Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **456,** 3886–3898 (2016) doi:10.1093/mnras/stv2925

# RCSLenS: a new estimator for large-scale galaxy–matter correlations

A. Buddendiek,[1⋆] P. Schneider,[1⋆] H. Hildebrandt,[1⋆] C. Blake,[2] A. Choi,[3] T. Erben,[1] C. Heymans,[3] L. van Waerbeke,[4] M. Viola,[5] J. Harnois-Deraps,[4] L. Koens[3] and R. Nakajima[1]

[1]*Argelander-Institut für Astronomie, University of Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[2]*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[3]*Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*
[4]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*
[5]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*

Accepted 2015 December 11. Received 2015 December 9; in original form 2015 October 24

**ABSTRACT**

We present measurements of the galaxy bias $b$ and the galaxy–matter cross-correlation coefficient $r$ for the Baryon Oscillation Spectroscopic Survey LOWZ luminous red galaxy sample. Using a new statistical weak lensing analysis of the Red Cluster Sequence Lensing Survey (RCSLenS), we find the bias properties of this sample to be higher than previously reported with $b = 2.45^{+0.05}_{-0.05}$ and $r = 1.64^{+0.17}_{-0.16}$ on scales between 3 and 20 arcmin. We repeat the measurement for angular scales of 20 arcmin $\leq \vartheta \leq$ 70 arcmin, which yields $b = 2.39^{+0.07}_{-0.07}$ and $r = 1.24^{+0.26}_{-0.25}$. This is the first application of a data compression analysis using a complete set of discrete estimators for galaxy–galaxy lensing and galaxy clustering. As cosmological data sets grow, our new method of data compression will become increasingly important in order to interpret joint weak lensing and galaxy clustering measurements and to estimate the data covariance. In future studies, this formalism can be used as a tool to study the large-scale structure of the Universe to yield a precise determination of cosmological parameters.

**Key words:** gravitational lensing: weak – methods: analytical – surveys.

## 1 INTRODUCTION

Since the discovery of the accelerated expansion of the Universe (Riess et al. 1998; Perlmutter et al. 1999), the origin and nature of dark energy remains unknown. Several possible explanations like a cosmological constant, quintessence, or a modification of gravity on cosmological scales have been suggested. Although the accelerated expansion has been confirmed using a combination of other cosmological probes like cosmic microwave background experiments (Hinshaw et al. 2013; Planck Collaboration XIII 2015), weak gravitational lensing (Schrabback et al. 2010; Heymans et al. 2013), galaxy clusters (Vikhlinin et al. 2009; Mantz et al. 2014), or baryon acoustic oscillations (BAO; Blake et al. 2012; Sánchez et al. 2013), the statistical power of these probes so far remains insufficient to reveal the true nature of dark energy. Statistical precision sufficient to distinguish a cosmological constant from a more dynamical nature of dark energy will only be reached by the next generation of cosmology experiments, like *Euclid* (Laureijs et al. 2011), the LSST (Ivezic et al. 2008), or *WFIRST* (Spergel et al. 2015). For this purpose, the *Euclid* satellite will not only map the whole extragalactic sky in the optical and the near-infrared, but it will also take near-infrared spectra of about 50 million galaxies up to a redshift of $z = 2$. Using this vast data set, the *Euclid* consortium will measure the geometry of the Universe using both BAO and cosmic shear.

Cosmic shear is the distortion of light bundles from distant sources caused by the intervening tidal gravitational field, caused by the large-scale matter distribution in the Universe, which is measured from the auto-correlation of galaxy shapes (e.g. Bacon, Refregier & Ellis 2000; Van Waerbeke et al. 2001; Hoekstra et al. 2002a; see Bartelmann & Schneider 2001 for a review). The gravitational lensing signal in the galaxy shapes contributes only a few per cent to the whole galaxy ellipticity; furthermore, these galaxies are intrinsically small, typically smaller than the point spread function (PSF) of ground-based observations, and correspondingly are measured over a limited number of CCD pixels. Correcting for PSF effects and pixelization still poses a challenge to the astronomical community (e.g. Kitching et al. 2012; Mandelbaum et al. 2015). Due to these technical difficulties, it is important to have multiple independent weak lensing probes to map the density field in our Universe. A particularly promising approach is to combine information from galaxy auto-correlations (e.g. Blake et al. 2012; Sánchez et al. 2013) and galaxy–matter correlations (e.g. van Uitert et al. 2011, 2012; Velander et al. 2014). Significant effort has been made to develop a new theoretical framework for these

⋆ E-mail: abuddend@astro.uni-bonn.de (AB); Peter@astro.uni-bonn.de (PS); hendrik@astro.uni-bonn.de (HH)

measurements (e.g. Leauthaud et al. 2011; Cacciato et al. 2012; Eriksen & Gaztanaga 2015; Coupon et al. 2015).

A further challenge in relating observed signals to theoretical predictions stems from the difficulty in understanding baryonic physics, such as cooling, star formation, and feedback. This affects the statistical properties of the large-scale structure on small scales. A particularly interesting approach was therefore suggested by Baldauf et al. (2010, henceforth B10) who introduced a new estimator $\Upsilon$ for clustering and lensing, which eliminates all small-scale contributions to the signals. This methodology was successfully applied to the Sloan Digital Sky Survey by Mandelbaum et al. (2013). The new estimator can be used to constrain cosmological parameters as well as the bias between galaxies and the dark matter distribution.

In this work, we show that the $\Upsilon$ statistic is a special case of the aperture mass formalism (Schneider 1996; Schneider et al. 1998). Using this information, we generalize the B10 approach; in particular, we define a complete set of estimators for a given range of scales which all are 'blind' to the correlation functions below a predescribed threshold. We expect that the first few elements of this discrete set contain all the relevant information, which thus leads to a substantial data compression and a lower dimensional covariance, similar to the Complete Orthogonal Sets of E-/B-mode Integrals (COSEBIs) statistic for cosmic shear (Schneider, Eifler & Krause 2010).

As a proof of concept, in this paper we fix the cosmology and use the new estimators to measure the galaxy bias of a particular galaxy sample. For this study, we use as lenses the galaxies from the Baryon Oscillation Spectroscopic Survey (BOSS) LOWZ sample (Eisenstein et al. 2011) and as sources photometrically selected background galaxies from the Red Cluster Sequence Lensing Survey (RCSLenS;[1] Hildebrandt et al., in preparation). In order to establish the accuracy of the estimator and to create the corresponding covariance matrix, we use mock catalogues, which are based on the simulations by Harnois-Deraps & van Waerbeke (2015).

Galaxy bias describes how galaxies trace the underlying dark matter field (Kaiser 1984). In this analysis, we concentrate on the linear bias factor $b$, which is defined as the square root of the ratio of the galaxy to dark matter power spectra, and the galaxy–matter cross-correlation coefficient $r$. The bias of the LOWZ sample was measured in Chuang et al. (2013) and in a different way by Parejko et al. (2013), the bias of the CMASS sample for example in Nuza et al. (2013). The WiggleZ sample was analysed in Marín et al. (2013) using three-point correlation functions. Measuring these parameters is crucial for redshift-space-distortion studies as well as many cosmological measurements where $b$ and $r$ represent nuisance parameters.

This is the first measurement of galaxy bias using galaxy–galaxy lensing in RCSLenS, a re-analysis of the Red Cluster Sequence Survey 2 (RCS2; Gilbank et al. 2011). In van Uitert et al. (2015), a similar measurement has been carried out on RCS2 using correlation functions instead of the advanced statistics we introduce here. Blake et al. (2015) present galaxy–galaxy measurements on the RCSLenS data to constrain modified gravity. The focus of this study is to introduce a new methodology and the benefits of data compression taking galaxy bias measurements as an example.

This paper is organized as follows. In Section 2, we introduce the B10 method, our generalization, and the approach to measure the galaxy bias. Section 3 describes the data analysis, and in Section 4 we give a detailed discussion. As the fiducial cosmology we use a flat $\Lambda$ cold dark matter cosmology constrained by *Planck* with $H_0 = 67.74\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.3089$, $\Omega_\Lambda = 0.6911$, and $\sigma_8 = 0.8159$ (Planck Collaboration XIII 2015). To test the sensitivity of our results with respect to cosmological parameters, we also use the cosmology obtained in Heymans et al. (2013): $H_0 = 73.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.271$, $\Omega_\Lambda = 0.729$, and $\sigma_8 = 0.799$.

## 2 METHOD

### 2.1 The $\Upsilon$ statistics interpreted as $M_\mathrm{ap}$

In B10 two new estimators were introduced, one in terms of the projected galaxy correlation function $\omega_\mathrm{p}$ and one in terms of the differential surface mass density $\Delta\Sigma$ around galaxies. This is measured using weak gravitational lensing, namely the tangential shear component $\gamma_\mathrm{t}$. These estimators are simultaneously analysed in order to recover information about the dark matter distribution. In this section, we will generalize these estimators, but instead of $\omega_\mathrm{p}$ and $\Delta\Sigma$ we will use the angular correlation function $\omega(\vartheta)$ and the tangential shear $\gamma_\mathrm{t}(\vartheta)$ around (foreground) galaxies. These quantities can be obtained from large photometric lensing surveys where spectroscopic redshift information is not available. When using only photometric redshifts, measuring $\omega_\mathrm{p}$ is not sensible. Nevertheless, for this proof-of-concept study, we make use of a spectroscopically selected galaxy sample. This simplifies the interpretation of the results since the spectroscopic sample has a well-defined redshift- and galaxy-type distribution. Furthermore, it is possible to measure the galaxy bias for such a sample by different means, like higher order clustering or redshift-space distortions. While measuring angular correlation functions for galaxies with spectroscopic redshifts might seem unnecessary, doing so makes this technique directly applicable to future photometric surveys that lack spectroscopy.

The estimator introduced by B10 in the case of the tangential shear $\gamma_\mathrm{t}$ is[2]

$$\hat{\Upsilon}(\vartheta, \vartheta_\mathrm{min}) = \gamma_\mathrm{t}(\vartheta) - \left(\frac{\vartheta_\mathrm{min}}{\vartheta}\right)^2 \gamma_\mathrm{t}(\vartheta_\mathrm{min}), \tag{1}$$

where $\vartheta_\mathrm{min}$ is the scale below which small-scale information is suppressed. There are two features in the definition of $\hat{\Upsilon}(\vartheta, \vartheta_\mathrm{min})$ which require attention. First, it is a continuous function of the scale $\vartheta > \vartheta_\mathrm{min}$. In any analysis, the signal needs to be measured in bins of $\vartheta$. This means that the angular scale needs to be discretized when comparing measurements with theoretical predictions. It is usually unclear how this discretization is optimized, as there is a balance between having enough points to include all relevant cosmological information on the one hand and to limit the number of points for a manageable covariance matrix on the other hand. A second feature is the occurrence of $\gamma_\mathrm{t}(\vartheta_\mathrm{min})$ for every $\vartheta$ in $\hat{\Upsilon}$, which means that any uncertainty in this quantity will affect $\hat{\Upsilon}(\vartheta, \vartheta_\mathrm{min})$ at all scales $\vartheta$. Furthermore, as the tangential shear at a fixed angular separation cannot be measured, but must be averaged over a finite interval, this can introduce systematics in the measurement of $\gamma_\mathrm{t}(\vartheta_\mathrm{min})$, and thus the $\hat{\Upsilon}(\vartheta, \vartheta_\mathrm{min})$. In fact, Mandelbaum et al. (2013) determined $\gamma_\mathrm{t}(\vartheta_\mathrm{min})$ by a power-law fit of the tangential shear (more precisely, of $\Delta\Sigma$) over a finite interval bracketing both sides of the minimum scale.

