Probing the darkness : the link between baryons and dark matter
Velliscig, M.

Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/36109 holds various files of this Leiden University dissertation

Author: Velliscig, Marco
Title: Probing the darkness : the link between baryons and dark matter
Issue Date: 2015-11-11

Probing the darkness:
The link between baryons and dark matter

Probing the darkness:
The link between baryons
and dark matter

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 11 november 2015
klokke 10.00 uur

door

Marco Velliscig

geboren te Udine, Italië
in 1985

Promotiecommissie

Promotor: Prof. dr. Joop Schaye
Co-promotor: Dr. Marcello Cacciato

Overige leden: Prof. dr. Koen Kuijken   University of Leiden
                     Prof. dr. Sarah Bridle   University of Manchester
                     Dr. Henk Hoekstra   University of Leiden
                     Dr. Benjamin Joachimi   University College London

Alla mia famiglia
senza la quale non sarei niente

Contents

1

Introduction

1.1 Numerical simulations in astrophysics

In a field of research such as astrophysics that studies objects beyond our reach, the role of direct experimentation is very limited. Simulations often take the role of experimentation in astrophysics, and allow us to conduct trials while modifying parameters in a manner similar to experiments. For an experiment to be successful, it needs to validate a hypothesis by confronting the predictions with real data. In the same way, simulations need to be compared against observations to understand to which degree they can reproduce the observables in the Universe.

If we were able to correctly simulate every process that has astrophysical significance, we could create a simulation that reproduces every possible observable, and we would be confident in our understanding of the physics governing our Universe. Of course, such a scenario is out of our reach and will likely remain so for many years. In the meantime, simulations can be used to guide the interpretation of observations, and for the design of new observational campaigns and instruments.

1.1.1 Dark matter-only simulations

Even the most basic cosmological simulations need to include the most important process for astrophysics: gravity. For this reason, N-body simulations were historically the first attempt to simulate our Universe. In N-body simulations all of the mass in the universe is only able to interact gravitationally. This approach, even if simplistic, has proven to be very successful for predictions of the large-scale distribution of matter, the halo abundance in a unit volume (the halo mass function), and the matter density profiles of haloes.

Moreover, even in recent years, only with dark matter simulations is it possible to simulate large volumes of the Universe with high resolution, or many realisations of small volumes with different cosmological parameters. For this reason, dark matter haloes in N-body simulation are often 'painted' with galaxies following simple prescriptions often calibrated to reproduce observables, an approach called semi-analytical modelling.

Even in this basic approach, some assumptions need to be made. For example, the particle mass of the simulated dark matter is many orders of magnitude greater than any theoretical expectation. This simplification is necessary because, even when neglecting hydrodynamics, the number of calculations required to evolve a system scales with the number of particles. Therefore, to keep the computational times reasonable the number of particles must be limited.

For a complete review on the topic please see Frenk & White (2012).

1.1.2 Hydrodynamical simulations

The next ingredient that needs to be included in simulations is the physics of the gas. However, before the processes that regulate the physical state of the gas can be included, the gas

itself should be modelled to follow the basic hydrodynamical equations. There are usually two different approaches to solve the hydrodynamical equations in simulations: the Eulerian approach and the Lagrangian approach. In the Eulerian approach, the cosmological volume is divided in small, discrete sub volumes, inside which the properties of the gas are simulated. In order to achieve higher resolutions, these volumes are subdivided in smaller parts where needed. In the Lagrangian approach, the mass of the gas is divided into particles that can interact with each other, and these are followed in time. The Lagrangian approach offers the advantage that the mass resolution is fixed and in turn the most dense parts of the Universe are naturally simulated with more particles. In this thesis, we use Lagrangian simulations realized using a technique known as smooth particle hydrodynamics (SPH, for a recent review see Springel 2010).

Baryonic physics in simulations

The physical state of the gas can be influenced by many processes, most of which are related to star formation and evolution. Stars produce radiation which can heat the gas and exert radiation pressure. Moreover the radiation from stars can change the ionisation balance of the gas and hence the rate at which it cools. Massive stars explode as supernovae (SNe) which can drive small-scale turbulence in the gas and produce large-scale outflows, a process referred to as SN feedback. Stars also produce heavy elements and dust which change the rate at which gas cools. The effect of star formation on the gas can also influence the star formation itself, making very difficult to predict it without self-consistently simulating both the star formation and the physical processes that alter the physical state of the gas. A similar role is played by the formation and evolution of supermassive black holes (SMBHs). Gas accretion onto SMBHs in the centres of galaxies can also result in radiative, thermal and mechanical feedback.

