



Universiteit  
Leiden  
The Netherlands

## Light weighed: on the statistics and systematics of weak gravitational lensing

Smit, D.M.

### Citation

Smit, D. M. (2021, December 8). *Light weighed: on the statistics and systematics of weak gravitational lensing*. Retrieved from <https://hdl.handle.net/1887/3245825>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3245825>

**Note:** To cite this publication please use the final published version (if applicable).



---

# AMICO galaxy clusters in KiDS-DR3: The impact of estimator statistics on the luminosity-mass scaling relation

---

As modern-day precision cosmology aims for statistical uncertainties of the percent level or lower, it becomes increasingly important to reconsider estimator assumptions at each step of the process, along with their consequences on the statistical variability of the scientific results.

We compare  $L^1$  regression statistics to the weighted mean, the canonical  $L^2$  method based on Gaussian assumptions, to infer the weak gravitational shear signal from a catalog of background ellipticity measurements around a sample of clusters, which has been a standard step in the processes of many recent analyses.

We use the shape measurements of background sources around 6925 AMICO clusters detected in the KiDS third data release. We investigate the robustness of our results and the dependence of uncertainties on the signal-to-noise ratios of the background source detections. Using a halo model approach, we derive lensing masses from the estimated excess surface density profiles.

The highly significant shear signal allows us to study the scaling relation between the  $r$ -band cluster luminosity,  $L_{200}$ , and the derived lensing mass,  $M_{200}$ . We show the results of the scaling relations derived in 13 bins in  $L_{200}$ , with a tightly constrained power-law slope of  $\sim 1.24 \pm 0.08$ . We observe a small, but significant, relative bias of a few percent in the recovered excess surface density profiles between the two regression methods, which translates to a  $1\sigma$  difference in  $M_{200}$ . The efficiency of  $L^1$  is at least that of the weighted mean and increases with higher signal-to-noise shape measurements.

Our results indicate the relevance of optimizing the estimator for inferring the gravitational shear from a distribution of background ellipticities. The interpretation of measured relative biases can be gauged by deeper observations, and the increased computation times remain feasible.



## 5.1 Introduction

Statistics is an essential part of astronomy (Heck et al. 1985, Feigelson 1988, 2009, Feigelson & Babu 2013). The field relies on inferring physical properties, which cannot be determined directly, from observable quantities, which in turn need to be corrected for systematic effects as well as instrumental and observational biases. The key question that always needs to be answered when interpreting observations and results – before discussing how accurately these results can be constrained – is what one is actually seeing.

Weak gravitational lensing, caused by the deflection of light rays by density variations along the traveled path, has been a case in point for the last three decades. Gravitational lensing is a convex focusing effect that can magnify and shear affected background sources. The observed shapes and number counts can conversely yield information about these density variations but need to be disentangled statistically from the unknown intrinsic properties of background sources, such as distance, size (and luminosity), and shape.

The first detections of coherent alignments of galaxy shapes were observed in the background of clusters (Tyson et al. 1990), and subsequently in the emerging fields of galaxy-galaxy lensing (where the lensing “structure” is itself an ensemble of lenses; Brainerd et al. 1996) and cosmic shear (the weak lensing induced by large-scale structure; Wittman et al. 2000, Bacon et al. 2000, Kaiser et al. 2000, Van Waerbeke et al. 2000). Since then, techniques have progressed rapidly, and demands on accuracy have become increasingly stringent.

This is the second in a set of papers wherein we focus on the statistical aspects of inferring the lensing signal from the intrinsic shapes and the estimated lensing geometry, which depends on the distances between the observer, the moment of deflection, and the background sources. Assuming the cosmological principle, the intrinsic shapes of a sample<sup>1</sup> of background galaxies, including their orientation, are random, and the intrinsic galaxy shapes should average out from a sufficiently large sample, leaving the weak lensing signal as a net ellipticity. The common approach has been to take a weighted mean of galaxy ellipticities, which has computational and analytical advantages and, most importantly, is an unbiased estimator of the shear in the absence of pixel noise in the galaxy images (Seitz & Schneider 1997).

In practice, however, there are many sources of noise and the mean is known to be biased, underestimating the underlying shear signal (Melchior & Viola 2012, Viola et al. 2014, Sellentin et al. 2018, Mandelbaum 2018). The distribution of intrinsic galaxy shapes is well known to be non-Gaussian (Lambas et al. 1992, Rodríguez & Padilla 2013) and, in fact, centrally peaked. In Smit & Kuijken (2018, hereafter Paper I), we explored alternative estimators besides the mean that could potentially be better suited for such a cuspy distribution. It was found, using realistic simulated distributions and resampling of Canada-France-Hawaii Lensing Survey (CFHTLenS) shape measurements (Heymans et al. 2012b), that  $L^1$  norm regression, also known as Least Absolute Deviations (LAD), reduces bias from between  $\sim -4\%$  and  $\sim -4.5\%$  to between  $\sim +1\%$  and  $\sim -3\%$ , while at the same time reducing uncertainty by  $\sim 9\%$  to  $\sim 23\%$ .

In this paper we extend this study by applying these statistics to a weak lensing analysis of 6925 galaxy clusters in the Adaptive Matched Identifier of Clustered Objects (AMICO) cluster catalog (Bellagamba et al. 2011, Radovich et al. 2017, Bellagamba et al. 2018, Maturi

---

<sup>1</sup>There are several considerations involved in the proper selection of such a sample, as explained in Sects. 5.2 and 5.3.



et al. 2019, Bellagamba et al. 2019) of the third data release of the Kilo-Degree Survey (KiDS-450; de Jong et al. 2017). As opposed to Paper I, in this case the true lensing signal (here in the form of the excess surface density of the clusters) is unknown. We therefore study the relative biases and uncertainties between LAD and the mean, and we compare results to our findings in Paper I.

An important application is then to study the relation between the observable properties of clusters and groups and the physical quantities derived from the lensing signal (i.e., the matter distribution) to better our understanding of galaxy and cluster formation and cosmological models (e.g., Kautsch et al. 2008, Leauthaud et al. 2010, Lesci et al. 2020). We calculate halo masses from the obtained lensing signals and derive a scaling relation between the observed  $r$ -band luminosity and the lensing mass, investigating the impact of estimator choice on the resulting constraints.

The order of magnitude of this estimated bias in the weak lensing results can be dominant compared to other sources of uncertainty in the process. Developments in the field have led to current constraints of the multiplicative bias in shape measurements on the order of  $\sim 1\%$  (Bernstein & Jarvis 2002, Hirata & Seljak 2003, Heymans et al. 2006, Massey et al. 2007, Miller et al. 2007, Kitching et al. 2008, Bridle et al. 2010, Voigt & Bridle 2010, Bernstein 2010, Kitching et al. 2012, Kacprzak et al. 2012, Melchior & Viola 2012, Refregier et al. 2012, Heymans et al. 2012a, Mandelbaum et al. 2015, Viola et al. 2015, Fenech Conti et al. 2017). The uncertainty in the lensing geometry between the observer, lens, and background sources, introduced by the estimation of the photometric redshift probability distributions, can be a few percent (Hildebrandt et al. 2017, Bellagamba et al. 2019, and Appendix 5.A.1). The broad category of selection biases, for example those introduced by intrinsic alignments, contamination of the background sample by cluster member galaxies, blending, detection, and subsequent selection effects, typically accumulate up to a few percent (Miyatake et al. 2015, van Uitert et al. 2017, Bellagamba et al. 2019) for cluster weak lensing. For instance, estimations on background selection yield a foreground contamination on the order of 2%, which can be partly corrected for, but does increase the uncertainty (Dvornik et al. 2017, Bellagamba et al. 2019, and Appendix 5.A.2). In this study we investigate the usability of background sources to radii smaller than in Bellagamba et al. (2019).

These demands on accuracy and precision become higher as the data yield, and therefore the statistical power of surveys, increases dramatically (Mandelbaum 2018), as achieved by COSMOS<sup>2</sup> (Leauthaud et al. 2007), CFHTLenS<sup>3</sup> (Heymans et al. 2012b), RCSLenS<sup>4</sup> (Hildebrandt et al. 2016), KiDS<sup>5</sup> (de Jong et al. 2013), and DES<sup>6</sup> (Dark Energy Survey Collaboration et al. 2016), and foreseen for future surveys such as LSST<sup>7</sup> (Ivezić et al. 2019) and Euclid<sup>8</sup> (Laureijs et al. 2011). While these two future surveys will require constraints on systematic uncertainty of order  $\leq 2 \times 10^{-3}$  (Mandelbaum 2018), we show that, even for weak lensing analyses in the last decade, the bias in shear inference can dominate other sources, such as the aforementioned multiplicative shape measurement bias that is commonly corrected for, as in Viola et al. (2015), Dvornik et al. (2017), and Bellagamba et al. (2019).

---

<sup>2</sup><http://cosmos.astro.caltech.edu/>

<sup>3</sup><http://www.cfhtlens.org>

<sup>4</sup><http://www.rcslens.org/>

<sup>5</sup><http://kids.strw.leidenuniv.nl/>

<sup>6</sup><http://www.darkenergysurvey.org/>

<sup>7</sup><https://www.lsst.org/>

<sup>8</sup><http://www.euclid-ec.org/>



Several other approaches have been made to address this, including analytic modeling of the bias (e.g., Viola et al. 2014), weight corrections and priors (Bonnet & Mellier 1995, Van Waerbeke et al. 2000, Bernstein & Jarvis 2002), or nulling techniques (Herbonnet et al. 2017). The calculation of the main observable, the shapes of lensed background sources, itself relies on statistical methods. These are based mainly on surface brightness moments (Kaiser et al. 1995, Rhodes et al. 2000) or model fitting (Kuijken 1999, Bernstein & Jarvis 2002, Hirata & Seljak 2003, Refregier & Bacon 2003, Kuijken 2006, Miller et al. 2007, Kitching et al. 2008). This means the most common approaches are corrections on a statistic that remains fundamentally skewed (Sellentin et al. 2018, Mandelbaum 2018).

Promising alternative approaches by Bernstein & Armstrong (2014) and Schneider et al. (2015) do not reproduce individual background shapes, but directly determine the underlying shear field from ensembles of background sources, reconsidering these steps in the chain of statistical inference. While future lensing surveys will require innovative improvements, these methods and their priors need to be gauged by deep observations of high signal-to-noise, and it is of fundamental importance that these calibrations are well constrained and do not suffer from even subtle systematic biases. In other words, the comparison of several perspectives is paramount in determining what we actually see.

The remainder of this paper is organized as follows. We introduce the definitions of galaxy shapes and the weak lensing formalism in Sect. 2 and relate these to our statistical approach. Data, analysis methods, and selection criteria are described in Sect. 3, while Sect. 4 states our results and analysis. Section 5 gives a summary of our conclusions.

Throughout this paper we assume a Planck (Planck Collaboration et al. 2014) cosmology with  $\Omega_M = 0.315$ ,  $\Omega_\Lambda = 0.685$ , and  $H_0 = 100.0 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ . All measurements are in co-moving units, unless specifically noted otherwise, such as in Sect. 5.2.

## 5.2 Weak gravitational lensing statistics

We briefly review the principles of weak gravitational lensing and relate the central concepts to our statistical approach, introducing the terminology and notation conventions used in this paper. We refer the reader to excellent reviews, such as Bartelmann & Schneider (2001), Schneider (2006), Hoekstra & Jain (2008), and Bartelmann & Maturi (2017), for more in-depth approaches.

### 5.2.1 Principles of weak lensing

Rays of light are deflected by the curvature of space-time due to mass inhomogeneities along their path. A mass overdensity acts as a convex lens on the light rays from distant sources behind that lens to an observer. In this section, we use  $D_l$  to denote angular-diameter distances from the observer to the lens,  $D_{ls}$  from the lens to the background source, and  $D_s$  from the observer to the background source (see Fig. 5.1), and in the remainder of this paper we translate quantities to co-moving units where necessary.