[1] www.rcslens.org

[2] As mentioned before, B10 actually define $\Upsilon$ in terms of $\Delta\Sigma$. To be consistent throughout the paper, we use $\gamma_\mathrm{t}$. Thus, we denote the B10 statistics in terms of $\gamma_\mathrm{t}$ as $\hat{\Upsilon}$.

Here we address all these issues, by first relating the $\hat{\Upsilon}$ statistic to the aperture mass (Schneider 1996), which is defined as

$$M_{\rm ap} = \int_{\phi_{\rm min}}^{\phi_{\rm max}} {\rm d}\phi\, \phi\, \mathcal{U}(\phi)\, \kappa(\phi), \tag{2}$$

where $\kappa(\phi)$ is the convergence, azimuthally averaged over polar angle and over the foreground galaxy population, $\mathcal{U}$ is a compensated filter function, i.e.

$$\int_{\phi_{\rm min}}^{\phi_{\rm max}} {\rm d}\phi\, \phi\, \mathcal{U}(\phi) = 0, \tag{3}$$

and $\phi_{\rm min}$ and $\phi_{\rm max}$ the inner and outer scales on which the weight function is non-zero. The aperture mass can be expressed in terms of the azimuthally averaged tangential shear $\gamma_{\rm t}$, yielding

$$M_{\rm ap} = \int_{\phi_{\rm min}}^{\phi_{\rm max}} {\rm d}\phi\, \phi\, \mathcal{Q}(\phi)\, \gamma_{\rm t}(\phi), \tag{4}$$

where $\mathcal{Q}$ is related to $\mathcal{U}$ via

$$\mathcal{Q}(\phi) = \frac{2}{\phi^2} \int_0^{\phi} {\rm d}\phi'\, \phi'\, \mathcal{U}(\phi') - \mathcal{U}(\phi). \tag{5}$$

For every value of $\vartheta$, we can interpret $\hat{\Upsilon}$ as an aperture mass. Indeed, comparing equation (4) with equation (1), we see immediately that $\hat{\Upsilon}(\vartheta, \vartheta_{\rm min})$ is a special case of $M_{\rm ap}$ if we set $\phi_{\rm min} = \vartheta_{\rm min}$, $\phi_{\rm max} = \vartheta$, and

$$\mathcal{Q}(\phi) = +\frac{1}{\phi}\delta_{\rm D}(\phi - \vartheta) - \frac{\vartheta_{\rm min}}{\vartheta^2}\delta_{\rm D}(\phi - \vartheta_{\rm min}), \tag{6}$$

where $\delta_{\rm D}$ is the Dirac delta function. Inverting equation (5), we find

$$\mathcal{U}(\phi) = -\mathcal{Q}(\phi) + 2\int_{\phi}^{\infty} {\rm d}\phi' \frac{\mathcal{Q}(\phi')}{\phi'}, \tag{7}$$

which yields

$$\begin{aligned}\mathcal{U}(\phi) = &-\frac{1}{\phi}\delta_{\rm D}(\phi - \vartheta) + \frac{\vartheta_{\rm min}}{\vartheta^2}\delta_{\rm D}(\phi - \vartheta_{\rm min}) \\ &+ \frac{2}{\vartheta^2}\left[\mathcal{H}(\vartheta - \phi) - \mathcal{H}(\vartheta_{\rm min} - \phi)\right],\end{aligned} \tag{8}$$

where $\mathcal{H}$ is the Heaviside step function. This equation shows that the $\hat{\Upsilon}$ statistics is indeed insensitive to $\kappa(\vartheta)$ on scales $\vartheta < \vartheta_{\rm min}$, and thus allows the exclusion of small scales where theoretical predictions are currently uncertain.

### 2.2 Measuring $\Upsilon$ by using a set of orthogonal functions

The filter functions $\mathcal{U}$ and $\mathcal{Q}$ of the aperture mass depend on the scale $\vartheta$ of $\hat{\Upsilon}$. Instead of using a continuum of scales $\vartheta$, we can define a complete set of compensated filter functions $\mathcal{U}_n$ over the range of scales $\vartheta_{\rm min} \le \vartheta \le \vartheta_{\rm max}$, i.e. each filter function satisfies

$$\int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} {\rm d}\vartheta\, \vartheta\, \mathcal{U}_n(\vartheta) = 0. \tag{9}$$

The completeness ensures that the corresponding set of aperture masses contains the full information contained in $\hat{\Upsilon}(\vartheta, \vartheta_{\rm min})$ for $\vartheta_{\rm min} \le \vartheta \le \vartheta_{\rm max}$. In fact, we expect that most of the information is included in only the first few elements of this set, whereas the remaining ones contain essentially only noise. This is due to the fact that the weight functions $\mathcal{U}_n$ are ordered according to their number of roots, together with the fact that the galaxy–galaxy lensing signal is not expected to contain substantial small-scale structure. Working with a few numbers, instead of a continuous function, will ease the analysis, in particular the generation of covariances, due to the associated data compression, while keeping the essential features of $\hat{\Upsilon}$, i.e. suppression of small-scale influence.

Given the many other studies measuring galaxy bias for BOSS galaxies, it is clear that the data compression is not crucial for this kind of measurement. However, with future surveys becoming increasingly large and the desire to split the huge galaxy samples into many subsamples (in redshift, type, etc.), it will become more important to minimize the size of the data vector. Since mock catalogues need to be used to estimate covariances, their required number directly scales with the number of elements in the data vector. This study represents a simple test case that can be directly compared to the literature in order to validate the method.

We choose the filter functions to be orthogonal, i.e.

$$\int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} {\rm d}\vartheta\, \mathcal{U}_n(\vartheta)\, \mathcal{U}_m(\vartheta) = 0 \quad \text{for} \quad m \neq n. \tag{10}$$

The Legendre polynomials $\mathcal{P}_n$ form a complete orthogonal set of functions on $[-1, 1]$, which we can use to find a set of suitable filter functions. We decide to use the Legendre polynomials as they already have many of the desired properties for the filter functions. For this to work, we define the transformation used in Schneider et al. (2010)

$$x = \frac{2(\vartheta - \bar{\vartheta})}{\Delta\vartheta}, \tag{11}$$

with $\Delta\vartheta = \vartheta_{\rm max} - \vartheta_{\rm min}$, $\bar{\vartheta} = (\vartheta_{\rm min} + \vartheta_{\rm max})/2$, and ${\rm d}\vartheta = \frac{\Delta\vartheta}{2}{\rm d}x$. This maps the interval $[\vartheta_{\rm min}, \vartheta_{\rm max}]$ on to $[-1, 1]$. Setting

$$\mathcal{U}_n(\vartheta) = \frac{1}{(\Delta\vartheta)^2} u_n\left(\frac{2(\vartheta - \bar{\vartheta})}{\Delta\vartheta}\right), \tag{12}$$

where we explicitly impose the dependence on $x$ and normalize by $1/(\Delta\vartheta)^2$, so that the $\mathcal{U}_n$ have correct units. This transforms the compensation and orthogonality conditions into

$$\int_{-1}^{1} {\rm d}x\, \left(\frac{x\Delta\vartheta}{2} + \bar{\vartheta}\right)\, u_n(x) = 0 \tag{13}$$

and

$$\int_{-1}^{1} {\rm d}x\, u_n(x)\, u_m(x) = \delta_{nm}, \tag{14}$$

where in the latter case we fixed the normalization of the filter functions. The Legendre polynomials can be defined via the recurrence relation

$$\mathcal{P}_{n+1}(x) = \frac{1}{n+1}\left[(2n+1)\, x\, \mathcal{P}_n(x) - n\, \mathcal{P}_{n-1}(x)\right], \tag{15}$$

with $\mathcal{P}_0(x) = 1$ and $\mathcal{P}_1(x) = x$. We first try to find dimensionless filters $u_n(x)$ which are proportional to the $\mathcal{P}_n(x)$; these can then be transformed into the $\mathcal{U}_n(\vartheta)$ according to equation (12). The first function to fulfil our two conditions is a first-order polynomial of the form $u_1(x) = a_1 x + a_0$, where the two coefficients $a_i$ are determined from the two conditions, to yield

$$u_1(x) = \frac{3Gx - 1}{\sqrt{2(1 + 3G^2)}}, \tag{16}$$

where we defined $G = 2\bar{\vartheta}/\Delta\vartheta$. Since

$$\int_{-1}^{1} {\rm d}x\, \mathcal{P}_n(x)\, x^m = 0 \tag{17}$$

for $m < n$ and because the Legendre polynomials are orthogonal, we can choose for $n \ge 2$ the filter functions

$$u_n(x) = \sqrt{\frac{2n+1}{2}}\, \mathcal{P}_n(x)\mathcal{H}(1 - x^2), \tag{18}$$

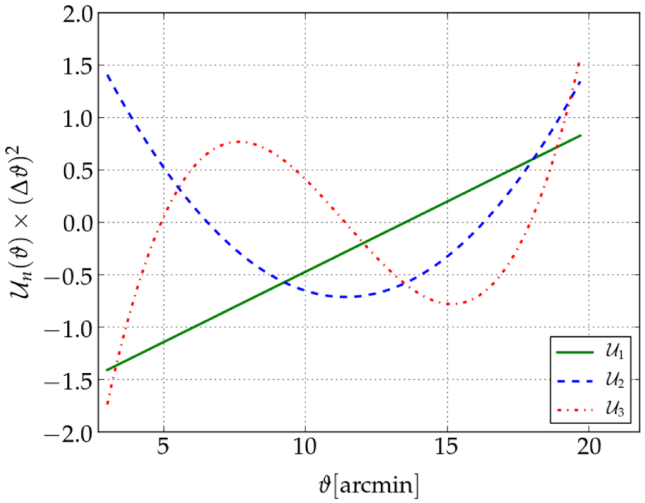


**Figure 1.** Filter functions $\mathcal{U}_n(\vartheta)$ for clustering defined by equations (19) and (20) for $\vartheta_{\rm min} = 3$ arcmin and $\vartheta_{\rm max} = 20$ arcmin. These enter the clustering estimator via equation (23).

