Università degli Studi di Studi di Padova

DIPARTIMENTO DI FISICA E ASTRONOMIA

Corso di Laurea Magistrale in Fisica

Tesi di laurea magistrale

Measuring the slope of the intrinsic alignment amplitude as a function of mass using simulation and real data

Candidato:

Davide Piras

Matricola 1114287

Relatore:

Prof. Sabino Matarrese

Supervisor:

Dr Benjamin Joachimi
Dept. of Physics and Astronomy
UCL, United Kingdom

Controrelatore:

Prof.ssa Giulia Rodighiero

Anno Accademico 2016–2017

To all the people who need Science
to become aligned with our Universe.

Contents

Chapter 1

Introduction

In a statistically isotropic universe [1], namely a universe which looks the same in all directions on average, one would expect galaxy images not to have preferred orientations; however, we observe that galaxy shapes are locally correlated with the surrounding large-scale structure, and therefore with each other. Several mechanisms have been proposed to explain this, such as the accretion of new material along favoured directions, and the effect of gravitational tidal fields from the surrounding dark matter distribution. In this latter picture, sketched in Fig. 1.1, the shape of luminous structures is affected by tidal interactions by the hosting dark matter halo, and, consequently, galaxies and clusters align with each other [2].

This phenomenon is known as intrinsic alignment (henceforth IA; see [2, 3, 4, 5] for recent reviews): besides carrying information about galaxy formation, intrinsic alignments have an impact on the measurements of the cosmological weak lensing effect. Galaxy image shapes show in fact induced correlations due to the gravitational lensing, namely the change in the path of light due to the the mass distribution along the line of sight, and IA must therefore be taken into consideration when analysing lensing surveys, such as LSST [6] and Euclid [7]. [8] and [9] first predicted the non-negligible contamination of the weak lensing signal due to correlations in the intrinsic shapes of galaxies, and a number of works afterwards (e.g. [10, 11, 12]) confirmed that this effect must be accounted for in order not to bias the results of the observations.

The intrinsic alignment signal has been measured in N-body simulations [10, 13, 14, 15, 16, 17], in hydrodynamical simulations [18, 19, 20, 21, 22], and in observations [23, 24, 25, 26, 27]. Results from hydrodynamical simulations and observations, in particular, claimed that massive red galaxies point towards matter overdensities, while blue galaxies do not show any clear sign of alignment [28, 29].

The amplitude of the IA signal has been found to increase with mass in observations. For example, a large sample of galaxy clusters from the SDSS DR7 was studied in [24], finding a dependence of the alignment with the mass of the brightest cluster galaxy; moreover, an increasing IA amplitude with luminosity, or the corresponding mass, has been identified for luminous red galaxies (LRGs) [23, 26] and for galaxy clusters [27].

Simulations agree with this picture: using a 512^3-particle N-body simulation, [30], for example, found that the alignment increases with mass over three orders of magnitude, up to about $10^{13} M_{\odot}$. In addition, [14] used data from the Millennium

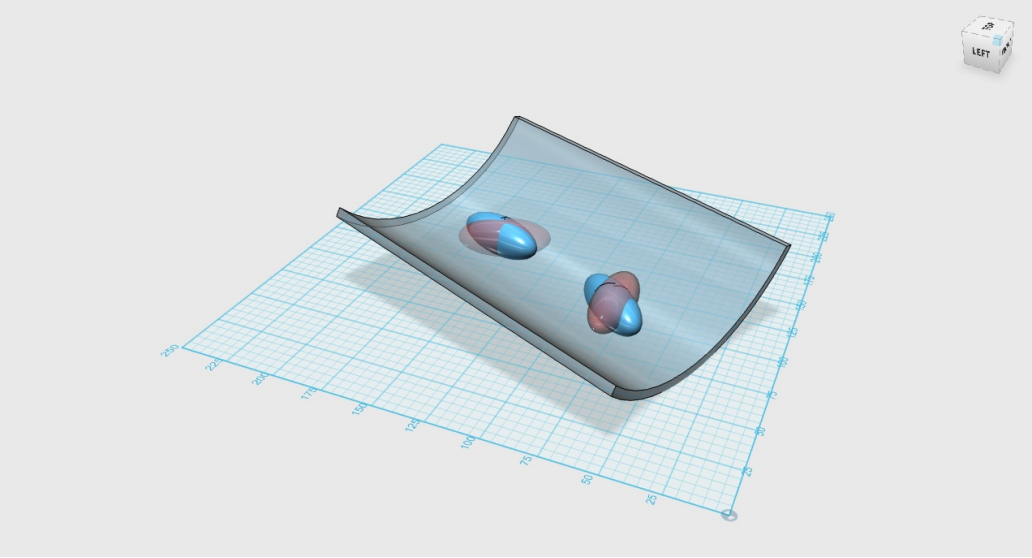


Figure 1.1: Sketch of the intrinsic alignment of luminous matter caused by tidal gravitational forces, taken from [2]. In this picture, elliptical galaxies experience correlated tidal shears and react to the tidal fields by changing their initially uncorrelated shapes (red), to correlated shapes (blue), because of the surrounding gravitational potential of the halo (grey sheet). In this work, we study the intrinsic alignment of dark matter haloes due to the presence of other dark matter haloes, which are used as tracers of the surrounding density field.

simulation, and claimed stronger correlations with higher mass over two mass bins around $10^{12} M_{\odot}$, while [17], using the same set of data, found an increasing trend over the same mass range only for early-type galaxies (whose shapes were assumed to follow those of the underlying haloes), while no dependence on luminosity or particular trend with mass or luminosity for late-type galaxies (whose orientation was determined by the halo spin, and which roughly correspond to the blue galaxies mentioned above) was identified.

These results motivate us to search for a universal relation between the alignment strength and the mass of dark matter haloes. This could lend support to an IA mechanism that successfully explains the trends; moreover, we need to reduce the degrees of freedom in modelling the IA signal to obtain tighter cosmology constraints from lensing observations.

To achieve this, we delve into the dependence of the amplitude of the intrinsic alignment signal on the mass of the halo, which is postulated to be the main driver of the shape and orientation of the hosted galaxies (according to the so-called halo model, see e.g. [4]). We first derive the expected scaling from the theory, and then test our predictions through a Bayesian analysis using data from two N-body simulations, mimicking the observational approach in order to be able to straightforwardly compare our results with real data.

We give details about our theoretical model and derive the expected scaling of the IA with halo mass in the tidal alignment paradigm in Chapter **??**. We then present the simulations that we use (Sect. 3.1), and how we define the shapes of the

dark matter haloes they contain (Sect. 3.2). In Sect. 4.1 we explain how we measure the intrinsic alignment signal, and then (Sect. 4.2) we describe our mass-dependent IA model. We finally show our results and compare them with our theoretical predictions (Sect. 5.1) and real data (Sect. 5.2).

Chapter 2

Theoretical background

In a self-gravitating system that is in virial equilibrium with a velocity dispersion σ^2, such as an elliptical galaxy or a cluster of galaxies, tidal gravitational forces from the surrounding large-scale structure distort the system's gravitational potential. The particles of the system remain in virial equilibrium and fill up the distorted potential along an isocontour of the gravitational potential, which results in a change in the shape of the system. This shape modification follows the orientation and the magnitude of the tidal gravitational fields, and the magnitude of the change in shape depends on how tightly the system is bound.

In isolated virialised systems the Jeans-equation holds [31]:

$$\frac{1}{\varrho}\frac{\partial}{\partial r}(\varrho\sigma^2) + \frac{2}{r}\beta_{\mathrm{aniso}}\sigma^2 = -\frac{\partial\Phi}{\partial r}\ , \tag{2.1}$$

with ϱ the particle density, r the distance, Φ the gravitational potential and β_{aniso} the anisotropy parameter, which we set to $\beta_{\mathrm{aniso}} = 0$ because our first goal is to derive the scaling behaviour, and not the orientation, of the alignment amplitude. In the case of vanishing anisotropy and a constant velocity dispersion, the Jeans-equation can be solved to yield an exponential dependence $\varrho \propto \exp(-\Phi/\sigma^2)$ between the density and the gravitational potential.

We work in the limit that the gravitational potential distortion is well-described by a second-order Taylor-expansion relative to the centre of the galaxy at $\boldsymbol{r}_0 = 0$. Therefore, the potential Φ relevant for the motion of galaxies is given by

$$\Phi(\boldsymbol{r}) \to \Phi(\boldsymbol{r}) + \frac{1}{2}\sum_{i,j=1}^{3}\frac{\partial^2\Phi(\boldsymbol{r}_0)}{\partial r_i \partial r_j} r_i r_j. \tag{2.2}$$

with i, j indicating the spatial dimensions. Consequently, the density of particles changes according to

$$\varrho \propto \exp\left(-\frac{\Phi(\boldsymbol{r})}{\sigma^2}\right) \times \left(1 - \frac{1}{2\sigma^2}\sum_{i,j=1}^{3}\frac{\partial^2\Phi(\boldsymbol{r}_0)}{\partial r_i \partial r_j} r_i r_j\right), \tag{2.3}$$

under the assumption of a weak distortion, so that a Taylor-expansion of the exponential to first order is sufficient.

The projected ellipticity is calculated through the tensor of second brightness moments [32], which we define using the density ϱ:

$$q_{ij} = \int \mathrm{d}A \; \varrho(x) x_i x_j \; , \tag{2.4}$$

where the integral is calculated over the plane of the projected sky, with coordinates x_i, x_j and $i, j \in \{1, 2\}$. If we consider the distortion of the density described in Eq. 2.3, we obtain:

$$\begin{aligned} \widetilde{q_{ij}} = \int \mathrm{d}A \; \varrho(x) \left(1 - \frac{\partial^2_{a,b}\Phi(x)}{2\sigma^2} x_a x_b\right) x_i x_j \simeq \\ \simeq \int \mathrm{d}A \; \varrho(x) x_i x_j - \frac{\partial^2_{a,b}\Phi}{2\sigma^2} \int \mathrm{d}A \; \varrho(x) x_a x_b x_i x_j = q_{ij} + \psi_{ij} \; , \end{aligned} \tag{2.5}$$

where the equivalence is not exact due to the approximation that $\partial^2_{a,b}\Phi(x)/\sigma^2$ is constant and can be thus taken out of the integral. From this latter equation, we can see that because of the tidal shear fields the second moments of the brightness distribution get a correction ψ_{ij}: this term is small in the limit of weak tidal fields, which is characterised by $R^2\partial^2\Phi/\sigma^2 \ll 1$ with R being the size of the halo and providing a bound for r.

If we then define the complex ellipticity[1] as in [32]:

$$\varepsilon = \frac{q_{11} - q_{22}}{q_{11} + q_{22}} + 2i \frac{q_{12}}{q_{11} + q_{22}} \; , \tag{2.6}$$

and consider the correction obtained in Eq. 2.5, we can write:

$$\begin{aligned} \widetilde{\varepsilon} = \frac{(q_{11} + \psi_{11}) - (q_{22} + \psi_{22})}{(q_{11} + \psi_{11}) + (q_{22} + \psi_{22})} + 2i \frac{q_{12} + \psi_{12}}{(q_{11} + \psi_{11}) + (q_{22} + \psi_{22})} \simeq \\ \simeq \frac{q_{11} - q_{22}}{q_{11} + q_{22}} + 2i \frac{q_{12}}{q_{11} + q_{22}} + \frac{\psi_{11} - \psi_{22}}{q_{11} + q_{22}} + 2i \frac{\psi_{12}}{q_{11} + q_{22}} = \varepsilon + \psi_\epsilon \; , \end{aligned} \tag{2.7}$$

where we assumed that ψ_{ij} is a small correction. We thus showed that a measurement of the projected ellipticity of a tidally distorted halo through the second moments of ϱ yields a correction which is proportional to $R^2\partial^2\Phi/\sigma^2$, in accordance with the linear alignment model for elliptical galaxies, where the tidally induced ellipticity is proportional to the magnitude of the tidal fields [33, 34, 35].

In the case of virialised systems it is possible to relate the velocity dispersion σ^2 with the size of the object R: the virial relationship $\sigma^2 = GM/R$ assumes a proportionality between the specific kinetic energy, $\sigma^2/2$, and the magnitude of the specific potential energy, GM/R. With the scaling $M \propto R^3$ one expects $\sigma^2 \propto M^{2/3}$, such that R^2/σ^2 is constant. Therefore, any scaling of the ellipticity with mass is entirely due to the dependence of tidal gravitational fields with the mass-scale, and more massive systems are subjected to stronger tidal interactions because of the stronger fluctuations in the tidal fields that they are experiencing.

[1] This definition follows the gravitational lensing convention, and slightly differs from the one that we use in our analysis (Eq. 3.6 and Eq. 3.7); nonetheless, the calculation can be easily repeated with our definition, without altering the conclusion.

The variance of tidal shear fields can be inferred from the variance of the matter density by the Poisson equation, $\Delta\Phi = 3\Omega_\mathrm{m}/(2\chi_H^2)\delta$, with Ω_m the matter density, $\chi_H = c/H_0$ the Hubble-distance, and δ the density field. Computing the tidal shear fields $\partial^2\Phi$ shows that they must have the same fluctuation statistics as δ: in Fourier-space, the solution to the Poisson equation is $\Phi \propto \delta/k^2$ with the wave vector k, and the tidal shear fields become $k_i k_j \Phi \propto k_i k_j/k^2\,\delta$. Therefore, the power spectrum $P_{\partial^2\Phi}(k)$ of the tidal shear fields is proportional to the power spectrum $P_\delta(k)$ of the density fluctuations.

Consequently, one can derive the variance of the tidal shear fields from the variance of the density fluctuations, i.e. from the standard cold dark matter (CDM) power spectrum $P_\delta(k)$. For doing that, one can relate a mass scale M to the wave vector k by requiring that the mass M should be contained in a sphere of radius R,

$$M = 4\pi\Delta\varrho_\mathrm{crit}R^3/3 = 32\pi^4\Delta\varrho_\mathrm{crit}/3k^3 \tag{2.8}$$

with $k = 2\pi/R$, $\varrho_\mathrm{crit} = 3H_0^2/8\pi G$ the critical density and Δ an overdensity factor[2], which is typically 200 [2]. This defines a scale k in the power spectrum which is proportional to $M^{-1/3}$. This implies that on galaxy and cluster scales, where the CDM-spectrum scales $\propto k^\gamma$, with $\gamma \in [-3,-2]$, one obtains for the standard deviation of the tidal shear field a behaviour $\propto M^{\beta_\mathrm{M}}$, with $\beta_\mathrm{M} \in [1/3, 1/2]$; in particular, for cluster-size objects ($R = 0.5$ Mpc–1.5 Mpc), where the non-linear matter power spectrum is proportional to $k^{-2.2}$ [36], our prediction is $\beta_\mathrm{M} \simeq 0.36$.