For the purposes of this work, the extent of the lensing mass along the line of sight, compared to the distances from the observer to the lens and from the lens to the background source, can be considered negligible. In this so-called thin-lens approximation, the deflection of light rays by a deflection angle,  $\hat{\alpha}$ , leads to an effective angular displacement (again, see



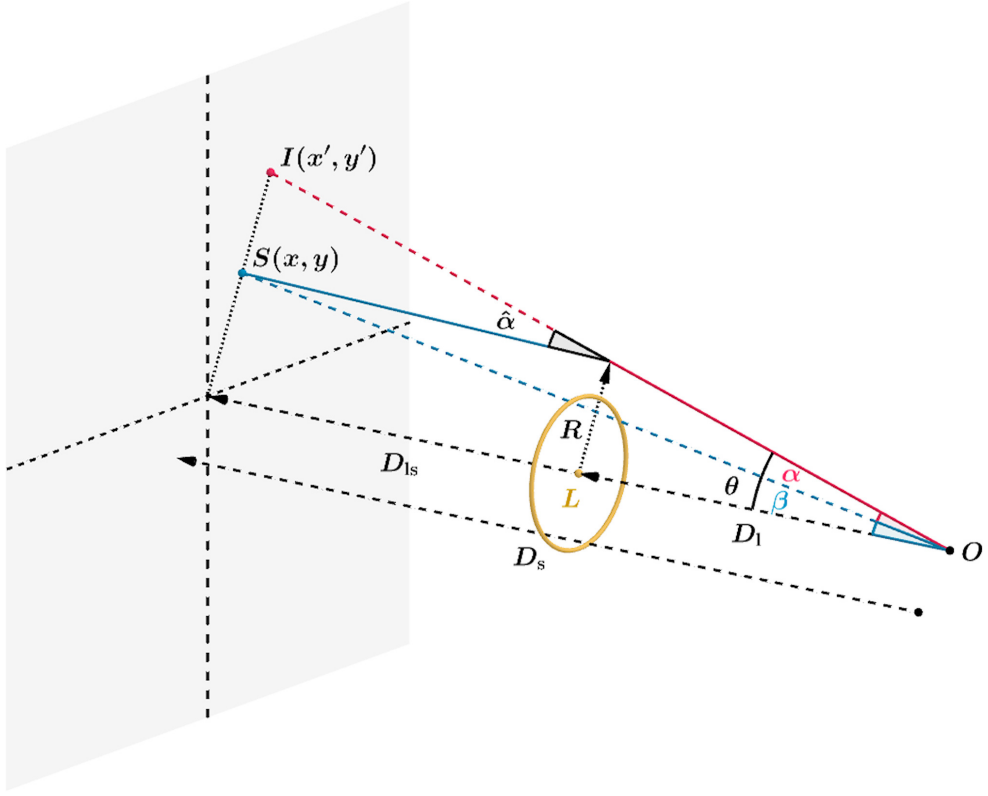


Figure 5.1: Representation of a gravitational lens system, showing the displacement of a source at position  $S(x, y)$  to an image at position  $I(x', y')$ , where we take the origin of the source plane to be collinear with the position of the lens,  $L$ , and the observer,  $O$ .

Fig. 5.1),

$$\vec{\alpha} = -\frac{D_{ls}}{D_s} \hat{\alpha}, \quad (5.1)$$

also called the reduced deflection angle, which relates the observed position,  $\vec{\theta}$ , of a distant point source to its unlensed position,  $\vec{\beta}$ , by the lens equation

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}. \quad (5.2)$$

It can be shown through the relation between  $\hat{\alpha}$  and the three-dimensional gravitational potential,  $\Phi$ , that this displacement is then described by  $\vec{\alpha} = \vec{\nabla}_\theta \psi$ , where

$$\psi = \frac{2}{c^2} \frac{D_{ls}}{D_l D_s} \int \Phi dz \quad (5.3)$$

is called the (two-dimensional) lensing potential.

The differential effect of the deflection of light on the images,  $I(x, y)$ , of extended background sources can to first order be described as a coordinate transformation by taking the



derivatives in the lens equation (Eq. 5.2) of the original angular position,  $\beta$ , with respect to the observed position,  $\theta$ . Substituting  $\vec{\nabla}_\theta \psi$  for  $\vec{\alpha}$ , we obtain the Jacobian matrix,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 - \psi_{11} & -\psi_{12} \\ -\psi_{21} & 1 - \psi_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (5.4)$$

with

$$\psi_{ij} = \frac{\partial^2 \psi}{\partial \theta_i \partial \theta_j}, \quad (5.5)$$

resulting in the lensed image  $I(x', y')$ , which is the key observable in our weak lensing study.

### Critical surface mass density

To interpret the effect on the source image, we note that such a transformation can be decomposed into three parts, namely the identity ( $\mathbf{I}$ ), an isotropic part that describes a multiplication, and an anisotropic traceless part that describes a shearing of the image:

$$\mathbf{I} - \frac{1}{2}(\psi_{11} + \psi_{22})\mathbf{I} + \begin{pmatrix} -\frac{1}{2}(\psi_{11} - \psi_{22}) & -\psi_{12} \\ -\psi_{21} & \frac{1}{2}(\psi_{11} - \psi_{22}) \end{pmatrix}. \quad (5.6)$$

To relate  $\psi_{ij}$  with the density of the lensing mass, we start with the isotropic term, which is half the Laplacian of the lensing potential:  $\frac{1}{2}(\psi_{11} + \psi_{22}) = \frac{1}{2}\nabla_\theta^2 \psi$ . From Eq. 5.3, we obtain

$$\frac{1}{2}\nabla_\theta^2 \psi = \frac{1}{c^2} \frac{D_1 D_{\text{ls}}}{D_s} \int 4\pi G \rho dz, \quad (5.7)$$

which is a dimensionless quantity. Defining the surface mass density as

$$\Sigma \equiv \int \rho dz \quad (5.8)$$

and gathering the rest of the right-hand side into

$$\frac{4\pi G}{c^2} \frac{D_1 D_{\text{ls}}}{D_s} \equiv \Sigma_{\text{cr}}^{-1}, \quad (5.9)$$

with  $\Sigma_{\text{cr}}$  being the critical surface mass density, we find that the isotropic term can be written as

$$\kappa \equiv \frac{1}{2}\nabla_\theta^2 \psi = \frac{\Sigma}{\Sigma_{\text{cr}}}, \quad (5.10)$$

with  $\kappa$  a normalized dimensionless surface mass density. Recognizing that  $\nabla_\theta^2 \psi = \vec{\nabla}_\theta \cdot \vec{\alpha}$  is the divergence of the deflection of the light rays (i.e., the manner in which those light rays converge due to the lensing effect),  $\kappa$  is simply the convergence.



### Shear and intrinsic ellipticity

The shear matrix in Eq. 5.6 has two independent components, simply called the shear  $\gamma = \gamma_1 + i\gamma_2$ , with  $\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22})$  and  $\gamma_2 = \psi_{12} = \psi_{21}$ . Equation 5.4 then becomes

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (5.11)$$

This transformation leads to the magnification and distortion of the light distribution of background sources. In this work, we focus on the most commonly used net distortion or reduced shear  $g = g_1 + ig_2 \equiv (\gamma_1 + i\gamma_2)/(1 - \kappa)$ ,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = (1 - \kappa) \begin{pmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (5.12)$$

where the transformation is written as a multiplication of  $(1 - \kappa)$  and a distortion matrix describing the alignment of lensed sources in the foreground potential.

The effect on a circular source is a shearing into an ellipse with axis ratio  $q = \frac{b}{a}$  as

$$q = \frac{1 - |g|}{1 + |g|} \Leftrightarrow |g| = \frac{1 - q}{1 + q} = \frac{a - b}{a + b} \quad (5.13)$$

and position angle  $\varphi$  via

$$g = |g|(\cos 2\varphi + i \sin 2\varphi). \quad (5.14)$$

As mentioned before, we do not measure this gravitational distortion directly. Background sources have an intrinsic shape distribution, and we effectively measure the combined effect of their intrinsic shape and a weak lensing distortion. It is adequate to describe images by their quadrupole brightness moments or their ellipticities as well as the respective response to weak shear distortions. It is straightforward to use the common definition<sup>9</sup> of ellipticity, defined as the reduced shear needed to create the intrinsic shape  $\epsilon = \epsilon_1 + i\epsilon_2$  of a source from an image with circular isophotes (Bernstein & Jarvis 2002, Kuijken 2006). The resulting ellipticity,  $\epsilon$ , after transforming an image with intrinsic<sup>10</sup> ellipticity  $\epsilon^I$  by a distortion,  $g$ , is then given by (Seitz & Schneider 1997)

$$\epsilon = \frac{\epsilon^I + g}{1 + g^* \epsilon^I} \quad \text{for } |g| \leq 1, \quad (5.15)$$

with  $g^*$  the complex conjugate of  $g$ .

The intrinsic shape distribution is called the shape noise and, assuming no preferred direction on the sky, should average to zero:  $\langle \epsilon^I \rangle = 0$ . This way, each background shape measurement,  $\epsilon$ , is then an independent estimate of the underlying reduced shear,  $g$ .

In this paper we make use of the fact that the lensing signal is weak (i.e.,  $\kappa \ll 1$ ) and assume  $g \approx \gamma$ .

<sup>9</sup> An alternative definition of ellipticity is often denoted as  $|\chi| = \frac{1-q^2}{1+q^2}$ , related to the geometrical eccentricity, and called polarization (e.g., Seitz & Schneider 1995, Viola et al. 2014).

<sup>10</sup> We note that our notation differs from Paper I. Here, the measured ellipticity is denoted as  $\epsilon$ , instead of  $\tilde{e}$  (Paper I), and the intrinsic ellipticity is denoted as  $\epsilon^I$ , instead of  $e$ .



### 5.2.2 Estimation of the surface density profile

The shear induced by gravitational lensing is sensitive to the density contrast. For an axisymmetric lens, we can write  $|\gamma|(R) = \bar{\kappa}(\leq R) - \kappa(R)$ , where  $\bar{\kappa}$  is the average convergence within radius  $R$ . In fact, this relation holds for other mass distributions if we average azimuthally around the lens. In this work, we study the stacked signal of many lenses and assume a net axisymmetry (see, e.g., Evans & Bridle 2009, Oguri et al. 2010, Clampitt & Jain 2016, van Uitert et al. 2017, for weak lensing studies on elliptical lenses).

Since it can be seen from Fig. 5.1 that the gravitational shear acts in the radial direction, we define the tangential and cross components of the shear as

$$\begin{pmatrix} \gamma_+ \\ \gamma_\times \end{pmatrix} = \begin{pmatrix} -\cos(2\phi) & -\sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}, \quad (5.16)$$

with  $\phi$  the counterclockwise angle between the positive  $x$  axis<sup>11</sup> and the vector from lens to source. This gives

$$\bar{\Sigma}(\leq R) - \Sigma(R) \equiv \Delta\Sigma(R) = \gamma_+(R) \Sigma_{\text{cr}}, \quad (5.17)$$

with  $\Delta\Sigma(R)$  the excess surface density (ESD) at a radius  $R$  around the lensing mass. In axisymmetric lenses, the cross component of the shear cannot arise from gravitational lensing and should average to zero, if only produced by intrinsic source orientations, and can therefore be used as an indication of systematic effects, such as imperfect corrections for the point-spread function (PSF; Schneider 2003, and Appendix 5.A.3).

The ESD is then estimated using the observed ellipticities of an ensemble of sources around the lens

$$\Delta\Sigma(R) = \langle \epsilon_+ \Sigma_{\text{cr,ls}} \rangle (R), \quad (5.18)$$

with each  $\epsilon_+ \Sigma_{\text{cr,ls}}$  an independent, albeit noisy, estimate of the ESD. Here,  $\langle \cdot \rangle$  denotes a weighted average, with weights to be specified.

The  $\Sigma_{\text{cr,ls}}$  behaves as a geometric scaling factor, indicating the lensing efficiency for each lens-source combination. Since the variance of the noise in  $\Delta\Sigma$  is then affected by  $\Sigma_{\text{cr,ls}}^2$ , the relative precision, or inverse variance, carried by each  $\epsilon_+$  scales as  $\Sigma_{\text{cr,ls}}^{-2}$ .