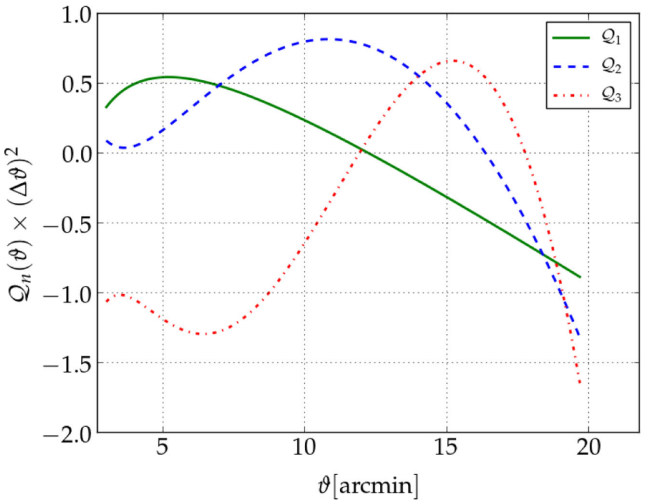


**Figure 2.** Filter functions $\mathcal{Q}_n(\vartheta)$ for lensing defined by equation (21) for $\vartheta_{\rm min} = 3$ arcmin and $\vartheta_{\rm max} = 20$ arcmin. These enter the galaxy–galaxy lensing estimator via equation (22).

which has the correct normalization, and we explicitly included the finite interval of support for the $u_n$. Using equation (12), we then find

$$\begin{aligned}\mathcal{U}_n(\vartheta) &= \frac{1}{(\Delta\vartheta)^2} u_n(x) \\ &= \frac{1}{(\Delta\vartheta)^2}\sqrt{\frac{2n+1}{2}}\mathcal{P}_n\left(\frac{2(\vartheta-\bar{\vartheta})}{\Delta\vartheta}\right) \\ &\quad \times \mathcal{H}(\vartheta-\vartheta_{\rm min})\mathcal{H}(\vartheta_{\rm max}-\vartheta),\end{aligned} \tag{19}$$

for $n \geq 2$ and

$$\begin{aligned}\mathcal{U}_1(\vartheta) &= \frac{1}{(\Delta\vartheta)^2}\frac{3G\left(\frac{2(\vartheta-\bar{\vartheta})}{\Delta\vartheta}\right)-1}{\sqrt{2(1+3G^2)}} \\ &\quad \times \mathcal{H}(\vartheta-\vartheta_{\rm min})\mathcal{H}(\vartheta_{\rm max}-\vartheta).\end{aligned} \tag{20}$$

The $\mathcal{Q}_n(\vartheta)$ follow immediately as

$$\mathcal{Q}_n(\vartheta) = \frac{2}{\vartheta^2}\int_0^{\vartheta} {\rm d}\vartheta'\ \vartheta'\,\mathcal{U}_n(\vartheta') - \mathcal{U}_n(\vartheta). \tag{21}$$

The final estimators for galaxy–galaxy lensing then become

$$\Upsilon_{\rm gm}(n) = \int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} {\rm d}\vartheta\ \vartheta\ \ \mathcal{Q}_n(\vartheta)\,\gamma_{\rm t}(\vartheta). \tag{22}$$

We want to compare the clustering of galaxies with the galaxy–galaxy lensing signal, to learn about the biasing of galaxies and the cross-correlation coefficient between the galaxies and the underlying matter distribution. Thus, we define integrals of the galaxy angular correlation function that have the same angular dependence as the filter functions for the convergence $\kappa$, i.e.

$$\Upsilon_{\rm gg}(n) = \int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} {\rm d}\vartheta\ \vartheta\ \ \mathcal{U}_n(\vartheta)\,\omega(\vartheta). \tag{23}$$

Note that the clustering signal will be measured using the lens sample from galaxy–galaxy lensing in order to probe the same density field. During our analysis, we will make use of only the first three orders of the filter functions; for our data set, those should contain all relevant information. The corresponding filter functions for $\vartheta_{\rm min} = 3$ arcmin and $\vartheta_{\rm max} = 20$ arcmin are displayed in Figs 1 and 2.

### 2.3 Connecting observables to theory

In order to constrain cosmological parameters or to measure the bias factor, we need to know how the observables $\Upsilon_{ij}(n)$ are connected to predictable theoretical quantities like the three-dimensional dark matter power spectrum $\mathcal{P}_{\rm 3D}(k, w)$, where $k$ is the comoving wavenumber and $w$ the comoving distance, characterizing the cosmic epoch. This is now shown for the case where the lens sample has a rather broad redshift distribution, as for increasingly small distributions the following approximations for the angular correlations diverge and are not valid any more.

In the following, we assume that the bias is linear and can be described by

$$\hat{b}^2 = \frac{\mathcal{P}_{\rm gg}}{\mathcal{P}_{\rm 3D}}, \tag{24}$$

with $\mathcal{P}_{\rm gg}(k, w)$ being the galaxy power spectrum. This assumption is valid on large scales which we explicitly limit ourselves to with the $\Upsilon$ formalism. Furthermore, we define the cross-correlation coefficient

$$\hat{r} = \frac{\mathcal{P}_{\rm gm}}{\sqrt{\mathcal{P}_{\rm gg}\mathcal{P}_{\rm 3D}}}, \tag{25}$$

where $\mathcal{P}_{\rm gm}(k, w)$ is the cross-power spectrum between matter and galaxies. $\hat{r}$ is important for determining the galaxy–matter cross-correlations.

The angular correlation function of galaxies is related to $\mathcal{P}_{\rm 3D}$ through (Hoekstra et al. 2002b)

$$\begin{aligned}\omega(\vartheta) &= \frac{1}{2\pi}\int {\rm d}w\ \left(\frac{p_{\rm lw}(w)}{f_k(w)}\right)^2 \\ &\quad \times \int {\rm d}\ell\ \ell\ \hat{b}^2(\ell, z)\mathcal{P}_{\rm 3D}\left(\frac{\ell}{f_k(w)}; w\right) J_0(\ell\vartheta),\end{aligned} \tag{26}$$

where $\hat{b}(\ell, z)$ is the galaxy bias as a function of angular wavenumber $\ell = k f_k(w)$ and redshift $z$, $w$ the comoving distance, $f_k(w)$ the comoving angular diameter distance, $p_{\rm lw}(w)$ the lens probability distribution in terms of $w$, and $J_0$ the zeroth-order Bessel function of the first kind. Changing the order of integration and replacing the probability distribution with respect to $w$, $p_{\rm lw}(w)$, by the observable

redshift distribution, using $p_{\rm lz}(z)\mathrm{d}z = p_{\rm lw}(w)\mathrm{d}w$, yields

$$\omega(\vartheta) = \frac{1}{2\pi}\int \mathrm{d}\ell\ \ell\ J_0(\ell\vartheta) \times \int \mathrm{d}w \left(\frac{p_{\rm lz}(z)}{f_k(w)}\right)^2 \left(\frac{\mathrm{d}z}{\mathrm{d}w}\right)^2 \hat{b}^2(\ell, z)\mathcal{P}_{\rm 3D}\left(\frac{\ell}{f_k(w)}; w\right), \tag{27}$$

with

$$\frac{\mathrm{d}z}{\mathrm{d}w} = \frac{H_0\sqrt{(1+z)^2(1+z\Omega_{\rm m}) - z(2+z)\Omega_\Lambda}}{c}.$$

By inserting equation (27) into equation (23), we obtain an expression for $\Upsilon_{\rm gg}(n)$, which depends quadratically on the galaxy bias

$$\Upsilon_{\rm gg}(n) = \frac{b^2}{2\pi}\int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} \mathrm{d}\vartheta\ \vartheta\ \mathcal{U}_n(\vartheta) \times \int \mathrm{d}\ell\ \ell\ J_0(\ell\vartheta) \int \mathrm{d}w \left(\frac{p_{\rm lz}(z)}{f_k(w)}\right)^2 \times \left(\frac{\mathrm{d}z}{\mathrm{d}w}\right)^2 \mathcal{P}_{\rm 3D}\left(\frac{\ell}{f_k(w)}; w\right). \tag{28}$$

Here we defined $b$ as a weighted average of the bias $\hat{b}(\ell, z)$ over $\ell$ and $z$, where the weight is given by the factors in the second integral in equation (27). We point out that $b$ still depends on the order $n$ (due to the dependence of the angular weight function $\mathcal{U}_n$ on $\vartheta$), which we do not write out explicitly.[3] The connection between $\mathcal{P}_{\rm 3D}$ and $\gamma_{\rm t}(\vartheta)$ has been shown to be (Kaiser 1992; Guzik & Seljak 2001)

$$\gamma_{\rm t}(\vartheta) = \frac{3\ \Omega_{\rm m}}{4\pi}\left(\frac{H_0}{c}\right)^2 \int \mathrm{d}w\ \frac{g(w)p_{\rm lw}(w)}{a(w)f_k(w)} \times \int \mathrm{d}\ell\ \ell\ \hat{b}(\ell, z)\,\hat{r}(\ell, z)\mathcal{P}_{\rm 3D}\left(\frac{\ell}{f_k(w)}; w\right) J_2(\ell\vartheta), \tag{29}$$

where $\hat{r}$ is the cross-correlation coefficient, $a(w)$ is the cosmic scale factor, and $g(w)$ is the mean of angular diameter distances (e.g. Schneider et al. 1998)

$$g(w) = \int_w^{w_{\rm H}} \mathrm{d}w'\ p_{\rm sw}(w')\frac{f_k(w'-w)}{f_k(w')}, \tag{30}$$

with $p_{\rm sw}(w)$ the source distance probability distribution in terms of $w$. Again, by changing the order of integration and inserting the redshift probability distribution into equation (22), one finds

$$\Upsilon_{\rm gm}(n) = \frac{3\ \Omega_{\rm m}}{4\pi}\left(\frac{H_0}{c}\right)^2 b\,r \int_{\vartheta_{\rm min}}^{\vartheta_{\rm max}} \mathrm{d}\vartheta\ \vartheta\ \mathcal{Q}_n(\vartheta) \times \int \mathrm{d}\ell\ \ell\ J_2(\ell\vartheta) \int \mathrm{d}w\ \frac{g(w)p_{\rm lz}(z)}{a(w)f_k(w)}\frac{\mathrm{d}z}{\mathrm{d}w} \times \mathcal{P}_{\rm 3D}\left(\frac{\ell}{f_k(w)}; w\right). \tag{31}$$

As before, we use the weighted average of $\hat{b}$ and $\hat{r}$ over $\ell$, $z$, and $\vartheta$. When measuring $\Upsilon_{\rm gm}(n)$ and $\Upsilon_{\rm gg}(n)$ from the data, we can simultaneously fit the models to both signals. In this way, we can either

(i) fix the cosmology and constrain $b$ and $r$,
(ii) fix $b$ and $r$ and constrain the cosmology,
(iii) set $r = 1$ and fit $b$ and the cosmology simultaneously,
(iv) or constrain $b$, $r$, and the cosmology simultaneously.

[3] When constraining $b$ later on, we will actually constrain an average over $n$, $\ell$, and $z$.

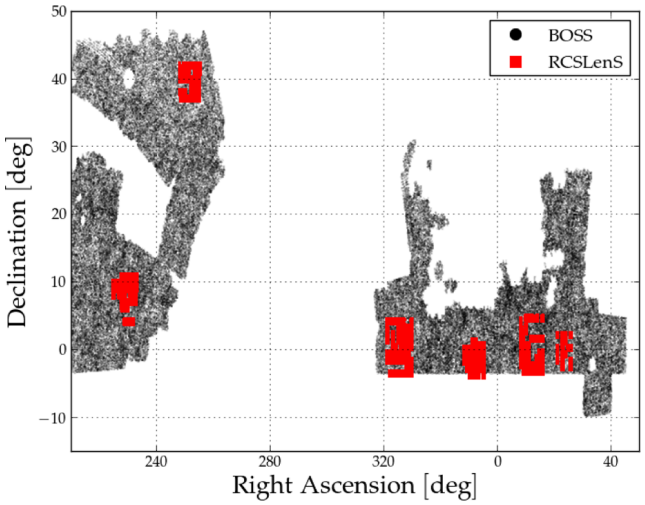


**Figure 3.** RCSLenS regions that were used, and the galaxies from BOSS. The RCSLenS regions are non-contiguous because of the lack of four-band data, which is needed for photometric redshifts.