We also take a step further: given the premises above, for the amplitude of the intrinsic alignment A_IA we can write:

$$A_\mathrm{IA} \propto \partial^2\Phi \propto \sqrt{P_{\partial^2\Phi}(k)} \propto \sqrt{P_\delta(k)} \propto \sqrt{P_\delta(M)}\ , \tag{2.9}$$

where in the last step we assume that Eq. 2.8 holds and provides the link between the wave vector k and the mass M. This approach overcomes the previous approximation that the power spectrum follows a power-law, and, consequently, it does not provide any predictions for β_M; we therefore investigate its goodness with a straightforward χ^2 analysis, described in Sect. 4.2. We then compare these very general theoretical predictions with our results in Chapter 5.

[2]Here, the overdensity is relative to the critical density, but hereinafter we sometimes consider overdensities relative to the underlying matter density. In those cases, we just rescale the mass by a factor of Ω_m, which is the ratio between the mean matter density and the critical density.

Chapter 3

Data

3.1 Simulations

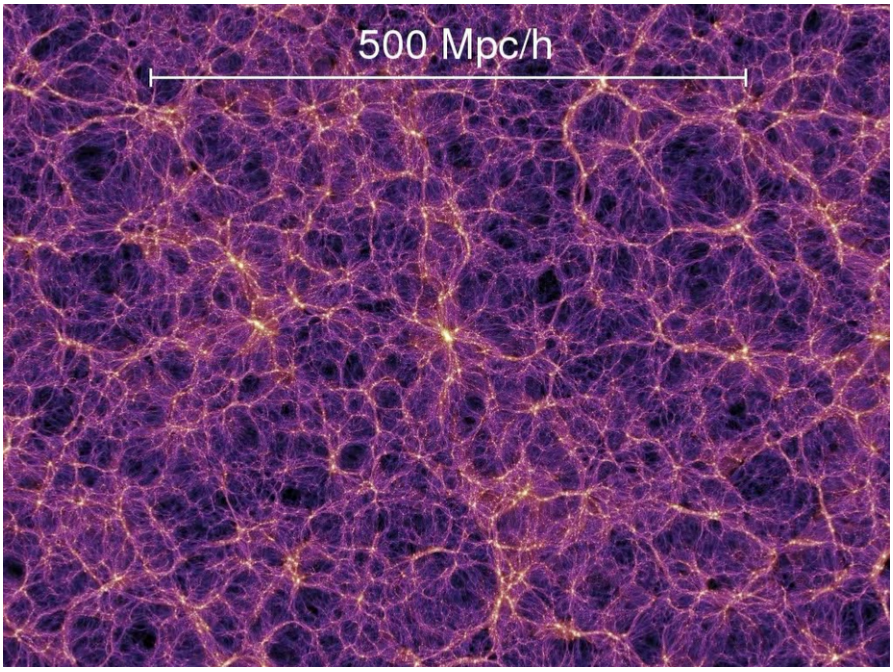
In this work we consider haloes from two different simulations:

1. the **Millennium simulation (MS)**, first presented in [37], which uses 2160^3 dark matter particles of mass $m_{\rm P}^{\rm MS} = 8.6 \times 10^8 M_\odot/h$ enclosed in a 500 Mpc/h-side box to sample dark matter haloes and study their growth. In particular, we consider 2 of its 64 snapshots, i.e. the one at $z = 0$ (snapshot 63, illustrated in Fig. 3.1), for our baseline, and the one at $z \simeq 0.46$ (snapshot 49), to study a potential redshift dependence of our results. Dark matter haloes are identified as in [16] and references therein: a simple "friends-of-friends" group-finder (FOF, [38]) is run first to spot virialised structures, followed by the `SUBFIND` algorithm [39, 37] to identify sub-haloes, some of which are then treated as separate haloes if they are only temporarily close to the halo. We consider haloes with a minimum number of particles $N_{\rm P} = 300$.

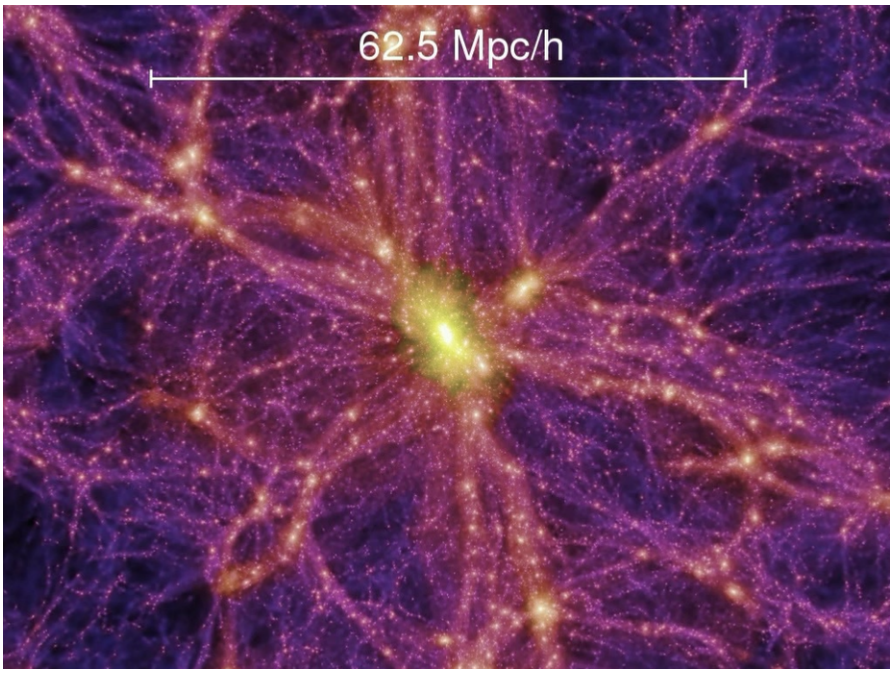
2. The **Millennium-XXL simulation (MXXL)**, which samples 6720^3 dark matter particles of mass $m_{\rm P}^{\rm MXXL} = 6.174 \times 10^9 M_\odot/h$ confined in a cubic region of 3000 Mpc/h on a side [40]. In this case, we consider only one snapshot, at $z = 0$. Haloes are selected using an ellipsoidal overdensity algorithm, as described in [41] and [42]: a traditional spherical overdensity algorithm [43] gives an initial estimate of the true shape and orientation of the halo, which is then improved by building up an ellipsoid using the previously selected particles.

We define the mass of the objects in the catalogues as the mass within a region where the density is 200 times the critical density at the redshift corresponding to the respective snapshot ($M_{\rm 200c}$). Note that for the MS we first convert the halo mass from $M_{\rm Dhalo}$, as defined in [44], to $M_{\rm 200c}$ using the median line in [44, figure 2]; this transformation is necessary in order to have a consistent definition of the mass of the haloes in the simulations, but its impact on our results is negligible. The number density distribution of the haloes used in our analysis is shown in Fig. 3.2.

In the simulations the same set of cosmological parameters is adopted, namely they both assume a spatially flat ΛCDM universe with the total matter density $\Omega_{\rm m} = \Omega_{\rm b} + \Omega_{\rm dm} = 0.25$, where $\Omega_{\rm b} = 0.045$ indicates the baryon density parameter



(a) The whole Millennium simulation.



(b) Detail of the Millennium simulation.

Figure 3.1: The dark matter distribution in the universe at the present time, based on the Millennium simulation, which studies more than 10^{10} dark matter particles. From these pictures, we can appreciate the morphology of the structures on different scales, and the large dynamic range of the simulation. In (a) the whole simulation box is shown to highlight the filaments and void distribution, while in (b) details of the central structures are zoomed in. Credit: [37].

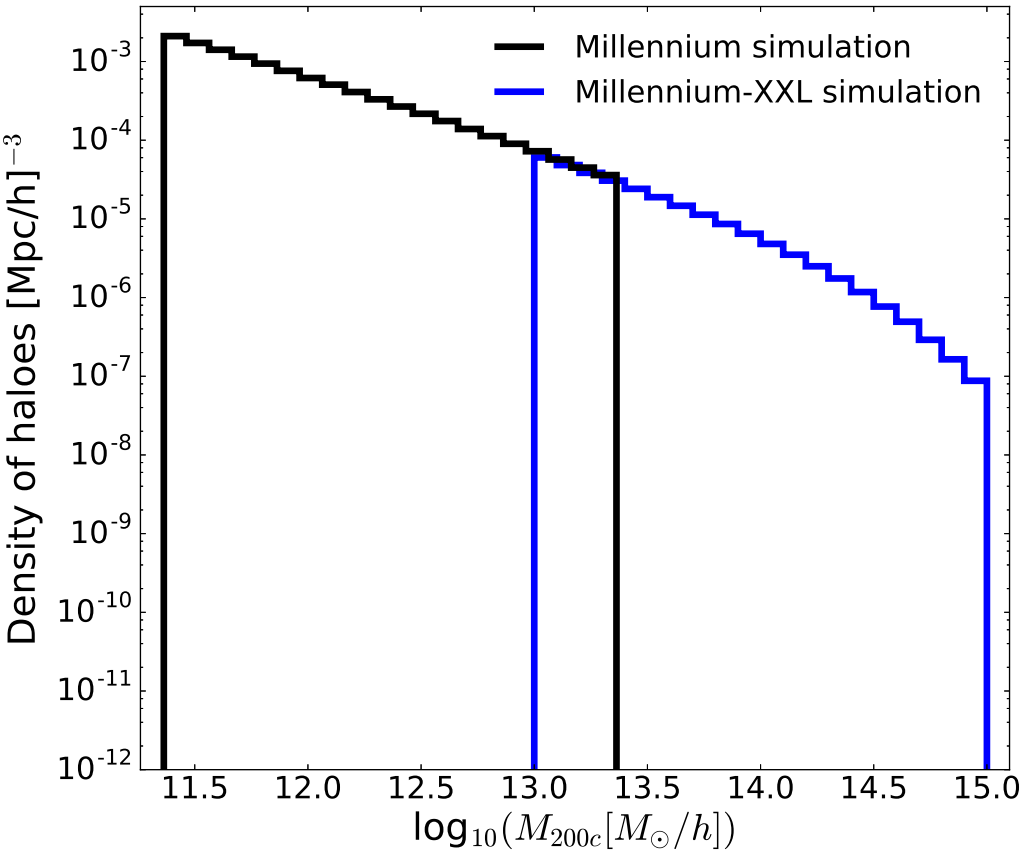


Figure 3.2: Histogram showing the number density distribution of the mass of the haloes in the two simulations. Each logarithmic mass bin has a size of 0.1 dex. In this case, the mass of the halo is defined as M_{200c}, the mass within a region where the density exceeds 200 times the critical density; more details about the chosen mass ranges are presented in Sect. 4.2. In the range where the selected bins overlap, the trend of the density agrees for the two simulations.

and $\Omega_{\rm dm} = 0.205$ represents the dark matter density parameter, a cosmological constant $\Omega_\Lambda = 1 - \Omega_{\rm m} = 0.75$, the dimensionless Hubble parameter $h = 0.73$, the scalar spectral index $n_{\rm s} = 1$, and the density variance in spheres of radius 8 Mpc/h, $\sigma_8 = 0.9$.

3.2 Halo shapes

We define the simple inertia tensor[1], whose eigenvalues and eigenvectors describe the shape of the halo, as:

$$\boldsymbol{M}_{\mu\nu} \propto \sum_{i=1}^{N_{\rm P}} r_{i,\mu} r_{i,\nu} \ , \tag{3.1}$$

where $N_{\rm P}$ is the total number of particles within the halo, $\mu, \nu \in \{1, 2, 3\}$, and $\boldsymbol{r}_i$ is the vector that indicates the position of the $i-$th particle with respect to the centre of the halo, i.e. the location of the gravitational potential minimum, in the reference frame of the simulation box. For the MS only, we also consider a reduced inertia tensor, which is defined as [46]:

$$\boldsymbol{M}^{\rm red}_{\mu\nu} \propto \sum_{i=1}^{N_{\rm P}} \frac{r_{i,\mu} r_{i,\nu}}{r_i^2} \ , \tag{3.2}$$

with r_i^2 the square of the three-dimensional distance of the i-th particle from the centre of the halo. The reduced inertia tensor is more weighted towards the centre of the halo, and may yield a more reliable approximation of the shape of the galaxy at its centre [17, 20].

The eigenvectors and eigenvalues define an ellipsoid, which we project onto one of the faces of the simulation box along the z-axis: the resulting ellipse is the projected shape of the halo. We proceed as in [16] to define the ellipticity ϵ of the objects, adopting their procedure for early-type galaxies. We denote the eigenvalues of the inertia tensor as ω_μ, and the three eigenvectors as $\boldsymbol{s}_\mu = \left\{s_{x,\mu}, s_{y,\mu}, s_{z,\mu}\right\}^\intercal$, $\mu \in \{1, 2, 3\}$, where the coordinates refer to the Cartesian system of the simulation box. The projected ellipse is given by the points $\boldsymbol{x}$ which satisfy $\boldsymbol{x}^\intercal \boldsymbol{W}^{-1} \boldsymbol{x} = 1$, defining a symmetric tensor

$$\boldsymbol{W}^{-1} = \sum_{\mu=1}^{3} \frac{\boldsymbol{s}_{\perp,\mu} \boldsymbol{s}^\intercal_{\perp,\mu}}{\omega_\mu^2} - \frac{\boldsymbol{\kappa}\boldsymbol{\kappa}^\intercal}{\alpha^2} \ , \tag{3.3}$$

with $\boldsymbol{s}_{\perp,\mu} = \left\{s_{x,\mu}, s_{y,\mu}\right\}^\intercal$,

$$\boldsymbol{\kappa} = \sum_{\mu=1}^{3} \frac{s_{z,\mu} \boldsymbol{s}_{\perp,\mu}}{\omega_\mu^2} \ , \tag{3.4}$$

and

$$\alpha^2 = \sum_{\mu=1}^{3} \left(\frac{s_{z,\mu}}{\omega_\mu}\right)^2 \ . \tag{3.5}$$

[1]MS and MXXL use two different tensor definitions to describe the shape, but they result in the same halo ellipticity (see also [45] for further details).

We compute the two Cartesian components of the ellipticity [32]:

$$\epsilon_1 = \frac{W_{11} - W_{22}}{W_{11} + W_{22} + 2\sqrt{\det \boldsymbol{W}}} \ , \tag{3.6}$$

$$\epsilon_2 = \frac{2W_{12}}{W_{11} + W_{22} + 2\sqrt{\det \boldsymbol{W}}} \ , \tag{3.7}$$

which we then translate in the radial (+) component, following the IA sign convention[2]:

$$\epsilon_+ = \epsilon_1 \cos(2\varphi) + \epsilon_2 \sin(2\varphi) \ , \tag{3.8}$$

where φ is the polar angle of the line that connects a halo pair with respect to the x-axis, in the reference frame of the simulation box. We show how we use ϵ_+ to measure the correlation between halo shapes in the next section.