In this paper we study the ESD profile in co-moving radial bins, and we therefore use the co-moving critical surface density

$$\Sigma_{\text{cr,com}} = (1 + z_l)^2 \Sigma_{\text{cr,prop}}. \quad (5.19)$$

In practice, the distance to each background source is not known exactly and is estimated by its redshift probability distribution,  $p(z_s)$ . Taking this into account, we estimate the co-moving critical surface density via

$$\langle \Sigma_{\text{cr,ls}}^{-1} \rangle = \frac{4\pi G}{c^2} D(z_l) (1 + z_l)^2 \int \frac{D(z_l, z_s)}{D(z_s)} p(z_s) dz_s. \quad (5.20)$$

<sup>11</sup>Of the coordinate system in which  $\gamma_1$  and  $\gamma_2$  are defined.



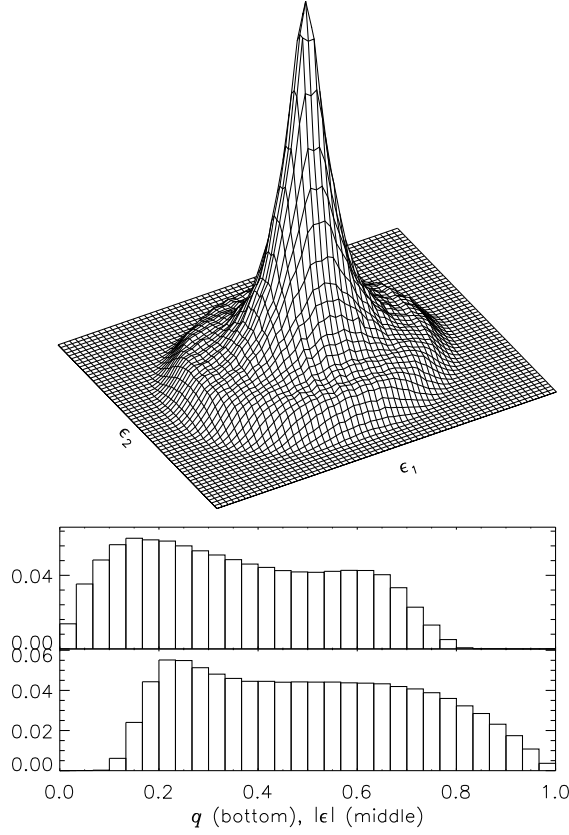


Figure 5.2: Estimated ellipticity and axis ratio distributions of sources in the KiDS-450 catalog. Top: Two-dimensional histogram of ellipticities. Middle: Histogram of the absolute ellipticity,  $|\epsilon|$ . Bottom: Histogram of the ellipse axis ratio,  $q$ .

### 5.2.3 Statistical framework

In this section we discuss the estimation of  $\langle \epsilon_+ \Sigma_{\text{cr},\text{ls}} \rangle$ . We refer to Paper I for a complementary discussion.

Important aspects of a good estimator,  $\hat{\epsilon}$ , are: (i) minimal bias, defined as the difference between the expected value of the estimator,  $\langle \hat{\epsilon} \rangle$ , and the value of the quantity being estimated, for instance the shear ( $\gamma$ ) or, in this case,  $\Delta\Sigma$ ; (ii) high efficiency, proportional to the inverse variance of the estimator,  $\sigma_{\hat{\epsilon}}^{-2}$ ; and (iii) robustness, meaning the estimator retains these properties for a sufficient range of likely parameter distributions.

#### Bias

Even though the measured ellipticity,  $\epsilon$ , is not a linear combination of the intrinsic shape,  $\epsilon^I$ , and the shear,  $\gamma$ , it can be shown (Seitz & Schneider 1997) that, in the absence of further uncertainties, the mean  $\mu(\epsilon)$  is an unbiased estimator for the underlying shear and that this



is independent of the intrinsic ellipticity distribution,  $P(\epsilon^l)$ . In the canonical approach, the ESD is therefore estimated as a weighted mean of an ensemble of lens-source combinations,

$$\Delta\Sigma(R) = \frac{\sum_{ls} w_{ls} \epsilon_{+,ls} \Sigma_{cr,ls}}{\sum_{ls} w_{ls}}, \quad (5.21)$$

where we use

$$w_{ls} = w_s \left\langle \Sigma_{cr,ls}^{-1} \right\rangle^2. \quad (5.22)$$

Here, the weight  $w_s$  is assigned to each measured ellipticity, scaled by the estimated lensing efficiency, as explained in the previous section (see, e.g., Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019).

In practice, there are various sources of uncertainty at each step of the process, such as source selection bias, distortion by the PSF, and biases due to the measurement pipeline. These lead to convolutions of the ellipticity distribution, before and after the gravitational lensing effect. The result is a bias in the mean as an estimate of the ESD (Melchior & Viola 2012, Refregier et al. 2012, Kacprzak et al. 2012, Viola et al. 2014, Kacprzak et al. 2014). In this case, the intrinsic shape distribution will play a role.

The weighted mean,  $\mu$ , is a statistic that, for a set of measurements  $\epsilon_i$  with weights  $w_i$ , finds the estimate of  $\gamma$  that minimizes the loss function

$$S_\mu = \sum_i w_i \left[ (\epsilon_{i,1} - \gamma_1)^2 + (\epsilon_{i,2} - \gamma_2)^2 \right], \quad (5.23)$$

that is, it is a least squares (LSQ) or  $L^2$  norm regression method and arises naturally as the optimal estimator for Gaussian distributions.

Figure 5.2 shows that the measured ellipticity distribution,  $P(\epsilon)$ , displays crucial differences with a Gaussian distribution, showing a sharp peak and a slower decline, including a higher number of high ellipticities,  $|\epsilon|$ . This central peak is an unbiased tracer of the underlying shear (Paper I). By Eq. 5.23, the mean is sensitive to outliers and therefore not robust when inferring the shear.

In contrast, the LAD or  $L^1$  norm regression minimizes the loss function,

$$S_{\text{LAD}} = \sum_i w_i \sqrt{(\epsilon_{i,1} - \gamma_1)^2 + (\epsilon_{i,2} - \gamma_2)^2}. \quad (5.24)$$

The LAD estimate is also known as the median in one dimension or the spatial median in higher dimensions. LAD is more sensitive to the peak and less sensitive to high ellipticity outliers. Where the mean is expected to be biased low (Melchior & Viola 2012, Refregier et al. 2012, Kacprzak et al. 2012, Viola et al. 2014, Kacprzak et al. 2014), we expect this to be less so for the LAD (Paper I).

## Efficiency

The formal definition of efficiency,  $\tilde{\eta}$ , relates the inherent (Fisher) information,  $\mathcal{I}$ , of a sample to the statistical variability around the expected value of the estimator, usually taken to be the variance,  $\sigma_\epsilon^2$ , of the estimator:

$$\tilde{\eta} = \frac{1}{\mathcal{I} \cdot \sigma_\epsilon^2}. \quad (5.25)$$



Since the variance of an unbiased estimator cannot be less than the reciprocal of the information,  $\mathcal{I}^{-1} \leq \sigma_{\hat{\epsilon}}^2$ , we have  $0 \leq \tilde{\eta} \leq 1$  (Rao 1945, Cramér 1946).

As we are comparing two estimators with unknown bias, it is appropriate to use the relative efficiency,

$$\eta = \frac{\sigma_{\mu}^2}{\sigma_{\text{LAD}}^2}, \quad (5.26)$$

with  $\eta < 1$  indicating a higher efficiency for the mean, and vice versa.

Paper I showed that the LAD consistently performed better than the mean, with both higher efficiency and less bias, for various cusped intrinsic ellipticity distributions, including the shear catalog from CFHTLenS. In what follows, we take the CFHTLenS shear distribution shape to be representative of KiDS data as well since both surveys were processed with the THELI pipeline (Erben et al. 2013) and the shape measurement pipeline *lensfit* (Miller et al. 2007, Kitching et al. 2008, Miller et al. 2013).

### 5.2.4 Halo model

Studying the effects of estimator choice on the weak lensing signal forms the technical core of this paper. The scientific goal, however, is to assess the relevance on the inference of physical quantities, such as the derivation of a lensing halo mass from an ESD profile. Since we calculate the stacked signal for an ensemble of clusters with some common (observable) property (here a range in  $r$ -band luminosity), of interest is the scaling relation between the observable and derived lensing mass,  $M_{200}$ , where we use the definition with respect to the mean density of the universe.

To do so, we modeled the lens density profile the same way as Dvornik et al. (2017), using the halo model (Seljak 2000, Peacock & Smith 2000, Cooray & Sheth 2002, van den Bosch et al. 2013, Cacciato et al. 2013, Mead et al. 2015). The initial lens density profile is described by a Navarro-Frenk-White profile (Navarro et al. 1996). We used the mass-concentration relation given by Duffy et al. (2008) and allowed for a re-normalization factor,  $f_c$  (Viola et al. 2015).

A dominant source of systematic bias in stacked weak lensing analyses is a miscentering of the lenses, which can be due to an offset of the cluster halo with the visible distribution of galaxies (see, e.g., George et al. 2012) or the resolution of the cluster detection method (less than  $0.1 \text{ Mpc } h^{-1}$  for AMICO; see Bellagamba et al. 2018). Following Johnston et al. (2007) as well as numerous subsequent works (e.g., Oguri et al. 2010, Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019, Giocoli et al. 2021), we allowed a fraction,  $p_{\text{off}}$ , of clusters to be offset from the center of the galaxy distribution, effectively smoothing the central stacked  $\Delta\Sigma$  profile with a characteristic radius,  $\mathcal{R}_{\text{off}}$ .

At large radii, typically beyond a few megaparsecs, the clustering of dark matter halos starts to dominate the signal. This “two-halo” term depends on the halo bias,  $b$  (Dvornik et al. 2017), and is modeled following Tinker et al. (2010). At small radii, the baryonic component of central galaxies can contribute to the signal, which is adequately described by a point mass,  $M_*$ , in the model (Viola et al. 2015, Dvornik et al. 2017).

In Table 5.1 we summarize these six free parameters for our halo model implementation, analogous to Dvornik et al. (2017).

In the AMICO cluster sample with 6925 lenses in 440 square degrees, many background sources are lensed by more than one cluster, contributing to the estimate of the ESD profile in



Table 5.1: Summary of the halo model fitting parameters and priors.

Parameter		Prior
$f_c$		[0.0, 8.0]
$p_{\text{off}}$		[0.0, 1.0]
$\mathcal{R}_{\text{off}}$	[ $h^{-1}$ Mpc ]	[0.0, 1.0]
$b$		[0.0, 10.0]
$\log(M_\star)$	$\log[ h^{-1} M_\odot ]$	[9.5, 12.5]
$\log(M_{200})$	$\log[ h^{-1} M_\odot ]$	[11.0, 17.0]

various radial bins of different clusters. In the model fitting, we took the covariance between the ESD estimates into account, as described in Sect. 5.3.3 (Viola et al. 2015, Dvornik et al. 2017, Bellagamba et al. 2019).

### 5.3 Data and analysis

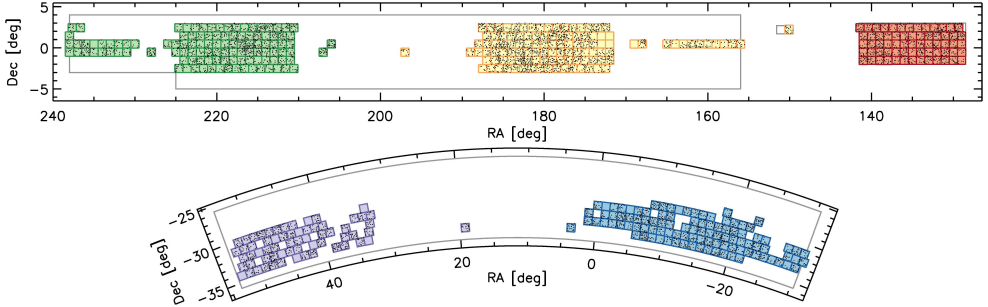


Figure 5.3: Overview of the KiDS-450 observations, with the KiDS-N (upper) and KiDS-S (lower) patches. The solid gray lines represent the planned KiDS survey area. Overplotted are the observed 1 square degree tiles, color coded with respect to their correspondence with the GAMA survey patches (G9 red, G12 yellow, G15 green, G23 blue, and GS purple; see Hildebrandt et al. 2017, for more details). The AMICO clusters analyzed in this work are represented by black dots.