Many of the processes that are included in hydrodynamical simulations such as radiative cooling, photo-heating, star formation, metal enrichment are happening on scales that are orders of magnitude smaller than the scales which can be reasonably resolved in simulations. For this reason, such processes are usually taken into account using so called 'subgrid' models. As in the case of N-body simulations, the gas is simulated in parcels that represent millions of solar masses of gas. With current computational capabilities it would be unthinkable to model gas at the atomic or molecular level. Therefore, we make use of models that translate molecular processes to much larger mass scales.

The simultaneous simulation of all these physical processes has a large impact on the computational cost of the simulations. Therefore, hydrodynamical simulations are often carried out in smaller boxes or with lower resolution than their N-body counterparts. On the other hand, hydrodynamical simulations allow one to self-consistently simulate the baryonic physics that ultimately produces the luminous matter that we see and study in our Universe.

OWLS and EAGLE hydrodynamical cosmological simulations

The analysis performed in this thesis is mostly based on two sets of hydrodynamical cosmological simulations: OWLS and EAGLE. OWLS stands for the OverWhelmingly Large Simulations project (OWLS; Schaye et al. 2010) and it consists of a suite of cosmological simulations with varying simulation volume and resolutions. OWLS is well-suited to study baryonic effects as it consists of a wide range of models that were run from identical initial conditions, but employing a wide variety of recipes for the uncertain baryonic processes. OWLS also includes a N-body only version that allow us to directly compare with the results obtained with the inclusion of baryon physics. In order to extend the mass range that can

be studied, we often employed a larger volume, lower resolution versions of a subset of the OWLS models (called cosmo-OWLS; see Le Brun et al. 2014; McCarthy et al. 2014).

EAGLE stands for Evolution and Assembly of GaLaxies and their Environments (Schaye et al. 2015; Crain et al. 2015). A new approach adopted by EAGLE is to calibrate the feedback efficiency to reproduce the observed number of galaxies per unit volume (the galaxy stellar mass function), since most of the radiative losses are due to unresolved physics and so can not be simulated by first principles. This results in a successful mapping between the stellar mass and the halo mass, a task that has proven to be difficult to achieve by simulations.

Combining OWLS and EAGLE provides us the unique opportunity to span a wide range of masses and spatial scales. Moreover, we have access to sufficient cosmological volume and resolution to study the reliability and the applicability of our results to the different astrophysical questions addressed in this thesis.

The interaction between baryons and dark matter

For many purposes it is a reasonable approximation to assume that the effect of baryon physics on the matter distribution is negligible. This approximation is valid for instance, in the intergalactic medium and on the outskirts of galaxy clusters where the gas has a long cooling time and approximately traces the dark matter. However, it is clear that on small scales and in lower mass haloes, where the gas condenses to high densities due to cooling, baryonic processes such as galactic winds can have important effects on the matter distribution. Moreover, the distribution of the dark matter will itself adjust to the resulting change in the gravitational potential.

In **Chapter 2** we examine the effect of baryon physics on the masses and profiles of haloes and on the halo mass function as inferred by hydrodynamical simulations. The study of how the inclusion of baryonic physics can alter the properties of cosmological haloes represents the first step in exploring the connection between baryons and their host haloes.

1.2   Gravitational lensing

Gravitational lensing (GL) represents a powerful tool to measure what cannot be seen directly but what is an essential part of the mass of the Universe: the dark matter.

Here we give an introduction to gravitational lensing, focusing mostly on weak lensing and on the aspects and applications most relevant for this thesis. We refer to Bartelmann & Schneider (2001) for a more comprehensive review of the vast field of weak gravitational lensing.