As a sanity check, we also calculate the intrinsic ellipticity dispersion, namely $\sigma_\epsilon = \sqrt{\sum_{i=1}^{N_\mathrm{h}} (\epsilon_1^2 + \epsilon_2^2)/N_\mathrm{h}}$ for several mass bins, with N_h the number of haloes (these details about the bins are reported in Sect. 4.2 and in Table 4.1). The magnitude of the dispersion is in good agreement with that of early-type galaxies [16] and clusters (0.13 to 0.19 for the cluster samples used in [27]). The dispersion increases with mass, as found in other simulations (e.g. [47, 48]), with a small excess for the MXXL samples due to the lower sampling of the halo shapes.

[2]Our definition of ϵ_+, as commonly done in IA works, has an opposite sign with respect to, for example, [32].

Chapter 4

Methodology

4.1 Measurements

To measure the alignment of every dark matter halo with other haloes, we mimic the standard analysis that is usually performed with observations (e.g. [27]), i.e. we use the halo catalogue as the tracer of the underlying density field. Instead of fitting physical models to the alignment signal and dividing out the galaxy/cluster bias dependence afterwards, though, we take a shortcut to remove the bias factor dependence, which is described in the next paragraphs .

We first define an estimator as a function of the comoving transverse separation R_p and the line-of-sight distance Π:

$$\hat{\xi}_{\mathrm{g}+}(R_\mathrm{p}, \Pi) = \frac{S_+ D}{DD} \ , \tag{4.1}$$

where S_+D represents the raw correlation between halo shapes (ϵ_+) and the density sample, and DD the number of halo shape-density pairs. The halo shape sample is always used also as density sample. We then integrate along the line of sight to obtain the total projected intrinsic alignment signal:

$$\hat{w}_{\mathrm{g}+}(R_\mathrm{p}) = \int_{-\Pi_\mathrm{max}}^{\Pi_\mathrm{max}} \mathrm{d}\Pi \ \hat{\xi}_{\mathrm{g}+}(R_\mathrm{p}, \Pi) \ . \tag{4.2}$$

Throughout this work, we adopt $\Pi_\mathrm{max} = 60$ Mpc$/h$, a value large enough not to miss part of the signal, but small enough not to pick up too much noise. We describe the intrinsic alignment signal by simplifying the model in Eq. 5 of [27], namely we assume:

$$w_{\mathrm{g}+}(R_\mathrm{p}, M) = A_\mathrm{IA}(M) \ b_\mathrm{h}(M) \ w_{\delta+}^\mathrm{model}(R_\mathrm{p}) \ , \tag{4.3}$$

with $A_\mathrm{IA}(M)$ the amplitude of the intrinsic alignment signal, $b_\mathrm{h}(M)$ the halo bias, and $w_{\delta+}^\mathrm{model}(R_\mathrm{p})$ a function in which we include the dependence on R_p. In the tidal alignment paradigm, $w_{\delta+}^\mathrm{model}$ is independent of halo mass, since it is fully determined by the properties of the dark matter distribution, assuming that any mass dependence of the response of a halo shape to the tidal gravitational field is captured by $A_\mathrm{IA}(M)$.

We evaluate the expression in Eq. 4.3 in the interval which covers 10 Mpc$/h <$ $R_\mathrm{p} < 20$ Mpc$/h$, denoted by R_p^*, to remove the dependence on R_p; in other words, we define the halo-pair weighted average in R_p^* as:

$$w_{\mathrm{g}+}(M) \equiv \langle w_{\mathrm{g}+}(R_\mathrm{p}, M) \rangle_{R_\mathrm{p}^*} \ . \tag{4.4}$$

We report a detailed discussion about the choice of this interval in Appendix A.

In order to constrain $b_{\rm h}(M)$ in Eq. 4.3, we use the LS estimator [49] to calculate the clustering signal:

$$\hat{\xi}_{\rm gg}(R_{\rm p},\Pi) = \frac{DD - 2DR + RR}{RR} \ , \tag{4.5}$$

where DD represents the number of halo pairs, DR the number of halo-random point pairs, and RR the number of random point pairs. To measure DR and RR, we generate random catalogues that contain objects uniformly distributed between the minimum and maximum value of the x, y and z coordinates of each sub-box[1]. These catalogues normally are three times denser, but in some cases, when a sub-box encloses very few objects, we switch to random catalogues which are ten times denser; we always use Eq. 4.5 since we re-normalise the estimators according to the sample size. The size of these random catalogues is further discussed in Appendix B.

We then integrate along the line of sight to obtain the total projected clustering signal:

$$\hat{w}_{\rm gg}(R_{\rm p}) = \int_{-\Pi_{\rm max}}^{\Pi_{\rm max}} {\rm d}\Pi \ \hat{\xi}_{\rm gg}(R_{\rm p},\Pi) \ . \tag{4.6}$$

and describe this with a simple model:

$$w_{\rm gg}(R_{\rm p},M) = b_{\rm h}^2(M) \ w_{\delta\delta}^{\rm model}(R_{\rm p}) + C_{\rm IC} \ , \tag{4.7}$$

with $w_{\delta\delta}^{\rm model}(R_{\rm p})$ a function in which we include the dependence on $R_{\rm p}$ [27, equation 9], and $C_{\rm IC}$ the integral constraint, which accounts for the bias in the observed clustering signal that is caused by the use of a finite survey area. We report all the details about how the integral constraint is calculated in Appendix B. Note that, since we are dealing with dark matter clustering, $w_{\delta\delta}^{\rm model}(R_{\rm p})$ does not depend on the mass of the halo. Again, we average the previous expression in $R_{\rm p}^*$, obtaining:

$$w_{\rm gg}(M) \equiv \langle w_{\rm gg}(R_{\rm p},M)\rangle_{R_{\rm p}^*} \ . \tag{4.8}$$

In observational analyses, the bias factor is usually measured from a fit to $w_{\rm gg}$ and then included in the modelling of $w_{\rm g+}$; equivalently, to remove the mass dependence of the halo bias $b_{\rm h}(M)$, we define:

$$r_{\rm g+}(M) = \frac{w_{\rm g+}(M)}{\sqrt{w_{\rm gg}(M)}} \simeq \frac{A_{\rm IA}(M)\langle w_{\delta+}^{\rm model}(R_{\rm p})\rangle_{R_{\rm p}^*}}{\sqrt{\langle w_{\delta\delta}^{\rm model}(R_{\rm p})\rangle_{R_{\rm p}^*}}} \propto A_{\rm IA}(M) \ , \tag{4.9}$$

where we assume that the clustering signal $w_{\rm gg}(M)$ is positive (see Sect. 4.2 and Chapter 5 for further discussion). The equivalence is not exact due to the small correction of the integral constraint $C_{\rm IC}$ in the clustering signal. We stress that, under our assumptions, this quantity depends only on the mass of the halo M.

[1]For further information about the sub-boxes, see Sect. 4.2.

4.2 Modelling

The goal of this work is to study the halo mass dependence of the intrinsic alignment amplitude $A_{\rm IA}(M)$, which we constrain by fitting $r_{\rm g+}(M)$. To achieve this goal, we select the haloes from the catalogues described in Sect. 3.1 in $n_{\rm M} = 4$ logarithmic mass bins between $10^{11.36} M_\odot/h$ and $10^{13.36} M_\odot/h$ for the MS and between $10^{13} M_\odot/h$ and $10^{15} M_\odot/h$ for the MXXL (see Table 4.1). We split them in $N_{\rm sub} = 3^3 = 27$ sub-boxes based on their positions inside the cube of the respective simulation[2], and calculate $w_{\rm g+}$ and $w_{\rm gg}$ for each sub-sample by replacing the integrals in Eq. 4.2 and 4.6 with a sum over 20 line-of-sight bins, each $2\Pi_{\rm max}/20 = 6$ Mpc$/h$ wide, defining the line of sight along the z-axis.

As a first approach, we adopt the following model for $r_{\rm g+}(M)$:

$$r_{\rm g+}(M) = A \cdot \left(\frac{M}{M_{\rm p}}\right)^{\beta_{\rm M}} , \tag{4.10}$$

with A a generic amplitude which we will treat as a nuisance parameter, $M_{\rm p} = 10^{13.5} M_\odot/h_{70}$ a pivot mass, and $\beta_{\rm M}$ a free power-law index, which we intend to compare with the value predicted in Chapter **??**.

We estimate the covariance matrix from the data:

$$\boldsymbol{C}_{\mu\nu} = \frac{1}{N_{\rm sub}(N_{\rm sub}-1)} \sum_{j=1}^{N_{\rm sub}} (d_{j,\mu} - \overline{d}_\mu)(d_{j,\nu} - \overline{d}_\nu) , \tag{4.11}$$

with $\mu, \nu \in \{1, \dots, n_{\rm M}\}$, $\overline{d}_\mu = \frac{1}{N_{\rm sub}} \sum_{j=1}^{N_{\rm sub}} d_{j,\mu}$, and $d_{j,\mu} = r_{\rm g+}(M)$ for each sub-box and each mass bin, as defined in Eq. 4.9. We then invert the covariance matrix and correct the bias on the inverse to obtain an unbiased estimate of the precision matrix [50], given by:

$$\boldsymbol{C}^{-1}_{\rm unbiased} = \frac{N_{\rm sub} - n_{\rm M} - 2}{N_{\rm sub} - 1} \, \boldsymbol{C}^{-1} , \tag{4.12}$$

where $N_{\rm sub} > n_{\rm M} + 2$ clearly holds [50].

We then perform a likelihood analysis over the data to infer the posteriors of A and $\beta_{\rm M}$: according to Bayes' theorem, if $\boldsymbol{d}$ is the vector of the data and $\boldsymbol{p}$ the vector of the parameters,

$$P(\boldsymbol{p}|\boldsymbol{d}) \propto P(\boldsymbol{d}|\boldsymbol{p}) \; P(\boldsymbol{p}) \propto e^{-\frac{1}{2}\chi^2} P(\boldsymbol{p}) , \tag{4.13}$$

with $P(\boldsymbol{p}|\boldsymbol{d})$ the posterior probability, $P(\boldsymbol{d}|\boldsymbol{p})$ the likelihood function, $P(\boldsymbol{p})$ the prior probability and $\chi^2 = (\boldsymbol{d} - \boldsymbol{m})^T \boldsymbol{C}^{-1} (\boldsymbol{d} - \boldsymbol{m})$, with $\boldsymbol{m}$ the vector of the model and $\boldsymbol{C}^{-1}$ the precision matrix, the inverse of the covariance matrix $\boldsymbol{C}$. We assume uninformative flat priors in the fit with ranges $\log_{10} \frac{A}{(\mathrm{Mpc}/h)^{1/2}} \in [-1.3; -0.6]$ and $\beta_{\rm M} \in [0.2; 0.7]$.

The choice of $n_{\rm M}$ and $N_{\rm sub}$ is constrained by several factors: first of all, if $N_{\rm sub}$ is too large, the single values of $w_{\rm gg}$ (and $w_{\rm g+}$) tend to fluctuate around the mean, thus increasing the error bar and sometimes dropping below 0, which is unacceptable for our choice of $r_{\rm g+}(M)$; see Eq. 4.9. Furthermore, we want $n_{\rm M}$ to be large enough to

[2] We discuss possible biases in this procedure in Appendix C.

Table 4.1: Mass range, number of haloes $N_{\rm h}$, and ellipticity dispersion σ_ϵ for each of the 4 bins of both the Millennium and the Millennium-XXL simulations. The masses are defined as $M_{200{\rm c}}$, the mass within a region where the density exceeds 200 times the critical density.

	Mass range $[M_\odot/h]$	$N_{\rm h}$	σ_ϵ
Millennium simulation	$10^{11.36} - 10^{11.86}$	916109	0.15
	$10^{11.86} - 10^{12.36}$	328364	0.15
	$10^{12.36} - 10^{12.86}$	113917	0.16
	$10^{12.86} - 10^{13.36}$	37511	0.18
Millennium-XXL simulation	$10^{13} - 10^{13.5}$	5450501	0.22
	$10^{13.5} - 10^{14}$	1618707	0.24
	$10^{14} - 10^{14.5}$	370911	0.26
	$10^{14.5} - 10^{15}$	48714	0.28

be capable of displaying the trend of the signals along the whole mass range chosen. Finally, we need to take $n_{\rm M} \ll N_{\rm sub}$ to avoid divergences related to the fact that we estimate the covariance from a finite number of samples [50].

As a second approach, we use the CLASS (Cosmic Linear Anisotropy Solving System, [36]) algorithm to generate a non-linear matter power spectrum as a function of the wave number k. We then use Eq. 2.8 to convert the wave number to the corresponding mass M, and calculate the square root of the values of the power spectrum $P_\delta(M)$, as in Eq. 2.9. We adopt the following model for $r_{\rm g+}(M)$:

$$r_{\rm g+}(M) = A_{\rm PS} \cdot \sqrt{P_\delta(M)} \ , \tag{4.14}$$

with $A_{\rm PS}$ a generic amplitude which we will treat as a nuisance parameter.

We test the goodness of this approach by generating $N_{\rm PS} = 1000$ models as in Eq. 4.14, varying the amplitude $\log_{10}\left(A_{\rm PS}/h{\rm Mpc}^{-1}\right)$ between -2.5 and -0.5, and then by calculating for all of them the χ^2 :

$$\chi^2(A_{\rm PS}) = \sum_{i,j=1}^{n_{\rm PS}} \left(r_{\rm g+}(M_i) - A_{\rm PS}\sqrt{P_\delta(M_i)}\right) \boldsymbol{C}_{ij}^{-1} \left(r_{\rm g+}(M_j) - A_{\rm PS}\sqrt{P_\delta(M_j)}\right) \tag{4.15}$$

where the inverse of the covariance matrix is corrected as in Eq. 4.12, and $n_{\rm PS} = 2n_{\rm M}$ is the total number of bins, to include both simulations in the analysis. Our best model is the model which minimises the χ^2.

Chapter 5

Results

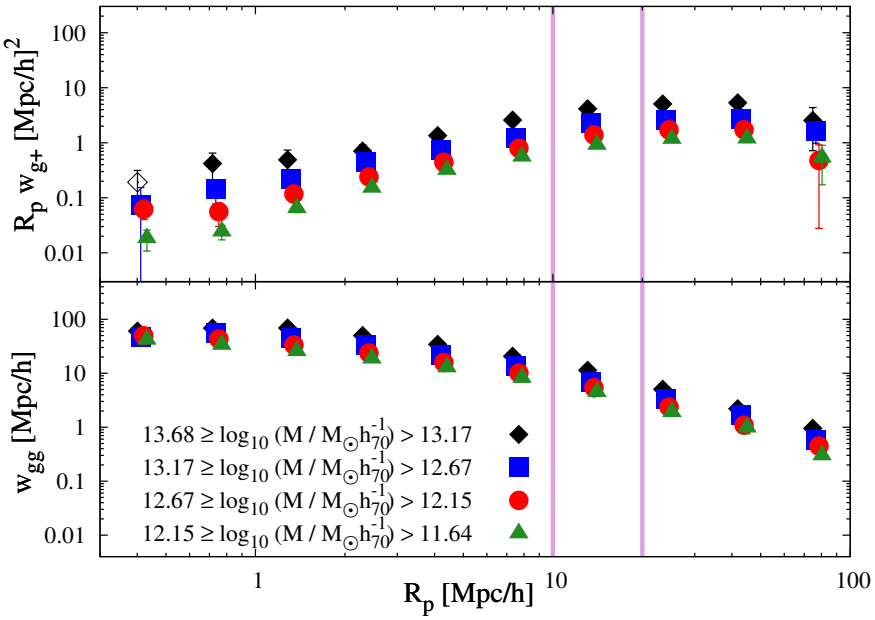
5.1 Simulation data

To be able to directly compare our results with those in [27, figure 7], we further convert the halo mass definition from $M_{200\mathrm{c}}$ to $M_{200\mathrm{m}}$, defined as the mass enclosed in a region inside of which the density is 200 times the mean density at the redshift corresponding to the respective snapshot.

