Astronomy & Astrophysics

KiDS+VIKING+GAMA: Testing semi-analytic models of galaxy evolution with galaxy–galaxy–galaxy lensing

Laila Linke[1], Patrick Simon[1], Peter Schneider[1], Thomas Erben[1], Daniel J. Farrow[2], Catherine Heymans[3,4], Hendrik Hildebrandt[4], Andrew M. Hopkins[5], Arun Kannawadi[6,7], Nicola R. Napolitano[8], Cristóbal Sifón[9], and Angus H. Wright[4]

[1] Argelander-Institut für Astronomie, Rheinische Friedrich-Wilhelms-Universität, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: llinke@astro.uni-bonn.de
[2] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse 1, 85748 Garching, Germany
[3] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
[4] Ruhr-Universität Bochum, Astronomisches Institut, German Centre for Cosmological Lensing (GCCL), Universitätsstr. 150, 44801 Bochum, Germany
[5] Australian Astronomical Optics, Macquarie University, 105 Delhi Rd, North Ryde, NSW 2113, Australia
[6] Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08544, USA
[7] Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands
[8] School for Physics and Astronomy, Sun Yat-sen University, Zhuhai Campus, Guangzhou 519082, PR China
[9] Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4059, Valparaíso, Chile

ABSTRACT

Context. Several semi-analytic models (SAMs) try to explain how galaxies form, evolve, and interact inside the dark matter large-scale structure. These SAMs can be tested by comparing their predictions for galaxy–galaxy–galaxy lensing (G3L), which is weak gravitational lensing around galaxy pairs, with observations.
Aims. We evaluate the SAMs by Henriques et al. (2015, MNRAS, 451, 2663, hereafter H15) and by Lagos et al. (2012, MNRAS, 426, 2142, hereafter L12), which were implemented in the Millennium Run, by comparing their predictions for G3L to observations at smaller scales than previous studies and also for pairs of lens galaxies from different populations.
Methods. We compared the G3L signal predicted by the SAMs to measurements in the overlap of the Galaxy And Mass Assembly survey (GAMA), the Kilo-Degree Survey (KiDS), and the VISTA Kilo-degree Infrared Galaxy survey (VIKING) by splitting lens galaxies into two colour and five stellar-mass samples. Using an improved G3L estimator, we measured the three-point correlation of the matter distribution with "mixed lens pairs" with galaxies from different samples, and with "unmixed lens pairs" with galaxies from the same sample.
Results. Predictions by the H15 SAM for the G3L signal agree with the observations for all colour-selected samples and all but one stellar-mass-selected sample with 95% confidence. Deviations occur for lenses with stellar masses below $9.5\,h^{-2}\,M_\odot$ at scales below $0.2\,h^{-1}$ Mpc. Predictions by the L12 SAM for stellar-mass selected samples and red galaxies are significantly higher than observed, while the predicted signal for blue galaxy pairs is too low.
Conclusions. The L12 SAM predicts more pairs of low stellar mass and red galaxies than the H15 SAM and the observations, as well as fewer pairs of blue galaxies. This difference increases towards the centre of the galaxies' host halos. Likely explanations are different treatments of environmental effects by the SAMs and different models of the initial mass function. We conclude that G3L provides a stringent test for models of galaxy formation and evolution.

Key words. gravitational lensing: weak – cosmology: observations – large-scale structure of Universe – galaxies: evolution

1. Introduction

One important goal of extragalactic astronomy and cosmology is understanding galaxy formation and evolution. The following two different types of simulations try to reproduce the observed galaxy and matter distribution: full hydrodynamical simulations (e.g. Vogelsberger et al. 2014; Crain et al. 2015; Kaviraj et al. 2017; Nelson et al. 2019) and dark-matter-only *N*-body simulations with galaxies inserted according to semi-analytic models (SAMs) of galaxy formation and evolution.

Multiple SAMs, with different assumptions on small-scale physics, such as the gas cooling time, the star formation rate, or supernovae feedback, have been proposed (e.g. Bower et al. 2006; Guo et al. 2011; Lagos et al. 2012; Henriques et al. 2015). These models must be assessed by comparing their predictions to observations of galaxy statistics. Previous tests of SAMs

included galaxy-galaxy lensing (GGL; e.g. Saghiha et al. 2017) and galaxy clustering (e.g. Henriques et al. 2017).

A more sensitive test than GGL is comparing the galaxy–galaxy–galaxy lensing (G3L) signal predicted by the SAMs to observations. The G3L effect, which was first discussed by Schneider & Watts (2005), describes the weak gravitational lensing of pairs of background galaxies around foreground galaxies (lens-shear-shear correlation) and of individual background galaxies around pairs of foreground galaxies (lens-lens-shear correlation). Unlike GGL or galaxy clustering, G3L depends on the galaxy-matter three-point correlation and the halo occupation distribution of galaxy pairs. In principle, it also depends on the ellipticity of dark matter halos as well as mis-alignments between the galaxy and matter distribution because the galaxy pair orientation introduces a preferred direction.

The lens-lens-shear correlation was measured for lens pairs separated by several megaparsecs (Mpc) in order to detect inter-cluster filaments (Mead et al. 2010; Clampitt et al. 2016; Epps & Hudson 2017; Xia et al. 2020). However, for the evaluation of SAMs, it is more suitable to study the correlation at smaller, sub-Mpc scales. At these scales, the G3L signal is more sensitive to the small-scale physics that vary between different SAMs, because it depends primarily on galaxy pairs with galaxies in the same dark matter halo. For lens pairs with galaxies of a similar stellar mass or colour, the small-scale lens-lens-shear correlation was determined by Simon et al. (2008) in the Red Cluster Sequence survey (Hildebrandt et al. 2016) and Simon et al. (2013) in the Canada-France-Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012). The G3L measured in CFHTLenS was compared to predictions by multiple SAMs that were implemented in the Millennium Run (MR; Springel et al. 2005) by Saghiha et al. (2017) and Simon et al. (2019). They demonstrate that G3L is more effective in evaluating SAMs than GGL and that the SAM by Henriques et al. (2015, H15 hereafter) is in agreement with the observations in CFHTLenS, while the SAM by Lagos et al. (2012, L12 hereafter) predicts G3L signals that are too large.

Nonetheless, these previous measurements of G3L at small scales only used photometric data with imprecise redshift estimates for the lens galaxies. Therefore, lens galaxy pairs with galaxies separated along the line-of-sight (chance pairs) were treated the same as lens galaxy pairs with galaxies close to each other (true pairs). As the G3L signal of chance pairs is much weaker, this lowers the signal-to-noise ratio (S/N).

However, Linke et al. (2020, L20 hereafter) demonstrate that the S/N could be improved substantially by weighting each lens galaxy pair according to the line-of-sight separation between its galaxies to reduce the impact of chance pairs. We used this improved estimator to test the H15 and the L12 SAMs with state-of-the-art observational data, consisting of the photometric Kilo-Degree Survey (KiDS) and VISTA Kilo-degree Infrared Galaxy survey (VIKING) as well as the spectroscopic Galaxy And Mass Assembly survey (GAMA). We used the shapes of galaxies observed by KiDS as shear estimates, while GAMA provides lens galaxies with precise spectroscopic redshifts. These spectroscopic redshifts allowed us to employ the redshift weighting suggested by L20. Furthermore, we extended the angular range at which we measured the G3L signal to lower scales with the adaptive binning scheme for the G3L three-point correlation function proposed by L20. Thereby, we could assess the SAMs deeper inside dark-matter halos.

As of now, the lens-lens-shear correlation has only been measured for lens pairs with galaxies from the same colour or stellar-mass sample (unmixed lens pairs) and not for lens pairs with galaxies from different samples (mixed lens pairs). However, comparing the measurements for G3L with mixed pairs is a compelling new test of SAMs, because this signal depends on the correlation of different galaxy populations inside halos. For example, the mixed pair G3L signal would be higher for two fully correlated galaxy populations than for two uncorrelated populations, while the GGL signal would stay the same. Therefore, we can assess the predictions of SAMs for the correlation between different galaxy populations with the G3L of mixed lens pairs. Accordingly, we measure not only the G3L signal for lens pairs from the same population but also the signal for mixed lens pairs, with galaxies from different colour- or stellar-mass samples.

This paper is structured as follows: in Sect. 2 we review the basics of G3L and introduce the third-order aperture statistics, which are the G3L observables throughout this work. Section 3
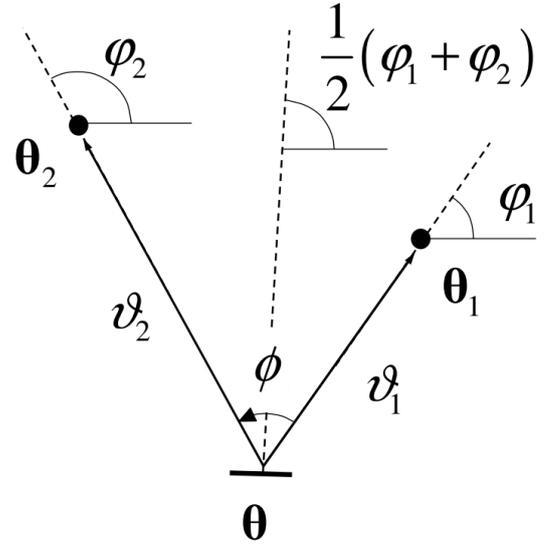
Fig. 1. Geometry of a G3L configuration with one source and two lens galaxies; adapted from Schneider & Watts (2005). The two lens galaxies are at angular positions θ_1 and θ_2 on the sky; the source galaxy is at θ. The separation vectors ϑ_1 and ϑ_2 of the lenses from the source have lengths ϑ_1 and ϑ_2, as well as polar angles φ_1 and φ_2. The angle between ϑ_1 and ϑ_2 is the opening angle ϕ. The tangential shear of the source galaxy is measured with respect to the dashed line, which is the bisector of ϕ.

discusses our estimators for the three-point correlation function and the aperture statistics. The SAMs and the creation of the simulated and observational data sets are described in Sect. 4. We present the G3L signals measured in the observation and the simulation in Sect. 5 and discuss our findings in Sect. 6.

Throughout this paper we assume a flat ΛCDM cosmology with matter density $\Omega_{\rm m} = 0.25$, baryon density $\Omega_{\rm b} = 0.045$, dark energy density $\Omega_\Lambda = 0.75$, Hubble constant $H_0 = 73\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and power spectrum normalisation $\sigma_8 = 0.9$. These parameters were used in the creation of the MR and differ from more recent constraints (e.g. Planck Collaboration I 2020). However, weak gravitational lensing is most sensitive to the combination of the matter density and the power spectrum normalisation $S_8 = \sigma_8\,\sqrt{\Omega_{\rm m}/0.3}$. This parameter is almost the same in the MR and the most recent *Planck* measurements; it is $S_{8,\rm MR} = 0.822$ for the MR and $S_{8,\rm Planck} = 0.825 \pm 0.011$ in Planck Collaboration I (2020).

