm on the mock, but rather used their observed location in the redshift-colour space $(u-g, g-r, r-i, i-z)$ to select them in the mock.

We first apply a broad colour selection using the MICE `z_cgal` 'spectroscopic' redshift. After assigning the photo-z to our mocks, we repeat the selection replacing the spectroscopic redshift with the photometric one. The photometric redshifts are designed to reproduce the distributions reported in Sect. 3.3 of Vakili et al. (2020). To do so, we draw a random value from a Student's t–distribution centred on $z_{\rm spec} - \mu\sigma_z$ and with the scale parameter equal to $s\sigma_z$, with μ, ν and s the Student's t–parameters fitted to the full distribution (of the real data). We remind the reader that ν defines the peakiness of the distribution, μ its mean and s sets the width. In the limit of the Student's t–distribution approaching a Gaussian ($\nu \to \infty$), s can be interpreted as the standard deviation of the distribution.

We note that our samples differ from Vakili et al. (2020), since we excluded the galaxies that overlap with the `luminous` sample from the `dense`

[22] `http://maia.ice.cat/mice/`

sample. We therefore recomputed the parameters of the Student's t–distributions specifically for our samples and report these in Table 6. Some care has to be taken when assigning σ_z to the mocks. The per-galaxy σ_z of the LRG samples correlates with the magnitude of the galaxy. We, therefore, identify the closest real galaxy in the (z, m_r) space to each galaxy in the mock, and assign it the corresponding σ_z. We repeat the process for one iteration, replacing the 'spectroscopic' redshift with the preliminary estimate of the photometric one. We note that this procedure results in multiple assignments of the same σ_z to the mock galaxies, but this is not a concern as we do not require it to be unique.

Since we require a high fidelity reproduction of the line-of-sight distribution of our galaxies, we divide our samples and their corresponding mock catalogues in thin redshift slices and match the galaxy number density per slice. At this step, we do not require a perfect match. In this way, we still have enough galaxies to apply the same $m_r^{\rm pivot}(z)$ cut as for our real data. We repeat these steps iteratively until the number densities are matched between the samples. We tested that the final $p(z^{\rm spec}|z^{\rm phot})$ of our mocks are in good agreement with the data $p(z^{\rm spec}|z^{\rm phot})$ (see Fig. 14) and that the resulting clustering signal at large scales reproduces the one in our data.

We generate two sets of mock catalogues: a magnified one and one without magnification. We use these for the calibration of α as discussed in Appendix A4.