In this paper we use a lensing cluster catalog and a background source catalog from KiDS-450 (de Jong et al. 2017), an optical wide-field imaging survey with OmegaCAM (Kuijken 2011) on the VLT Survey Telescope (VST; Capaccioli & Schipani 2011, de Jong et al. 2013). KiDS-450 consists of two patches, KiDS-N and KiDS-S (see Fig. 5.3), with 454 tiles of imaging data, for a total of  $449.7 \text{ deg}^2$ , in four optical filters, *ugri*. The survey was designed for lensing, ensuring a stable PSF, low seeing ( $< 0.96''$ , with an average of  $0.66''$  in *r*), and good photometric redshifts (photo-*z*; Hildebrandt et al. 2017).

The KiDS data were reduced with ASTRO-WISE (Valentijn et al. 2007, Verdoes Kleijn et al. 2012, Begeman et al. 2013, McFarland et al. 2013), as described in de Jong et al. (2015), Hildebrandt et al. (2017). Photometric redshifts, also termed  $z_B$ , were determined



using a Bayesian photo- $z$  estimation (BPZ; Benítez 2000, Coe et al. 2006) with PSF-matched photometry, as described in Hildebrandt et al. (2012), Kuijken et al. (2015), Hildebrandt et al. (2017).

### 5.3.1 Lenses

We made use of the galaxy cluster catalog derived with AMICO (Bellagamba et al. 2011, Radovich et al. 2017, Bellagamba et al. 2018), extracted from 440 tiles of KiDS-450 data and described in Maturi et al. (2019) and Bellagamba et al. (2019). For each cluster, the luminosity  $L_{200}$  is defined<sup>12</sup> as the sum of  $r$ -band luminosities of bright candidate member galaxies, weighted by membership probability (see Maturi et al. 2019). We selected galaxies with  $k$ -corrected  $r$ -band magnitudes brighter than  $m^*(z_1) + 1$  within  $R_{200}(z_1)$ , where  $z_1$  is the estimated cluster redshift and  $R_{200}$  is derived from the adopted cluster model and is used in the construction of the cluster detection filter, as defined in Maturi et al. (2019). In this sense,  $L_{200}$  is defined analogously to the apparent richness,  $\lambda^*$ , which is a sum of membership probabilities of galaxies with  $m < m^* + 1.5$ , within  $R_{200}$ .

We selected clusters in the range  $0.1 \leq z_1 \leq 0.6$ . We excluded clusters below  $z = 0.1$  due to their unfavorable lensing geometry and above  $z = 0.6$  due to the low density of background sources. For some clusters, no lens-source pairs were found, due to source selection criteria or masking. Our final selection comprises 6925 clusters, divided over the KiDS-450 survey area as shown in Fig 5.3 and described in Table 5.2. The redshift distribution of these clusters is shown in Fig. 5.4, with a median redshift of  $z_1 = 0.39$ .

Table 5.2: Summary of the survey patches as described in Hildebrandt et al. (2017), with corresponding numbers of KiDS mosaic tiles and analyzed clusters.

KiDS field	Subfield	Tiles	Clusters
North	G9	65	1039
	G12	113	1778
	G15	112	1737
South	G23	101	1517
	GS	63	854

We divided the clusters into 13 bins of cluster  $L_{200}$ . The limits of these bins were chosen so that the signal-to-noise ratios of the ESD measurements were approximately the same in each bin. We give an overview of these bins, together with the estimated  $M_{200}$ , in Table 5.3.

### 5.3.2 Sources

We selected an initial sample of background sources using the same photometric redshift criteria as Hildebrandt et al. (2017),  $0.1 < z_B \leq 0.9$ , to reduce the outlier rate. We also applied the cut  $z_1 + \Delta z < z_B$ , following Dvornik et al. (2017). Here,  $\Delta z = 0.2$  is an offset between the redshift estimation,  $z_1$ , of the cluster by AMICO and the photometric redshift,

<sup>12</sup>We note that this does not take intracluster light into account.



$z_B$ , of the source to sufficiently lessen the contamination of the background sample by cluster member galaxies (see also Appendix 5.A.2).

Our selection of AMICO clusters is deeper than the lenses from the Galaxy And Mass Assembly (GAMA Driver et al. 2011, Robotham et al. 2011) catalog used in Dvornik et al. (2017). As can be seen in Fig. 5.4, the redshift distributions of lenses and background sources significantly overlap, and the cut at  $\Delta z = 0.2$  reduces the number density severely for clusters at higher redshift. Following Bellagamba et al. (2019), we also selected background sources using the color selection proposed by Oguri et al. (2012):

$$g - r < 0.3 \quad \vee \quad r - i > 1.3 \quad \vee \quad r - i > g - r. \quad (5.27)$$

In Fig. 5.5 we show the photometric redshift distribution of this cut in the KiDS-450 catalog and compare it to the photometric and spectroscopic redshift distribution of the same cut in the spectroscopic redshift (spec- $z$ ) catalog used in Hildebrandt et al. (2017). Based on this analysis, we additionally required  $z_B \geq 0.6$  for this selection to reduce contamination by sources at low redshift and find that 98 % of the galaxies in this color selection have  $z_{\text{spec}} > 0.6$ .

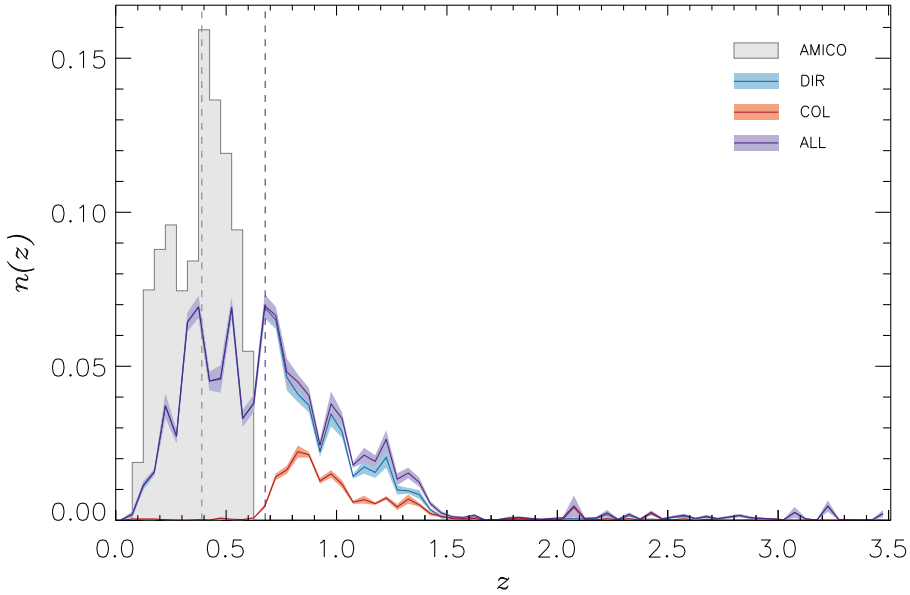


Figure 5.4: Redshift distribution of AMICO clusters (gray), with a median redshift of  $z_l = 0.38$  (dashed gray), and KiDS-450 background sources (purple), with a median redshift of  $z = 0.68$  (dashed purple). In blue (DIR), we show the initial selection following Hildebrandt et al. (2017) and Dvornik et al. (2017). In red, we show the estimated redshift distribution of the  $gri$  color selection (COL), corresponding to the bottom panel of Fig. 5.5.

### Redshift distribution

To estimate the redshift distribution of background galaxies, we did not directly use the individual redshift probability distribution,  $p(z_s)$ , per source galaxy. Instead, we applied a



weighted direct calibration method (DIR), as motivated by Hildebrandt et al. (2017).

For each cluster, we used the spec- $z$  catalog described in Hildebrandt et al. (2017) to select objects using the same selection criteria as described above. We then used the normalized spectroscopic redshift distribution,  $n(z_s)$ , of this sample to calculate the co-moving critical surface density analogous to Eq. 5.20:

$$\langle \Sigma_{\text{cr},1}^{-1} \rangle = \frac{4\pi G}{c^2} D(z_l) (1 + z_l)^2 \int_{z_l + \Delta z}^{\infty} \frac{D(z_l, z_s)}{D(z_s)} n(z_s) dz_s. \quad (5.28)$$

The resulting redshift distribution for selected sources from the full KiDS-450 catalog is shown in Fig. 5.4.

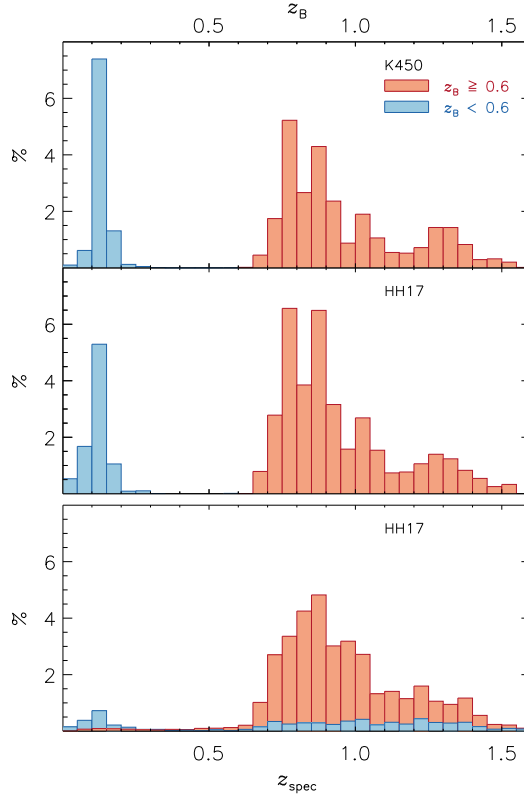


Figure 5.5: Redshift distribution of background sources selected by color. The upper panel shows the distribution of the photometric redshift,  $z_B$ , of the sources in the KiDS-450 catalog that satisfy the color cut of Eq. 5.27, of which we select the sources with  $z_B \geq 0.6$  (red) and discard those with  $z_B < 0.6$  (blue). The bottom two panels show the same selection applied to the spec- $z$  catalog (Hildebrandt et al. 2017), plotted in terms of photometric redshift (middle) and spectroscopic redshift (lower). We find the contamination of sources with  $z_B \geq 0.6$  and  $z_{\text{spec}} < 0.6$  is  $\sim 2\%$ .



## Shape measurements

For shape measurements, the  $r$ -band data were reduced using the THELI pipeline, developed to meet the requirements for weak gravitational lensing analyses (Erben et al. 2005, 2009, Schirmer 2013, Erben et al. 2013). Galaxy shapes in the KiDS-450 catalog were then measured by *lensfit* (Miller et al. 2007, Kitching et al. 2008, Miller et al. 2013, Fenech Conti et al. 2017).

For each source, *lensfit* produces the ellipticity ( $\epsilon_1, \epsilon_2$ ), an approximately inverse-variance weight  $w_s$  (see Miller et al. 2013), and a fitting quality parameter. We excluded sources with unreliable ellipticities from our source sample, using the same *lensfit* selection criteria as described in Hildebrandt et al. (2017).

In Paper I we compared the performance of estimators for ellipticity measurements in the CFHTLenS data with a subset of that catalog, selecting sources on the signal-to-noise ratio parameter  $\nu_{\text{SN}}$  output by *lensfit*. We repeated that approach for a qualitative comparison here, using two subsets of the selected KiDS-450 sources. The first set selects sources with  $\nu_{\text{SN}} \geq 20$ , similar to Paper I, retaining  $\sim 30\%$  of the full background sample. The second set is a more stringent cut of the first set, additionally selecting objects with  $w_s \geq 14.5$ , comprising  $\sim 20\%$  of the full sample.