2. Theory of galaxy–galaxy–galaxy lensing

G3L is a weak gravitational lensing effect (see e.g. Bartelmann & Schneider 2001). It comprises the correlation of individual lens galaxies with the shear of source galaxy pairs, as well as the correlation of lens galaxy pairs with the shear of individual source galaxies. Here, we study the second effect, the lens-lens-shear correlation. Figure 1 illustrates the geometric configuration of lens and source galaxies for this correlation.

2.1. Three-point correlation function

The main observable for the lens-lens-shear correlation is the three-point correlation function $\tilde{G}$. This function correlates the projected lens galaxy number density and the tangential gravitational lensing shear $\gamma_{\rm t}$, measured with respect to the bisector of the angle ϕ between the lens-source separations ϑ_1 and ϑ_2

(see Fig. 1). For unmixed lens pairs, whose galaxies have the projected number density $N(\boldsymbol{\vartheta})$, $\tilde{\mathcal{G}}$ is

$$\tilde{\mathcal{G}}(\boldsymbol{\vartheta_1}, \boldsymbol{\vartheta_2}) = \frac{1}{\overline{N}^2} \langle N(\boldsymbol{\theta} + \boldsymbol{\vartheta_1}) N(\boldsymbol{\theta} + \boldsymbol{\vartheta_2}) \gamma_t(\boldsymbol{\theta}) \rangle, \tag{1}$$

where $\overline{N}$ is the mean number density of lenses. For mixed lens pairs, where the lenses are from samples with number densities N_1 and N_2, the correlation function is

$$\tilde{\mathcal{G}}(\boldsymbol{\vartheta_1}, \boldsymbol{\vartheta_2}) = \frac{1}{\overline{N_1}\,\overline{N_2}} \langle N_1(\boldsymbol{\theta} + \boldsymbol{\vartheta_1}) N_2(\boldsymbol{\theta} + \boldsymbol{\vartheta_2}) \gamma_t(\boldsymbol{\theta}) \rangle. \tag{2}$$

Instead of measuring $\tilde{\mathcal{G}}$, we estimated a redshift-weighted correlation function $\tilde{\mathcal{G}}_Z$, which includes a redshift weighting function Z, which depends on the redshift difference $\Delta z_{12} = z_1 - z_2$ of the lenses in a pair. We chose the redshift weighting such that it is large for small Δz_{12} and vanishes for large Δz_{12}. It, therefore, weights true pairs with small redshift differences higher than chance pairs with large redshift differences and increases the S/N (L20). To define the redshift-weighted correlation function $\tilde{\mathcal{G}}_Z$, we used that the projected lens number densities $N_{1,2}$ are related to the three-dimensional number densities $n_{1,2}(\boldsymbol{\theta}, z)$ at redshift z by the selection functions $\nu_{1,2}(z)$,

$$N_{1,2}(\boldsymbol{\theta}) = \int \mathrm{d}z \, \nu_{1,2}(z) \, n_{1,2}(\boldsymbol{\theta}, z). \tag{3}$$

The selection functions give the fraction of galaxies at redshift z included in the lens sample. For a flux-limited galaxy sample, this corresponds to the fraction of galaxies brighter than the magnitude limit. The selection functions $\nu_{1,2}(z)$ are related to the galaxies' redshift distributions $p_{1,2}(z)$ by

$$\nu_{1,2}(z) = p_{1,2}(z) \, \frac{\int_A \mathrm{d}^2\theta \, N_{1,2}(\boldsymbol{\theta})}{\int_A \mathrm{d}^2\theta \, n_{1,2}(\boldsymbol{\theta}, z)} \tag{4}$$

with Eq. (3), the redshift-weighted correlation function is

$$\tilde{\mathcal{G}}_Z(\boldsymbol{\vartheta_1}, \boldsymbol{\vartheta_2})$$
$$= \left[\int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, \nu_1(z_1) \, \nu_2(z_2) Z(\Delta z_{12}) \, \bar{n}_1(z_1) \, \bar{n}_2(z_2) \right]^{-1}$$
$$\times \int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, \nu_1(z_1) \, \nu_2(z_2) \, Z(\Delta z_{12}) \tag{5}$$
$$\times \langle n_1(\boldsymbol{\theta} + \boldsymbol{\vartheta_1}, z_1) \, n_2(\boldsymbol{\theta} + \boldsymbol{\vartheta_2}, z_2) \, \gamma_t(\boldsymbol{\theta}) \rangle,$$

where $\Delta z_{12} := z_1 - z_2$ is the redshift difference between lenses in a pair and

$$\bar{n}_{1,2}(z) = \langle n_{1,2}(\boldsymbol{\theta}, z) \rangle = \frac{1}{A} \int_A \mathrm{d}^2\theta \, n_{1,2}(\boldsymbol{\theta}, z). \tag{6}$$

Additionally, we also measured the physical correlation function $\tilde{\mathcal{G}}_{\mathrm{phys}}(\boldsymbol{r_1}, \boldsymbol{r_2})$, which gives the projected excess mass around lens pairs with physical lens-source separations $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ projected on a plane midway between the lenses. To find $\tilde{\mathcal{G}}_{\mathrm{phys}}$, instead of averaging the tangential shear γ_t, we averaged the projected excess mass density $\Delta\Sigma$ around lens pairs with

$$\tilde{\mathcal{G}}_{\mathrm{phys}}(\boldsymbol{r_1}, \boldsymbol{r_2})$$
$$= \left[\int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, \nu_1(z_1) \, \nu_2(z_2) \, Z(\Delta z_{12}) \, \bar{n}_1(z_1) \, \bar{n}_2(z_2) \right]^{-1}$$
$$\times \int \mathrm{d}z_1 \int \mathrm{d}z_2 \, \nu_1(z_1) \, \nu_2(z_2) \, Z(\Delta z_{12})$$
$$\times \langle n_1\left(\boldsymbol{\theta} + D_{12}^{-1}\, \boldsymbol{r_1}, z_1\right) n_2\left(\boldsymbol{\theta} + D_{12}^{-1}\, \boldsymbol{r_2}, z_2\right) \Delta\Sigma(\boldsymbol{\theta}, z_{12}) \rangle, \tag{7}$$

where

$$z_{12} = \frac{z_1 + z_2}{2}, \tag{8}$$

and

$$D_{12} = D_{\mathrm{A}}(0, z_{12}), \tag{9}$$

with the angular diameter distance $D_{\mathrm{A}}(z_1, z_2)$ between redshifts z_1 and z_2. The projected excess mass density $\Delta\Sigma$ is

$$\Delta\Sigma(\boldsymbol{\theta}, z_{\mathrm{d}}) = \frac{\gamma_t(\boldsymbol{\theta})}{\overline{\Sigma}_{\mathrm{crit}}^{-1}(z_{\mathrm{d}})}, \tag{10}$$

with the source-averaged inverse critical surface mass density $\overline{\Sigma}_{\mathrm{crit}}^{-1}$. Using the source redshift distribution $p(z_{\mathrm{s}})$, $\overline{\Sigma}_{\mathrm{crit}}^{-1}$ is

$$\overline{\Sigma}_{\mathrm{crit}}^{-1}(z_{\mathrm{d}}) = \int_{z_{\mathrm{d}}}^\infty \mathrm{d}z_{\mathrm{s}} \, p(z_{\mathrm{s}}) \, \frac{4\pi G}{c^2} \, \frac{D_{\mathrm{A}}(z_{\mathrm{d}}, z_{\mathrm{s}}) \, D_{\mathrm{A}}(z_{\mathrm{d}})}{D_{\mathrm{A}}(z_{\mathrm{s}})}, \tag{11}$$

where $D_{\mathrm{A}}(z) := D_{\mathrm{A}}(0, z)$. This critical surface mass density is not the comoving critical surface mass density $\overline{\Sigma}_{\mathrm{crit, com}}$, defined by

$$\overline{\Sigma}_{\mathrm{crit,com}}^{-1}(z_{\mathrm{d}}) = \int_{z_{\mathrm{d}}}^\infty \mathrm{d}z_{\mathrm{s}} \, p(z_{\mathrm{s}}) \, \frac{4\pi G}{c^2} \, \frac{D_{\mathrm{A}}(z_{\mathrm{d}}, z_{\mathrm{s}}) \, D_{\mathrm{A}}(z_{\mathrm{d}})}{(1 + z_{\mathrm{d}}) \, D_{\mathrm{A}}(z_{\mathrm{s}})},$$

used in some weak lensing studies. Appendix C in Dvornik et al. (2018) discusses the implications of different definitions of the critical surface mass density.

The correlation functions $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{\mathrm{phys}}$ depend only on the lens-source distances ϑ_1, ϑ_2 and r_1, r_2, and the opening angle ϕ between the lens-source separations, because of the statistical isotropy of the matter and galaxy density fields. Therefore, we define

$$\tilde{\mathcal{G}}_Z(\vartheta_1, \vartheta_2, \phi) := \tilde{\mathcal{G}}_Z(\boldsymbol{\vartheta_1}, \boldsymbol{\vartheta_2}), \tag{12}$$

and

$$\tilde{\mathcal{G}}_{\mathrm{phys}}(r_1, r_2, \phi) := \tilde{\mathcal{G}}_{\mathrm{phys}}(\boldsymbol{r_1}, \boldsymbol{r_2}). \tag{13}$$

2.2. Aperture statistics

The three-point correlation function $\tilde{\mathcal{G}}$ contains second- and third-order statistics. This can be seen by writing $\tilde{\mathcal{G}}$ as

$$\tilde{\mathcal{G}}(\vartheta_1, \vartheta_2, \phi) = \langle \kappa_g(\boldsymbol{\theta} + \boldsymbol{\vartheta_1}) \, \kappa_g(\boldsymbol{\theta} + \boldsymbol{\vartheta_2}) \, \gamma_t(\boldsymbol{\theta}) \rangle$$
$$+ \langle \kappa_g(\boldsymbol{\theta} + \boldsymbol{\vartheta_1}) \, \gamma_t(\boldsymbol{\theta}) \rangle + \langle \kappa_g(\boldsymbol{\theta} + \boldsymbol{\vartheta_2}) \, \gamma_t(\boldsymbol{\theta}) \rangle, \tag{14}$$

where $\kappa_g(\boldsymbol{\vartheta}) = N(\boldsymbol{\vartheta})/\bar{N} - 1$ is the two-dimensional galaxy number density contrast. The second and third term in Eq. (14) are GGL statistics which correspond to the shear around individual lens galaxies.