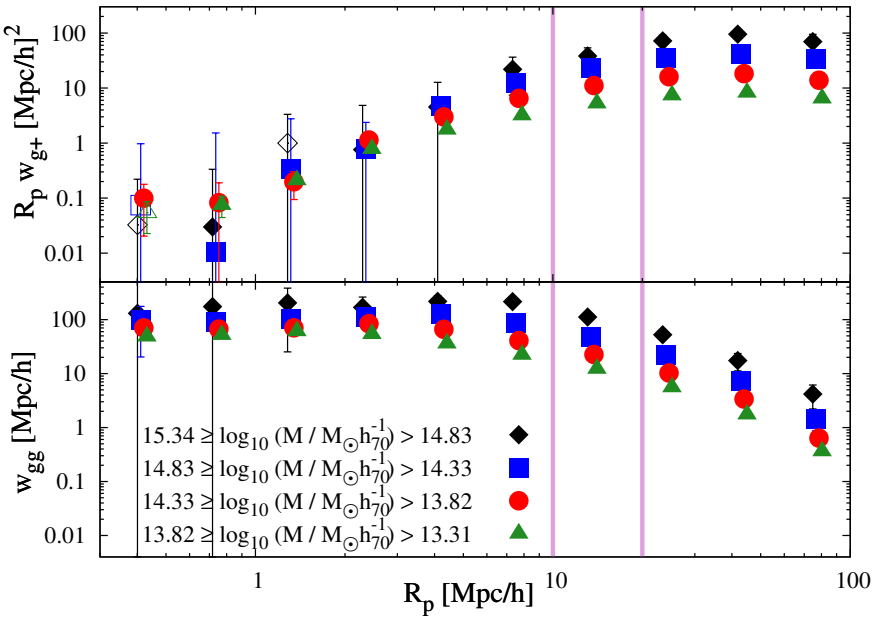
The trend of $w_{\mathrm{g}+}$ and w_{gg} with R_{p} for the four mass bins for each simulation is shown in Fig. 5.1. Note that in the chosen interval w_{gg} is always positive within the error bar, and even though fluctuations may result in negative values, our choice of n_{M} and N_{sub} and our large-bin average ensure that Eq. 4.9 always returns a real value. The points shown in Fig. 5.1 are the arithmetic mean of the N_{sub} values for each mass bin, while the error bars are the standard errors of the mean values. The overall behaviour of $w_{\mathrm{g}+}$ and w_{gg} agrees with literature works [23, 27]; in particular, we clearly detect positive intrinsic alignment in all samples, implying that the projected ellipticities of dark matter haloes tend to point towards the position of other haloes. Also, it is worth noting that both signals increase with increasing mass.

We then study the dependence of $w_{\mathrm{g}+}, w_{\mathrm{gg}}$ and $r_{\mathrm{g}+}$ on the mass of the halo. In Fig. 5.2 we also include, for the Millennium simulation only, two more results: grey diamonds represent the signal from the objects at redshift $z = 0.46$, while open black diamonds represent the signal from the objects at $z = 0$ obtained using the reduced inertia tensor (*rit*, as in Eq. 3.2) to measure the shapes of the haloes. As one can see, despite the use of two different halo finders and two different quantities to measure the shapes, the MS and the MXXL follow the same trend, and yield consistent results in the small mass range where they overlap; furthermore, all three $w_{\mathrm{g}+}, w_{\mathrm{gg}}$ and $r_{\mathrm{g}+}$ increase with increasing mass. We find that the use of the reduced inertia tensor leads to lower alignment signals, as found in [17], and that, on the other hand, these signals increase with increasing redshift.

We proceed by showing the results of the likelihood analysis described in Sect. 4.2: Fig. 5.3(a) shows the results from the single catalogues and from the joint analysis of the two simulations, obtained by multiplying the likelihood functions and assuming the same flat priors on the parameters. The most stringent bounds come from the MXXL, while the MS yields larger errors on the parameters, albeit consistent with the results of the MXXL. The joint analysis returns a value for the slope in agreement (within 3σ) with the prediction made in Chapter **??** for a DM-only



(a) w_{g+} and w_{gg} for the Millennium simulation.



(b) w_{g+} and w_{gg} for the Millennium-XXL simulation.

Figure 5.1: The intrinsic alignment signal w_{g+} and the clustering signal w_{gg} as a function of the comoving transverse separation R_p for (a) the Millennium and (b) the Millennium-XXL simulation. The pink lines indicate the $10 < R_p/\ h^{-1}\text{Mpc} < 20$ interval, the range used to model the signal. The mass ranges are displayed considering M_{200m} as the mass of the halo. In the graph, points are slightly horizontally shifted, so that they do not overlap; negative values are displayed in absolute value with open symbols of the same colour. The MS clustering signal is corrected to account for the integral constraint as explained in Appendix B. An increasing trend with mass is clear in each panel separately, and comparing the two panels as well.

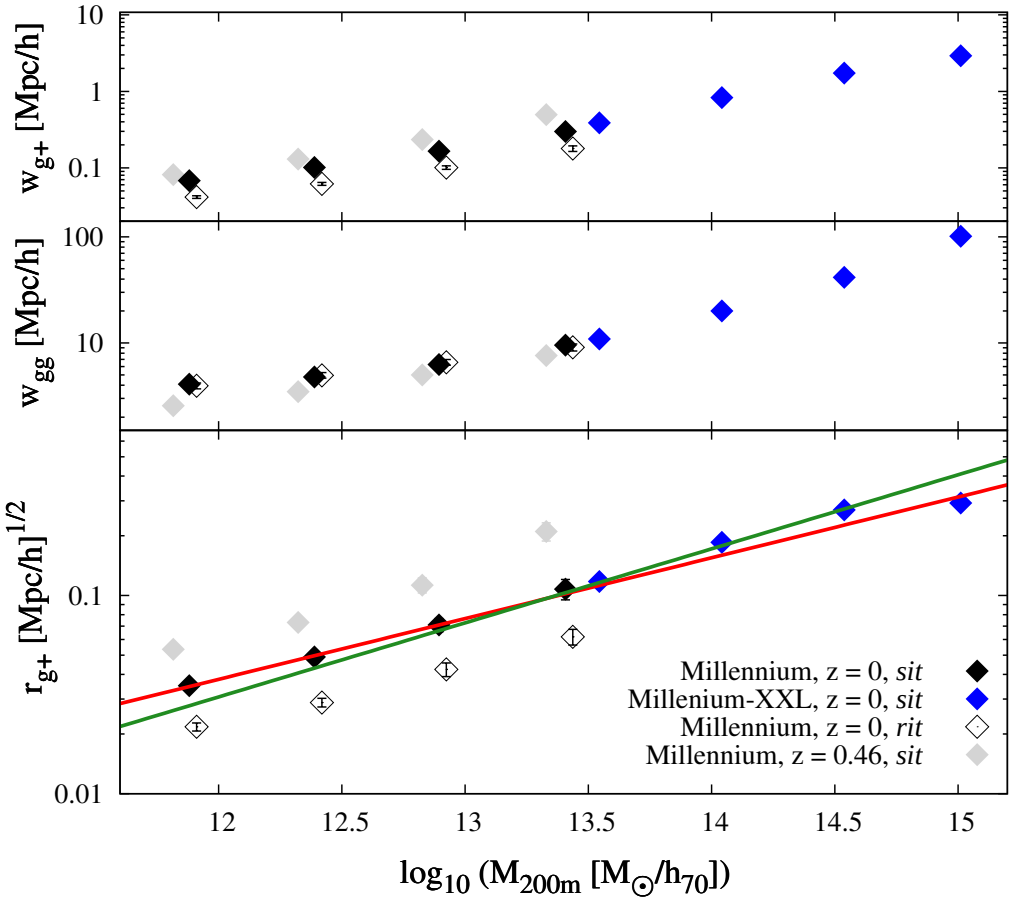


Figure 5.2: The intrinsic alignment signal w_{g+}, the clustering signal w_{gg} and r_{g+} as defined in Eq. 4.9 as a function of halo mass M_{200m} for the Millennium and the Millennium-XXL simulations. The label *sit* stands for simple inertia tensor, while *rit* means reduced inertia tensor; note that for the MXXL data only the simple inertia tensor is available, as mentioned in Sect. 3.2. The points are not placed at the midpoint of the bin, but at the value corresponding to the arithmetic mean of the mass of the objects. The red line and the green line represent the best-fit lines for the MS and MXXL likelihood analyses respectively; they are drawn using the parameters reported in Table 5.1. Points showing the results from the *rit* choice are horizontally shifted by a small amount, so that they do not overlap with the corresponding *sit* dots.

Table 5.1: Mean and 68% confidence interval of the fit parameters of the IA model from the likelihood analysis over the Millennium simulation, the Millennium-XXL simulation, their joint contribution, the Milennium simulation at $z = 0.46$, the Millennium simulation using the reduced inertia tensor (*rit*) and real data. Note that the values from the snapshot at different redshift and from the reduced inertia tensor assumption are compatible with the outcomes of the MS only. A detailed discussion about the reasons why the reduced χ^2 values obtained considering the Millennium-XXL simulation significantly differ from unity is presented in the text.

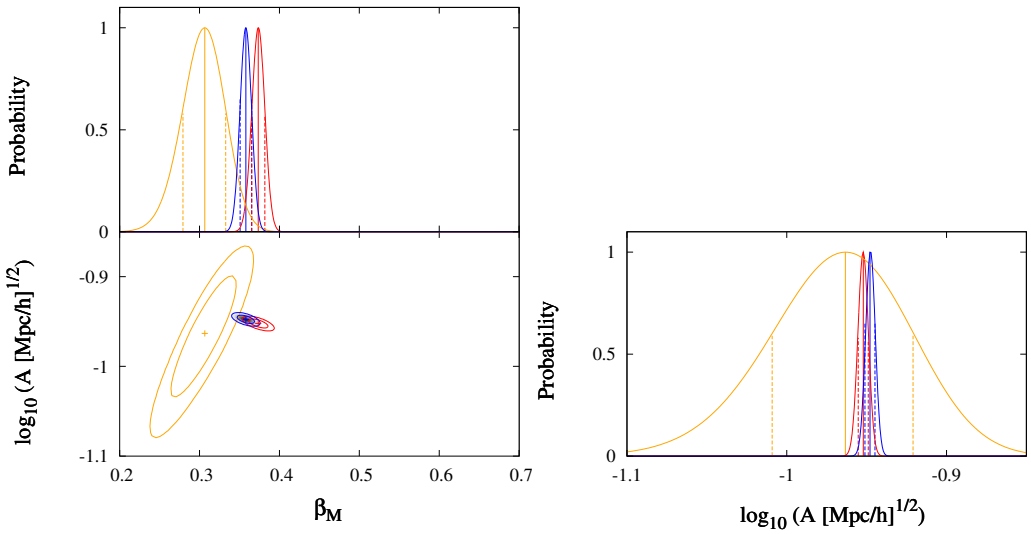
	MS only	MXXL only	Joint	MS, $z = 0.46$	MS, *rit*
$\beta_{\rm M}$	$0.31^{+0.03}_{-0.03}$	$0.37^{+0.01}_{-0.01}$	$0.36^{+0.01}_{-0.01}$	$0.35^{+0.03}_{-0.03}$	$0.29^{+0.02}_{-0.02}$
$\log_{10}(A\ [\mathrm{Mpc}/h]^{1/2})$	$-0.96^{+0.04}_{-0.05}$	$-0.95^{+0.03}_{-0.03}$	$-0.95^{+0.03}_{-0.03}$	$-0.72^{+0.04}_{-0.04}$	$-1.21^{+0.04}_{-0.04}$
χ^2/dof	0.27	28.06	11.72	2.27	0.70
	Real data				
$\beta_{\rm M}$	$0.56^{+0.05}_{-0.05}$				
$\log_{10} A_r$	$0.61^{+0.03}_{-0.04}$				
χ^2/dof	1.68				

universe ($\beta_{\rm M} = 1/3$); in particular, the tightest constraints are obtained with cluster-size objects, for which we predicted $\beta_{\rm M} \simeq 0.36$, in excellent agreement with our MXXL high-mass results ($\beta_{\rm M} = 0.37^{+0.01}_{-0.01}$). Moreover, we find that neither the inertia tensor definition nor the chosen redshift for our default analysis have significant impact on our conclusions for $\beta_{\rm M}$, although we note that the alignment amplitude increases with z. We report all the best-fit values, together with their respective errors and reduced χ^2, in Table 5.1.

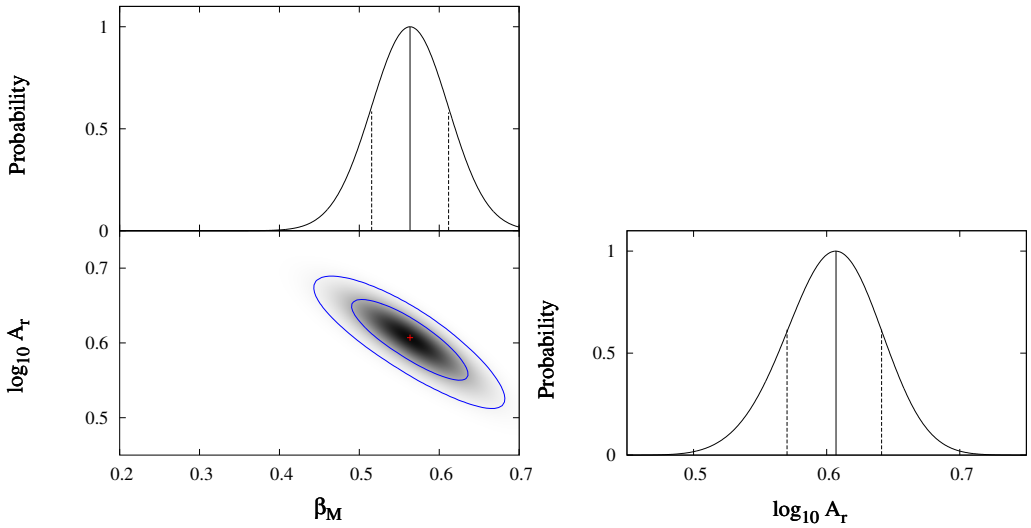
We note that the high value of the reduced χ^2 for the MXXL (and, consequently, for the joint analysis) can be attributed to the behaviour of the alignment signal at the high-mass end, which leads to a value of $r_{\rm g+}$ which differs more than 4σ from the fitted one; this issue is evident in Fig. 5.2. We calculate that the reduced χ^2 for the MXXL can be lowered to 7.71 by excluding the highest-mass point in the fit without significantly affecting the value of the slope; also, in this last bin the error on the data point is probably underestimated due to the low number density of the most massive haloes, as one can see from Fig. 3.2.