The latter is possible by combining galaxy clustering and galaxy–galaxy lensing with a cosmic shear signal, weighted by the same kernel functions $\mathcal{U}_n(\vartheta)$. Since the scope of this work is to prove the concept, we will use a fixed cosmology and constrain the galaxy bias $b$ and the cross-correlation coefficient $r$.

## 3 DATA ANALYSIS

We choose to apply our new methodology to determine a large-scale bias measurement of the BOSS LOWZ sample. This sample is well suited for this first analysis, as there are already measurements and it is less complicated compared to a whole cosmological study.

### 3.1 Data sets

#### *3.1.1 BOSS LOWZ*

We measure the weak lensing signal around galaxies from BOSS (Eisenstein et al. 2011), using the 10th Data Release (Ahn et al. 2014). We select galaxies following Chuang et al. (2013) and Sánchez et al. (2013) to select a spectroscopic redshift sample with $0.15 \leq z \leq 0.43$. This yields 9102 galaxies within the RCSLenS footprint. For the lensing measurements, we only use the BOSS galaxies that lie within the BOSS–RCSLenS overlap; however, for the clustering measurement, the whole LOWZ sample is used, which is spread over a much larger area ($\sim$5000 deg$^2$; Tojeiro et al. 2014) and consists of 218 891 galaxies. In this way, we can make use of the much better statistics arising from the larger sample. This is a valid approach as in Section 3.4 we show that the signals measured for both samples are consistent with each other. The BOSS and RCSLenS overlapping area is shown in Fig. 3. The summed $p_{\rm lz}(z)$ derived from spectroscopic redshifts of the lenses can be seen in Fig. 4. For the clustering measurements, we make use of the weights, $\Theta$, provided by the BOSS collaboration, which account for fibre collisions as explained in Anderson et al. (2014).

#### *3.1.2 RCSLenS*

RCSLenS (Hildebrandt et al., in preparation) is an analysis of the original RCS2 using the Canada–France–Hawaii Telescope Lensing

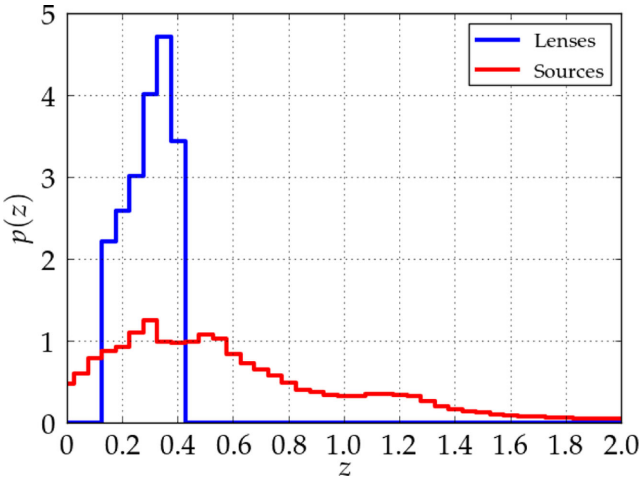


**Figure 4.** $p_{\mathrm{l}z}(z)$ of lenses (blue) and $p_{\mathrm{s}z}(z)$ of sources (red). For the lenses, we use the spectroscopic redshifts to estimate $p_{\mathrm{l}z}(z)$, whereas for the sources we make use of the stacked full $p(z)$ of every source galaxy, which is estimated by the photometric redshift code. The distributions are normalized so that $\sum p(z)\Delta z = 1$. Additionally, we weight the distributions using the weights described in Section 3.

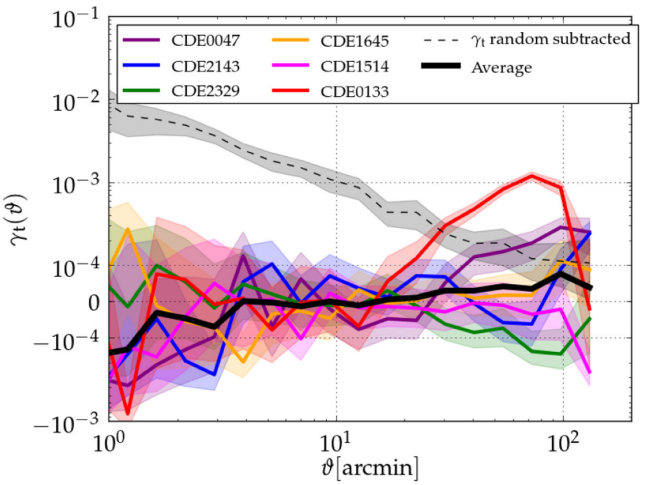


**Figure 5.** The lensing signal around random points. The coloured lines show the signal for every region, whereas the black solid line shows the average. Furthermore, we display the measured signal of $\gamma_{\mathrm{t}}$ around BOSS LOWZ galaxies as the dashed black line. The shaded regions correspond to the $1\sigma$ errors. The strongest random signal corresponds to CDE0133, which is the smallest region in the area we use, and thus it has the smallest impact on the total signal. For the measurements, we subtract the signal for each region separately.

Survey (CFHTLenS) pipeline (Heymans et al. 2012; Hildebrandt et al. 2012; Erben et al. 2013; Miller et al. 2013) to reduce the data and create shape and photometry catalogues. The survey was carried out using Megacam at the Canada–France–Hawaii Telescope (CFHT) and has only one exposure per band per pointing. It covers roughly 500 deg$^2$ in the $g'$, $r'$, $i'$, and $z'$ bands and with an additional 250 deg$^2$ with three or fewer bands. The $r'$ band is used as the lensing band with a $5\sigma$ point source limiting magnitude of $m_{\mathrm{lim}} = 24.3$ and a median seeing of 0.71 arcsec (Gilbank et al. 2011). Galaxy shapes are measured using *lens*fit (Miller et al. 2013). As described in Blake et al. (2015) we use the *lens*fit weights $\eta$ and the BOSS weights $\Theta$ for the lensing analysis. We take both weights in order to use the same weighting scheme in the lensing as well as in the clustering analysis. The resulting estimator is

$$\langle\gamma_{\mathrm{t}}(\vartheta)\rangle = \frac{\sum_{i,\mathrm{sources}}\sum_{j,\mathrm{lenses}}\eta_i\Theta_j e_{\mathrm{t},i,j}}{\sum_{i,\mathrm{sources}}\sum_{j,\mathrm{lenses}}\eta_i\Theta_j}. \tag{32}$$

Here $\eta_i$ denotes the *lens*fit weight of the $i$th source galaxy and $\Theta_j$ the BOSS weight of the $j$th lens galaxy, whereas $e_{\mathrm{t},i,j}$ is the tangential ellipticity of the $i$th source with respect to the $j$th lens. For selecting source galaxies, we only use the six RCSLenS regions that have four-band photometry and sufficient overlap with BOSS. Those are CDE0133, CDE0047, CDE1645, CDE2329, CDE1514, and CDE2143. This leaves us with about 170 deg$^2$ in area and 4657 415 source galaxies. As sources we select all galaxies with a *lens*fit weight $\eta > 0$ that are outside of masks. We use the posterior redshift distribution for each source galaxy, estimated with the photometric redshift code BPZ (Benítez 2000), to find the summed $p_{\mathrm{s}z}(z)$ of the sources, which is displayed in Fig. 4.

The shear measurements for RCSLenS suffer from a multiplicative as well as an additive bias so that

$$\langle e_{\mathrm{obs}}\rangle = (1 + \langle m\rangle)\langle e_{\mathrm{true}}\rangle + c, \tag{33}$$

as explained for example in Miller et al. (2013). Here $e_{\mathrm{obs}}$ is the observed ellipticity of a galaxy image, $e_{\mathrm{true}}$ the sheared intrinsic ellipticity, $1 + m$ the correction factor for the multiplicative bias ($m$-correction), and $c$ the correction for the additive bias ($c$-correction). We correct the measured shapes of galaxies for the multiplicative bias using the factor $(1 + m)$ determined for every galaxy (for more details see e.g. Miller et al. 2013 or Hildebrandt et al., in preparation). We apply the $m$-correction as an ensemble correction in order to avoid correlations between the correction and the intrinsic shape of the galaxy (Miller et al. 2013)

$$\langle\gamma_{\mathrm{t}}^{\mathrm{cal}}(\vartheta)\rangle = \frac{\langle\gamma_{\mathrm{t}}(\vartheta)\rangle}{1 + K(\vartheta)}, \tag{34}$$

where

$$1 + K(\vartheta) = \frac{\sum\eta_i\Theta_j(1 + m_i)}{\sum\eta_i\Theta_j}. \tag{35}$$

As before, $\eta_i$ denotes the *lens*fit weight of the $i$th source galaxy and $\Theta_j$ the BOSS weight of the $j$th lens galaxy. The sums are taken over all lens–source pairs separated by the angle $\vartheta$. The correction $1 + K(\vartheta)$ is of the order of 0.95 for all scales used. As common in galaxy–galaxy lensing studies (e.g. Mandelbaum et al. 2006), we do not apply an additive $c$-correction but subtract the $\gamma_{\mathrm{t}}$ signal measured around random points, which is equivalent to a direct $c$-correction for galaxy–galaxy lensing measurements. To determine this correction, the number of random points used depends on the region size and differs between ∼100 000 and ∼180 000. The measured signal around random points is consistent with zero on scales below 30–40 arcmin and rises out to larger scales, where for $\vartheta >$ 40 arcmin it can reach an amplitude of a few times $10^{-4}$ for some regions. We subtract this signal for every region separately as it would average out when combined from all regions. The signals are shown in Fig. 5. The region with the strongest random signal is CDE0133, which is the smallest in area and thus contributes the least to the total signal.

For the weighted average source density, we find ∼5.1 galaxies/arcmin$^2$ when using

$$n_{\mathrm{eff}} = \frac{1}{A_{\mathrm{eff}}}\frac{\left(\sum\eta_i\right)^2}{\sum(\eta_i)^2}, \tag{36}$$

as defined in Heymans et al. (2012), where $A_{\mathrm{eff}} = 174.32$ deg$^2$ is the total unmasked area in the BOSS–RCSLenS overlap. We use this definition to account for the fact that we use the *lens*fit weight in the analysis. The RCSLenS catalogues are also subject to

a blinding scheme. In order to avoid confirmation bias, the galaxy ellipticities exist in four versions A, B, C, and D. One of them is the true measured one, whereas the rest have been changed by a small factor as described in Hildebrandt et al. (in preparation) for RCSLenS and in Kuijken et al. (2015) for the Kilo Degree Survey. This analysis has been performed four times using the different ellipticity versions. After the analysis had been finished, the lead author contacted the external blinder, Matthias Bartelmann, who revealed which catalogue was the truth. We then used the results of the true measured ellipticities only. No changes were made after 'unblinding'. For more information about RCSLenS and the data production process, we refer to Hildebrandt et al. (in preparation).

### 3.2 Mock catalogues

In order to estimate the covariance of the $\Upsilon$s, we make use of the simulations described in Harnois-Deraps & van Waerbeke (2015). Those have box sizes of $505\,h^{-1}$ Mpc, $1536^3$ particles each and are on $3072^3$ grids, which are projected on to $12\,288^2$ pixels. The light cones are then extracted from those on to $6000^2$ pixels grids. The cosmology used is $\Omega_\mathrm{m} = 0.2905$, $\Omega_\Lambda = 0.7095$, $\sigma_8 = 0.826$, and $H_0 = 68.98\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The slight difference to the cosmologies we use will introduce a small systematic error in the covariance, which we will neglect in this study.