When studying G3L we are interested in the excess shear around lens pairs, which is given only by the first term in Eq. (14). Therefore, we converted $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{\mathrm{phys}}$ to the third-order aperture statistics $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}} \rangle$ and $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}} \rangle_{\mathrm{phys}}$, which only include the third-order statistics, as shown in L20. For this, we use a compensated filter function with an aperture scale θ,

$$U_\theta(\vartheta) = \frac{1}{\theta^2} u\left(\frac{\vartheta}{\theta}\right), \tag{15}$$

which fulfills

$$\int \mathrm{d}\vartheta \, \vartheta \, U_\theta(\vartheta) = 0 \tag{16}$$

with this filter function, the third-order aperture statistics are defined as

$$\langle \mathcal{N} \mathcal{N} M_{\mathrm{ap}} \rangle (\theta_1, \theta_2, \theta_3)$$

$$= \left[\int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, Z(\Delta z_{12}) \, \bar{n}_1(z_1) \, \bar{n}_2(z_2) \right]^{-1}$$

$$\times \int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, Z(\Delta z_{12}) \left[\prod_{i=1}^3 \int \mathrm{d}^2\vartheta_i \, \frac{1}{\theta_i^2} \, u\left(\frac{\vartheta_i}{\theta_i}\right) \right]$$

$$\times \langle n_1(\boldsymbol{\vartheta_1}, z_1) \, n_2(\boldsymbol{\vartheta_2}, z_2) \, \kappa(\boldsymbol{\vartheta_3}) \rangle, \qquad (17)$$

and

$$\langle \mathcal{N} \mathcal{N} M_{\mathrm{ap}} \rangle_{\mathrm{phys}} (r_1, r_2, r_3)$$

$$= \left[\int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, Z(\Delta z_{12}) \, \bar{n}_1(z_1) \, \bar{n}_2(z_2) \right]^{-1}$$

$$\times \int_0^\infty \mathrm{d}z_1 \int_0^\infty \mathrm{d}z_2 \, Z(\Delta z_{12}) \left[\prod_{i=1}^3 \int \mathrm{d}^2 x_i \, \frac{1}{D_{12}^{-2} r_i^2} \, u\left(\frac{x_i}{r_i}\right) \right]$$

$$\times \langle n_1(D_{12}^{-1} \boldsymbol{x_1}, z_1) \, n_2(D_{12}^{-1} \boldsymbol{x_2}, z_2) \, \Sigma(D_{12}^{-1} \boldsymbol{x_3}, z_{12}) \rangle, \qquad (18)$$

with the lensing convergence $\kappa(\boldsymbol{\theta})$ and the surface mass density Σ, given by

$$\Sigma(\boldsymbol{\theta}, z) = \frac{\kappa(\boldsymbol{\theta})}{\bar{\Sigma}_{\mathrm{crit}}^{-1}(z)}. \qquad (19)$$

We use the exponential filter function

$$u(x) = \frac{1}{2\pi} \left(1 - \frac{x^2}{2} \right) \exp\left(-\frac{x^2}{2} \right), \qquad (20)$$

for which the aperture statistics can be calculated from $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{\mathrm{phys}}$ with

$$\langle \mathcal{N} \mathcal{N} M_{\mathrm{ap}} \rangle (\theta_1, \theta_2, \theta_3)$$

$$= \int_0^\infty \mathrm{d}\vartheta_1 \vartheta_1 \int_0^\infty \mathrm{d}\vartheta_2 \vartheta_2 \int_0^{2\pi} \mathrm{d}\phi \, \tilde{\mathcal{G}}_Z(\vartheta_1, \vartheta_2, \phi)$$

$$\times \mathcal{A}_{\mathcal{N} \mathcal{N} M}(\vartheta_1, \vartheta_2, \phi \,|\, \theta_1, \theta_2, \theta_3). \qquad (21)$$

and

$$\langle \mathcal{N} \mathcal{N} M_{\mathrm{ap}} \rangle_{\mathrm{phys}} (r_1, r_2, r_3)$$

$$= \int_0^\infty \mathrm{d}x_1 x_1 \int_0^\infty \mathrm{d}x_2 x_2 \int_0^{2\pi} \mathrm{d}\phi \, \tilde{\mathcal{G}}_{\mathrm{phys}}(x_1, x_2, \phi)$$

$$\times \mathcal{A}_{\mathcal{N} \mathcal{N} M}(x_1, x_2, \phi \,|\, r_1, r_2, r_3). \qquad (22)$$

The kernel function $\mathcal{A}_{\mathcal{N} \mathcal{N} M}$ is defined in the appendix of Schneider & Watts (2005).

2.3. Galaxy bias

The aperture statistics can be used to constrain the galaxy bias, which is the relation between the galaxy number density contrast and the matter density contrast (e.g Schneider & Watts 2005). The simplest assumption for this bias is a linear deterministic relation,

$$\kappa_{\mathrm{g}}(\boldsymbol{\theta}) = b \, \kappa(\boldsymbol{\theta}), \qquad (23)$$

where b is a scale-independent bias factor (Kaiser 1984). A larger bias factor b implies a higher galaxy number density for a given matter overdensity. For two galaxy populations with bias

factors b_1 and b_2, this simple model predicts for the aperture statistics

$$\langle \mathcal{N}_1 \mathcal{N}_2 M_{\mathrm{ap}} \rangle \propto b_1 \, b_2. \qquad (24)$$

From this follows, that

$$R := \frac{\langle \mathcal{N}_1 \mathcal{N}_2 M_{\mathrm{ap}} \rangle}{\sqrt{\langle \mathcal{N}_1 \mathcal{N}_1 M_{\mathrm{ap}} \rangle \langle \mathcal{N}_2 \mathcal{N}_2 M_{\mathrm{ap}} \rangle}} = \frac{b_1 b_2}{\sqrt{b_1^2 b_2^2}} = 1. \qquad (25)$$

We measure R in the observation and simulation to assess the assumption of linear deterministic bias.

3. Methods

3.1. Estimating the three-point correlation function

To measure $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{\mathrm{phys}}$, we used the estimators from L20 for N_{s} source and N_{d} lens galaxies. These estimators measure the correlation functions by averaging the ellipticities of the source galaxies over all lens-lens-source galaxy triplets. For $\tilde{\mathcal{G}}_Z$, in the bin B of $(\vartheta_1, \vartheta_2, \phi)$, the estimator is the real part of

$$\tilde{\mathcal{G}}_{Z,\mathrm{est}}(B)$$

$$= -\frac{\sum_{i,j=1}^{N_{\mathrm{d}}} \sum_{k=1}^{N_{\mathrm{s}}} w_k \, \epsilon_k \, \mathrm{e}^{-\mathrm{i}(\varphi_{ik} + \varphi_{jk})} \left[1 + \omega_Z(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|) \right] Z(\Delta z_{ij}) \, \Delta_{ijk}(B)}{\sum_{i,j=1}^{N_{\mathrm{d}}} \sum_{k=1}^{N_{\mathrm{s}}} w_k \, Z(\Delta z_{ij}) \, \Delta_{ijk}(B)}$$

$$(26)$$

$$:= -\frac{\sum_{ijk} w_k \, \epsilon_k \, \mathrm{e}^{-\mathrm{i}(\varphi_{ik} + \varphi_{jk})} \left[1 + \omega_Z\left(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j|\right) \right] Z(\Delta z_{ij}) \, \Delta_{ijk}(B)}{\sum_{ijk} w_k \, Z(\Delta z_{ij}) \, \Delta_{ijk}(B)}, \qquad (27)$$

where w_k is the weight of the source ellipticity ϵ_k, ω_Z is the redshift-weighted angular two-point correlation function of the lens galaxies, and

$$\Delta_{ijk}(B) = \begin{cases} 1 & \text{for } (|\boldsymbol{\theta}_k - \boldsymbol{\theta}_i|, |\boldsymbol{\theta}_k - \boldsymbol{\theta}_j|, \phi_{ijk}) \in B \\ 0 & \text{otherwise} \end{cases}. \qquad (28)$$

The angles φ_{ik} and φ_{jk} are the polar angles of the lens-source separation vectors $\boldsymbol{\theta}_i - \boldsymbol{\theta}_k$ and $\boldsymbol{\theta}_j - \boldsymbol{\theta}_k$ (corresponding to φ_1 and φ_2 in Fig. 1) and $\phi_{ijk} = \varphi_{ik} - \varphi_{jk}$ is the opening angle between the lens-source separation vectors (corresponding to ϕ in Fig. 1).

Source galaxies with more precise shape measurements receive a higher ellipticity weight w_k. The weight, therefore, increases the contribution of source galaxies with more exact shapes to the estimator. For the simulated shear data, we set the weights to $w_k = 1$ for all sources.

We obtained the lens two-point correlation function ω_Z, with the estimator by Szapudi & Szalay (1998) which is

$$\omega_Z(\theta) = \frac{N_{\mathrm{r}_1} N_{\mathrm{r}_2}}{N_{\mathrm{d}_1} N_{\mathrm{d}_2}} \frac{D_1 D_{2Z}(\theta)}{R_1 R_{2Z}(\theta)} - \frac{N_{\mathrm{r}_1}}{N_{\mathrm{d}_1}} \frac{D_1 R_{2Z}(\theta)}{R_1 R_{2Z}(\theta)} - \frac{N_{\mathrm{r}_2}}{N_{\mathrm{d}_2}} \frac{D_2 R_{1Z}(\theta)}{R_1 R_{2Z}(\theta)} + 1, \qquad (29)$$

for two different observed lens samples with N_{d_1} and N_{d_2} galaxies and two "random samples". These random samples contain N_{r_1} and N_{r_2} unclustered galaxies following the same selection function as the observed galaxies.