We finally show the results for the second approach, in order to understand whether the high values of the reduced χ^2 above can be attributed to the approximation that the power spectrum is not perfectly described by a power-law; to be consistent with the definition of the mass, we choose $\Delta = 200\Omega_{\rm m}$ in Eq. 2.8. In the fit, we exclude the highest-mass result, therefore $n_{\rm PS} = 2n_{\rm M} - 1 = 7$.

We calculate the χ^2 for each model: the lowest value we obtain is $\chi^2/\mathrm{d.o.f.} = 6.16$ for $\log_{10}\left(A_{\rm PS}/h\mathrm{Mpc}^{-1}\right) = -1.54$. We report our results with this model fitted in Fig. 5.4. It is evident that with this second approach problems rise at both mass ends, where our model is not capable of fitting the results: this issue may be related to the fact that we need a more complicated equation than Eq. 2.8 to link the mass of our objects to the wave number k. This second approach does not improve significantly the value of the reduced χ^2, and therefore further investigation is needed.



(a) Single and joint likelihood analysis for the simulations.



(b) Likelihood analysis for the real data.

Figure 5.3: Posterior on the fit parameters of the IA model, obtained for (a) the Millennium simulation, the Millennium-XXL simulation, joint MS and MXXL, and (b) real data. Note that the ranges of the prefactors are quite different. The bottom-left graph in the right panel shows the contour lines and the 2-D greyscale posterior for the real data, while the bottom-left graph in the left panel shows the contour lines of the 2-D posteriors for all the simulations (single and joint), but the 2-D greyscale posterior for the joint analysis only. All other sub-panels show the marginalized 1-D posterior normalized to a peak amplitude of 1. Contour lines enclose the 68% and 95% confidence intervals, crosses and vertical solid lines indicate the best-fitting values, while dashed lines represent the 1-σ confidence interval. We note that the results are consistent for the two catalogues, while real data yield a value of the slope which is incompatible with the ones from the simulation data. The mean values and 68% confidence intervals of A, A_r and $\beta_{\rm M}$ are listed in Table 5.1.

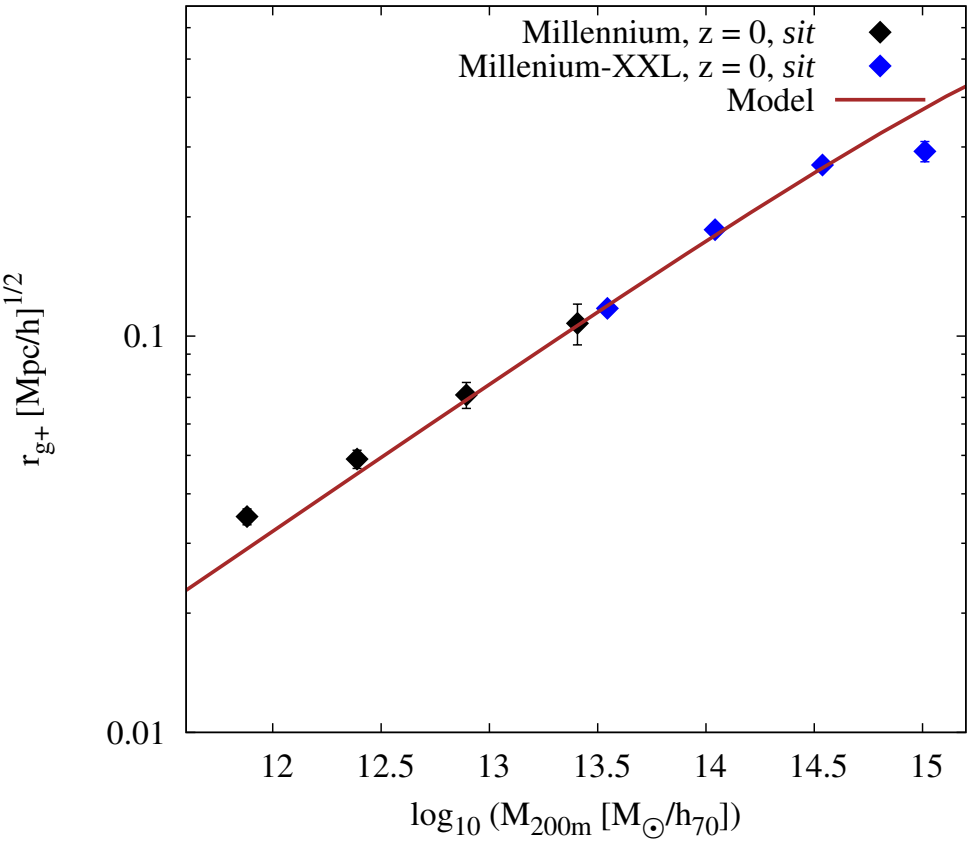


Figure 5.4: Best-fit model for our second approach, in which we consider the full power spectrum without approximating it with a power-law. The results for the Millennium and Millennium-XXL simulation are the same as in Fig. 5.2, but in this fit we exclude the highest-mass bin point. The fitted model yields χ^2/d.o.f. = 6.16; reasons to explain why it differs from unity are presented in the text.

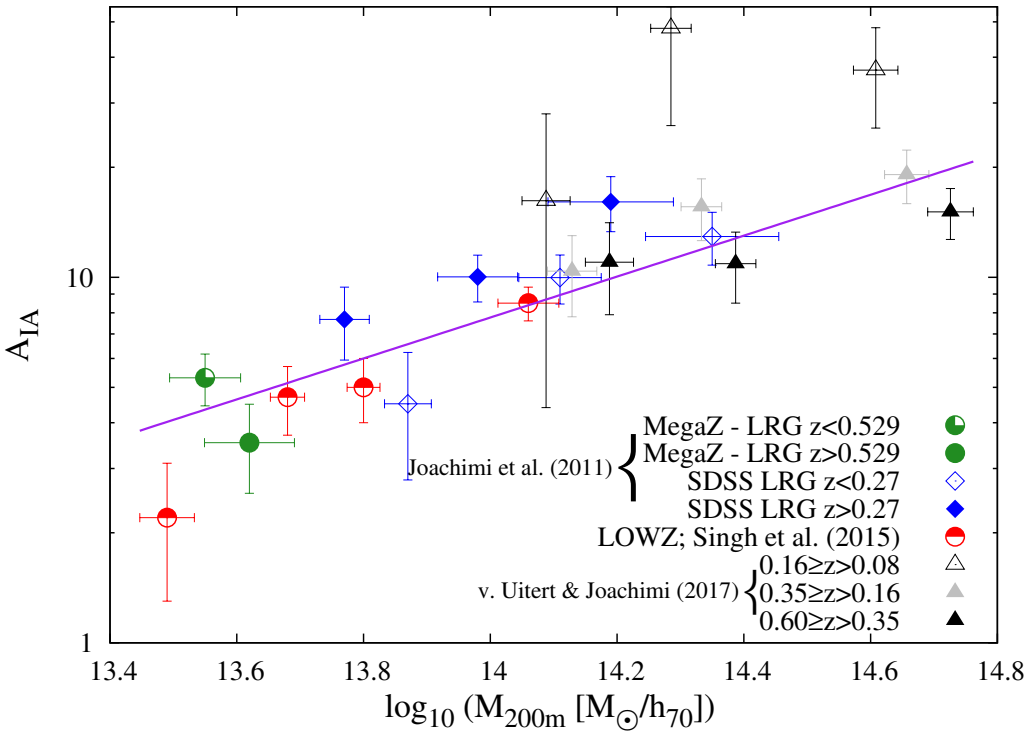


Figure 5.5: Amplitude of the intrinsic alignment model as a function of halo mass from observations. The green, blue and red symbols represent the results for luminous red galaxies ([23, 26]), while the triangles indicate the signal for galaxy clusters [27]. The solid purple line shows the best-fit line from our likelihood analysis. The exact values of the parameters are shown in Table 5.1.

5.2 Observation data

To compare our results with real data, we re-analyse the results presented in [27]: in that work, the clusters contained in the redMaPPer catalogue [51] version 6.3 were used to constrain the intrinsic alignment amplitude $A_{\rm IA}$, which was then studied as a function of halo mass.

We repeat the likelihood analysis for the collection of observational datasets shown in [27, figure 7]: we consider all 21 data points, which include literature results for LRGs from [26] and [23], as well as results for clusters from [27]. We neglect the error bars on the mass, which are smaller than the errors on $A_{\rm IA}$ and whose impact is subdominant [27], and treat all the data as independent, as in [27]. In this way, the covariance matrix is diagonal. We show the points, together with the best-fitting power-law from our analysis, in Fig. 5.5.

We adopt the model of Eq. 4.10 but with a different prefactor A_r, which has an altered meaning and is now dimensionless, thus making it impossible to directly compare its value to the one obtained with simulation data. We also assume a different range for the flat prior in the fit for this new parameter, namely $\log_{10} A_r \in [0.4; 0.9]$.

The outcomes of our analysis are shown in Fig. 5.3(b) and in Table 5.1: we observe that the value of the reduced chi-square for this latter analysis can be further

improved to 1.36 by excluding the high-redshift SDSS results (filled-blue diamonds) without affecting the value of the slope in a significant way.

The incompatibility between the values of the slope $\beta_{\rm M}$ between simulation and real data is significant. This discrepancy may be attributed to the additional mass dependence of the response of a galaxy ellipticity to the tidal field, and, more generally, to the fact that, while we observe luminous matter, the simulations and the theory model only consider dark matter. Moreover, a bias in $\beta_{\rm M}$ could be determined in observational data if, with increasing halo mass, there is also a trend in redshift, which however is not the case of the data collection of Fig. 5.5.

Chapter 6

Discussion and conclusion

In this work we studied the dependence of the intrinsic alignment amplitude on the mass of dark matter haloes, using data from the Millennium and Millennium-XXL N-body simulations.

We derived the intrinsic alignment amplitude scaling with mass in the tidal alignment paradigm for a dark matter-only universe. Our analytical estimate assumed a virialised system which is weakly perturbed by ambient tidal shear fields: in this model, the magnitude of the tidal distortion of a halo is determined entirely by the amount of fluctuations of the tidal fields on the mass scale of the halo. The model predicts a scaling with halo mass $\propto M^{\beta_{\rm M}}$, with $\beta_{\rm M} \in [1/3, 1/2]$, if we approximate the matter power spectrum with a power-law. We also tested our model with a second approach, which does not rely on this latter assumption but requires a link between the halo mass and the wave number.

We mimicked the observational approach to measure the halo shape-position alignments, and we first performed a Bayesian analysis on the mass dependence of the alignments with a simple power-law model. We found that the results from the two simulation data sets agree well with each other and, more noticeably, with the theoretical prediction; in particular, the joint analysis yields $\beta_{\rm M} = 0.36^{+0.01}_{-0.01}$, and for cluster-size scales, for which we predicted $\beta_{\rm M} = 0.36$, we found $\beta_{\rm M} = 0.37^{+0.01}_{-0.01}$. Furthermore, there is no dependence of the slope on redshift or on the definition of the inertia tensor which describes the shape of the halo.

As a second approach, we performed a χ^2 analysis on the same results. In this case, our model depended only on one parameter, and we found that it does not provide a good fit of the data at both mass ends, resulting in a very high value of the reduced χ^2 (6.16). This approach needs further investigation, though, since it relies on the assumption that our relation between the mass and the wave number, which is valid for spherical objects, holds for the ellipsoidal haloes in the simulations as well.

We repeated our likelihood analysis using observational data, inferring a value of $\beta_{\rm M} = 0.56^{+0.05}_{-0.05}$, which is not compatible with the simulation results. The incompatibility can be attributed to the fact that simulations consider a dark-matter only universe, while we observe luminous matter. The presence of a significant amount of misalignment between the orientations of LRGs and their dark matter haloes was proposed in [52], and hydrodynamical simulations suggest that the presence of baryons could significantly affect the shape of the host halo, especially in

the inner regions ([2], and references therein). For example, [53] found an increase in halo sphericity that diminishes at larger radii (and thus with larger mass values), in agreement with [54], while more recently [55] and [56] discovered that galaxies are more misaligned with the hosting halo at lower masses.

All these papers suggest that it is safe to assume that galaxies and clusters of galaxies are less aligned with each other than haloes at lower mass values, which means that the slope $\beta_{\rm M}$ expected from the observations should be higher than in the case of a dark matter-only scenario, in agreement with what we found in our analysis. Possible future works in this sense would be then to measure the slope $\beta_{\rm M}$ considering a reduced inertia tensor for the Millennium-XXL simulation, to better describe the shape of the inner part of the halo, as well as using a hydrodynamical simulation large enough to contain clusters, which could then account for the additional effects of baryons and gas.

Appendices

Appendix A

The choice in the comoving transverse separation bin

In Sect. 4.1 we explained that we evaluate our estimators w_{g+} and w_{gg} in a wide R_p bin in order to focus on the dependence of our signals on the mass of the haloes. The initial choice was 6 Mpc/h < R_p < 30 Mpc/h, for two reasons: first, 6 Mpc/h is the minimum scale usually adopted in observational papers, below which the bias becomes non-linear [57, 58]; second, since we are studying local correlations, we need to place an upper threshold, and with 30 Mpc/h we do not lose much information, as the signal-to-noise ratio of the IA signal is small at larger scales.

We report the results that we obtain with this wider bin in Fig. A.1: with the exception of an increase in both the w_{g+} and w_{gg} signals with a narrower interval, the overall trend is not affected by this choice, and the results of the analysis are consistent with the two bins, as clear from Table A.1. However, we chose to consider a narrower range to avoid problems at the high- and low-mass ends: if we invert Eq. 2.8 we can calculate that the highest mass objects, for example, have a size of up to 3 Mpc/h, and it becomes therefore clear that it is not correct to look for the correlation between two haloes which almost merge with each other. On the other hand, the lowest mass objects have a size of about 0.1 Mpc/h, and for these haloes the signal-to-noise ratio could be already quite small below 30 Mpc/h.

Since we calculate a weighted average inside our R_p bin, it is thus possible that with the choice of a wider range we could bias the results at the two mass ends, underweighting the MS signals at high R_p and the MXXL signals at low R_p. Therefore, we reach a compromise by switching to 10 Mpc/h < R_p < 20 Mpc/h, even though this choice slightly increases the error bars, and perform our analysis in this interval.