Based on these simulations, we use a set of mock catalogues designed to match the properties of the RCSLenS sources and the BOSS LOWZ lenses. They specifically match the ellipticity and redshift distributions of RCSLenS and the clustering properties of the LOWZ sample. We apply photometric redshift scatter to the mock sources through a $z_\mathrm{spec}$–$z_\mathrm{phot}$ matrix calibrated from the BPZ redshift probability distributions. The mock LOWZ lenses are added to the simulation using a halo occupation distribution (HOD) approach calibrated by matching the observed clustering amplitude. In total, we use 360 mock catalogues, which are $60\,\mathrm{deg}^2$ each. The size of the region used for the mocks is just determined by the size of the simulations themselves. We do not aim to simulate the whole survey area, but for practicality we area-scale the covariance from the $60\,\mathrm{deg}^2$ outputs. Using six of the mocks, we can create one mock survey, assuming that each of the six RCSLenS regions fits within the $60\,\mathrm{deg}^2$. This then results in 60 mock realizations of RCSLenS. Whenever the regions are too big, we use as much area as possible and scale the covariance accordingly by using the ratio of the area of the mock region and the real region. Furthermore, for the covariance estimation we use only the BOSS–RCSLenS overlap for the measurements of the clustering signal, whereas for the data we use the whole BOSS area. In order to account for this, we rescale the clustering part of the covariance with the ratio of the two areas. Additionally, we set the cross-covariance between $\Upsilon_\mathrm{gg}$ and $\Upsilon_\mathrm{gm}$ to 0, as the BOSS–RCSLenS overlap is just a small fraction of the whole BOSS area. This has been shown to be a valid approach by More et al. (2015), who conduct similar measurements with BOSS and the CFHTLenS catalogues. In the end, we have 60 mock surveys, to which we apply the same masks as for the data set. For this, we neglect that the mocks assume a flat sky, as the resulting differences are clearly negligible compared to the statistical error of our measurements given the small extent of each region.

### 3.3 Measuring two-point correlations

Before we can determine the compressed observables $\Upsilon_{ij}(n)$, we first need to measure the corresponding galaxy–galaxy lensing and galaxy clustering signals. We choose to measure these in two intervals

(i) $3\,\mathrm{arcmin} \leq \vartheta \leq 20\,\mathrm{arcmin}$,
(ii) $20\,\mathrm{arcmin} \leq \vartheta \leq 70\,\mathrm{arcmin}$

in 200 linear bins. The centre of the first range corresponds to comoving length of $\sim$3 Mpc at a redshift of $z \approx 0.29$, and the second one to a comoving length of $\sim$12 Mpc. These are both large scale, which will enable us to measure the large-scale bias of the LOWZ sample. As a cross-check, we also determine these signals for a larger angular scale in larger logarithmic bins. The 200 linear bins will later be used for determining the $\Upsilon$. For $\omega(\vartheta)$, we use the Landy–Szalay estimator (Landy & Szalay 1993). We show the mean signals for $\gamma_\mathrm{t}$ and $\omega$ measured in the mocks together with the real data in Fig. 6. Those measurements are in good agreement.

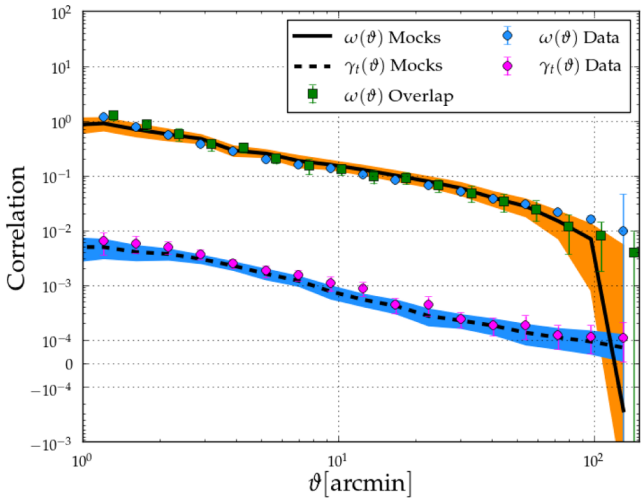


**Figure 6.** Comparison of galaxy clustering and galaxy–galaxy lensing signals in the mocks and data. The black lines show the mean; the $1\sigma$ standard deviation is indicated by the blue and orange shaded regions. The measurement from the data is shown as the blue and magenta points. They are in good agreement with the mocks. Additionally, the clustering signal measured just for the BOSS–RCSLenS overlap is displayed as the green points. This is consistent with the signal from the whole LOWZ sample.

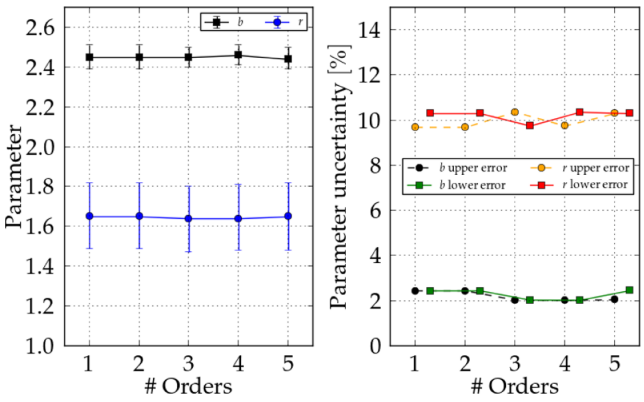


**Figure 7.** Left: the measured parameter values as a function of maximum $\Upsilon$-order $n$ for $b$ and $r$. No significant difference in the values is visible, from which we conclude that the data compression is indeed working and only a few orders contain all the information from the measured signals. Right: the parameter uncertainty in per cent for $b$ and $r$, again as a function of $n$. Here, we do not find a significant difference, which again shows that the data compression of $\Upsilon(n)$ is robust.

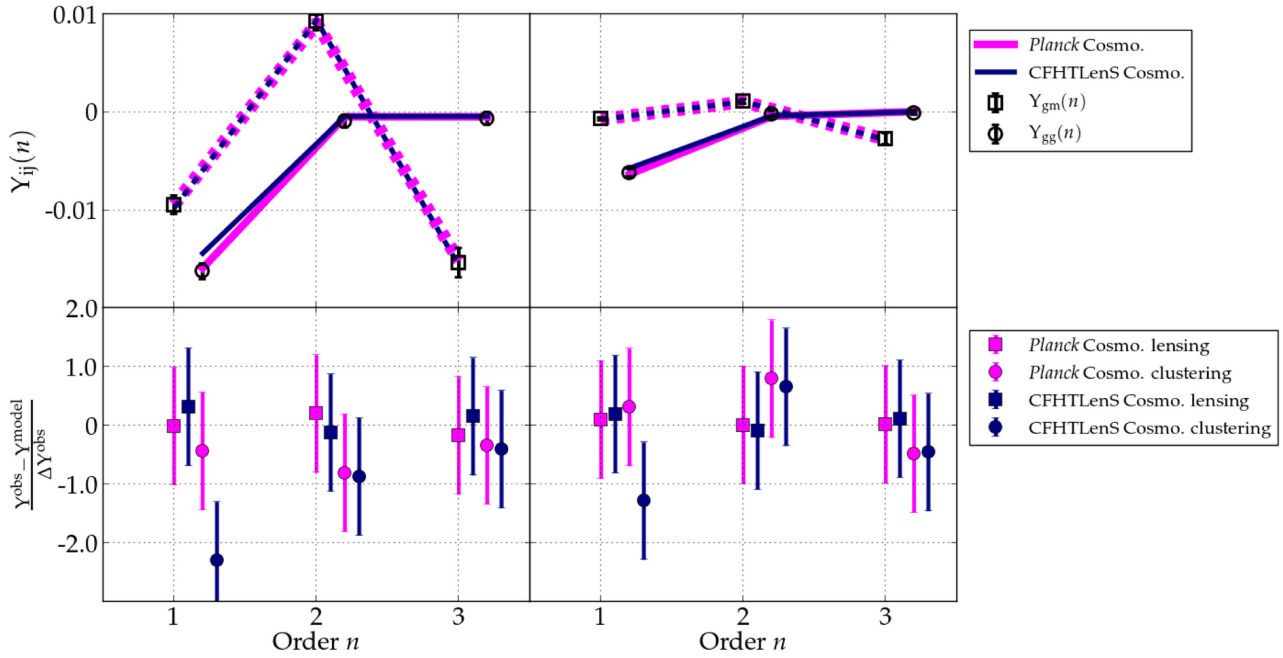


**Figure 8.** The top panels show the measured $\Upsilon_{gm}$ and $\Upsilon_{gg}$ and the best fit using one of the two cosmologies. The magenta and dark blue lines are the connections between the predicted data points using the *Planck* or the CFHTLenS cosmology. In the bottom panels, we show the residuals $(\Upsilon^{obs} - \Upsilon^{model})/\Delta\Upsilon^{obs}$, where $\Delta\Upsilon^{obs}$ is the uncertainty in the measured $\Upsilon$. Left: measurements for the 3−20 arcmin interval. Clearly, on these scales the model we adopt to describe the galaxy bias is not a good description of the data shown here, especially the clustering data. For more details, see Section 3.4 Right: measurements for the 20−70 arcmin interval.

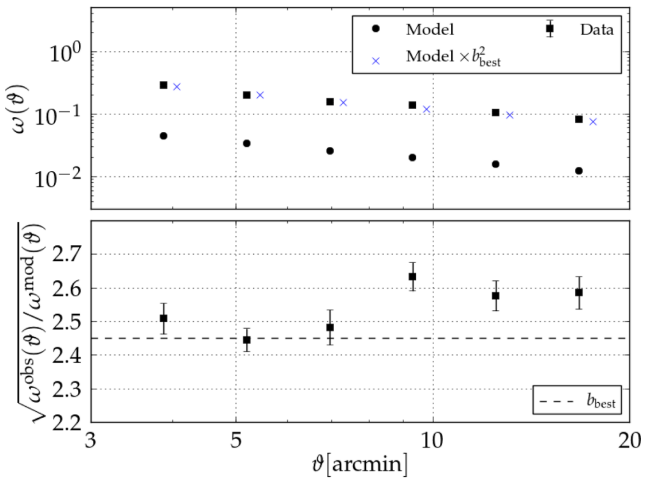


**Figure 9.** Top: the angular correlation function within the small-scale interval, the corresponding model for $b = 1$, and the best-fitting model. This best-fitting model has been determined from a joint fit of $\Upsilon_{gm}$ and $\Upsilon_{gg}$. Bottom: the square root of the ratio between the measured $\omega(\vartheta)$ and the model one, which is an estimator for $b$. Apparently, in contradiction to our assumption, there is a scale dependence of $b$. This is why the data shown in the left-hand panel of Fig. 8 are not well described by the model. A variation of $b$ of about 5 per cent within this interval would already be enough to reconcile the data and the model. The data shown here correspond to the *Planck* cosmology measurements.