The $D_1 D_{2Z}$, $D_1 R_{2Z}$, $D_2 R_{1Z}$, and $R_1 R_{2Z}$ are the pair counts of observed and random galaxies. For two equal lens samples and

$DD = D_1 D_2$, $DR = D_1 R_2 = D_2 R_1$, and $RR = R_1 R_2$, the estimator in Eq. (29) reduces to the usual Landy–Szalay estimator (Landy & Szalay 1993)

$$\omega_Z(\theta) = \frac{N_r^2}{N_d^2} \frac{DD_Z(\theta)}{RR_Z(\theta)} - 2 \frac{N_r}{N_d} \frac{DR_Z(\theta)}{RR_Z(\theta)} + 1. \tag{30}$$

Usually, pair counts are defined as the number of pairs within a certain angular separation. However, we used redshift-weighted pair counts to account for the stronger clustering of true lens pairs due to the redshift weighting function Z. They are defined for a bin centred at θ with bin size $\Delta\theta$ with the Heaviside step function Θ_H as

$$D_1 D_{2Z}(\theta) = \sum_{i=1}^{N_{d_1}} \sum_{j=1}^{N_{d_2}} \Theta_H \left(\theta + \Delta\theta/2 - |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right)$$
$$\times \Theta_H \left(-\theta + \Delta\theta/2 + |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right) Z(\Delta z_{ij}), \tag{31}$$

$$R_1 R_{2Z}(\theta) = \sum_{i=1}^{N_{r_1}} \sum_{j=1}^{N_{r_2}} \Theta_H \left(\theta + \Delta\theta/2 - |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right)$$
$$\times \Theta_H \left(-\theta + \Delta\theta/2 + |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right) Z(\Delta z_{ij}), \tag{32}$$

and

$$D_a R_{bZ}(\theta) = \sum_{i=1}^{N_{d_a}} \sum_{j=1}^{N_{r_a}} \Theta_H \left(\theta + \Delta\theta/2 - |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right)$$
$$\times \Theta_H \left(-\theta + \Delta\theta/2 + |\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right) Z(\Delta z_{ij}), \tag{33}$$

with $a, b \in 1, 2$.

Following L20, we chose a Gaussian weighting function,

$$Z(\Delta z_{12}) = \exp\left(-\frac{\Delta z_{12}^2}{2\sigma_z^2} \right), \tag{34}$$

with width $\sigma_z = 0.01$. This width is larger than typical galaxy correlation lengths and redshifts induced by the peculiar motion of galaxies. Accordingly, true lens pairs are not affected by the redshift weighting, while chance pairs are down-weighted. Choosing a different σ_z influences the magnitude of the measured aperture statistics as well as the S/N of the measurement. Nonetheless, as long as the same width is chosen for the observation and the simulation, their G3L signals can be compared.

The estimator for $\tilde{\mathcal{G}}_{phys}$ for the bin B of (r_1, r_2, ϕ) is

$$\tilde{\mathcal{G}}_{est,phys}(B) =$$
$$- \frac{\sum_{ijk} w_k \, \epsilon_k \, e^{-i(\varphi_{ik} + \varphi_{jk})} \left[1 + \omega_Z \left(|\boldsymbol{\theta_i} - \boldsymbol{\theta_j}| \right) \right] Z(\Delta z_{ij}) \, \bar{\Sigma}_{crit}^{-1}(z_{ij}) \, \Delta_{ijk}^{ph}(B)}{\sum_{ijk} w_k \, \bar{\Sigma}_{crit}^{-2}(z_{ij}) \, Z(\Delta z_{ij}) \, \Delta_{ijk}^{ph}(B)},$$
$$\tag{35}$$

with

$$\Delta_{ijk}^{ph}(B) = \begin{cases} 1 & \text{for } \left(D_{ij} |\boldsymbol{\theta_k} - \boldsymbol{\theta_i}|, D_{ij} |\boldsymbol{\theta_k} - \boldsymbol{\theta_j}|, \phi_{ijk} \right) \in B \\ 0 & \text{otherwise} \end{cases}. \tag{36}$$

We measured $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{phys}$ initially for $128 \times 128 \times 128$ bins, which were linearly spaced along ϕ and logarithmically spaced along $\vartheta_{1,2}$ and $r_{1,2}$. For $\tilde{\mathcal{G}}_Z$, the $\vartheta_{1,2}$ are between $0'.15$ and $200'$ for the observed and between $0'.15$ and $320'$ for the simulated data. For $\tilde{\mathcal{G}}_{phys}$, we chose $r_{1,2}$ between $0.02\,\text{Mpc}$ and $40\,\text{Mpc}$.

We then applied the adaptive binning scheme of L20, by which the parameter space was tessellated to remove bins for which no galaxy triplet is in the data.

The correlation function was measured individually for 24 tiles of the observational data of size $2.5° \times 3°$ and 64 fields-of-view of the MR of size $4° \times 4°$, leading to estimates $\tilde{\mathcal{G}}_{est}^i$ and $\tilde{\mathcal{G}}_{est, ph}^i$ for each tile and field-of-view, respectively. The division into small patches allowed us to project the observational measurements to Cartesian coordinates and to estimate the uncertainty of the measurement with jackknife resampling. For each data set, the individual estimates were combined to form the total correlation functions with

$$\tilde{\mathcal{G}}_{est}(B) = \frac{\sum_{i=1}^{N} \tilde{\mathcal{G}}_{est}^i(B) \, W^i(B)}{\sum_{i=1}^{N} W^i(B)}, \tag{37}$$

where

$$W^i(B) = \sum_{ijk} w_k \, Z(\Delta z_{ij}) \, \Delta_{ijk}(B), \tag{38}$$

and

$$\tilde{\mathcal{G}}_{est, ph}(B) = \frac{\sum_{i=1}^{N} \tilde{\mathcal{G}}_{est, ph}^i(B) \, W_{ph}^i(B)}{\sum_{i=1}^{N} W_{ph}^i(B)}, \tag{39}$$

where

$$W_{ph}^i(B) = \sum_{ijk} w_k \, \bar{\Sigma}_{crit}^{-2}(z_{ij}) \, Z(\Delta z_{ij}) \, \Delta_{ijk}^{ph}(B). \tag{40}$$

3.2. Computing aperture statistics

To compute $\langle \mathcal{NNM}_{ap} \rangle$ and $\langle \mathcal{NNM}_{ap} \rangle_{phys}$, we integrated over $\tilde{\mathcal{G}}_Z$ using Eqs. (21) and (22). We numerically approximated the integrals by summing over all N_{bins} of $\tilde{\mathcal{G}}_Z$ after tessellation with

$$\langle \mathcal{NNM}_{ap} \rangle(\theta_1, \theta_2, \theta_3)$$
$$= \sum_{i=1}^{N_{bin}} V(b_i) \, \tilde{\mathcal{G}}_{Z,est}(b_i) \, \mathcal{A}_{NNM}(b_i \,|\, \theta_1, \theta_2, \theta_3), \tag{41}$$

and

$$\langle \mathcal{NNM}_{ap} \rangle_{phys}(r_1, r_2, r_3)$$
$$= \sum_{i=1}^{N_{bin}} V(b_i) \, \tilde{\mathcal{G}}_{est, phys}(b_i) \, \mathcal{A}_{NNM}(b_i \,|\, r_1, r_2, r_3). \tag{42}$$

Here, b_i is the ith bin, which has size $V(b_i)$, and $\mathcal{A}_{NNM}$ is the kernel function evaluated at the tessellation seed of bin b_i.

The aperture statistics were measured only for equal scale radii $\theta_1 = \theta_2 = \theta_3$, so we abbreviate

$$\langle \mathcal{NNM}_{ap} \rangle(\theta, \theta, \theta) = \langle \mathcal{NNM}_{ap} \rangle(\theta), \tag{43}$$
$$\langle \mathcal{NNM}_{ap} \rangle_{phys}(r, r, r) = \langle \mathcal{NNM}_{ap} \rangle_{phys}(r). \tag{44}$$

We estimated the statistical uncertainty of the aperture statistics in the observational data with jackknife resampling. For this we assumed that the 24 tiles are statistically independent. Although this assumption is not correct for noise due to sample variance, we expect our noise to be dominated by shape noise which is independent for each tile. In the jackknife resampling, we combined the $\tilde{\mathcal{G}}_Z^i$ of all N tiles to the total $\tilde{\mathcal{G}}_Z$, and also create N jackknife samples, for which all but one tile are combined. The aperture statistics $\langle \mathcal{NNM}_{ap} \rangle(\theta)$ are calculated for the

total $\tilde{\mathcal{G}}_Z$, as well as for each of the N jackknife samples to get N $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k(\theta)$. The covariance matrix of $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle(\theta)$ was then estimated with

$$
C_{ij} = \frac{N}{N-1} \sum_{k=1}^{N} \left[\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k(\theta_i) - \overline{\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k}(\theta_i) \right]
$$
$$
\times \left[\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k(\theta_j) - \overline{\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k}(\theta_j) \right], \tag{45}
$$

where $\overline{\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k}(\theta_i)$ is the average of all $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_k(\theta_i)$.

As discussed by Hartlap et al. (2007) and Anderson (2003), the inverse of this estimate of the covariance matrix is not an unbiased estimate of the inverse covariance matrix. Following their suggestion, we instead estimated the inverse covariance matrix with

$$
C_{ij}^{-1} = \frac{N}{N-p-1} \left(C_{ij} \right)^{-1}, \tag{46}
$$

where p is the number of data points. This gives an unbiased estimate of the inverse covariance matrix if the realisations are statistically independent and have Gaussian errors.

With this estimate of the inverse covariance matrix, we calculated the S/N of our observational measurement with

$$
S/N = \left[\sum_{i,j=1}^{p} \langle \mathcal{NN}M_{\mathrm{ap}} \rangle(\theta_i) \, C_{ij}^{-1} \, \langle \mathcal{NN}M_{\mathrm{ap}} \rangle(\theta_j) \right]^{1/2}. \tag{47}
$$

We also use C_{ij}^{-1} to perform a χ^2-test, evaluating the agreement of the observational measurement $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{obs}}$ with the SAMs prediction $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{sim}}$. For this, we calculate the reduced χ^2_{redu} as

$$
\chi^2_{\mathrm{redu}} = \frac{1}{p} \sum_{i,j}^{p} \left(\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{obs}}(\theta_i) - \langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{sim}}(\theta_i) \right)
$$
$$
\times C_{ij}^{-1} \left(\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{obs}}(\theta_j) - \langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{sim}}(\theta_j) \right). \tag{48}
$$

Aside from $\langle \mathcal{NN}M_{\mathrm{ap}} \rangle_{\mathrm{phys}}$, we also estimated the so-called B-mode $\langle \mathcal{NN}M_\perp \rangle$, given by applying Eq. (22) not on $\tilde{\mathcal{G}}_{\mathrm{phys}}$ but on

$$
\tilde{\mathcal{G}}_\perp(B)
$$
$$
= -\frac{\sum_{ijk} w_k \, \epsilon_k^* \, \mathrm{e}^{\mathrm{i}(\varphi_{ik}+\varphi_{jk})} \left[1 + \omega_Z \left(|\boldsymbol{\theta}_i - \boldsymbol{\theta}_j| \right) \right] Z(\Delta z_{ij}) \, \bar{\Sigma}_{\mathrm{crit}}^{-1} \, \Delta_{ijk}^{\mathrm{ph}}(B)}{\sum_{ijk} w_k \, \bar{\Sigma}_{\mathrm{crit}}^{-2} \, Z(\Delta z_{ij}) \, \Delta_{ijk}^{\mathrm{ph}}(B)},
$$
$$
\tag{49}
$$

where ϵ_k^* is the complex conjugate of the galaxies ellipticity. If there are no dominating systematic effects inducing a parity violation, the B-mode has to vanish (Schneider 2003). We therefore tested for such systematics by estimating $\langle \mathcal{NN}M_\perp \rangle$.