Table A.1: Mean and 68% confidence interval of the fit parameters of the IA model from the likelihood analysis over the Millennium simulation, the Millennium-XXL simulation, their joint contribution, the Milennium simulation at $z = 0.46$, and the Millennium simulation using the reduced inertia tensor (*rit*). In this case, we consider a wider bin in the comoving transverse separation $R_{\rm p}$, without significantly changing the results of the analysis.

	MS only	MXXL only	Joint	MS, $z = 0.46$	MS, *rit*
$\beta_{\rm M}$	$0.29^{+0.02}_{-0.02}$	$0.36^{+0.01}_{-0.01}$	$0.35^{+0.01}_{-0.01}$	$0.34^{+0.02}_{-0.02}$	$0.29^{+0.02}_{-0.02}$
$\log_{10}(A\ [{\rm Mpc}/h]^{1/2})$	$-0.85^{+0.03}_{-0.03}$	$-0.82^{+0.02}_{-0.03}$	$-0.82^{+0.03}_{-0.01}$	$-0.57^{+0.03}_{-0.04}$	$-1.06^{+0.03}_{-0.03}$
χ^2/dof	0.85	43.07	16.7	1.37	0.40

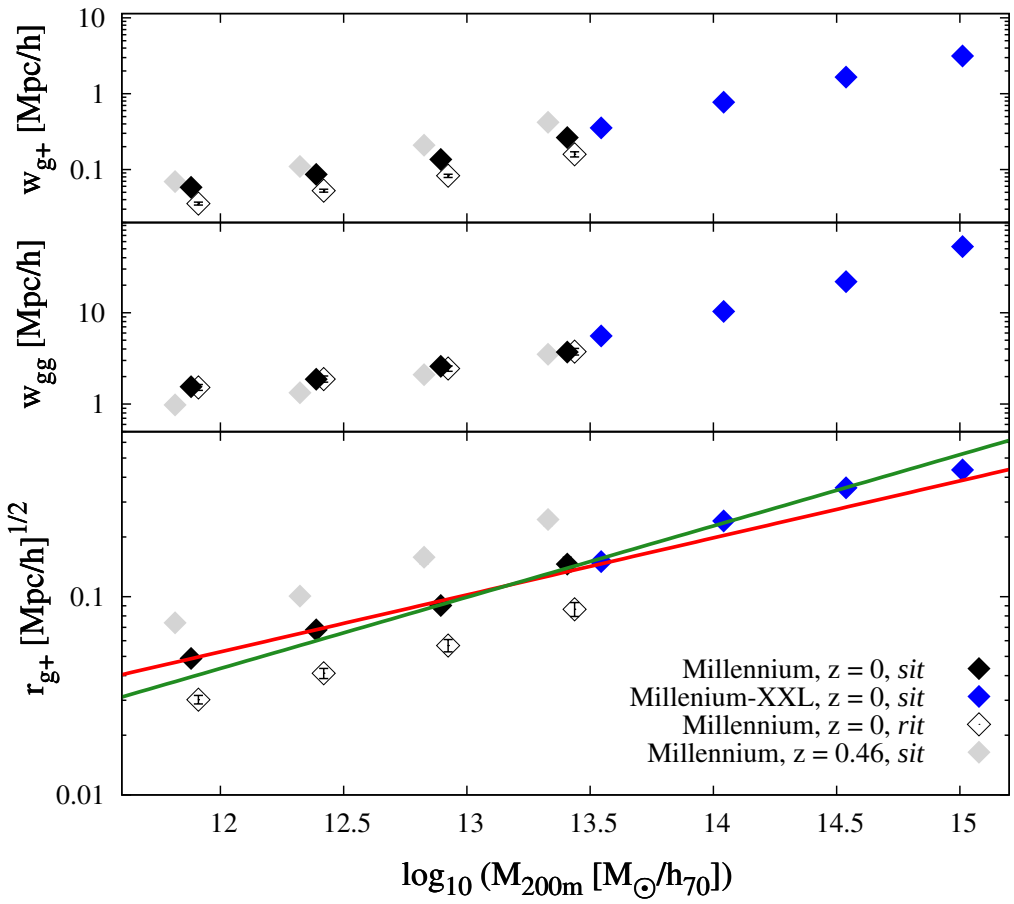


Figure A.1: The intrinsic alignment signal $w_{\rm g+}$, the clustering signal $w_{\rm gg}$ and $r_{\rm g+}$ as defined in Eq. 4.9 as a function of halo mass $M_{\rm 200m}$ for the Millennium and the Millennium-XXL simulations, considering a wider bin in the comoving transverse separation $R_{\rm p}$. The results of the analysis do not change significantly, but with this wider range we may introduce a bias at the two ends of the mass range, as explained in detail in the text.

Appendix B

Determination of the integral constraint

The observed clustering signal $w_{\rm gg}$ is affected by the restricted area of the simulation box: the integral over all pair counts relative to randoms has to yield zero over the whole sky, but we look at smaller boxes, and it is thus necessary to add an offset, which we call integral constraint (IC), not to bias our results [59]. Our first approach in determining the integral constraint consists in following [59]: we describe $w_{\rm gg}$ as a function of $R_{\rm p}$ with a power-law,

$$w_{\rm gg}(R_{\rm p}) = AR_{\rm p}^{-0.8} \ , \tag{B.1}$$

and then state that

$$w_{\rm gg}(R_{\rm p}) = A(R_{\rm p}^{-0.8} - C_{\rm IC}) \ , \tag{B.2}$$

with

$$C_{\rm IC} = \frac{\sum N_{\rm rr}(R_{\rm p}) R_{\rm p}^{-0.8}}{\sum N_{\rm rr}(R_{\rm p})} \ , \tag{B.3}$$

where $N_{\rm rr}(R_{\rm p})$ is the number of random halo pairs in each $R_{\rm p}$ bin. The results are reported in Fig. B.1: even though this correction is clearly non-negligible, it is evident that the inferred values for $\beta_{\rm M}$ with and without the IC are perfectly compatible.

To assess how the finite size of the simulation boxes affects our results, then, we change the way in which the random catalogues are built: after a first attempt with variable-size random catalogues, we normalise with regard to the full simulation box, rather than to the sub-box catalogues. In other words, rather than using random catalogues whose size is three times the corresponding size of each sub-box, we consider random catalogues all with the same number of objects fixed to the 3/27 of the number of haloes in the whole simulation, for each mass bin. We test for changes in the clustering signal with these two different approaches, and from Fig. B.2 it is evident that there is no impact on the amplitudes of $w_{\rm gg}$ in the MS and MXXL. For this reason, we decide not to correct the clustering signal with the procedure described above, and always calculate $w_{\rm gg}$ with fixed-size random catalogues.

However, the boxes of the two simulations do have a different size, as mentioned in Sect. 3.1. To test the effects of this difference we select both catalogues in the same mass bin ($10^{13} M_\odot/h - 10^{13.5} M_\odot/h$) and calculate the clustering signal $w_{\rm gg}$ as

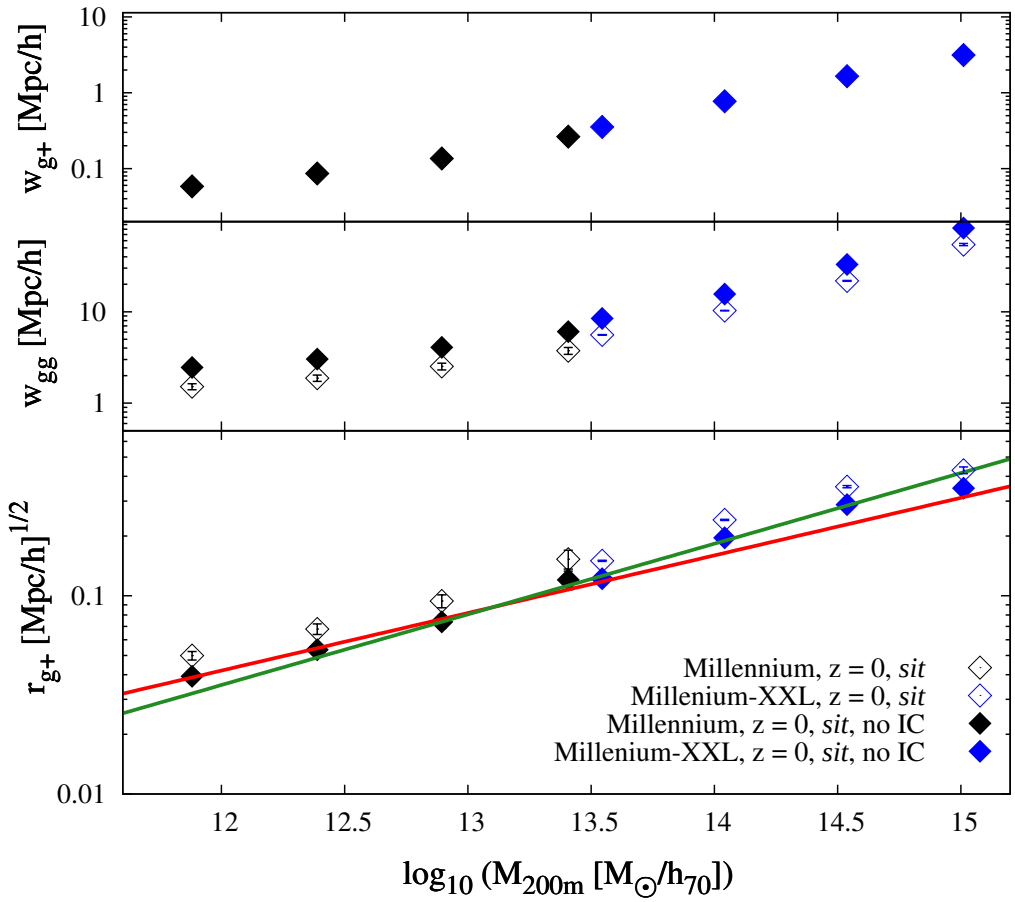
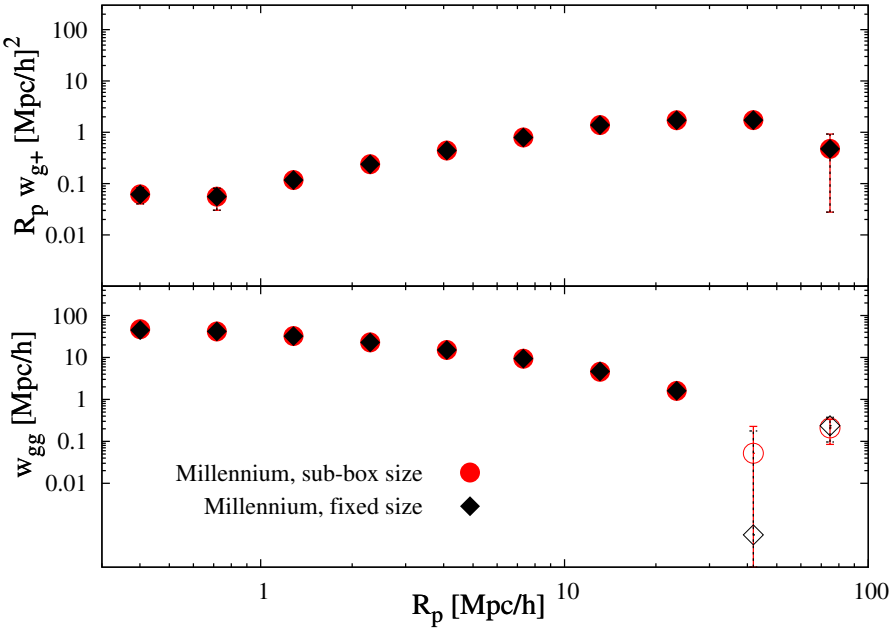
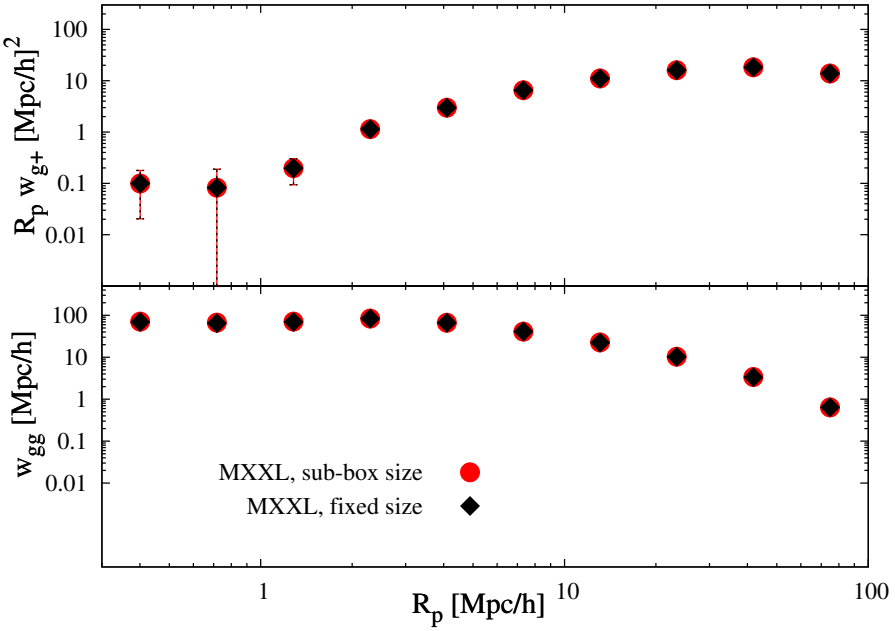


Figure B.1: The intrinsic alignment signal w_{g+}, the clustering signal w_{gg} and r_{g+} as defined in Eq. 4.9 as a function of halo mass M_{200m} for the Millennium and the Millennium-XXL simulations, considering 6 Mpc/h < R_p < 30 Mpc/h. In particular, in this plot we want to assess the impact of the integral constraint on our results: open (filled) diamonds represent the signals with (without) this correction. While it is clear that w_{gg} is significantly corrected with a negative offset, the inferred value of the slope β_B is not affected by this subtraction.



(a) Clustering signal comparison for the Millennium simulation.



(b) Clustering signal comparison for the Millennium-XXL simulation.

Figure B.2: The intrinsic alignment signal w_{g+} and the clustering signal w_{gg} as a function of the comoving transverse separation R_p for (a) the Millennium and (b) the Millennium-XXL simulation, for the second mass bin of each simulation. In this case, we test if the size of the random catalogues in calculating the clustering signal has an impact on w_{gg}: as one can see, no significant differences are present at any R_p. The w_{g+} signal does not depend on the random catalogues, and thus does not change with the two different approaches. Note that in the main analysis we always use fixed-size random catalogues.