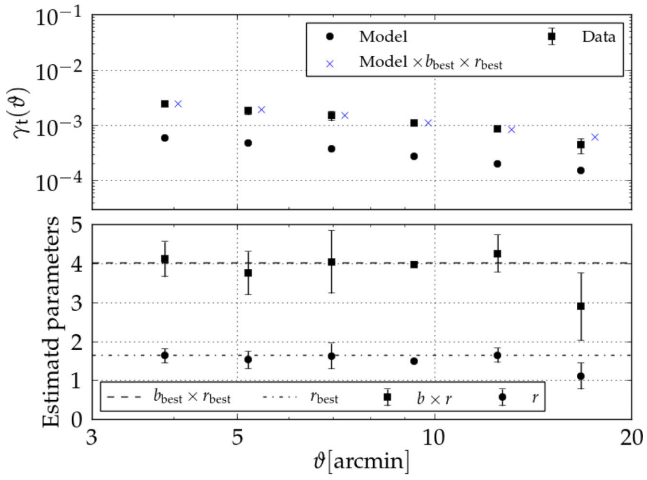


**Figure 10.** Top: the tangential shear function within the small-scale interval, the corresponding model for $b$, $r = 1$, and the best-fitting model. This best-fitting model has been determined from a joint fit of $\Upsilon_{gm}$ and $\Upsilon_{gg}$. Bottom: the ratio between the measured $\gamma_t(\vartheta)$ and the model one, which is an estimator for $b \times r$. Due to the larger uncertainties, the shear measurements do not show a preference for scale-dependent bias. Additionally, we also show the corresponding estimate for $r$, if we use $b$ as estimated in Fig. 9. The data shown here correspond to the *Planck* cosmology measurements.

### 3.4 $\Upsilon_{gm}(n)$ and $\Upsilon_{gg}(n)$

We use $\gamma_t(\vartheta)$ and $\omega(\vartheta)$ measured in the 200 linear bins and integrate them using equations (22) and (23) in order to find $\Upsilon_{gm}(n)$ and $\Upsilon_{gg}(n)$. Here we only compute the first three orders. At the end of our analysis, we tested how the parameter constraints on $b$ and $r$ changed with the number of $\Upsilon$ orders used. We found no significant difference for up to five orders and decided to use three orders, which yields a sufficient number of data points for our analysis and still benefits from a low-dimensional covariance. The fact that we do not find a decrease of parameter uncertainty with increasing number of orders shows that the first few orders indeed contain all the relevant information (see Fig. 7 for more details). The measured data points for both angular intervals are presented in Fig. 8. Unlike

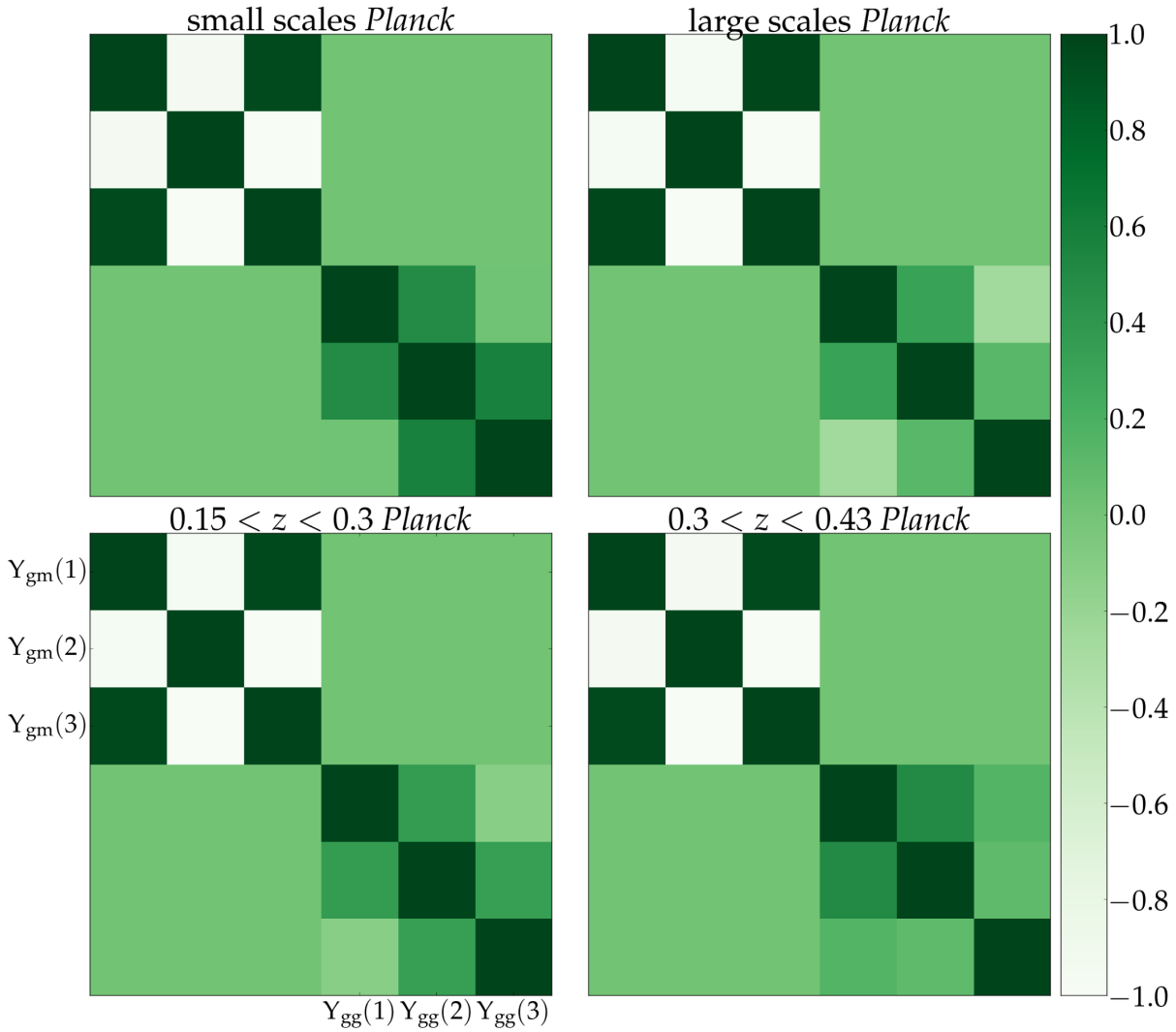


**Figure 11.** The correlation matrices for $\Upsilon_{gm}$ and $\Upsilon_{gg}$ for the measurements using the *Planck* cosmology. As they only depend on the cosmology used in the mocks, we do not show the correlation matrices for the CFHTLenS cosmology measurements. The upper-left part of the matrices corresponds to galaxy–galaxy lensing, and the bottom-right part to galaxy clustering. The cross-covariance terms are set to 0 as the area for the lensing measurement is only a small fraction of the clustering area, which makes those measurements independent. In the order left to right, top to bottom we show the matrix for the 3−20 arcmin interval and the *Planck* cosmology, the 20−70 arcmin interval and the *Planck* cosmology, the 3−20 arcmin interval and the $0.15 < z < 0.3$ lens sample, and the 3−20 arcmin interval and the $0.3 < z < 0.45$ lens sample.

correlation function measurements, these $\Upsilon$ data points cannot be interpreted easily. However, it is clear that in the large-scale interval the model (see Section 2) is a very good fit to the data regardless of the cosmological parameters used.[4] This is not the case for the smaller scale interval, where one of the clustering data points is several $\sigma$ away from the best-fitting model. Clearly, the assumption of linear bias on these non-linear scales is not valid for the clustering data. This is partly due to the small uncertainties in these measurements as well as the fact that we neglect the model uncertainties. As can be seen in Fig. 9, a change of $b$ of about 5 per cent within this interval would already be enough to reconcile the data with the model. If model uncertainties had been included in this figure, it is likely that data and model would be in line again. Another possible explanation of the discrepancy between data and model in this case could be that the fiducial cosmology is wrong. Furthermore, we also investigate if an indication for a scale-dependent bias can be found in the $\gamma_t(\vartheta)$ data in Fig. 10 and find no such preference.

From the 60 mock realizations, we compute the $\Upsilon_{gm}$ and $\Upsilon_{gg}$ covariance matrix by measuring the signals for each mock survey. For the inverse covariance, we take into account the correction factor from Hartlap, Simon & Schneider (2007), which prevents us from underestimating the uncertainty in the parameter estimates. The correlation matrices for all measurements are shown in Fig. 11. The covariance matrix is then used for a maximum likelihood analysis, in which we simultaneously fit theoretical predictions to $\Upsilon_{gm}$ and $\Upsilon_{gg}$ with the galaxy bias $b$ and the cross-correlation coefficient $r$ as free parameters. We compute the predictions from equations (28) and (31) using the 3D matter power spectrum computed with NICAEA (Kilbinger et al. 2009), which uses the recipe from Smith et al. (2003). The resulting likelihood contours are displayed in Fig. 12. We perform this fit twice using the *Planck* cosmology as well as the best-fitting cosmology from CFHTLenS, constrained in

[4] Note that the data points are highly covariant (see also Fig. 11).