## Effective source density

The KiDS-450 catalog includes a filtering on general object detection and quality flags, for example, possibly blended sources or artifacts, as described by Kuijken et al. (2015) and Hildebrandt et al. (2017), and we discarded objects that lie in a mask. This removed approximately  $\sim 12\%$  of the sources. Our final selection comprises 14124197 sources, which translates to an effective number density of  $n_{\text{eff}} \approx 8.23 \text{ arcmin}^{-2}$ , as defined in Heymans et al. (2012b),

$$n_{\text{eff}} = \frac{1}{A} \frac{(\sum_i w_i)^2}{\sum_i w_i^2}, \quad (5.29)$$

with  $A$  the effective surface area, excluding masked regions.

## 5.3.3 Implementation

### ESD estimation

Following Bellagamba et al. (2019), we measured the ESD in 14 logarithmic bins between  $0.1 \text{ Mpc } h^{-1}$  and  $3.16 \text{ Mpc } h^{-1}$ . Not only does this make for an easy comparison of the results, but it has several other practical advantages.

We avoided radii smaller than the AMICO detection pixel size, which has a median size of  $0.1 \text{ Mpc } h^{-1}$ , to lessen the chance of a mismodeling the halo miscentering (Sect. 5.2.4). Here, the line of sight is also most contaminated by cluster members, which can lead to an overabundance by incorrectly including ellipticity measurements that carry no lensing signal, or by an obscuring and blending of background sources, which leads to an under-abundance of sources. While these effects may partially cancel out in the number counts, the effects on the ESD measurements do not cancel out, as the first leads to a diluted signal and the second to a very poor signal-to-noise ratio (see Appendix 5.A.2 for an assessment of cluster member contamination).



At large radii, systematic additive biases can start to play a role (see, e.g., Dvornik et al. 2017, for this data set), which may differ for each KiDS survey patch (Fenech Conti et al. 2017, Hildebrandt et al. 2017). Another concern at larger separations is that the two-halo term becomes the dominant contribution to the ESD signal, which means we would need to properly constrain the halo bias, and we explain below how our approach does not fully take the clustering of dark matter halos into account.

The combination of background selection criteria from Dvornik et al. (2017) and Belagamba et al. (2019) allows us to retain the three inner radial bins between 0.1 and 0.2 Mpc  $h^{-1}$ . We justify this inclusion in Appendix 5.A.2, where we repeat the tests of Dvornik et al. (2017).

Each lens-source pair was then assigned a combined weight of

$$w_{ls} = w_s \Sigma_{cr,l}^{-2}, \quad (5.30)$$

as motivated in Sect. 5.2.2.

For LAD optimization, that is, the estimator that minimizes the  $L^1$  norm (Eq. 5.24), there exists no general analytic solution. The problem can, however, be formulated as a linear optimization, which can be solved iteratively (e.g., with simplex-based methods; Barrodale & Roberts 1973). In our weak lensing analyses, we find that convergence is robust.

To derive the covariance matrices for the ESD estimates using the mean and LAD in the same way, we can therefore also not employ the analytical prescription of Viola et al. (2015) used in earlier KiDS analyses (e.g., Sifón et al. 2015a, van Uitert et al. 2016, Brouwer et al. 2016). Instead, we used a bootstrap approach.

Since the cluster bins of highest  $r$ -band luminosity,  $L_{200}$ , contain only a small number of clusters, covering only a small fraction of the KiDS-450 tiles, we cannot use the same bootstrap approach as Viola et al. (2015) and Dvornik et al. (2017) by bootstrapping  $1 \text{ deg}^2$  tiles with replacement. Instead, we bootstrapped the source catalog, in accordance with Belagamba et al. (2019).

This means that we are not sensitive to the clustering effect of dark matter halos, which justifies our choice of radial lens-source separation mentioned above. To assess the accuracy of these assumptions, we estimated the covariance matrix of the full 6925 cluster sample by bootstrapping the sources and by bootstrapping by KiDS-450 tiles in Appendix 5.A.4. We conclude that our bootstrapping method yields a good estimate of the covariance matrix.

## Halo model fitting

Having produced the LSQ and LAD shear profiles for the stacked clusters, we fit a halo model to the results. We used the fitting procedure described in Dvornik et al. (2017), producing the full posterior probabilities by a Bayesian inference technique, via a Monte Carlo Markov chain (MCMC) maximum likelihood approach. We assumed a Gaussian likelihood and made use of the full covariance between radial bins:

$$\mathcal{L} \propto \exp \left[ -\frac{1}{2} \mathbf{R}^T \mathbf{C}^{-1} \mathbf{R} \right], \quad (5.31)$$

where the  $\mathbf{R}$  are the residuals and  $\mathbf{C}$  is the covariance matrix.

We used the `emcee` Python package (Foreman-Mackey et al. 2013) for the MCMC procedure, setting flat priors for all parameters. For the evaluation of the power spectrum and the halo mass function, we used the median redshift for each cluster luminosity bin.



## 5.4 Results

We present our results, starting with the derived ESD profiles obtained with the mean and LAD estimators, discussing potential biases and efficiency. Then, we show the results of the halo model fitting (i.e.,  $M_{200}$  for each luminosity bin) and conclude with the scaling relation between  $L_{200}$  and  $M_{200}$ . We visualize the results for the case in which all 6925 clusters are stacked together, giving the numerical results of the 13 luminosity bins in Table 5.3.

### 5.4.1 ESD profiles

We calculated the ESD profiles using  $10^4$  bootstraps with replacement. We estimated the ESD signal in the 14 radial bins, using both the mean and the LAD estimators, for each bootstrapped sample, preserving the bootstrap order of all 28 values throughout the whole process.

We find the estimator distribution to be almost perfectly normal, as expected from the central limit theorem. The correlation between the 14 bins of the full stack of clusters is shown in Fig. 5.6 and is given by

$$\rho_{ij} \equiv \frac{\text{Cov}_{ij}}{\sigma_i \sigma_j}, \quad (5.32)$$

where  $i$  and  $j$  denote the radial bin subscripts.

The upper-left part of the matrix shows the correlation between the LAD estimates of the radial bins, and the lower-right part the shows the mean results. Although the correlation between bins is very low, it is clear that the overall trends are the same for the two estimators.

The signal-to-noise ratio of the recovered ESD profile of the full stack, which is shown in Fig. 5.7, is high enough to allow us to notice the difference between the estimators, which indicates a small relative bias. The blue points show the LAD estimates, and the red points represent the mean estimates, with error bars in both cases defined as the square root of the diagonal elements of the covariance matrices (i.e., the classical standard deviation).

Tests for systematic effects, such as the cross signal, and a test for systematic additive noise around random points were already conducted by Dvornik et al. (2017) and Bellagamba et al. (2019). In Appendix 5.A we repeat these tests for completeness since we use the KiDS-S field and an extended source selection with respect to Dvornik et al. (2017) and use a different source selection and three smaller radial bins with respect to Bellagamba et al. (2019). Our results show no residual systematic effects, in accordance with these papers.

### 5.4.2 Bias and efficiency

A possible bias is expected to depend on the strength of the underlying shear field since a zero lensing signal would imply no bias. In that case, the expected relevant distributions, tangential ellipticities or noise, are symmetric around zero ellipticity.

To quantify the difference between the ESD estimates, which we call the relative bias,  $\Delta\Sigma_{\text{Mean}} - \Delta\Sigma_{\text{LAD}}$ , we assumed<sup>13</sup> to first order

$$\Delta\Sigma_{\text{Mean}} - \Delta\Sigma_{\text{LAD}} = m \cdot \Delta\Sigma, \quad (5.33)$$

<sup>13</sup>This assumption is only made here to quantify the bias and is not used elsewhere in the paper.



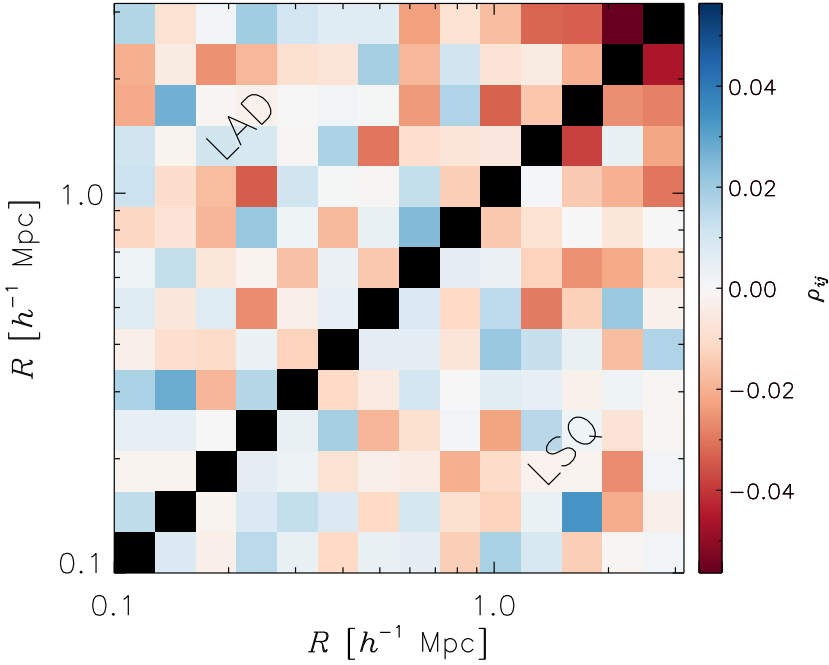


Figure 5.6: Correlation matrix between the stacked ESD signals in different radial bins, using the full AMICO cluster catalog and  $10^4$  bootstraps. The upper-left triangle shows the correlation,  $\rho_{\text{LAD}}$ , between the LAD estimations. The lower-right triangle shows the correlation,  $\rho_{\text{mean}}$ , between the ESD estimates using a weighted mean. We note the general similarities in the two patterns.

where we arbitrarily<sup>14</sup> use  $\Sigma_{\text{LAD}}$  for  $\Sigma$ .

We used the full stack for its high signal-to-noise ratio, using the 15.9th and 84.1th percentiles of the differences in all bootstrap results to calculate uncertainties for each bin. We find  $m = -0.088 \pm 0.020$ . In Fig. 5.8 we show this relative bias, plotting for visualization purposes

$$\frac{\Delta\Sigma_{\text{Mean}} - \Delta\Sigma_{\text{LAD}}}{\Delta\Sigma} \quad (5.34)$$

and a horizontal line at  $m = -0.88$  to give a more intuitive impression of the relative error bars.

We reiterate that it is impossible to determine the absolute bias of each estimator as we did in Paper I, as we have no knowledge of the true ESD. However, the overall trend between the mean and LAD is similar in sign and order of magnitude, as we found for the CFHTLenS data in Paper I.

In Fig. 5.9 we show the derived relative efficiency (Eq. 5.26)  $\eta = 1.047 \pm 0.006$ , which is in accordance with the findings for CFHTLenS data in Paper I. The measured<sup>15</sup> ellipticity distribution is expected to differ for shape measurements with a higher signal-to-noise ratio.

<sup>14</sup>We find no qualitative difference in our results when we use  $\Delta\Sigma_{\text{Mean}}$  instead.

<sup>15</sup>i.e., the combination of the intrinsic distribution and the various effects before and after the lensing by AMICO clusters, which affects the observation and measurement of the source ellipticities.