Gravitational lensing is the effect of the gravitational potential on the light path of photons travelling trough it. "Gravitational" refers to the fact that this effect depend solely on the gravitational potential, making it insensitive to the dynamical state or the type of matter that generates it. "Lensing" stands for the fact that this effect acts on the light in the same way as a optical lens does. The light travelling from a distant source is lensed by the intervening gravitational field, and the image on the sky appears to be displaced, magnified and distorted by the interaction of the photons with the gravitational potential. Gravitational lensing can be exploited to measure different properties of the matter in the Universe, some of which we are going to briefly review in the next sections.
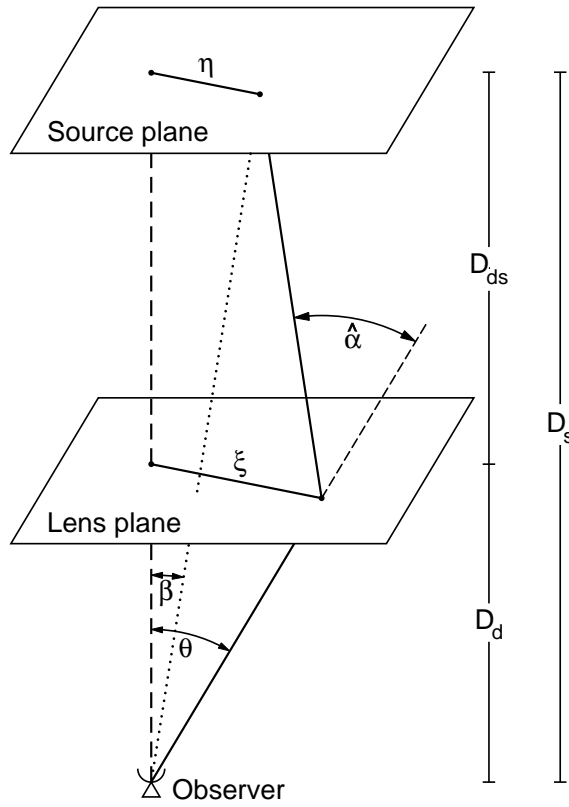
Figure 1.1: Basic schematics of a lensed light ray (from Bartelmann & Schneider 2001).

1.2.1   Single lens

The basic concept of GL can be reviewed by analysing the effect of a single lens on the light path of a single light ray. The thin lens approximation is commonly applied, which is valid if the mass distribution is extended over a region that is small compared to the distance between source and lens, and between lens and observer. Using this approximation, we can assume that the total mass of the lens is lying in a plane, and that a light ray passing through will be deflected once and instantaneously. The light rays, which are smoothly curved in the neighbourhood of the lens, can therefore be replaced by two straight rays that change their direction only in the plane of the lens. In Fig. 1.1 we show a schematic representation of a lens system. The source of photons lies in the source plane at the position η and at the angular diameter distance D_s (or redshift z_s) away from the observer. The lens lies in the lens plane at a distance D_l from the observer and D_{ls} from the source. In the case where the lens is not present, the light ray emitted from the source would be observed at the angle β. The effect of the lens is to deflect the light ray, at the impact parameter ξ, by an angle of $\hat{\alpha}$. This results in the light ray being seen by the observer at an angle θ. The deflection is then:

$$\vec{\hat{\alpha}} = \frac{4GM}{c^2 \, \vec{\xi}} \, , \tag{1.1}$$

where M is the mass of the lens, G is the gravitational constant and $\vec{\xi}$ is the impact parameter vector. Therefore, the deflection angle not only depends on the mass of the lens but also on the impact parameter.

From this equation it is possible to understand the basic effect of gravitational lensing. Due to the presence of the lens, more light rays are converged towards the observer than otherwise, resulting in a magnification of the source. Since the deflection angle depends on the impact parameter, light rays that are passing at different distances from the source are deflected at different angles, producing a distortion in the original shape of the source. In extreme cases when the source is perfectly aligned with the lens this will produce a circular image, known as Einstein ring.