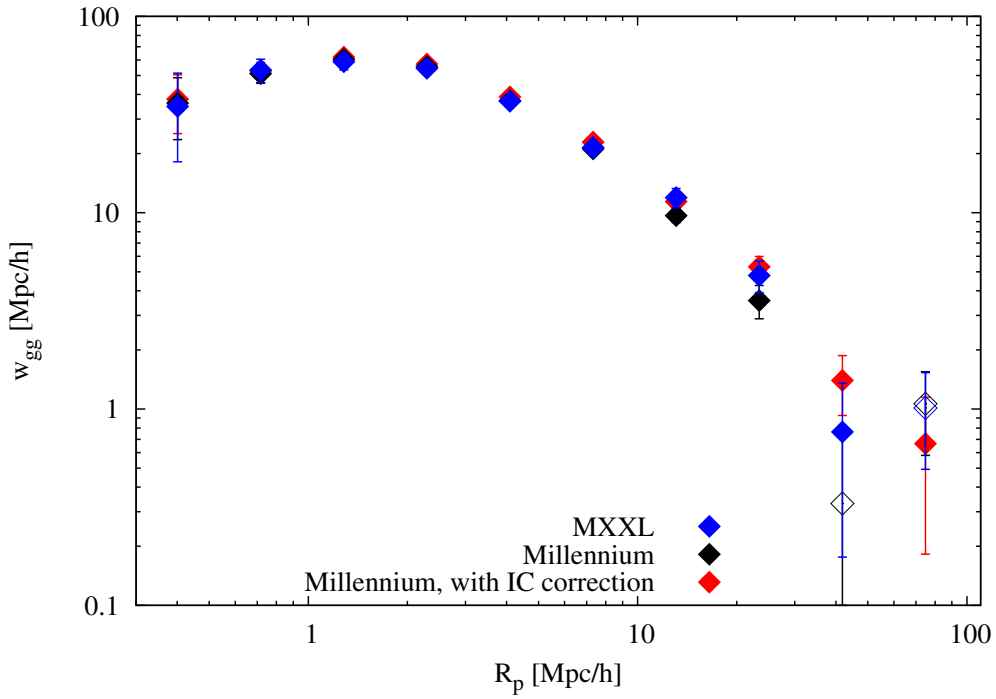


Figure B.3: The clustering signal $w_{\rm gg}$ as a function of the comoving transverse separation $R_{\rm p}$, in the same mass bin, for both the simulations. It is clear that the Millennium signal tends to become negative at smaller scales than the Millennium-XXL, due to the different size of the simulation boxes. For this reason, we perform a simple χ^2 analysis, described in the text, to find the best value of the integral constraint to correct the MS point; this corrected signal is represented by the red-filled diamonds. As always, negative values are displayed in absolute value with open symbols of the same colour.

a function of $R_{\rm p}$. Since the MS has a smaller simulation box, the signal tends to become negative at smaller values of the separation $R_{\rm p}$, as displayed in Fig. B.3.

To overcome this issue, we assume that the MXXL values are not affected by finite-size box effects, as explained above, and that they are the "true" values of the clustering signal. Then, we fit the MS signal to the MXXL points by stating that it is possible to describe the MXXL signal by adding a constant factor IC to the MS points, and calculate the χ^2:

$$\chi^2({\rm IC}) = \sum_{i=1}^{10} \frac{\left(w_{\rm gg}^{\rm MS}(R_{{\rm p},i}) + {\rm IC} - w_{\rm gg}^{\rm MXXL}(R_{{\rm p},i})\right)^2}{\sigma^2[w_{\rm gg}^{\rm MS}(R_{{\rm p},i})]} \,, \tag{B.4}$$

where we neglect any correlations between the points, and the index i runs over the 10 points of Fig. B.3. The best value of the IC is the one that minimises the χ^2, and

can be calculated straightforwardly:

$$\frac{\mathrm{d}\chi^2(\mathrm{IC})}{\mathrm{d}(\mathrm{IC})} \stackrel{!}{=} 0 \rightarrow \mathrm{IC} = -\sum_{i=1}^{10} \frac{w_{\mathrm{gg}}^{\mathrm{MS}}(R_{\mathrm{p},i}) - w_{\mathrm{gg}}^{\mathrm{MXXL}}(R_{\mathrm{p},i})}{\sigma^2[w_{\mathrm{gg}}^{\mathrm{MS}}(R_{\mathrm{p},i})]} \Big/ \sum_{i=1}^{10} \frac{1}{\sigma^2[w_{\mathrm{gg}}^{\mathrm{MS}}(R_{\mathrm{p},i})]} \ . \quad \text{(B.5)}$$

We obtain IC = 1.73 Mpc/h.

The last step in our procedure is to determine the IC in each mass bin of the Millennium simulation. If we look at Fig. 5.1, it is safe to write $w_{\mathrm{gg}}(M, R_{\mathrm{p}}) = \mathrm{A}_{\mathrm{gg}}(M)\tilde{w}_{\mathrm{gg}}(R_{\mathrm{p}})$, in order to separate the dependence on the mass and on the comoving transverse distance: in fact, the trend looks the same for all the mass bins, with the only difference given by an amplitude factor depending on the mass. Following the definition in [59], then:

$$\mathrm{IC}(M) = \frac{1}{\Omega^2} \int\int w_{\mathrm{gg}}(M, R_{\mathrm{p}}) \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 = \frac{\mathrm{A}_{\mathrm{gg}}(M)}{\Omega^2} \int\int \tilde{w}_{\mathrm{gg}}(R_{\mathrm{p}}) \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 \ , \quad \text{(B.6)}$$

where Ω is the area of a slice of the simulation box. If we denote the $10^{13} M_{\odot}/h - 10^{13.5} M_{\odot}/h$ mass bin as M^*, we know that $\mathrm{IC}(M^*) = \frac{\mathrm{A}_{\mathrm{gg}}(M^*)}{\Omega^2} \int\int \tilde{w}_{\mathrm{gg}}(R_{\mathrm{p}}) \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 =$ 1.73 Mpc/h, and considering 10 Mpc/h $< R_{\mathrm{p}} <$ 20 Mpc/h, we can calculate $w_{\mathrm{gg}}(M^*, R_{\mathrm{p}}^*) = (8.40 \pm 0.75)$ Mpc/h; therefore, $\mathrm{IC}(M) = w_{\mathrm{gg}}(M, R_{\mathrm{p}}^*) \frac{1.73}{8.40} \simeq 0.21 \times w_{\mathrm{gg}}(M, R_{\mathrm{p}}^*)$. The integral constraint is thus included in our analysis by incrementing all the MS clustering signals by a 21% factor.

Appendix C

On the error estimation when using the sub-boxes

In the analysis, we divide each simulation box in $N_{\rm sub} = 27$ sub-boxes, which we consider as independent realisations of our Universe, in each of which we calculate the intrinsic alignment ($w_{\rm g+}$) and clustering ($w_{\rm gg}$) signals; however, one may argue whether or not this procedure introduces any systematic errors in the covariance assigned to our measurements.

To test this possibility, we select both simulations in the same mass bin ($10^{13} M_\odot/h - 10^{13.5} M_\odot/h$), and divide the Millennium-XXL simulation in 27 sub-boxes of 500 Mpc/h on a side (which is the size of the Millennium simulation box), centred in the 1000 Mpc/h-side boxes used in the main analysis. We then repeat our measurements in each of these sub-boxes, and average the results that we obtain over all of them. We first compare the results for our signals as a function of the comoving transverse distance $R_{\rm p}$, and show them in Fig. C.1, with the error bars being the average of the errors: it is evident that MS-size $w_{\rm gg}$ measurements are systematically pushed down on large scales, as already discussed in Appendix B, but the errors are consistent between the MS and MS-in-MXXL signals.

To further confirm the absence of systematic errors, we measure the same signals in the $6 < R_{\rm p}/\ h^{-1}{\rm Mpc} < 30$ interval. We report our results in Table C.1 and in the corresponding Fig. C.2: the third, fourth and fifth columns all show the same mean value, which is the average of the 27 MS-in-MXXL sub-boxes, but with three different error bars, respectively the average of the errors (column 3), the standard deviation of the data (column 4), and the error on the mean (column 5).

From the first row of Fig. C.2, it is evident that the procedure of sub-dividing is not systematically biased; in the second row, we note general agreement between the MS and MS-in-MXXL results, which are however inconsistent with the MXXL measurements due to the lack of the integral constraint; this is easily fixed when applying the procedure of Appendix B.

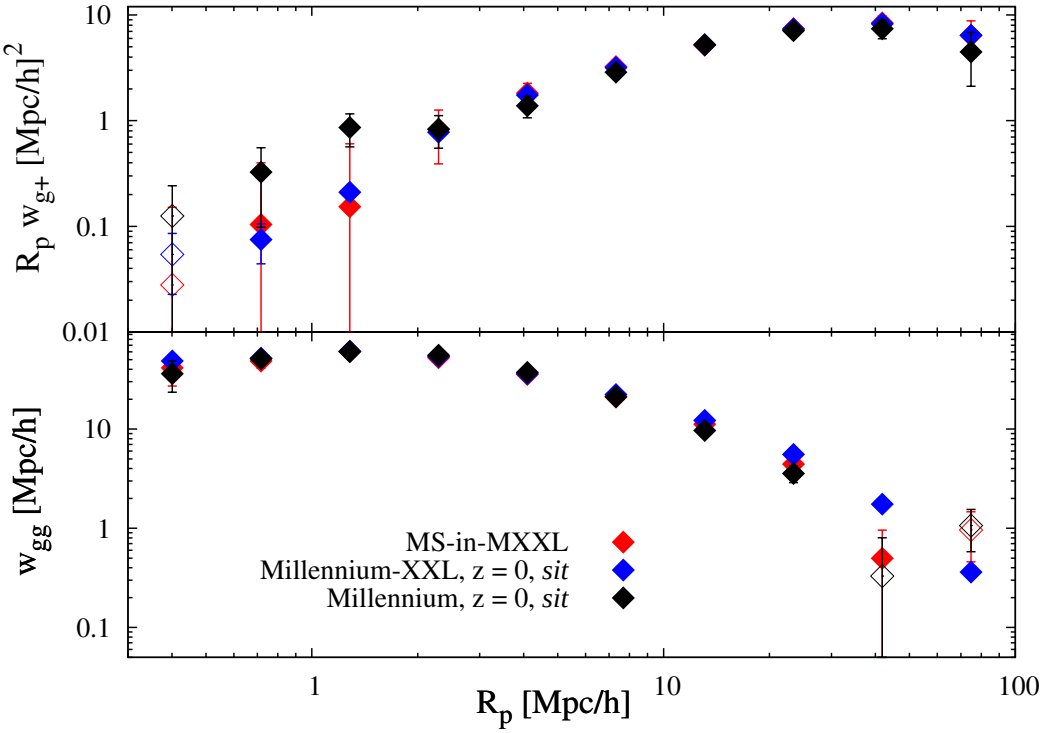


Figure C.1: The intrinsic alignment signal w_{g+} and the clustering signal w_{gg} as a function of the comoving transverse separation R_p in the $10^{13} M_\odot/h - 10^{13.5} M_\odot/h$ mass bin for the Millennium simulation, the Millennium-XXL simulation, and the MS-in-MXXL signals, whose derivation is described in the text. In this Appendix we want to test whether we bias our measurements when dividing the simulation box in 27 sub-boxes: while it is clear that the errors for the MS and the MS-in-MXXL signals are consistent, we note that the clustering measurements are systematically pushed down on large scales; we study this effect in Appendix B.

Table C.1: Results of the measurements in the $6 < R_{\rm p}/\ h^{-1}{\rm Mpc} < 30$ interval for the Millennium simulation, the Millennium-XXL simulation, and the MS-in-MXXL signals, whose derivation is described in the text, in the $10^{13}M_{\odot}/h - 10^{13.5}M_{\odot}/h$ mass bin. In particular, the third, fourth and fifth columns all show the same mean value, which is the average of the 27 MS-in-MXXL sub-boxes, but with three different errors, respectively the average of the errors (column 3), the standard deviation of the data (column 4), and the error on the mean (column 5). We note that $w_{\rm g+}$ is in perfect agreement over the different approaches, while $w_{\rm gg}$ is consistent with the MS-size measurements, which means that the procedure of dividing in sub-boxes is not introducing any biases. However, these latter signals are incompatible with the MXXL result; we further discuss this issue in Appendix B.

	MS	MXXL	MS-in-MXXL I	MS-in-MXXL II	MS-in-MXXL III
$w_{\rm g+}[{\rm Mpc}/h]$	0.35 ± 0.03	0.354 ± 0.002	0.36 ± 0.03	0.36 ± 0.03	0.36 ± 0.01
$w_{\rm gg}[{\rm Mpc}/h]$	6.81 ± 0.65	8.43 ± 0.06	7.53 ± 0.73	7.53 ± 0.79	7.53 ± 0.15

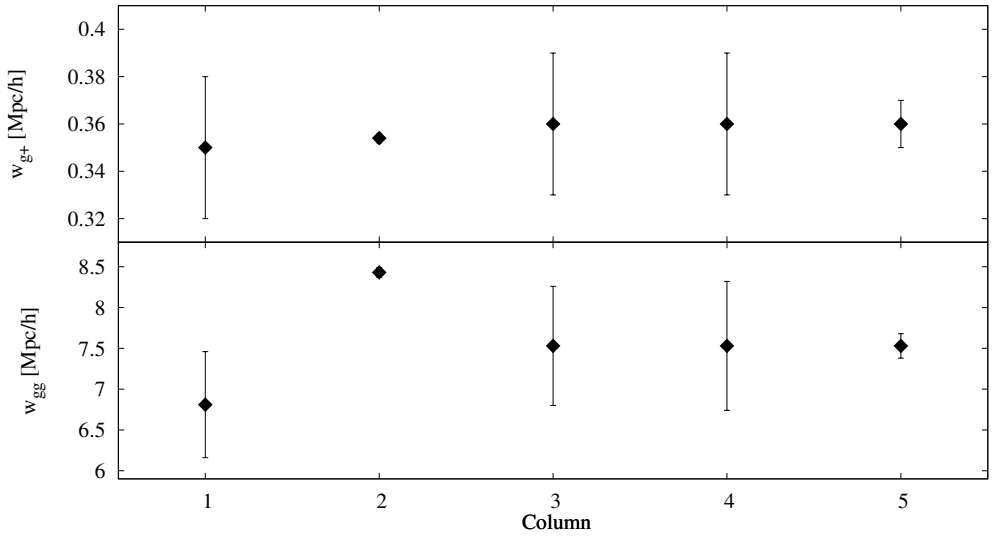


Figure C.2: Visualization of the results presented in Table C.1. The first realisation corresponds to the Millennium simulation, the second to the Millennium-XXL simulation, the third to the average of the 27 MS-size sub-boxes considering the average of the errors as the error bar, the fourth to the average of the 27 MS-size sub-boxes considering the standard deviation of data as the error bar, and the fifth to the average of the 27 MS-size sub-boxes considering the standard error on the mean as the error bar. Discussion is present in the caption of Table C.1 and in the text.