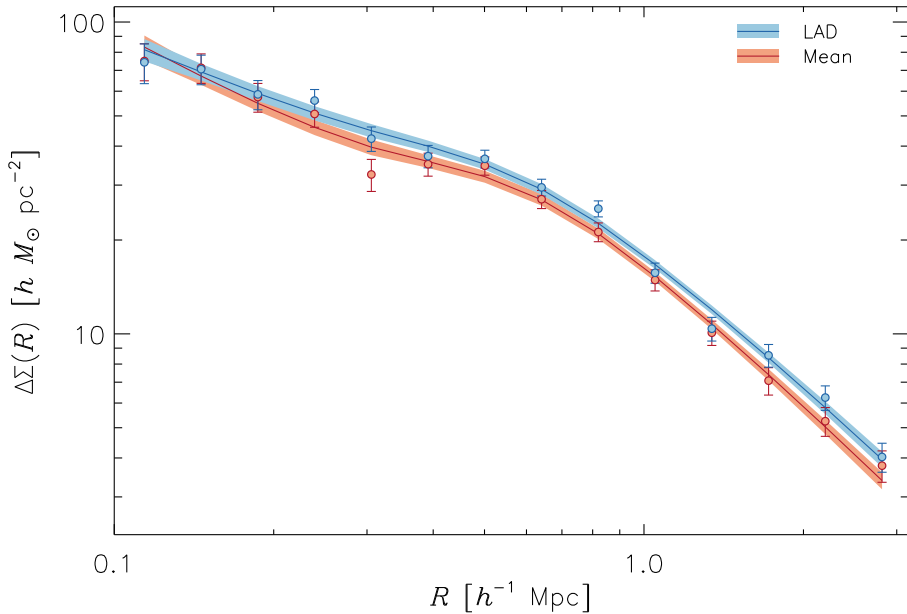


Figure 5.7: Estimated ESD profile from the full AMICO cluster catalog, using the LAD estimator (blue) and a weighted mean (red). The error bars are the square roots of the diagonal values of the respective covariance matrices. The solid lines represent the best fitting halo model obtained by the MCMC fit. The shaded regions show the 68.3% confidence bands, estimated using the 15.9th and 84.1th percentiles of the MCMC realizations.

For example, the cuspieness of the distribution shown in Fig. 5.2 can be smoothed out by noise convolutions. As the LAD estimation is more sensitive to the central peak, this will affect its precision.

In Paper I this was confirmed in the comparison of simulated data with and without noise, as well as in the results of the CFHTLenS sample with a stringent signal-to-noise selection compared with the full sample. As in Paper I, we compared the relative efficiencies for our selections with  $v_{\text{SN}} \geq 20$  and with  $w_i \geq 14.5$ , finding indeed a higher efficiency for less noisy shapes, namely  $\eta = 1.240 \pm 0.010$  and  $\eta = 1.386 \pm 0.018$ , respectively.

### 5.4.3 Halo masses

We ran MCMC chains of 120 000 samples, using 120 walkers with 1000 steps each. The resulting chains were fully converged after the first 200 steps, so we discarded the first 24 000 samples.

We summarize the  $M_{200}$  derived from the ESD estimation of the 13 luminosity bins in Table 5.3. Reduced  $\chi^2$ , estimated between 0.730 and 2.528, are fairly consistent between derived results for mean and LAD.



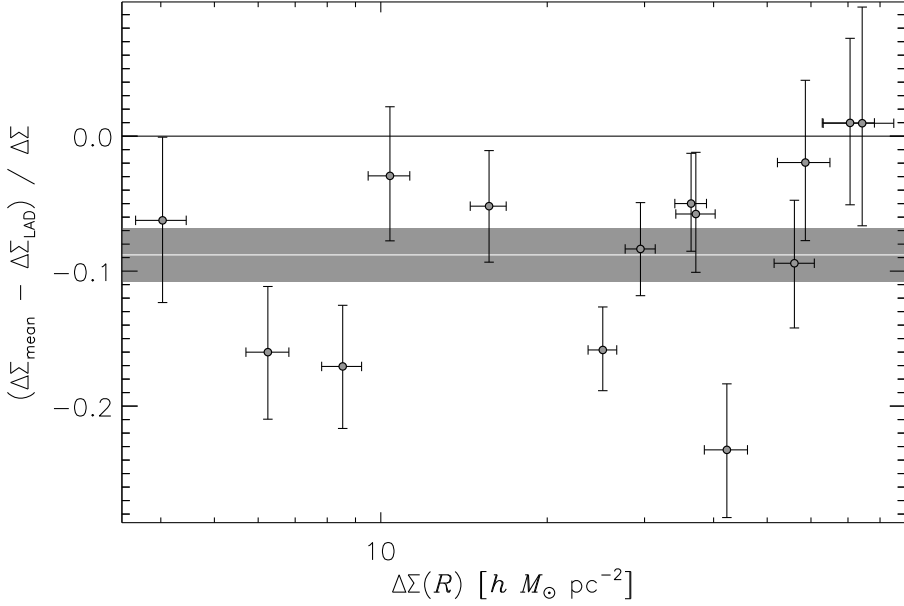


Figure 5.8: Difference between the recovered ESD signals in the radial bins, showing  $(\Delta\Sigma_{\text{mean}} - \Delta\Sigma_{\text{LAD}}) / \Delta\Sigma$ . The solid white line represents the average difference, with the shaded region showing the formal  $1\sigma$  error. As a possible bias in the recovered values is expected to increase with increasing shear, the differences are plotted versus the full ESD signal in each bin where we use  $\Delta\Sigma_{\text{LAD}}$ , but we note that the small variations in the individual points, when plotting against  $\Delta\Sigma_{\text{mean}}$  instead, give the same result, within statistical significance.

For the full stack of clusters, we derived

$$M_{200} = (0.453^{+0.030}_{-0.030}) \times 10^{14} h^{-1} M_{\odot}, \quad \chi^2_{\nu} = 1.25 \quad (\text{Mean}) \quad (5.35a)$$

$$M_{200} = (0.487^{+0.033}_{-0.036}) \times 10^{14} h^{-1} M_{\odot}, \quad \chi^2_{\nu} = 1.37 \quad (\text{LAD}). \quad (5.35b)$$

The confidence intervals are derived from the 15.9th and 84.1th percentiles of the posterior distributions.

The best fitting ESD models are shown in Fig. 5.7. The 68.3% confidence bands overlap at some radii and are in tension at other radii. While the difference in ESD is significant, the 68.3% confidence intervals for  $M_{200}$  just touch.

#### 5.4.4 $L_{200} - M_{200}$ scaling relation

We assumed a power-law relation between the derived halo masses and the median  $r$ -band luminosity of each cluster bin. We fit this relation in the form

$$\log\left(\frac{M_{200}}{M_{\text{piv}}}\right) = a + b \log\left(\frac{L_{200}}{L_{\text{piv}}}\right), \quad (5.36)$$

with  $a$  the intercept and  $b$  the slope, where  $M_{\text{piv}} \approx 10^{14.1} h^{-1} M_{\odot}$  and  $L_{\text{piv}} \approx 10^{11.8} h^{-2} L_{\odot}$  are typical pivotal values of the halo mass and luminosity, derived from the fit itself. The fit was



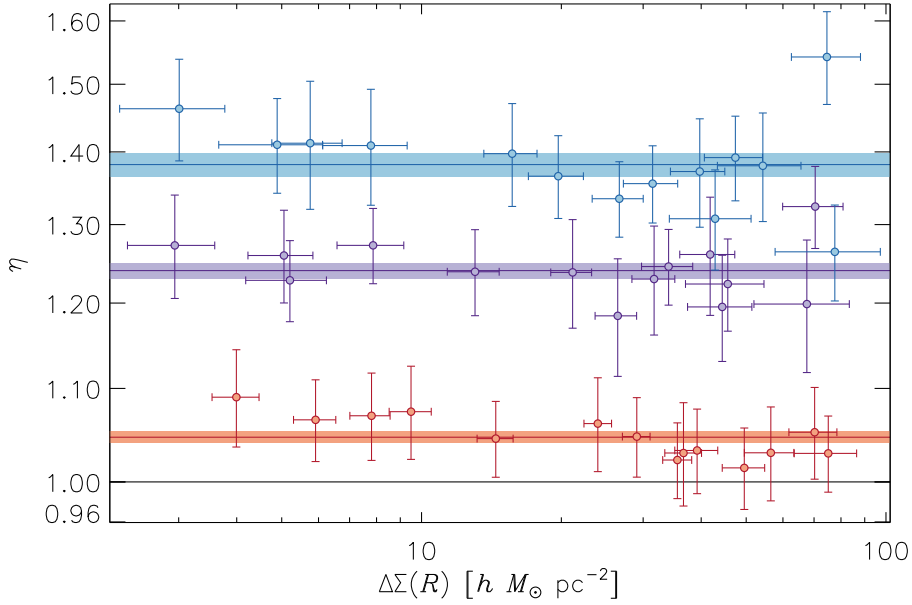


Figure 5.9: Relative efficiencies,  $\eta \equiv \sigma_{\text{Mean}}^2 / \sigma_{\text{LAD}}^2$ , defined as the ratio of the diagonal elements of the covariance matrices. The relative efficiencies using the full KiDS-450 source catalog are shown in red. The purple and blue represent the higher S/N and higher *lensfit* weight selections, respectively. The solid lines represent the average  $\eta$  for each selection, with the shaded regions showing the  $1\sigma$  errors.

done in log basis as the derived posterior distributions of the halo mass are log-normal.

We did not take a redshift dependence into account, as Bellagamba et al. (2019) showed only a marginal and not very steep dependence of the halo mass on redshift.

We obtained the scaling relations

$$\frac{M_{200}}{10^{14.1} h^{-1} M_{\odot}} = (0.97 \pm 0.06) \left( \frac{L_{200}}{10^{11.8} h^{-2} L_{\odot}} \right)^{(1.24 \pm 0.08)} \quad (\text{Mean}) \quad (5.37a)$$

$$\frac{M_{200}}{10^{14.1} h^{-1} M_{\odot}} = (1.03 \pm 0.05) \left( \frac{L_{200}}{10^{11.8} h^{-2} L_{\odot}} \right)^{(1.24 \pm 0.08)} \quad (\text{LAD}) \quad (5.37b)$$

and plot the results in Fig. 5.10. As with the derived ESD profiles and halo masses for the full stack of clusters, the 68.3% confidence bands just touch at the pivot point

$$(L_{200}, M_{200}) = \left( 10^{11.8} h^{-2} L_{\odot}, 10^{14.1} h^{-1} M_{\odot} \right), \quad (5.38)$$

recognizable as the narrowest parts of the confidence bands.

## 5.5 Summary and conclusions

We conducted a weak shear analysis of 6925 AMICO clusters in the KiDS-450 data. We derived a tightly constrained scaling relation between *r*-band luminosity,  $L_{200}$ , and average lensing masses,  $M_{200}$ , in concordance with earlier results in the literature.



Table 5.3: Properties and lensing results of the individual luminosity bins, with the median  $L_{200}$  and  $z_1$ . The errors on the median  $L_{200}$  and derived  $M_{200}$  in each luminosity bin are the differences with the 15.9th and 84.1th percentiles.  $L_{200}$  values are given in  $[10^{10}h^{-2}L_{\odot}]$ ,  $M_{200}$  in  $[10^{14}h^{-1}M_{\odot}]$ .