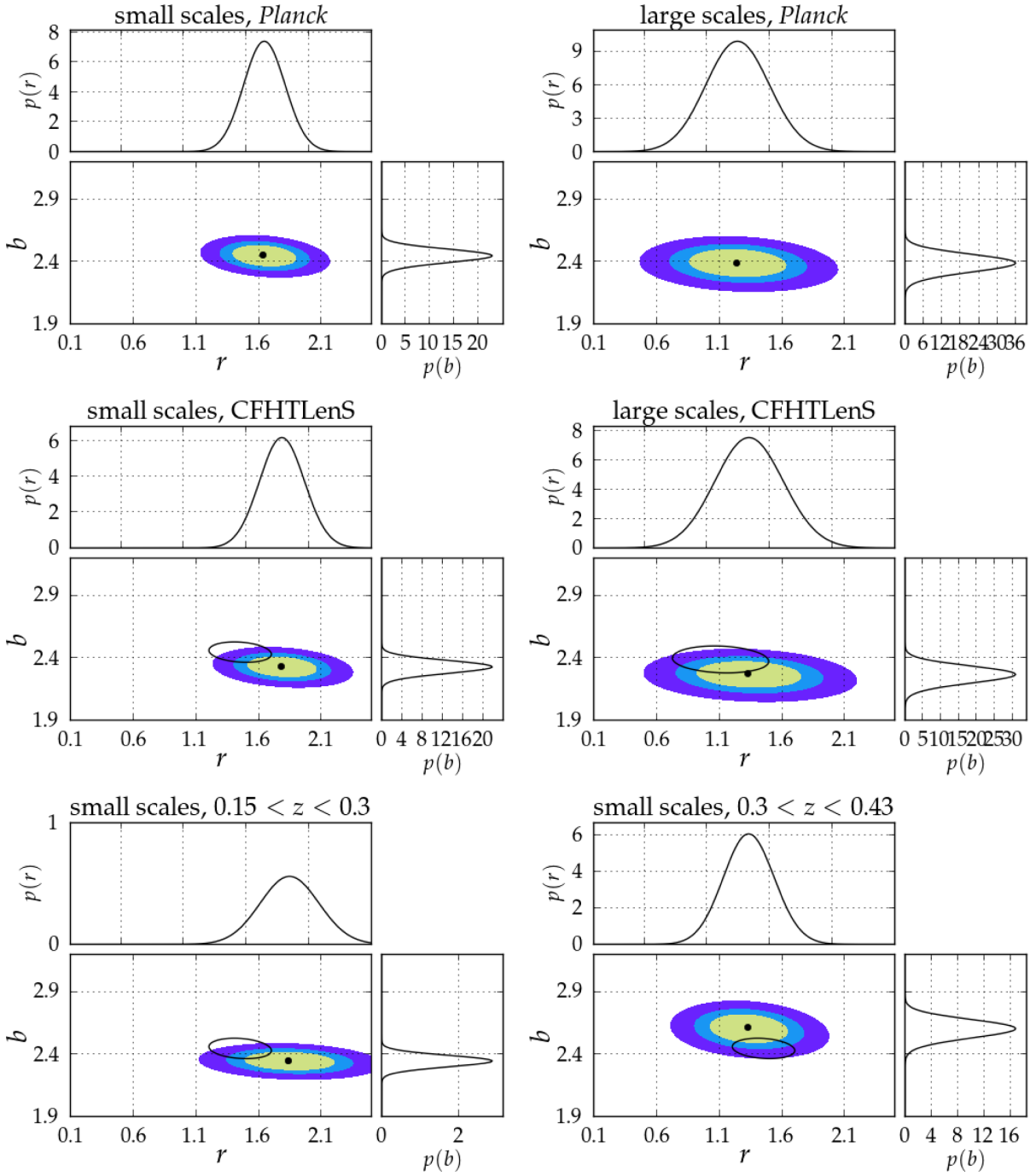


**Figure 12.** The $1\sigma$, $2\sigma$, and $3\sigma$ parameter constraints on the galaxy bias parameter $b$ and $r$ for the BOSS LOWZ galaxy sample as well as the marginalized likelihoods of $b$ and $r$. The black ellipse, if shown, is the $1\sigma$ contour of the corresponding measurement using a *Planck* cosmology from the upper two panels. These parameters were constrained by a maximum likelihood fit to the $\Upsilon_{\rm gm}$ and $\Upsilon_{\rm gg}$. All constraints agree within $1\sigma$. Top left: likelihood contours for the 3−20 arcmin interval using a *Planck* cosmology. Top right: likelihood contours for the 20−70 arcmin interval using a *Planck* cosmology. Middle left: likelihood contours for the 3−20 arcmin interval using the Heymans et al. (2013) cosmology. Middle right: likelihood contours for the 20−70 arcmin interval using the Heymans et al. (2013) cosmology. Bottom left: likelihood contours for the 3−20 arcmin interval using a *Planck* cosmology and the $0.15 < z < 0.3$ lens sample. Bottom right: likelihood contours for the 3−20 arcmin interval using a *Planck* cosmology and the $0.3 < z < 0.43$ lens sample.

Heymans et al. (2013), to test for the dependence of the parameters on different cosmologies. The results are presented in Table 1. For the maximum likelihood analysis, we assume a Gaussian likelihood function. Note that one cannot directly interpret the $\chi^2$/d.o.f. values since the model is non-linear and the data noisy (Andrae, Schulze-Hartung & Melchior 2010). We find $b = 2.45^{+0.05}_{-0.05}$ and $r = 1.64^{+0.17}_{-0.16}$ for the small-scale interval, and for angular scales of 20 arcmin $\leq \vartheta \leq$ 70 arcmin we find $b = 2.39^{+0.07}_{-0.07}$ and $r = 1.24^{+0.26}_{-0.25}$.

The estimated values for $b$ are slightly higher compared to the findings by Parejko et al. (2013), who determine the bias by fitting their projected clustering signal to HOD populated $N$-body simulations. Using their best-fitting model and the

**Table 1.** Parameter estimates for galaxy bias $b$ and cross-correlation coefficient $r$. In the case of the full sample, the second column indicates the cosmology used. For the samples used in Section 3.5, it indicates which subsample was used.

| Scale (arcmin) | | $b$ | $r$ | $\chi^2$/d.o.f. |
|---|---|---|---|---|
| 3–20 | *Planck* | $2.45^{+0.05}_{-0.05}$ | $1.64^{+0.17}_{-0.16}$ | 0.38 |
| 3–20 | CFHTLenS | $2.33^{+0.05}_{-0.05}$ | $1.78^{+0.18}_{-0.18}$ | 0.53 |
| 20–70 | *Planck* | $2.39^{+0.07}_{-0.07}$ | $1.24^{+0.26}_{-0.25}$ | 0.47 |
| 20–70 | CFHTLenS | $2.27^{+0.07}_{-0.07}$ | $1.33^{+0.28}_{-0.27}$ | 0.38 |
| 3–20 | $0.15 < z < 0.3$ | $2.35^{+0.04}_{-0.05}$ | $1.84^{+0.24}_{-0.23}$ | 2.01 |
| 3–20 | $0.3 < z < 0.45$ | $2.61^{+0.07}_{-0.08}$ | $1.33^{+0.21}_{-0.20}$ | 0.73 |

corresponding simulations, they predict the bias for the LOWZ sample as a function of physical scale. For 3 Mpc, which corresponds to about 11 arcmin at a redshift of 0.29, they find a bias of ∼2.2, whereas for 12 Mpc (∼45 arcmin) it corresponds to a bias of ∼2.1. This differs by ∼10 per cent from our results. The discrepancy could be explained by our approach of averaging over $\ell$ and $z$ and the corresponding weight functions, but as there are no error bars in Parejko et al. (2013), we cannot judge how significant the difference is. Chuang et al. (2013) also measure the bias for the LOWZ sample, finding a value of $b \times \sigma_8 = 1.102 \pm 0.039$ for scales between 24 and 200 $h^{-1}$ Mpc. This corresponds to a significantly smaller value of $b$ compared to the findings in this study. However, the two approaches of measuring the bias, as well as the scales used are very different. This discrepancy could therefore be resolved if we considered scale-dependent bias. Whereas one might have expected that the cross-correlation coefficient $r$ is close to unity on these scales, we instead find it to be $3\sigma$ away from unity. On large scales, however, we find $r$ to be close to unity as expected for deterministic large-scale bias. One should note that a measured $r > 1$ is possible, as was discussed in B10 and also found by Marian, Smith & Angulo (2015) in the Millennium simulations, as the angular galaxy correlation function is a shot-noise subtracted estimator. Furthermore, we point out that the values measured for different cosmologies differ by a few per cent which is smaller than the parameter uncertainties from statistical errors.

### 3.5 Redshift evolution test: splitting up the LOWZ sample

In Fig. 13, we show the measured signals for $\gamma_t(\vartheta)$ and $\omega(\vartheta)$ for the whole sample as well as for the two subsamples (described below). We also scale the expected signals for both with the constrained values of $b$ and $r$. The data are consistent with constant values of $b$ and $r$, and the values for both parameters obtained from the fit to the $\Upsilon$s are consistent with the signals of the correlation functions $\gamma_t(\vartheta)$ and $\omega(\vartheta)$. This means that the method introduced here is indeed capable of compressing the data while not losing information contained in the correlation functions.

Furthermore, we conduct a redshift evolution test where we split up the lens sample into two subsamples with $0.15 < z < 0.3$ and $0.3 < z < 0.43$. In this way, we can test if the model is capable of describing these measurements in a proper way. We then make the same measurements as before using the *Planck* cosmology and the $\vartheta \in [3\,\text{arcmin}, 20\,\text{arcmin}]$ interval. This yields two new estimates for $b$ and for $r$. We find $b = 2.35^{+0.04}_{-0.05}$ and $r = 1.84^{+0.24}_{-0.23}$ for the low-redshift sample and $b = 2.61^{+0.07}_{-0.08}$ and $r = 1.33^{+0.21}_{-0.20}$ for the high-redshift one. They are also shown in Table 1. The measured correlation functions are displayed in Fig. 13 and the likelihood contours in Fig. 12. We find that $r$ becomes smaller for the higher redshift sample, whereas $b$ gets larger. All estimates are, however, consistent with the parameters determined using the whole sample. In fact, the two subsample values for $b$ and $r$ bracket their whole sample counterparts.

## 4 DISCUSSION AND OUTLOOK

We introduced a new estimator for galaxy clustering, $\Upsilon_{gg}$, and for galaxy–galaxy lensing, $\Upsilon_{gm}$. Those are generalizations of the methods introduced and tested in B10 and Mandelbaum et al. (2013), respectively. The estimators are a discretization of the $\Upsilon(\vartheta, \vartheta_{min})$, which leads to substantial data compression and thus a lower dimensional covariance, while still eliminating the sensitivity to the matter distribution on small scales. Especially, lowering the dimension of the data covariance increases the accuracy in its measurement for a fixed number of mock realizations. Recall that the number of mock realizations needed to find a good estimate of the covariance increases with the number of data points. We applied this method to data using the BOSS LOWZ sample as lenses and galaxies from the RCSLenS as sources. While fixing the cosmology, we performed a simultaneous fit to $\Upsilon_{gg}$ and $\Upsilon_{gm}$ with $b$ and $r$ as free parameters. For different angular scales as well as different assumed cosmologies, we find $b$ slightly higher than the findings of Parejko et al. (2013) and Chuang et al. (2013). This tension could be resolved if our assumption of scale-independent bias was a poor approximation to the true galaxy bias of this sample, as both of the studies mentioned allow for scale-dependent bias.

On large angular scales, the cross-correlation coefficient $r$ is found to be compatible with unity, as expected for the corresponding spatial scales (e.g. B10). On the smaller angular scale interval, we find a value for $r$ that is significantly larger than unity, most likely due to a different scale and redshift dependence of the various power spectra that enter the $\Upsilon$s in equations (28) and (31), and our definition of the 'effective' coefficients $b$ and $r$ as an average of the three-dimensional bias and correlation coefficients $\hat{b}$ and $\hat{r}$. If one had already measured values for $b$ and $r$, this method can even be used for cosmological studies. In these studies, it will be necessary to find out how many orders of $\Upsilon$ are sufficient to extract all cosmological information from the signal. As in this work it was not possible to do so as all information is already contained in the first few orders, due to our simplified bias models. This might change in a cosmological analysis from substantially larger data sets with more complex models.

Summarizing, the new estimators presented in this paper are promising tools for future large-scale structure studies, especially given their advantageous abilities concerning data compression and the dimension of the data covariance.