Acknowledgements

We would like to thank thank Edo van Uitert and Björn Malte Schäfer for being always available for useful discussion, and Björn in particular for building up the main points of the theory section.

I personally thank Sabino, for his constant support, Benjamin, for patiently sharing his precious time, and the city of London and its people, for harbouring me for the last five months.

I must then thank Alessandro, *fenomeno della diagnostica*, Corrado, *piedipiatti oltraggiosamente latente in zone innevate alpine*, Debora, *telematicamente esalante gas*, Emma, *donna aponense nascostamente emigrata socialmente esemplare*, Federica, *colorati animali rari amante*, Federica, *doverosamente oggi riportata impunemente*, Giancarlo, *diligente acconciatore designato*, Gianluca, *bravamente recuperante ossigeno*, Giovanni, *opportunamente battente dita*, Irene, *gelata in Norvegia guardando estere regioni*, Maddalena, *felicemente osservante saggi testi egregiamente rilegati*, Marco, *operaio molto oberato*, Martina, *purtroppo ulteriormente zelantemente zona alternante*, Massimiliano, *attivamente governante eserciti*, Nevia, *madre ultra minuziosa*, Sofia, *per aver lottato estenuamente senza troppe risposte avere*, Stefano, *favoloso regista occasionalmente complottista impuro obiettivamente*, Laura, *per aver troppo atteso teneri abbracci*, and all the others.

Bibliography

[1] D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris, and J. D. McEwen. How Isotropic is the Universe? *Physical Review Letters*, 117(13):131302, September 2016.

[2] A. Kiessling, M. Cacciato, B. Joachimi, D. Kirk, T. D. Kitching, A. Leonard, R. Mandelbaum, B. M. Schäfer, C. Sifón, M. L. Brown, and A. Rassat. Galaxy Alignments: Theory, Modelling & Simulations. *Space Sci. Rev.*, 193:67–136, November 2015.

[3] M. A. Troxel and M. Ishak. The intrinsic alignment of galaxies and its impact on weak gravitational lensing in an era of precision cosmology. *Phys. Rep.*, 558:1–59, February 2015.

[4] B. Joachimi, M. Cacciato, T. D. Kitching, A. Leonard, R. Mandelbaum, B. M. Schäfer, C. Sifón, H. Hoekstra, A. Kiessling, D. Kirk, and A. Rassat. Galaxy Alignments: An Overview. *Space Sci. Rev.*, 193:1–65, November 2015.

[5] D. Kirk, M. L. Brown, H. Hoekstra, B. Joachimi, T. D. Kitching, R. Mandelbaum, C. Sifón, M. Cacciato, A. Choi, A. Kiessling, A. Leonard, A. Rassat, and B. M. Schäfer. Galaxy Alignments: Observations and Impact on Cosmology. *Space Sci. Rev.*, 193:139–211, November 2015.

[6] LSST Science Collaboration, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, and more. LSST Science Book, Version 2.0. *ArXiv e-prints*, December 2009.

[7] R. Laureijs, J. Amiaux, S. Arduini, J. . Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, and more. Euclid Definition Study Report. *ArXiv e-prints*, October 2011.

[8] A. Heavens, A. Refregier, and C. Heymans. Intrinsic correlation of galaxy shapes: implications for weak lensing measurements. *MNRAS*, 319:649–656, December 2000.

[9] R. A. C. Croft and C. A. Metzler. Weak-Lensing Surveys and the Intrinsic Correlation of Galaxy Ellipticities. *ApJ*, 545:561–571, December 2000.

[10] C. Heymans, M. White, A. Heavens, C. Vale, and L. Van Waerbeke. Potential sources of contamination to weak lensing measurements: constraints from N-body simulations. *MNRAS*, 371:750–760, September 2006.

[11] E. Semboloni, C. Heymans, L. van Waerbeke, and P. Schneider. Sources of contamination to weak lensing three-point statistics: constraints from N-body simulations. *MNRAS*, 388:991–1000, August 2008.

[12] D. Kirk, A. Rassat, O. Host, and S. Bridle. The cosmological impact of intrinsic alignment model choice for cosmic shear. *MNRAS*, 424:1647–1657, August 2012.

[13] M. Kuhlen, J. Diemand, and P. Madau. The Shapes, Orientation, and Alignment of Galactic Dark Matter Subhalos. *ApJ*, 671:1135–1146, December 2007.

[14] J. Lee, V. Springel, U.-L. Pen, and G. Lemson. Quantifying the cosmic web - I. The large-scale halo ellipticity-ellipticity and ellipticity-direction correlations. *MNRAS*, 389:1266–1274, September 2008.

[15] M. D. Schneider, C. S. Frenk, and S. Cole. The shapes and alignments of dark matter halos. *JCAP*, 5:030, May 2012.

[16] B. Joachimi, E. Semboloni, P. E. Bett, J. Hartlap, S. Hilbert, H. Hoekstra, P. Schneider, and T. Schrabback. Intrinsic galaxy shapes and alignments - I. Measuring and modelling COSMOS intrinsic galaxy ellipticities. *MNRAS*, 431:477–492, May 2013.

[17] B. Joachimi, E. Semboloni, S. Hilbert, P. E. Bett, J. Hartlap, H. Hoekstra, and P. Schneider. Intrinsic galaxy shapes and alignments - II. Modelling the intrinsic alignment contamination of weak lensing surveys. *MNRAS*, 436:819–838, November 2013.

[18] S. Codis, R. Gavazzi, Y. Dubois, C. Pichon, K. Benabed, V. Desjacques, D. Pogosyan, J. Devriendt, and A. Slyz. Intrinsic alignment of simulated galaxies in the cosmic web: implications for weak lensing surveys. *MNRAS*, 448:3391–3404, April 2015.

[19] M. Velliscig, M. Cacciato, J. Schaye, H. Hoekstra, R. G. Bower, R. A. Crain, M. P. van Daalen, M. Furlong, I. G. McCarthy, M. Schaller, and T. Theuns. Intrinsic alignments of galaxies in the EAGLE and cosmo-OWLS simulations. *MNRAS*, 454:3328–3340, December 2015.

[20] N. Chisari, S. Codis, C. Laigle, Y. Dubois, C. Pichon, J. Devriendt, A. Slyz, L. Miller, R. Gavazzi, and K. Benabed. Intrinsic alignments of galaxies in the Horizon-AGN cosmological hydrodynamical simulation. *MNRAS*, 454:2736–2753, December 2015.

[21] A. Tenneti, R. Mandelbaum, and T. Di Matteo. Intrinsic alignments of disc and elliptical galaxies in the MassiveBlack-II and Illustris simulations. *MNRAS*, 462:2668–2680, November 2016.

[22] S. Hilbert, D. Xu, P. Schneider, V. Springel, M. Vogelsberger, and L. Hernquist. Intrinsic alignments of galaxies in the Illustris simulation. *MNRAS*, 468:790–823, June 2017.

[23] B. Joachimi, R. Mandelbaum, F. B. Abdalla, and S. L. Bridle. Constraints on intrinsic alignment contamination of weak lensing surveys using the MegaZ-LRG sample. *A&A*, 527:A26, March 2011.

[24] J. Hao, J. M. Kubo, R. Feldmann, J. Annis, D. E. Johnston, H. Lin, and T. A. McKay. Intrinsic Alignment of Cluster Galaxies: The Redshift Evolution. *ApJ*, 740:39, October 2011.

[25] C. Li, Y. P. Jing, A. Faltenbacher, and J. Wang. The Detection of the Large-scale Alignment of Massive Galaxies at $z \sim 0.6$. *ApJ*, 770:L12, June 2013.

[26] S. Singh, R. Mandelbaum, and S. More. Intrinsic alignments of SDSS-III BOSS LOWZ sample galaxies. *MNRAS*, 450:2195–2216, June 2015.

[27] E. van Uitert and B. Joachimi. Intrinsic alignment of redMaPPer clusters: cluster shape-matter density correlation. *MNRAS*, 468:4502–4512, July 2017.

[28] C. M. Hirata, R. Mandelbaum, M. Ishak, U. Seljak, R. Nichol, K. A. Pimbblet, N. P. Ross, and D. Wake. Intrinsic galaxy alignments from the 2SLAQ and SDSS surveys: luminosity and redshift scalings and implications for weak lensing surveys. *MNRAS*, 381:1197–1218, November 2007.

[29] R. Mandelbaum, C. Blake, S. Bridle, F. B. Abdalla, S. Brough, M. Colless, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, K. Glazebrook, B. Jelliffe, R. J. Jurek, I.-H. Li, B. Madore, C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, and T. Wyder. The WiggleZ Dark Energy Survey: direct constraints on blue galaxy intrinsic alignments at intermediate redshifts. *MNRAS*, 410:844–859, January 2011.

[30] Y. P. Jing. Intrinsic correlation of halo ellipticity and its implications for large-scale weak lensing surveys. *MNRAS*, 335:L89–L93, October 2002.

[31] J. Binney and S. Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 2008.

[32] M. Bartelmann and P. Schneider. Weak gravitational lensing. *Phys. Rep.*, 340:291–472, January 2001.

[33] C. M. Hirata and U. Seljak. Intrinsic alignment-lensing interference as a contaminant of cosmic shear. *Phys. Rev. D*, 70(6):063526, September 2004.

[34] J. Blazek, M. McQuinn, and U. Seljak. Testing the tidal alignment model of galaxy intrinsic alignment. *JCAP*, 5:010, May 2011.

[35] J. Blazek, Z. Vlah, and U. Seljak. Tidal alignment of galaxies. *JCAP*, 8:015, August 2015.

[36] D. Blas, J. Lesgourgues, and T. Tram. The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes. *JCAP*, 7:034, July 2011.

[37] V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005.

[38] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394, May 1985.

[39] V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at $z = 0$. *MNRAS*, 328:726–750, December 2001.

[40] R. E. Angulo, V. Springel, S. D. M. White, A. Jenkins, C. M. Baugh, and C. S. Frenk. Scaling relations for galaxy clusters in the Millennium-XXL simulation. *MNRAS*, 426:2046–2062, November 2012.

[41] G. Despali, G. Tormen, and R. K. Sheth. Ellipsoidal halo finders and implications for models of triaxial halo formation. *MNRAS*, 431:1143–1159, May 2013.

[42] M. Bonamigo, G. Despali, M. Limousin, R. Angulo, C. Giocoli, and G. Soucail. Universality of dark matter haloes shape over six decades in mass: insights from the Millennium XXL and SBARBINE simulations. *MNRAS*, 449:3171–3182, May 2015.

[43] C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-body Simulations. *MNRAS*, 271:676, December 1994.

[44] L. Jiang, J. C. Helly, S. Cole, and C. S. Frenk. N-body dark matter haloes with simple hierarchical histories. *MNRAS*, 440:2115–2135, May 2014.

[45] P. Bett, V. Eke, C. S. Frenk, A. Jenkins, J. Helly, and J. Navarro. The spin and shape of dark matter haloes in the Millennium simulation of a Λ cold dark matter universe. *MNRAS*, 376:215–232, March 2007.

[46] M. J. Pereira, G. L. Bryan, and S. P. D. Gill. Radial Alignment in Simulated Clusters. *ApJ*, 672:825–833, January 2008.

[47] G. Despali, C. Giocoli, and G. Tormen. Some like it triaxial: the universality of dark matter halo shapes and their evolution along the cosmic time. *MNRAS*, 443:3208–3217, October 2014.

[48] T. Schrabback, S. Hilbert, H. Hoekstra, P. Simon, E. van Uitert, T. Erben, C. Heymans, H. Hildebrandt, T. D. Kitching, Y. Mellier, L. Miller, L. Van Waerbeke, P. Bett, J. Coupon, L. Fu, M. J. Hudson, B. Joachimi, M. Kilbinger, and K. Kuijken. CFHTLenS: weak lensing constraints on the ellipticity of galaxy-scale matter haloes and the galaxy-halo misalignment. *MNRAS*, 454:1432–1452, December 2015.

[49] S. D. Landy and A. S. Szalay. Bias and variance of angular correlation functions. *ApJ*, 412:64–71, July 1993.

[50] A. Taylor, B. Joachimi, and T. Kitching. Putting the precision in precision cosmology: How accurate should your data covariance matrix be? *MNRAS*, 432:1928–1946, July 2013.

[51] E. S. Rykoff, E. Rozo, M. T. Busha, C. E. Cunha, A. Finoguenov, A. Evrard, J. Hao, B. P. Koester, A. Leauthaud, B. Nord, M. Pierre, R. Reddick, T. Sadibekova, E. S. Sheldon, and R. H. Wechsler. redMaPPer. I. Algorithm and SDSS DR8 Catalog. *ApJ*, 785:104, April 2014.

[52] T. Okumura, Y. P. Jing, and C. Li. Intrinsic Ellipticity Correlation of SDSS Luminous Red Galaxies and Misalignment with Their Host Dark Matter Halos. *ApJ*, 694:214–221, March 2009.

[53] J. Bailin, D. Kawata, B. K. Gibson, M. Steinmetz, J. F. Navarro, C. B. Brook, S. P. D. Gill, R. A. Ibata, A. Knebe, G. F. Lewis, and T. Okamoto. Internal Alignment of the Halos of Disk Galaxies in Cosmological Hydrodynamic Simulations. *ApJ*, 627:L17–L20, July 2005.

[54] S. Kazantzidis, A. V. Kravtsov, A. R. Zentner, B. Allgood, D. Nagai, and B. Moore. The Effect of Gas Cooling on the Shapes of Dark Matter Halos. *ApJ*, 611:L73–L76, August 2004.

[55] A. Tenneti, R. Mandelbaum, T. Di Matteo, Y. Feng, and N. Khandai. Galaxy shapes and intrinsic alignments in the MassiveBlack-II simulation. *MNRAS*, 441:470–485, June 2014.

[56] M. Velliscig, M. Cacciato, J. Schaye, R. A. Crain, R. G. Bower, M. P. van Daalen, C. Dalla Vecchia, C. S. Frenk, M. Furlong, I. G. McCarthy, M. Schaller, and T. Theuns. The alignment and shape of dark matter, stellar, and hot gas distributions in the EAGLE and cosmo-OWLS simulations. *MNRAS*, 453:721–738, October 2015.

[57] A. Tasitsiomi, A. V. Kravtsov, R. H. Wechsler, and J. R. Primack. Modeling Galaxy-Mass Correlations in Dissipationless Simulations. *ApJ*, 614:533–546, October 2004.

[58] R. Mandelbaum, C. M. Hirata, M. Ishak, U. Seljak, and J. Brinkmann. Detection of large-scale intrinsic ellipticity-density correlation from the Sloan Digital Sky Survey and implications for weak lensing surveys. *MNRAS*, 367:611–626, April 2006.

[59] N. Roche and S. A. Eales. The angular correlation function and hierarchical moments of ˜70000 faint galaxies to R=23.5. *MNRAS*, 307:703–721, August 1999.