Range	$L_{200}$	Clusters	$z_1$	$M_{200}$ Mean	$\chi^2_{\nu}$	$M_{200}$ LAD	$\chi^2_{\nu}$
[0.4, 17.4[	$12.3^{+3.5}_{-4.6}$	2346	0.29	$0.187^{+0.029}_{-0.026}$	1.27	$0.191^{+0.031}_{-0.026}$	1.18
[17.4, 24.8[	$20.8^{+2.7}_{-2.3}$	1545	0.41	$0.416^{+0.061}_{-0.076}$	2.19	$0.445^{+0.055}_{-0.061}$	1.60
[24.8, 31.8[	$28.0^{+2.5}_{-2.2}$	1027	0.41	$0.371^{+0.048}_{-0.046}$	0.81	$0.385^{+0.050}_{-0.046}$	0.91
[31.8, 40.5[	$35.2^{+3.2}_{-2.4}$	685	0.42	$0.519^{+0.078}_{-0.078}$	1.34	$0.621^{+0.076}_{-0.074}$	1.29
[40.5, 49.0[	$44.4^{+2.9}_{-2.7}$	457	0.42	$1.076^{+0.182}_{-0.181}$	2.18	$0.998^{+0.150}_{-0.139}$	2.53
[49.0, 59.9[	$54.1^{+3.6}_{-3.9}$	305	0.40	$1.387^{+0.425}_{-0.335}$	1.05	$1.462^{+0.533}_{-0.353}$	0.98
[59.9, 72.9[	$65.2^{+5.2}_{-3.7}$	202	0.41	$1.318^{+0.267}_{-0.234}$	0.77	$1.309^{+0.233}_{-0.203}$	0.74
[72.9, 84.1[	$78.6^{+3.2}_{-3.9}$	135	0.38	$1.406^{+0.344}_{-0.215}$	1.22	$1.528^{+0.409}_{-0.254}$	1.30
[84.1, 102[	$91.8^{+5.7}_{-4.8}$	90	0.39	$2.438^{+0.486}_{-0.443}$	0.75	$2.472^{+0.479}_{-0.425}$	0.73
[102, 129[	$112^{+10}_{-7.2}$	60	0.40	$2.143^{+0.618}_{-0.417}$	0.77	$1.914^{+0.680}_{-0.373}$	0.79
[129, 160[	$138^{+11}_{-6.9}$	40	0.395	$3.999^{+1.957}_{-0.993}$	1.35	$3.715^{+1.557}_{-0.910}$	1.64
[160, 221[	$175^{+27}_{-11}$	26	0.37	$4.207^{+0.713}_{-0.610}$	1.47	$4.786^{+0.824}_{-0.637}$	1.18
[221, 400[	$277^{+106}_{-24}$	8	0.375	$7.638^{+2.293}_{-1.613}$	1.02	$9.141^{+3.105}_{-1.913}$	0.81

We investigated the impact of estimator choice for inferring the central moment of the cusped and skewed ellipticity distribution of background galaxies, finding a relative bias on the order of a few percent, as predicted in Paper I. We find that the constraints obtained via LAD regression are tighter than those obtained via LSQ regression, and they significantly improved as the signal-to-noise ratio of the shape measurements of the background galaxies increased. Complemented by simulations from Paper I, we give an alternative perspective on the problem of inferring the central shear value from the skewed distribution of background galaxy shapes, at the minor cost of increased, but still feasible, computation times for numerical iterative regression.

### 5.5.1 $L_{200} - M_{200}$ relation

Since the relative bias we found in both this research and in Paper I is approximately proportional to the ESD signal, it is expected that the LAD estimator will mainly have an effect on the intercept of the  $L_{200} - M_{200}$  scaling relation. This was confirmed by our results in Eq. 5.37.

The power-law index of the  $L_{200} - M_{200}$  scaling relation was constrained to  $1.24 \pm 0.08$  (Eq. 5.37), independent of estimator choice. This is in agreement with earlier work in the literature, such as Viola et al. (2015, and references therein), who cite  $1.16 \pm 0.13$ . This agreement is noteworthy since the AMICO clusters are derived from photometric redshifts, as opposed to the spectroscopically derived groups from GAMA (Driver et al. 2011, Robotham et al. 2011). The difference in confidence is explained by the increased number of lenses, which is



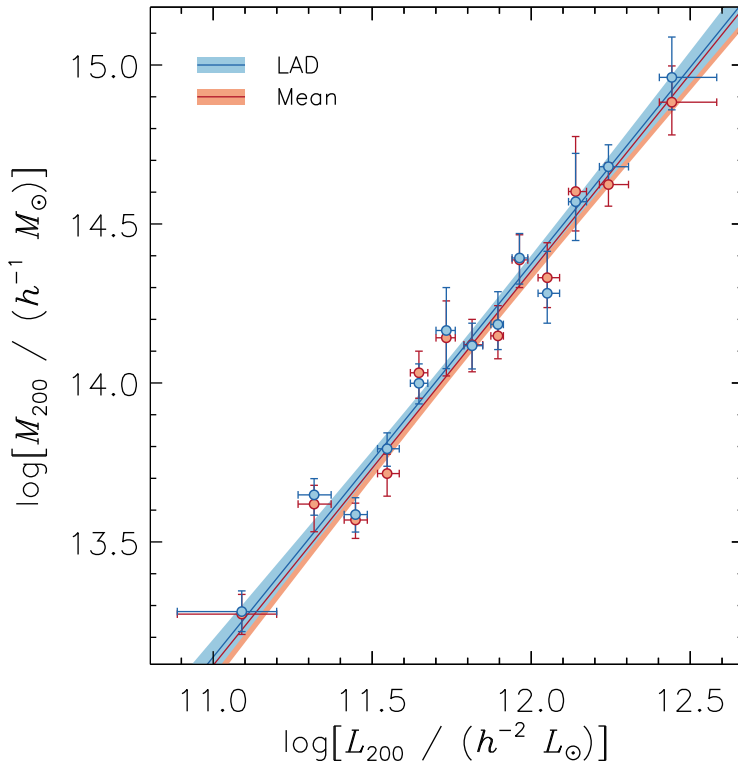


Figure 5.10:  $r$ -band luminosity-halo mass scaling relations, derived from the ESD profiles estimated using the weighted mean (red) and LAD (blue). At the pivot point, recognizable as the narrowest parts of the confidence bands, the relations just touch at the 68.3% confidence level.

not surprising given the large overlap in mass range: Viola et al. (2015) analyzed 1413 galaxy groups between  $z = 0.03$  and  $z = 0.33$ , with  $r$ -band luminosity bin limits between  $2.5 \times 10^9$  and  $5.0 \times 10^{12} h^{-2} L_{\odot}$ , deriving halo masses between  $1.4 \times 10^{13}$  and  $3.5 \times 10^{14} h^{-1} M_{\odot}$ .

The choice of estimator produces a difference in intercept at just the  $1\sigma$  level. Using a weighted mean to derive the ESD leads to an intercept of  $0.97 \pm 0.06$  in the scaling relation, while LAD gives an intercept of  $1.03 \pm 0.05$ . This is to be expected as the relative bias seems roughly constant, when normalized by the ESD (see Fig. 5.8), and in good agreement with the bias of  $\sim 5\%$  found using simulations in Paper I.

As in Bellagamba et al. (2019), we estimate that systematic effects mostly affect the intercept. While the derived intercept is in agreement with the aforementioned papers, we note that the chosen definition<sup>16</sup> of  $L_{200}$ , the difference in redshift range and definitions, and the completeness of group and cluster membership can account for possible differences on the same order of magnitude. This would not affect our conclusions on methodology, as the results from both estimators would be similarly affected.

In a further comparison with the scaling relation between richness and mass, cited in

<sup>16</sup>e.g., Viola et al. (2015), where the group  $r$ -band luminosities are calculated by summing over spectroscopically confirmed group members.



Bellagamba et al. (2019), we find similar significance in constraints on the slope. We define bins in luminosity instead of AMICO detection amplitude (Radovich et al. 2017) or richness, but since the cluster luminosity is tightly correlated with the richness,  $\lambda^*$ , this confirms our findings.

There are a few differences to consider. We chose not to account for a possible redshift dependence. This is motivated by Bellagamba et al. (2019) finding only a shallow dependence, which they point out may be driven mainly by the highest redshift bin. Our lensing analysis employs a slightly different background selection for an increased source density, combined with a different derivation of the associated redshift distribution,  $n(z)$ . Another difference is the inclusion of radial bins at  $0.1 \leq R < 0.2 \text{ Mpc } h^{-1}$ . This is expected to only have a minor effect as the contribution of the stellar mass is an order of magnitude lower than the halo term at these radii, while the contribution of miscentered halos only starts to become significant at larger radii (see also Rykoff et al. 2016, Oguri et al. 2018). In this sense, these findings are a confirmation of the robustness of the results across these papers.

### 5.5.2 Optimal estimators

Our results are in good agreement with Paper I, with a relative bias between the two estimators that shows the recovered lensing signal is higher with LAD, suggestive of a lower absolute bias. At the same time, LAD regression gives a small (albeit significant) gain in efficiency, giving a reduction in error bars of a few percent, and potentially up to 11% – 18% for shape measurements of a higher signal-to-noise ratio. Least absolute deviation regression comes at the cost of a higher computation time, but at a step in the analysis process that does not dominate the total computational cost.

Both simulations (Paper I) and analyses on real data (this article) cite quantitative results of significance while at the same time showing similar trends between estimators on a qualitative level. We have conducted a cautious and thorough investigation but can never exclude the unknown: biases arising due to assumptions in the simulations of Paper I or uncorrected systematic effects in this research, or, most likely, both. However, given the range and realism in simulated distributions and the similarities in findings among those simulations, this research, and other work in the literature, we are confident that the recovered differences in results between the two estimators are real.

We note some differences between the two analyses. In Paper I we analyzed the regression of a sample of ellipticities with a single underlying value of the shear and, for each type of simulation, a single intrinsic ellipticity distribution. In this research, the situation is more complex. We studied the stacked signal around samples of lenses and of samples of background sources at a range of redshifts. This means also stacking noise that has been scaled by a range in lensing geometries, quantified by  $\Sigma_{\text{cr}}^{-2}$ . Furthermore, in each radial bin of each luminosity bin, we assume: (i) a constant lensing effect, which is in reality the stacked average of a range in  $L_{200}$ , and therefore a range in  $M_{200}$ , confounded by intrinsic scatter between these two quantities, and (ii) radial distance  $R$  from the lens, combined with a miscentering of halos.

All these effects tend to convolve the intrinsic galaxy shape distribution, which makes the level of agreement and significance between the two papers in fact remarkably robust. In conclusion:



- The combination of Paper I and this research shows that LAD regression is more naturally suited to the cusped intrinsic ellipticity distribution of background galaxies.
- Our simulations in Paper I showed a lower bias for LAD regression than for LSQ regression in the presence of noise in the background source shape measurements, while this research confirmed the same relative bias between the two estimators.
- Constraints obtained via LAD regression are comparable with or tighter than constraints obtained via LSQ regression.

An optimal estimator is, from a principled point of view, more objective and better suited than corrections to an approach, which is known to mismatch the sample distribution. More practically, LAD regression provides a robust consistency check for shear inference, which has been and still remains a major investment in the field of weak lensing. Keeping different perspectives, such as exploring these alternative statistical approaches, is fundamental for determining the way forward.

## Acknowledgements

MSm acknowledges support from the Netherlands Organization for Scientific Research (NWO). AD acknowledges ERC Consolidator Grant (No. 770935). LM acknowledges the grants ASI-INAF n. 2018-23-HH.0 and PRIN-MIUR 2017 WSCC32 “Zooming into dark matter and proto-galaxies with massive lensing clusters”. MSe acknowledges financial contribution from contract ASI-INAF n.2017-14-H.0 and contract INAF mainstream project 1.05.01.86.10.

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 Target/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).

## 5.A Tests for systematics

For completeness, we repeated some of the tests for systematics that were already carried out in Dvornik et al. (2017) and Bellagamba et al. (2019) because of our difference in sky coverage, background selection, and estimated redshift distribution compared with those two studies.

### 5.A.1 Photometric redshift

We used the same method as Dvornik et al. (2017) to determine the co-moving critical density. There are two important differences that could affect the uncertainty in  $\Sigma_{\text{cr}}$ : We selected lenses at a significantly higher redshift, and we complemented our background source selection with the color selection described in Sect. 5.3.2.



We assessed the relative errors in  $\Sigma_{\text{cr}}$  by performing  $10^4$  bootstraps of the spectroscopic catalog of Hildebrandt et al. (2017). We find the median error on  $\Sigma_{\text{cr}}$  to be  $\sim 0.5\%$ , as shown in Fig. 5.11.