## ACKNOWLEDGEMENTS

We are grateful to the RCS2 team for planning the survey, applying for observing time, and conducting the observations. We would like to thank Matthias Bartelmann for being our external blinder, revealing which of the four catalogues analysed was the true unblinded catalogue at the end of this study. We would also like to thank Christopher Morrison and Patrick Simon for fruitful discussions. Also, we would like to thank our anonymous referee for constructive comments. We acknowledge use of the Canadian Astronomy Data Centre operated by the Dominion Astrophysical Observatory for the National Research Council of Canada's Herzberg Institute of

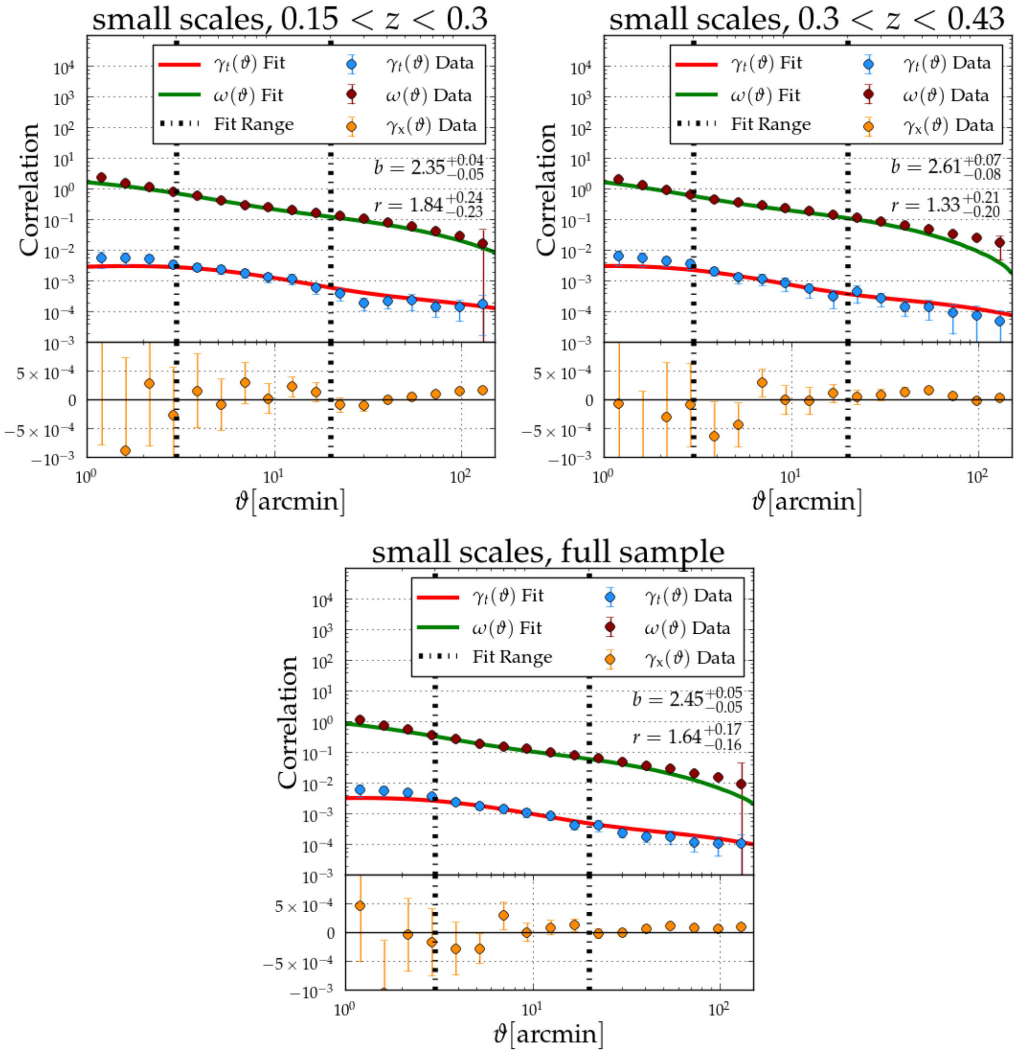


**Figure 13.** Galaxy clustering and galaxy–galaxy lensing signals with the best-fitting theoretical model for the 3−20 arcmin interval using a *Planck* cosmology. The two subsamples from the redshift evolution test are used as well as the full sample. The best-fitting lines were fitted to the $\Upsilon$s, not the signals shown here. Within the fitting range, the estimated parameter values for $b$ and $r$ appear to be in excellent agreement with the data. We also show $\gamma_\times$, which is consistent with zero in all cases. A non-zero $\gamma_\times$ points to systematic issues in the data.

Astrophysics. Computations for the *N*-body simulations were performed on the GPC supercomputer at the SciNet HPC Consortium. SciNet is funded by the Canada Foundation for Innovation under the auspices of Compute Canada, the Government of Ontario, Ontario Research Fund – Research Excellence, and the University of Toronto.

AB was supported for this research partly through a stipend from the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne and through funding from the Transregional Collaborative Research Centre 'The dark Universe' (TR 33) of the DFG. PS acknowledges support from the DFG grant SCHN 342-13. HH is supported by an Emmy Noether grant (No. Hi 1495/2-1) of the Deutsche Forschungsgemeinschaft. CB acknowledges the support of the Australian Research Council through the award of a Future Fellowship. CH and LK acknowledge support from the European Research Council under FP7 grant number 240185. JHD is supported by the NSERC of Canada. RN acknowledges support from the German Federal Ministry for Economic Affairs and Energy (BMWi) provided via DLR under project no. 50QE1103. MV acknowledges support from the European Research Council under FP7 grant number 279396 and the Netherlands Organisation for Scientific Research (NWO) through grants 614.001.103. For this work, we made use of the correlation codes ATHENA[5] (Kilbinger, Bonnett & Coupon 2014) and SWOT[6] (Coupon et al. 2012).

*Author contributions*. All authors contributed to the development and writing of this paper. The authorship list reflects the lead authors of this paper (AB, PS, and HH) followed by two alphabetical groups. The first alphabetical group includes key contributors to the science analysis and interpretation in this paper, the founding core team, and those whose long-term significant effort produced the final RCSLenS data product. The second group covers members of the RCSLenS team who made a significant contribution to the project, this paper, or both. HH led the RCSLenS collaboration.

[5] http://www.cosmostat.org/athena.html
[6] https://github.com/jcoupon/swot

## REFERENCES


Ahn C. P. et al., 2014, ApJS, 211, 17
Anderson L. et al., 2014, MNRAS, 439, 83
Andrae R., Schulze-Hartung T., Melchior P., 2010, preprint (arXiv:1012.3754)
Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625
Baldauf T., Smith R. E., Seljak U., Mandelbaum R., 2010, Phys. Rev. D, 81, 063531 ( B10)
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Benítez N., 2000, ApJ, 536, 571
Blake C. et al., 2012, MNRAS, 425, 405
Blake C. et al., 2015, preprint (arXiv:1507.03086)
Cacciato M., Lahav O., van den Bosch F. C., Hoekstra H., Dekel A., 2012, MNRAS, 426, 566
Chuang C.-H. et al., 2013, preprint (arXiv:1312.4889)
Coupon J. et al., 2012, A&A, 542, A5
Coupon J. et al., 2015, MNRAS, 449, 1352
Eisenstein D. J. et al., 2011, AJ, 142, 72
Erben T. et al., 2013, MNRAS, 433, 2545
Eriksen M., Gaztanaga E., 2015, MNRAS, 452, 2149
Gilbank D. G., Gladders M. D., Yee H. K. C., Hsieh B. C., 2011, AJ, 141, 94
Guzik J., Seljak U., 2001, MNRAS, 321, 439
Harnois-Deraps J., van Waerbeke L., 2015, MNRAS, 450, 2857
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Heymans C. et al., 2012, MNRAS, 427, 146
Heymans C. et al., 2013, MNRAS, 432, 2433
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Hinshaw G. et al., 2013, ApJS, 208, 19
Hoekstra H., Yee H. K. C., Gladders M. D., Barrientos L. F., Hall P. B., Infante L., 2002a, ApJ, 572, 55
Hoekstra H., van Waerbeke L., Gladders M. D., Mellier Y., Yee H. K. C., 2002b, ApJ, 577, 604
Ivezic Z. et al., 2008, preprint (arXiv:0805.2366)
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1992, ApJ, 388, 272
Kilbinger M. et al., 2009, A&A, 497, 677
Kilbinger M., Bonnett C., Coupon J., 2014, Astrophysics Source Code Library, record ascl:1402.026
Kitching T. D. et al., 2012, MNRAS, 423, 3163
Kuijken K. et al., 2015, MNRAS, 454, 3500
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, ApJ, 738, 45
Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715
Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544
Mandelbaum R. et al., 2015, MNRAS, 450, 2963
Mantz A. B., Allen S. W., Morris R. G., Rapetti D. A., Applegate D. E., Kelly P. L., von der Linden A., Schmidt R. W., 2014, MNRAS, 440, 2077
Marian L., Smith R. E., Angulo R. E., 2015, MNRAS, 451, 1418
Marín F. A. et al., 2013, MNRAS, 432, 2654
Miller L. et al., 2013, MNRAS, 429, 2858
More S., Miyatake H., Mandelbaum R., Takada M., Spergel D. N., Brownstein J. R., Schneider D. P., 2015, ApJ, 806, 2
Nuza S. E. et al., 2013, MNRAS, 432, 743
Parejko J. K. et al., 2013, MNRAS, 429, 98
Perlmutter S. et al., 1999, ApJ, 517, 565
Planck Collaboration XIII, 2015, preprint (arXiv:1502.01589)
Riess A. G. et al., 1998, AJ, 116, 1009
Sánchez A. G. et al., 2013, MNRAS, 433, 1202
Schneider P., 1996, MNRAS, 283, 837
Schneider P., van Waerbeke L., Jain B., Kruse G., 1998, MNRAS, 296, 873
Schneider P., Eifler T., Krause E., 2010, A&A, 520, A116
Schrabback T. et al., 2010, A&A, 516, A63
Smith R. E. et al., 2003, MNRAS, 341, 1311
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Tojeiro R. et al., 2014, MNRAS, 440, 2222
van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14
van Uitert E., Hoekstra H., Schrabback T., Gilbank D. G., Gladders M. D., Yee H. K. C., 2012, A&A, 545, A71
van Uitert E., Cacciato M., Hoekstra H., Herbonnet R., 2015, A&A, 579, A26
Van Waerbeke L. et al., 2001, A&A, 374, 757
Velander M. et al., 2014, MNRAS, 437, 2111
Vikhlinin A. et al., 2009, ApJ, 692, 1060


## APPENDIX A: ESTIMATING $\Upsilon_x$

For weak gravitational lensing measurements, it is important to check if the cross shear, $\gamma_x$, is consistent with zero. If not so, this points to systematic issues in the data. We can conduct a similar test for the $\Upsilon_{gm}(n)$, where we replace $\gamma_t$ with $\gamma_x$ in equation (22),

$$\Upsilon_x(n) = \int_{\vartheta_{min}}^{\vartheta_{max}} d\vartheta\ \vartheta\ \ Q_n(\vartheta)\, \gamma_x(\vartheta). \tag{A1}$$

As $\gamma_x$ needs to be zero on all scales, so does its compressed counterpart $\Upsilon_x$. We estimated $\Upsilon_x$ for all six measurements described in this paper and show the signal in Fig. A1. Indeed, we find it to be consistent with zero for all three orders.

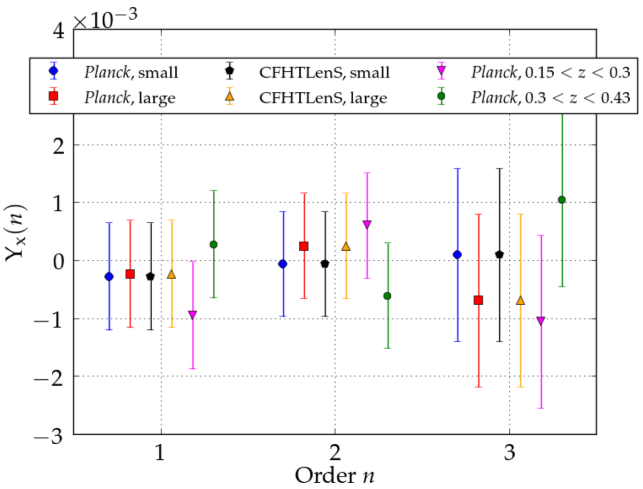


**Figure A1.** We present the $\Upsilon_x$ for all six measurements conducted in this paper. We find it to be always consistent with zero.

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