4. Data

4.1. Observational data

Our observational data is the overlap of KiDS, VIKING, and GAMA (KV450 × GAMA). This overlap encompasses approximately $180\,\mathrm{deg}^2$, divided into the three patches G9, G12 and G15, each with dimensions of $12 \times 5\,\mathrm{deg}^2$.

VIKING (Edge et al. 2013; Venemans et al. 2015) is a photometric survey in five near-infrared bands, conducted at the VISTA telescope in Paranal, Chile and covering approximately $1350\,\mathrm{deg}^2$. It covers the same area as KiDS (Kuijken et al. 2015; de Jong et al. 2015), an optical photometric survey conducted with the OmegaCAM at the VLT Survey Telescope. The data of KiDS and VIKING were combined to form the KV450 data set, described in detail in Wright et al. (2019), which we use in the following. KV450 has the same footprint as the third data release of KiDS (de Jong et al. 2017) and was processed by the same data reduction pipelines, described in detail in Hildebrandt et al. (2017). Data are processed by THELI (Erben et al. 2005; Schirmer 2013) and Astro-WISE (de Jong et al. 2015). Shears are measured with *lens*fit (Miller et al. 2013; Kannawadi et al. 2019). Photometric redshifts are obtained from PSF-matched photometry (Wright et al. 2019) and calibrated using external overlapping spectroscopic surveys (Hildebrandt et al. 2020).

We use the galaxies observed by KV450 with photometric redshift between 0.5 and 1.2 as source galaxies. Galaxies with a photometric redshift less than 0.5 are excluded because most of them are in front of our lens galaxies and therefore dilute and bias the lensing signal. The averaged inverse critical surface mass density $\bar{\Sigma}_{\mathrm{crit}}^{-1}$ is calculated as described in Sect. 3 by using the weighted direct calibration redshift distributions (DIR distributions) of the KV450 galaxies as the source distribution. These DIR distributions were obtained with in-depth spectroscopic surveys overlapping with KiDS and VIKING. The spectroscopic redshift distributions from these surveys were weighted according to the photometric data in KV450 to estimate the redshift distribution of KV450 galaxies. Details of this procedure are given in Hildebrandt et al. (2017, 2020). We neglect the uncertainties on the redshift distribution and the multiplicative bias of the shear estimate. However, as these uncertainties are small, we do not expect them to impact our conclusions.

GAMA (Driver et al. 2009, 2011; Liske et al. 2015) is a spectroscopic survey carried out at the Anglo Australian Telescope with the AAOmega spectrograph. We use the data management unit (DMU) `distanceFramesv14`, which contains positions and spectroscopic redshifts z of galaxies with a Petrosian observer-frame r-band magnitude brighter than 19.8 mag. The spectroscopic redshifts were flow-corrected to account for the proper motion of the Milky Way using the model by Tonry et al. (2000) according to the procedure in Baldry et al. (2012). We include all galaxies with a spectroscopic redshift lower than 0.5 and redshift quality flag `N_Q`≥ 3. For the calculation of the angular two-point correlation function of lenses, we use randoms from the DMU `randomsv02` (Farrow et al. 2015), which incorporates the galaxy selection function of GAMA while maintaining an unclustered galaxy distribution.

From the GAMA galaxies, we select lens samples according to their colour and stellar mass. Restframe photometry and stellar masses were obtained from the DMU `stellarMassesLambdarv20`. An overview of our samples is given in Table 1.

We select a "red" and "blue" lens sample, defined according to the galaxies' rest-frame $(g-r)_0$ colour. We use the colour cut by Farrow et al. (2015), according to which a galaxy is red if its rest-frame colour $(g-r)_0$ and its absolute Petrosian magnitude M_r in the r-band fulfil

$$
(g-r)_0 + 0.03 \, (M_r - 5 \log_{10} h + 20.6) > 0.6135. \tag{50}
$$

Otherwise, the galaxy is considered blue. This colour cut is chosen to yield approximately equal numbers of red and blue galaxies (93 524 red and 93 702 blue galaxies). Using a hard colour cut does not automatically produce two physically distinct galaxy populations (Taylor et al. 2015). However, as we apply the same

Table 1. Selection criteria for lens samples and number density N of selected galaxies per sample.

Sample	Selection criterion	N (GAMA) [arcmin^{-2}]	N (H15) [arcmin^{-2}]	N (L12) [arcmin^{-2}]
m1	$8.5 < \log_{10}(M^*/M_\odot\, h^{-2}) \leq 9.5$	0.037	0.040	0.059
m2	$9.5 < \log_{10}(M^*/M_\odot\, h^{-2}) \leq 10$	0.058	0.059	0.064
m3	$10 < \log_{10}(M^*/M_\odot\, h^{-2}) \leq 10.5$	0.099	0.096	0.095
m4	$10.5 < \log_{10}(M^*/M_\odot\, h^{-2}) \leq 11$	0.080	0.076	0.058
m5	$11 < \log_{10}(M^*/M_\odot\, h^{-2}) \leq 11.5$	0.014	0.011	0.009
Red	$(g{-}r)_0 + 0.03\,(M_r - 5\log_{10} h + 20.6) > 0.6135$	0.143	0.140	0.152
Blue	$(g{-}r)_0 + 0.03\,(M_r - 5\log_{10} h + 20.6) \leq 0.6135$	0.144	0.142	0.139

Notes. Lenses are selected either according to their stellar mass M^* or to their rest-frame $(g{-}r)_0$ colour and absolute r-band magnitude M_r and need to have $r < 19.8$ mag.

cuts in the observational and simulated data, we expect to obtain comparable "red" and "blue" galaxy samples.

Absolute magnitudes and rest-frame colours of the GAMA galaxies were obtained by Wright et al. (2016) using matched aperture photometry and the LAMBDAR code. These magnitudes were aperture corrected, using

$$M_{r,\text{tot}} = M_{r,\text{meas}} - 2.5 \log_{10} f + 5 \log_{10} h, \qquad (51)$$

where f is the flux scale, which is the ratio between the measured r-band flux and the total r-band flux inferred from fitting a Sérsic-profile to the galaxies photometry.

We define five stellar mass bins with the same cuts as Farrow et al. (2015), with M^* between $10^{8.5}\,h^{-2}\,M_\odot$ and $10^{11.5}\,h^{-2}\,M_\odot$. The stellar masses of GAMA galaxies were obtained by Wright et al. (2017), assuming the initial mass function by Chabrier (2003), stellar population synthesis according to Bruzual & Charlot (2003), and dust extinction according to Calzetti et al. (2000).

The estimator for $\tilde{\mathcal{G}}_Z$ and $\tilde{\mathcal{G}}_{\text{phys}}$ are defined in terms of Cartesian coordinates. Therefore, we project the right ascension α and the declination δ of the galaxies onto a tangential plane on the sky. For this, we divide the source and the lens galaxy catalogues into 24 tiles with a size of $2.5 \times 3\,\text{deg}^2$, which are also used for the jackknife resampling. We use the tile centres (α_0, δ_0) as projection points and find the Cartesian coordinates (x, y) with the orthographic projection

$$x = \cos(\delta) \sin(\alpha - \alpha_0), \qquad (52)$$
$$y = \cos(\delta_0) \sin(\delta) - \sin(\delta_0) \cos(\delta) \cos(\alpha - \alpha_0). \qquad (53)$$

4.2. Simulated data

We compare the results for the aperture statistics in KV450 × GAMA to measurements in the MR with two different SAMs.

The MR (Springel et al. 2005) is a dark-matter-only cosmological N-body-simulation. It traces the evolution of 2160^3 dark matter particles of mass $m = 8.6 \times 10^8\,h^{-1}\,M_\odot$ from redshift $z = 127$ to today in a cubic region with co-moving side length $500\,h^{-1}$ Mpc.

Maps of the gravitational shear γ, caused by the matter distribution in the MR, are created with the multiple-lens-plane ray-tracing algorithm by Hilbert et al. (2009). With this algorithm, we obtain 64 maps of γ on a regular mesh with 4096^2 pixels, corresponding to $4\times4\,\text{deg}^2$ on a set of redshift planes. For each field-of-view, we combine the shear on nine redshift planes between $z = 1.2$ and $z = 0.5$ by averaging it, weighted according to the redshift distribution of the KV450 source galaxies. From this,

we obtain shear maps, which have the same source galaxy distribution as the observational data. We use the DIR redshift distribution, whose creation we described in Sect. 4.1, and do not add any shape noise to the shear.

We obtain simulated lens galaxies from two SAMs implemented in the MR, the SAM by H15 and the SAM by L12. The H15 SAM assumes the same initial mass function by Chabrier (2003) as the observations, but a different stellar population model, that is, the one by Maraston (2005). The L12 SAM uses the initial mass function by Kennicutt (1983) and the stellar population model of Bruzual & Charlot (2003). While the magnitudes of the H15 SAM are given in AB-magnitudes, the magnitudes of the L12 SAM are originally in the Vega magnitude system. We convert the magnitudes to the AB-system with the conversion suggested by Blanton & Roweis (2007),

$$g_{\text{AB}} = g_{\text{Vega}} - 0.08, \qquad (54)$$
$$r_{\text{AB}} = r_{\text{Vega}} + 0.16. \qquad (55)$$

The lens galaxies are selected in the same way as the lenses in GAMA. We use all galaxies with redshifts less than 0.5 and brighter than $r = 19.8$ mag, which is the limiting magnitude of GAMA. With this criterion, we aim to mimic the selection function of GAMA galaxies and expect to obtain samples of similar lenses as in the observation. Systematic errors in the galaxy fluxes, for example, due to the dust modelling of either GAMA or the SAM galaxies, could invalidate this expectation, as different galaxies would be sampled. However, as shown in Fig. 2, the redshift distribution of selected simulated and observed lens galaxies agree well. This likely would not be the case if there were fundamental differences in the selection function for simulated and observed galaxies. The number density of simulated lenses is $0.282\,\text{arcmin}^{-2}$ for the H15 SAM and $0.291\,\text{arcmin}^{-2}$ for the L12 SAM, which are both close to the GAMA number density of $0.287\,\text{arcmin}^{-2}$. Consequently, we expect the lens samples in the simulated and observational data to be comparable.