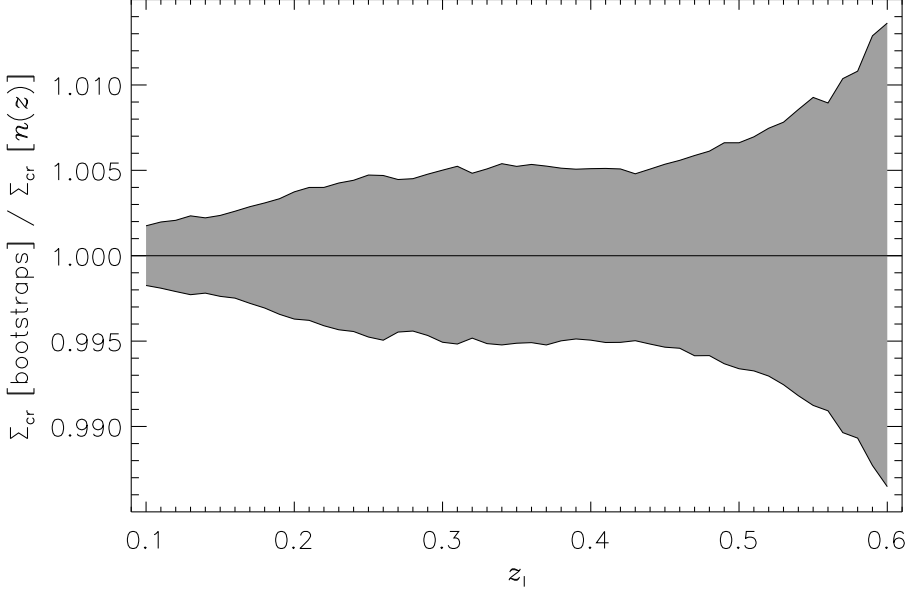


Figure 5.11: Relative errors in  $\Sigma_{\text{cr}}$ , estimated using  $10^4$  bootstraps of the spectroscopic catalog of Hildebrandt et al. (2017). The relative error on the ESD is negligible.

### 5.A.2 Contamination of the background sample by cluster galaxies

Dvornik et al. (2017) showed that an offset of  $\Delta z = 0.2$  is enough to avoid a significant contamination of the background sources by unidentified GAMA group members. For lenses at a higher redshift, this contamination increases, while at the same time the density of available background sources decreases due to the observed depth of KiDS-450.

We used the same test as Dvornik et al. (2017) to assess the source density around AMICO clusters in order to determine the necessary  $\Delta z$  offset between the lens and the sources. We find that  $\Delta z = 0.2$  is appropriate for our cluster selection (Fig. 5.12).

### 5.A.3 Individual bin ESD profiles and cross signals

In Fig. 5.14 we show the ESD profiles for the 13 cluster bins, including the cross signal, which is consistent with zero. We also show the derived halo model fits and their confidence intervals, comparing the fits using the full AMICO cluster catalog from Fig. 5.7.



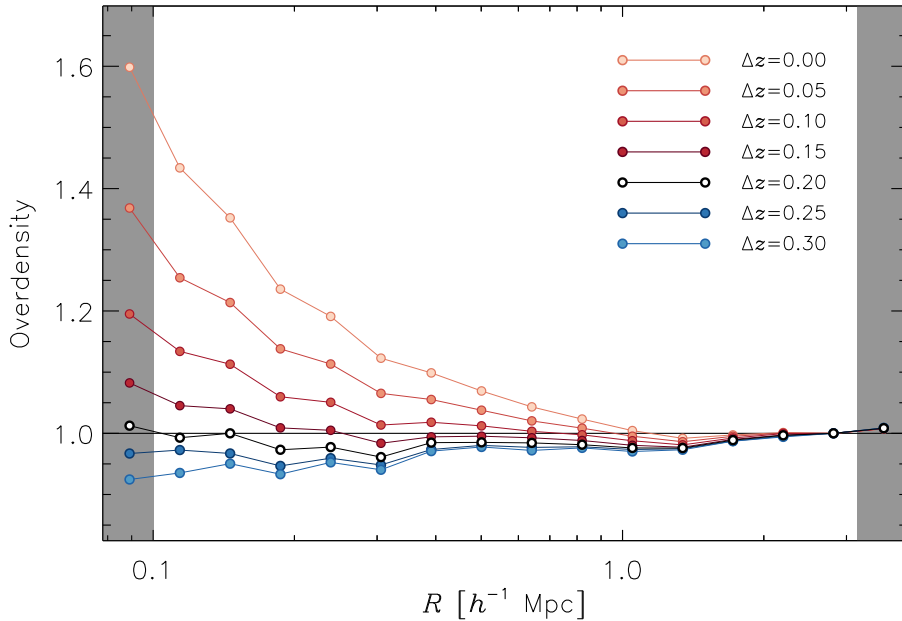


Figure 5.12: Relative source densities around AMICO clusters as a function of radius,  $R$ , and the photometric redshift offset,  $z_B \geq z_l + \Delta z$ , between the lens and the source. We note that some small under-densities around  $R = 0.2 h^{-1} \text{ Mpc}$  may be due to the relative normalization.

### 5.A.4 Tile bootstrap

As described in Sect. 5.3.3, we could not use the same bootstrap approach as Dvornik et al. (2017), due to the sparsity of lenses in the highest lens luminosity bins. Since our bootstrap approach described in Sect. 5.3.3 does not account for cosmic variance and is not sensitive to the clustering effect of dark matter halos, we compare the covariances derived by the two bootstrap methods for the ESD of the whole lens selection in Fig. 5.13 and find no significant differences or pattern beyond what is expected from statistical noise. Since we expect the contribution from cosmic variance to be even lower for subsets of lenses, we conclude that our bootstrap approach yields a good estimate of the covariance.

## 5.B Analysis of dependence on outer data points

In Fig. 5.10 it can be seen that the distribution of clusters in the two outermost luminosity bins is not symmetric. At the lower end, this is due to the selection criterion of  $\lambda^*$  in the AMICO catalog. At the higher end, we have only a few clusters.

We assessed the effect these two points have on the  $L_{200} - M_{200}$  scaling relation by repeating the fit without these bins. We find no difference within the statistical uncertainties, as given in Eq. 5.39:



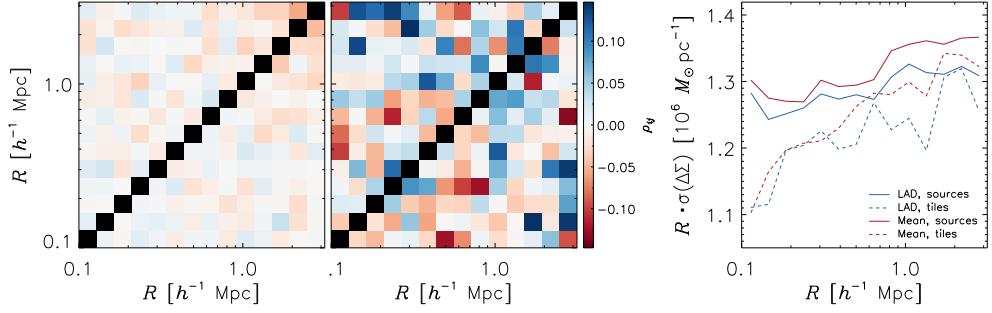


Figure 5.13: Left: Correlation derived from bootstrapping the signal by individual sources; same as Fig. 5.6, but with the color stretch adjusted to the middle plot. Middle: Correlation derived from bootstrapping the signal in  $1 \text{ deg}^2$  tiles. The upper-left corners show the correlations from LAD regression. The lower-right corners show the correlations from using the weighted mean. Right: Comparison of the errors obtained from bootstrapping sources (LAD: solid blue; mean: solid red) and bootstrapping  $1 \text{ deg}^2$  tiles (LAD: dashed blue; mean: dashed red).

$$\frac{M_{200}}{10^{14.1} h^{-1} M_{\odot}} = (0.98 \pm 0.06) \left( \frac{L_{200}}{10^{11.8} h^{-2} L_{\odot}} \right)^{(1.25 \pm 0.10)} \quad (\text{Mean}) \quad (5.39a)$$

$$\frac{M_{200}}{10^{14.1} h^{-1} M_{\odot}} = (1.02 \pm 0.06) \left( \frac{L_{200}}{10^{11.8} h^{-2} L_{\odot}} \right)^{(1.23 \pm 0.09)} \quad (\text{LAD}) \quad (5.39b)$$



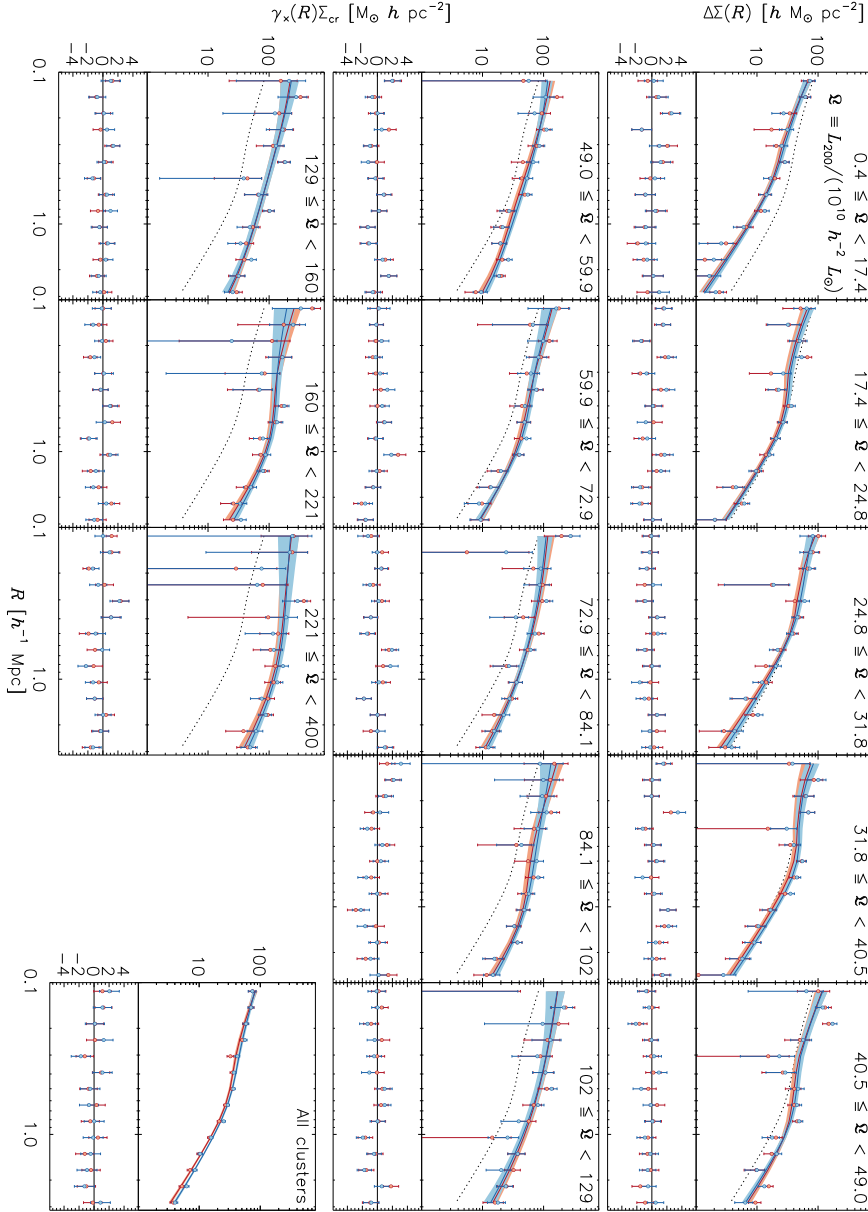


Figure 5.14: Estimated ESD profiles and cross signal of the 13 cluster bins, using the LAD estimator (blue) and a weighted mean (red). The error bars are the square roots of the diagonal values of the respective covariance matrices. The solid lines represent the best fitting halo model obtained by the MCMC fit. The shaded regions show the 68.3% confidence bands, estimated using the 15.9th and 84.1th percentiles of the MCMC realizations. Individual bins are indicated by the range in normalized luminosity,  $\vartheta \equiv L_{200}/(10^{10} h^{-2} L_{\odot})$ . The lower-right plot shows the ESD profile estimated from the full AMICO cluster catalog, also shown in Fig. 5.7. The average of the two best fitting halo models from Fig. 5.7 are shown in each panel as a dotted line for easy comparison.