We split the simulated lens galaxies into colour and stellar-mass samples by applying the same cuts as to the GAMA galaxies (Table 1). Figures 3 and 4 show the colour- and stellar mass distribution of observed and simulated galaxies. The colour distributions of GAMA and SAM galaxies have similar modes. However, the blue mode of the L12 SAM is more concentrated. The H15 SAM also predicts stellar mass distributions similar to the observation, while the L12 SAM predicts more galaxies with stellar masses below $9.5 \times 10^{10}\,M_\odot$ and fewer galaxies with stellar masses above $11 \times 10^{10}\,M_\odot$.
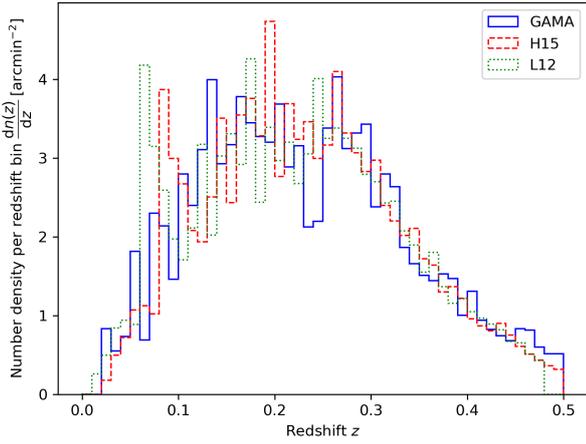
Fig. 2. Number density per redshift bin of GAMA (solid blue) H15 galaxies (dashed red), and L12 galaxies (dotted green) for the limiting magnitude of $r < 19.8$. The bin size is $\Delta z = 0.01$.
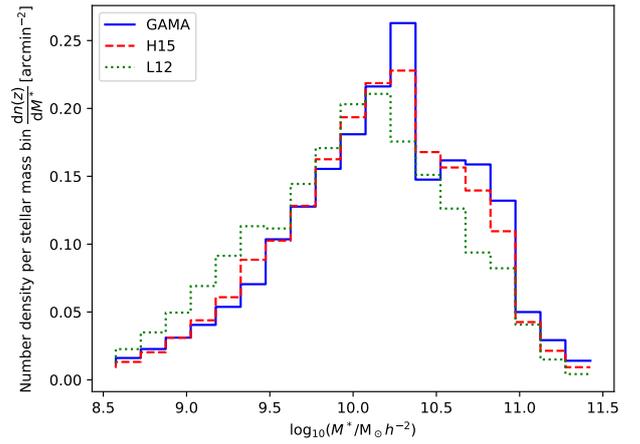
Fig. 4. Number density per stellar mass bin of GAMA (solid blue) H15 (dashed red), and L12 galaxies (dotted green) for the limiting magnitude of $r < 19.8$. The bin size is $\Delta \log(M^*/M_\odot h^{-2}) = 0.15$.
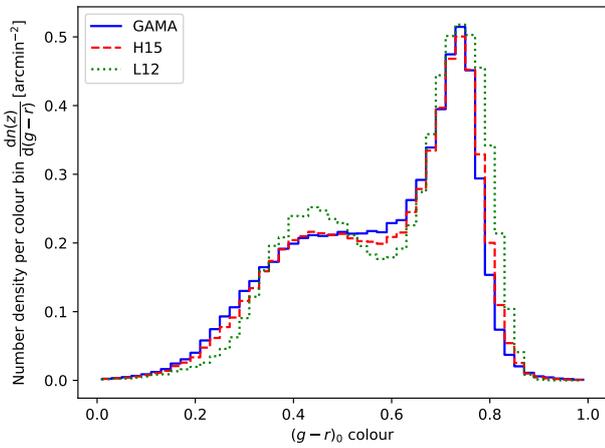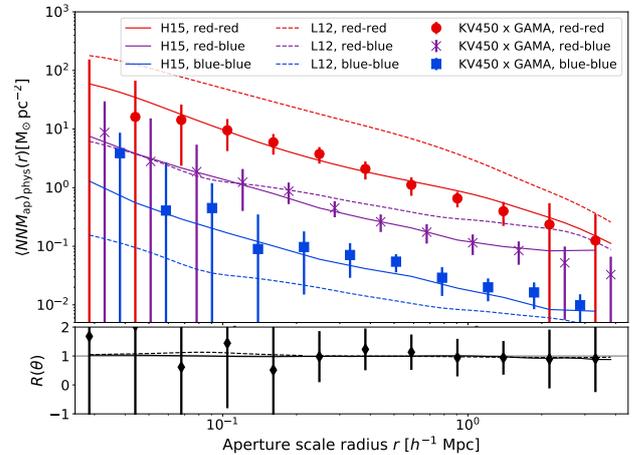

Fig. 3. Number density per colour bin of GAMA (solid blue) H15 (dashed red), and L12 galaxies (dotted green) for the limiting magnitude of $r < 19.8$. The bin size is $\Delta(g-r)_0 = 0.01$.

Fig. 5. *Upper panel*: physical aperture statistics for colour-selected lens samples of the H15 galaxies (solid lines), L12 galaxies (dashed lines) and KV450 × GAMA (points). The signal is shown for red–red lens pairs (red lines and filled circles), red–blue lens pairs (purple lines and crosses), and blue–blue lens pairs (blue lines and squares). Error bars on the observational measurements are the standard deviation from jack-knifing. *Lower panel*: ratio statistics R as given by Eq. (25) for the red and blue lens samples of KV450 × GAMA (points), the H15 SAM (solid line) and the L12 SAM (dashed line).

5. Results

In this section, we present our results for the physical aperture statistics $\langle \mathcal{N}\mathcal{N}M_{\text{ap}}\rangle_{\text{phys}}$, defined in Eq. (18). The measured angular aperture statistics $\langle \mathcal{N}\mathcal{N}M_{\text{ap}}\rangle$, which exhibit similar trends, are given in Appendix A.

The upper plot of Fig. 5 presents $\langle \mathcal{N}\mathcal{N}M_{\text{ap}}\rangle_{\text{phys}}$ for red–red, red–blue, and blue–blue lens pairs. For the observed and both simulated data sets, the signal for red–red lens pairs is larger than for red–blue and blue–blue lens pairs. Consequently, the linear deterministic bias model of Eq. (23) suggests that the bias factor b_{red} of red galaxies is larger than the bias factor b_{blue} of blue galaxies.

The linear deterministic bias model predicts that the aperture statistics for mixed red–blue lens pairs are the geometric mean of the aperture statistics for red–red and blue–blue lens pairs (see Eq. (25)). To test this prediction, we show R in the lower plot of Fig. 5. For the observed galaxies, R is consistent with unity, supporting the linear deterministic bias model. However, for scales below $0.2\,h^{-1}$ Mpc, the noise of the observed R is more than three times larger than R itself, which inhibits any meaningful deductions on the bias model at small scales. For the H15 model, the prediction by the linear bias model is fulfilled , while for the L12 model R is slightly larger than unity at scales below $0.2\,h^{-1}$ Mpc.

The SAMs give different predictions for the aperture statistics. While the $\langle \mathcal{N}\mathcal{N}M_{\text{ap}}\rangle_{\text{phys}}$ of the H15 SAM agrees well with the observations, the signals for red–red and blue–blue lens pairs of the L12 model differ markedly. The L12 SAM predicts much larger aperture statistics for red–red pairs than the observation and significantly smaller aperture statistics for blue–blue pairs. For red–blue pairs, the signal from the L12 SAM is similar to the observed one at small scales, but too high for $r > 0.3\,h^{-1}$ Mpc.

The difference between the SAMs is also visible in Table 2, whose upper part shows the χ^2_{redu} values for the different colour-selected lens pairs. We consider here $p = 12$ data points and define a tension between observation and simulation at the 95% confidence level (CL) if $\chi^2_{\text{redu}} > 1.75$. For the H15 SAM, χ^2_{redu} is smaller than this threshold for red–red, red–blue, and blue–blue lens pairs, so there is no tension between the observation and this model. The χ^2_{redu} for the L12 SAM, though, are notably higher than the threshold. Consequently, the predictions by the L12 SAM do not agree with the observations for these.

Table 2. χ^2_{redu} of $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$ for the H15 and L12 SAMs.

Lens pairs	χ^2_{redu} for H15	χ^2_{redu} for L12
Red–Red	1.08	**55.12**
Red–Blue	0.95	**3.13**
Blue–Blue	1.10	**2.19**
m1–m1	1.44	**32.72**
m1–m2	1.75	**42.96**
m1–m3	1.54	**46.83**
m1–m4	**2.84**	**45.33**
m1–m5	1.75	**64.22**
m2–m2	1.58	**16.04**
m2–m3	0.80	**17.11**
m2–m4	0.97	**10.04**
m2–m5	0.85	**47.10**
m3–m3	1.31	**51.21**
m3–m4	1.17	**41.01**
m3–m5	1.10	**8.60**
m4–m4	1.62	**2.56**
m4–m5	0.97	**8.54**
m5–m5	0.73	**6.93**

Notes. Samples are selected according to Table 1. Bold values indicate a tension at the 95% CL.

Figure 6 shows the measured $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$ for lenses split by their stellar mass. The amplitude of the aperture statistics increases with the stellar mass of galaxies in a pair. Consequently, the bias factor increases with stellar mass. This trend exists for observed and both kinds of simulated lenses. Nevertheless, the predictions of the SAMs differ notably, with the aperture statistics obtained from the L12 SAM being substantially higher than those from the H15 SAM. The L12 SAM also deviates strongly from the observational measurements in KV450×GAMA, that agree better with the H15 SAM. The deviation of the L12 SAM from the observations is strongest for lenses with $M^* \leq 10^{9.5}\,h^{-2}\,M_\odot$ and decreases for larger stellar masses.

To quantify the deviation, we list the χ^2_{redu} of the aperture statistics measured in the H15 and L12 SAM in the lower part of Table 3. Again, a $\chi^2_{\mathrm{redu}} > 1.75$ indicates a tension at the 95% CL. The L12 SAM disagrees with the observation for all lens samples. The χ^2_{redu} of the H15 SAM, though, are smaller than 1.75 for all but one correlation. The only tension exists for the correlation of lenses from stellar-mass samples m1 and m4, driven by differences at $r \lesssim 0.2\,h^{-1}$ Mpc, where the H15 SAM underestimates $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$.

Finally, we test for systematic effects by considering the B-mode $\langle \mathcal{N}\mathcal{N}M_\perp\rangle$. Table 3 compares the S/Ns, defined by Eq. (47), of $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$ with those of the B-modes $\langle \mathcal{N}\mathcal{N}M_\perp\rangle$ for all observed lens pairs. The S/Ns of $\langle \mathcal{N}\mathcal{N}M_\perp\rangle$ are considerably smaller than the S/Ns of the $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$, and they are consistent with a vanishing B-mode.

6. Discussion

We evaluated the SAM by H15 and the SAM by L12 by comparing their G3L predictions to measurements with KiDS, VIKING and GAMA. For this, we applied the improved estimator for the G3L three-point correlation function by L20 and measured aperture statistics for mixed and unmixed lens galaxy pairs from colour- or stellar-mass-selected lens samples.

Our measurements show a higher S/N than previous studies of G3L, due to the use of an improved estimator for G3L and new data. As shown in L20, redshift weighting increases the S/N of the aperture statistics by 35% on mock data with a similar lens- and redshift distribution as in our observation. We also extended the considered scales. Therefore, we could probe the predictions of the SAMs well inside of dark matter halos at lengths below $1\,h^{-1}$ Mpc. These ranges are particularly interesting for testing galaxy formation models because the principal variations between different SAMs are the assumptions on phenomena, whose effects are most substantial at small scales, such as star formation, stellar and AGN feedback and environmental processes (Guo et al. 2016).

The aperture statistics are larger for red–red lens pairs than for red–blue or blue–blue lens pairs, which indicates that red galaxies have higher bias factors than blue galaxies. We also found that the bias factor increases with stellar mass. These results support the general expectation that redder and more massive galaxies have higher bias factors, which has been found in multiple studies (e.g. Zehavi et al. 2002; Sheldon et al. 2004; Simon & Hilbert 2018; Saghiha et al. 2017).

The predictions by the H15 SAM for aperture statistics of colour-selected lens samples agree with the observations at the 95% CL. The signal predicted by the L12 SAM, though, deviates significantly from the observed G3L signal, being too high for red–red and red–blue, and too low for blue–blue pairs.

This deviation could be due to an overproduction of red galaxies in massive halos by the L12 SAM. As shown by Watts & Schneider (2005), the G3L signal increases if more lens pairs reside in massive halos, so the relatively high $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle_{\mathrm{phys}}$ indicates that in the L12 SAM massive halos contain too many pairs of red galaxies. This interpretation is supported by studies by Baldry et al. (2006) of the Bower et al. (2006) SAM, on which the L12 SAM is based. They compare the fraction of red galaxies in the SAM with observations by the SDSS and found that the SAM predicts too many red galaxies, especially in regions of high surface mass density.

Font et al. (2008) accredit the overproduction of red satellite galaxies to excessive tidal interactions and ram pressure stripping in the L12 SAM. This process decreases the amount of gas in satellite galaxies inside halos and thereby inhibits their star formation. Consequently, the stripped galaxies become redder, so the fraction of red galaxies increases, while the number of blue galaxies decreases. This effect could explain the low aperture statistics for blue–blue lens pairs in the Lagos et al. (2012) SAM, as fewer blue galaxies remain inside massive halos.

The aperture statistics for the stellar-mass-selected samples measured in the observation agree with the H15 SAM at the 95% CL except for one sample. This finding is consistent with the conclusion by Saghiha et al. (2017), although their study is limited to angular scales between 1′ and 10′, did not consider mixed lens pairs and had a lower S/N due to the effect of chance lens pairs.

The H15 SAM agrees with the observations at the 95% CL for all but the correlation of m1 and m4 lens galaxies. This difference is driven mainly by a low signal by the SAM at scales below $0.2\,h^{-1}$ Mpc. At these scales, the SAM also gives lower predictions for $\langle \mathcal{N}\mathcal{N}M_{\mathrm{ap}}\rangle$ than the observations for m1–m2, m1–m3, and m2–m2 lens pairs. This trend could indicate that the SAM underpredicts G3L at small scales for low stellar masses. A possible reason is the limited resolution of the MR. The MRs softening length is $5\,h^{-1}$kpc, so its spatial resolution is in the order of tens of kiloparsec (Vogelsberger et al. 2020). Therefore, the difference between the aperture statistics in the H15 SAM
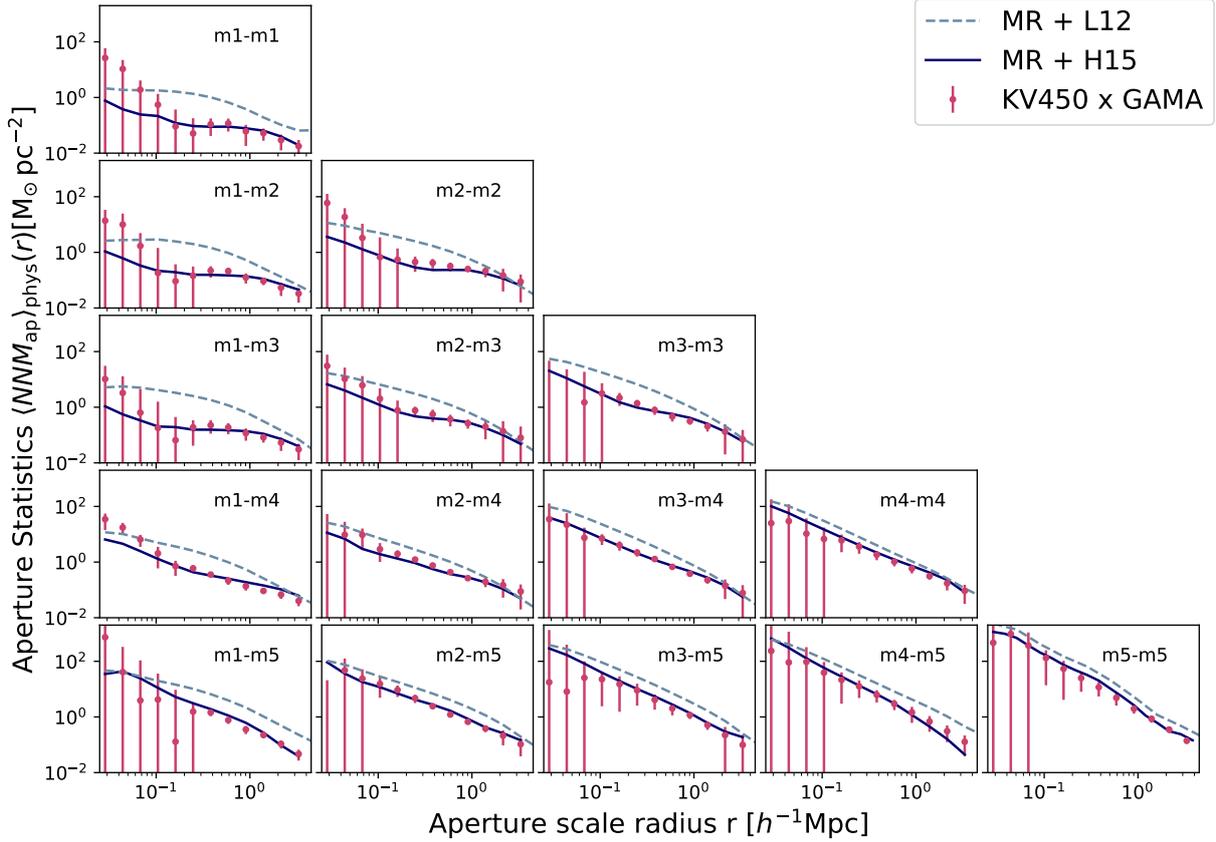
Fig. 6. Physical aperture statistics for stellar mass-selected lens samples in the MR with the H15 SAM (solid blue lines), the L12 SAM (dashed grey lines), and in GAMA with KV450 sources (pink points), using the mass bins defined in Table 1. Plots on the diagonal show the signal for unmixed lens pairs, while the other plots show the signal for mixed lens pairs. Error bars are the standard deviation from jackknife resampling.

and the observation at small scales might be due to the limited resolution.

The L12 SAM disagrees with the observations for all considered stellar-mass samples at the 95% CL, and its predicted signal is significantly larger. The tension increases for lenses with lower stellar mass and is more prominent at smaller scales.

This tension might be due to inaccurate stellar masses of the simulated lens galaxies. If the SAM assigns too low stellar masses, galaxies from a higher stellar mass bin are incorrectly assigned to a lower mass bin, for example into m2 instead of m3. The SAM then overestimates the aperture statistics, because the bias factors of galaxies with larger stellar masses are higher. The choice of initial mass function could cause different stellar-mass assignments by the SAMs. While the H15 SAM used the same initial mass function as the observations (Chabrier 2003), the L12 SAM assumes the initial mass function by Kennicutt (1983). Therefore, the stellar masses of the observation and the L12 might be inconsistent with each other.

Another cause for the tension of the L12 SAM with the observation could be an overproduction of satellite galaxies inside massive halos. This interpretation agrees with Saghiha et al. (2017), who find that the satellite fraction and mean halo masses for the L12 SAM is higher than for the H15 SAM. The tension between the L12 SAM and the observation increases for lower stellar masses and smaller scales, indicating that especially galaxies with low stellar mass are overproduced by the SAM and that their fraction rises closer to the centre of their dark matter halo. An excess of galaxies with small stellar masses would be consistent with excessive galaxy interactions

Table 3. S/N of observed aperture statistics for the *E*-mode in the middle column and the *B*-mode in the right column

Lens pairs	S/N of $\langle \mathcal{NNM}_{ap} \rangle_{phys}$	S/N of $\langle \mathcal{NNM}_\perp \rangle$
Red–red	7.3	1.7
Red–blue	5.1	1.4
Blue–blue	4.7	0.6
m1–m1	3.9	1.7
m1–m2	4.1	1.4
m1–m3	3.7	0.7
m1–m4	5.6	0.8
m1–m5	9.2	0.7
m2–m2	8.6	1.2
m2–m3	3.3	1.8
m2–m4	5.3	1.5
m2–m5	3.4	0.9
m3–m3	4.7	0.8
m3–m4	5.7	0.9
m3–m5	6.2	1.3
m4–m4	7.1	0.7
m4–m5	9.8	1.1
m5–m5	9.1	0.4

Notes. Samples are selected according to Table 1. *B*-mode is consistent with zero.

inside halos. This finding, therefore, fits with the interpretation of the high G3L signal for red–red lens pairs in the SAM as caused by excessive ram pressure stripping.

We presented the first measurements of G3L for mixed lens pairs and used the aperture statistics for red–blue lens pairs to test the linear deterministic bias model. This bias model predicts that the aperture statistics for mixed lens pairs is the geometric mean of the signals for equal lens pairs. Our observational measurements are consistent with this prediction, although the signal is too noisy at scales below $0.2\,h^{-1}$ Mpc for meaningful constraints on the bias model.

The aperture statistics for mixed lens pairs are also useful to constrain the correlations of different galaxy populations inside the same dark matter halos. For example, the measured aperture statistics for red–blue lens pairs indicate that lens galaxies of different samples co-populate the same halos, as the signal would decrease at sub-Mpc scales due to a vanishing 1-halo term. Modelling of mixed-pair G3L in the context of the halo model will provide further insights into the correlation of galaxy populations inside halos. In contrast, GGL, which is only sensitive to the mean number of lenses inside halos and hence blind to the way mixed lens pairs populate halos, cannot yield the same information.

A compelling future study would be investigating whether full hydrodynamical simulations predict G3L with the same accuracy as the H15 SAM. Such a study would complement previous comparisons of GGL in hydrodynamical simulations to observations, for example by Velliscig et al. (2017) for the EAGLE simulation to KiDS and GAMA data, or Gouin et al. (2019) for the Horizon-AGN simulation to CFHTLenS and the Baryon Oscillation Spectroscopic Survey. While these studies conclude that the GGL predictions of these simulations agree with the observations, the same is not necessarily true for G3L, which depends on the correlation of matter and galaxy pairs.

References

Anderson, T. W. 2003, An Introduction to Multivariate Statistical Analysis, 3rd edn. (Hoboken, NJ: Wiley-Interscience)

Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469
Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, MNRAS, 421, 621
Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291
Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Chabrier, G. 2003, PASP, 115, 763
Clampitt, J., Miyatake, H., Jain, B., & Takada, M. 2016, MNRAS, 457, 2391
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, A&A, 582, A62
de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, A&A, 604, A134
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, Astron. Geophys., 50, 12
Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971
Dvornik, A., Hoekstra, H., Kuijken, K., et al. 2018, MNRAS, 479, 1240
Edge, A., Sutherland, W., Kuijken, K., et al. 2013, The Messenger, 154, 32
Epps, S. D., & Hudson, M. J. 2017, MNRAS, 468, 2605
Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, Astron. Nachr., 326, 432
Farrow, D. J., Cole, S., Norberg, P., et al. 2015, MNRAS, 454, 2120
Font, A. S., Bower, R. G., McCarthy, I. G., et al. 2008, MNRAS, 389, 1619
Gouin, C., Gavazzi, R., Pichon, C., et al. 2019, A&A, 626, A72
Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101
Guo, Q., Gonzalez-Perez, V., Guo, Q., et al. 2016, MNRAS, 461, 3457
Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399
Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663
Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2017, MNRAS, 469, 2626
Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, MNRAS, 427, 146
Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31
Hildebrandt, H., Choi, A., Heymans, C., et al. 2016, MNRAS, 463, 635
Hildebrandt, H., Viola, M., Heymans, C., et al. 2017, MNRAS, 465, 1454
Hildebrandt, H., Köhlinger, F., van den Busch, J. L., et al. 2020, A&A, 633, A69
Kaiser, N. 1984, ApJ, 284, L9
Kannawadi, A., Hoekstra, H., Miller, L., et al. 2019, A&A, 624, A92
Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739
Kennicutt, R. C., Jr. 1983, ApJ, 272, 54
Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, MNRAS, 454, 3500
Lagos, C. D. P., Bayet, E., Baugh, C. M., et al. 2012, MNRAS, 426, 2142
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Linke, L., Simon, P., Schneider, P., & Hilbert, S. 2020, A&A, 634, A13
Liske, J., Baldry, I. K., Driver, S. P., et al. 2015, MNRAS, 452, 2087
Maraston, C. 2005, MNRAS, 362, 799
Mead, J. M. G., King, L. J., & McCarthy, I. G. 2010, MNRAS, 401, 2257
Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858
Nelson, D., Springel, V., Pillepich, A., et al. 2019, Comput. Astrophys. Cosmol., 6, 2
Planck Collaboration I. 2020, A&A, in press, https://doi.org/10.1051/0004-6361/201833880
Saghiha, H., Simon, P., Schneider, P., & Hilbert, S. 2017, A&A, 601, A98
Schirmer, M. 2013, ApJS, 209, 21
Schneider, P. 2003, A&A, 408, 829
Schneider, P., & Watts, P. 2005, A&A, 432, 783
Sheldon, E. S., Johnston, D. E., Frieman, J. A., et al. 2004, AJ, 127, 2544
Simon, P., & Hilbert, S. 2018, A&A, 613, A15
Simon, P., Watts, P., Schneider, P., et al. 2008, A&A, 479, 655
Simon, P., Erben, T., Schneider, P., et al. 2013, MNRAS, 430, 2476
Simon, P., Saghiha, H., Hilbert, S., et al. 2019, A&A, 622, A104
Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
Szapudi, I., & Szalay, A. S. 1998, ApJ, 494, L41
Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2015, MNRAS, 446, 2144
Tonry, J. L., Blakeslee, J. P., Ajhar, E. A., & Dressler, A. 2000, ApJ, 530, 625
Velliscig, M., Cacciato, M., Hoekstra, H., et al. 2017, MNRAS, 471, 2856
Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, MNRAS, 453, 2259
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518
Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, Nat. Rev. Phys., 2, 42
Watts, P., & Schneider, P. 2005, IAU Symp., 225, 243
Wright, A. H., Robotham, A. S. G., Bourne, N., et al. 2016, MNRAS, 460, 765
Wright, A. H., Robotham, A. S. G., Driver, S. P., et al. 2017, MNRAS, 470, 283
Wright, A. H., Hildebrandt, H., Kuijken, K., et al. 2019, A&A, 632, A34
Xia, Q., Robertson, N., Heymans, C., et al. 2020, A&A, 633, A89
Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172

Appendix A: Results for aperture statistics in angular units
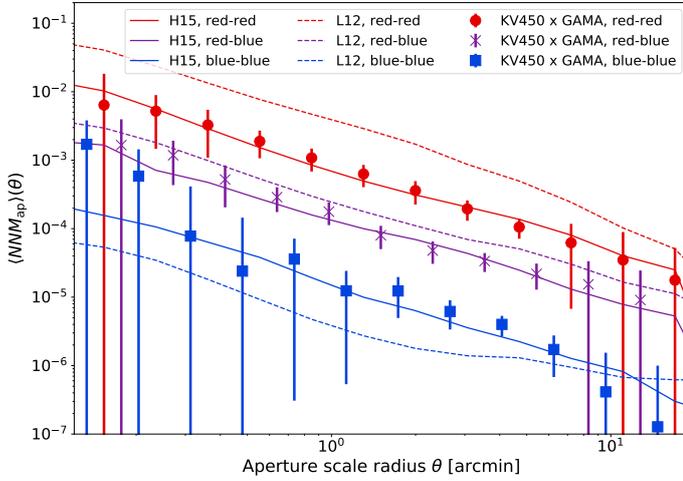
Fig. A.1. Aperture statistics in angular units for colour-selected lens samples of the H15 galaxies (solid lines), L12 galaxies (dashed lines) and KV450 × GAMA (points). The signal is shown for red–red lens pairs (red lines and filled circles), red–blue lens pairs (purple lines and crosses), and blue–blue lens pairs (blue lines and squares). Error bars on the observational measurements are the standard deviation from jackknifing.

For completeness, we show here our results for the angular aperture statistics $\langle\mathcal{NNM}_{\rm ap}\rangle$, for colour-selected lens samples (Fig. A.1) and stellar-mass-selected lens samples (Fig. A.2). The $\langle\mathcal{NNM}_{\rm ap}\rangle$ exhibit similar trends to the $\langle\mathcal{NNM}_{\rm ap}\rangle_{\rm phys}$ (see Sect 5). In particular, $\langle\mathcal{NNM}_{\rm ap}\rangle$ also increases with the lenses stellar masses and is larger for red–red than for red–blue or blue–blue lens galaxies. Furthermore, the predictions by the H15 SAM agrees well with the observed $\langle\mathcal{NNM}_{\rm ap}\rangle$, while the L12 SAM

Table A.1. $\chi^2_{\rm redu}$ of $\langle\mathcal{NNM}_{\rm ap}\rangle$ for H15 and L12 SAMs.

Lens pairs	$\chi^2_{\rm redu}$ for H15	$\chi^2_{\rm redu}$ for L12
Red–Red	1.33	**32.4**
Red–Blue	0.39	**1.92**
Blue–Blue	0.85	**2.31**
m1–m1	0.95	**27.0**
m1–m2	0.81	**28.9**
m1–m3	1.27	**50.3**
m1–m4	**3.69**	**22.13**
m1–m5	1.18	**5.29**
m2–m2	1.29	**10.28**
m2–m3	0.74	**17.13**
m2–m4	0.45	**7.90**
m2–m5	1.37	**21.66**
m3–m3	0.40	**60.61**
m3–m4	0.56	**18.57**
m3–m5	0.90	**27.14**
m4–m4	0.66	**3.15**
m4–m5	1.36	**11.43**

Notes. Bold values indicate a tension at the 95% CL.

expects too large aperture statistics, especially for low stellar-mass galaxies.

The agreement of the H15 SAM and the discrepancy of the L12 SAM with the observations is supported by the $\chi^2_{\rm redu}$ of the SAMs predictions for $\langle\mathcal{NNM}_{\rm ap}\rangle$, presented in Table A.1. The H15 SAM disagrees with the observations only for the correlation of m1 and m4 galaxies at the 95% CL, while the L12 SAM is in tension with the observation for all samples.

Note, that while the measurements of $\langle\mathcal{NNM}_{\rm ap}\rangle$ do not depend on the choice of cosmology, they change with the lens redshift distribution. Comparing $\langle\mathcal{NNM}_{\rm ap}\rangle$ measured in different observational surveys requires, therefore, careful consideration of the survey's selection functions.

Fig. A.2. Angular aperture statistics for stellar mass-selected lens samples in the MR with the H15 SAM (solid blue lines), the L12 SAM (dashed grey lines), and in GAMA with KV450 sources (pink points), using the mass bins defined in Table 1. Plots on the diagonal show the signal for unmixed lens pairs, while the other plots show the signal for mixed lens pairs. Error bars are the standard deviation from jackknife resampling.