The angular diameter distance is defined as the ratio between the projected separation and the angle under which it is seen in the sky, and so $\vec{\theta} = \vec{\xi}/D_l$ and $\vec{\beta} = \vec{\eta}/D_s$. The projected separation of the source depends on the the line of sight η, the impact parameter and the deflection angle as:

$$\vec{\eta} = \frac{D_s}{D_l} \, \vec{\xi} - D_{ls}\vec{\hat{\alpha}}(\vec{\xi}) \, . \tag{1.2}$$

By using the definition of $\vec{\beta}$ and $\vec{\theta}$, this can be rewritten as the commonly known lens equation:

$$\vec{\beta} = \vec{\theta} - \frac{D_{ls}}{D_s} \, \vec{\hat{\alpha}}(D_l\vec{\theta}) \equiv \vec{\theta} - \vec{\alpha}(\vec{\theta}) \, , \tag{1.3}$$

where $\vec{\alpha}(\vec{\theta})$ is the deflection angle scaled by the distances from the lens and observer and the lens and source. The convergence parameter κ is defined as:

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_d\vec{\theta})}{\Sigma_{cr}} \quad \text{with} \quad \Sigma_{cr} = \frac{c^2}{4\pi G} \frac{D_s}{D_d \, D_{ds}} \, , \tag{1.4}$$

where Σ_{cr} is the critical surface mass density (which depends on the redshifts of source and lens). The lens equation can have more than one $\vec{\theta}$ for a fixed $\vec{\beta}$ if the convergence parameter $\kappa > 1$, in which case the lensing is *strong*. In this thesis we focus on weak lensing for which $\kappa \ll 1$.

1.2.2   Statistical weak lensing

The net effect of gravitational lensing is, through the deflection of light rays, a re-mapping of the image of a source on to the plane of the observer. If the source image is small compared to the scale on which the properties of the lens change, the mapping can be described by the deflection matrix. The deflection matrix maps the unlensed surface brightness of the source $I(x, y)$ to the observed one $I'(x', y')$, as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = (1 - \kappa) \begin{bmatrix} 1 - g_1 & -g_2 \\ -g_2 & 1 + g_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \tag{1.5}$$

where (x, y) are the true coordinates, (x', y') are the distorted ones, and κ is the convergence. The reduced shear g is related to the shear by $g_1, g_2 = (\gamma_1, \gamma_2)/(1 - \kappa)$ and in the case of weak lensing, where $\kappa \ll 1$, $\gamma \approx g$. The shear describes the change in the observed ellipticity of a galaxy, ϵ^s, as:

$$\epsilon = \epsilon^s + \gamma. \tag{1.6}$$

The value of the shear can be estimated by averaging a large number of observed ellipticities, if galaxies are randomly oriented, as $\langle \epsilon \rangle = \gamma$. This means that gravitational lensing produces an apparent alignment of galaxy shapes. However, gravitational tidal forces can induce a distortion in the shape of galaxies that can produce an *intrinsic* alignment, which can be a non-negligible fraction of the *apparent* alignment induced by gravitational lensing. The intrinsic alignment can be important for all types of weak lensing analyses, since by its very nature the effect of weak lensing is too small to be detected using the shape of a single galaxy, and so it must be studied statistically. The need for a statistical study to obtain the weak lensing signal is at the very heart of why the intrinsic alignment of galaxies can be a source of nuisance for weak lensing studies.

Galaxy intrinsic and apparent alignments

Even if GL can be expressed in its simpler form as the effect of a single source on a single light ray, in reality the Universe is filled with light and matter. Nearly all of the light emitted by distant galaxies could be influenced by intervening matter, enabling the study of the matter distribution of the Universe. Moreover, by dividing sources in redshift, and by studying the distortion induced by gravitational lensing on their images, is possible to probe how the distribution of matter changes throughout cosmic time, providing us with the opportunity to constrain the cosmological parameters of the Universe. This application of weak lensing is called *cosmic shear*. In practice, cosmic shear connects the lensing power spectrum to the matter power spectrum, by measuring the galaxy shapes correlations induced by GL (see Kaiser 1992; Bartelmann & Schneider 2001; Schneider 2005; Hoekstra & Jain 2008, for reviews on the applications of weak lensing to constrain cosmological parameters).

Here we will not go into the details of cosmic shear, but instead highlight one of the main complications for the cosmic shear measurements: the intrinsic alignment.

Cosmic shear measurements are obtained in the form of projected 2-point correlation functions (or their equivalent angular power spectra) between galaxy shapes.

$$\langle \epsilon \epsilon \rangle = \langle \gamma \gamma \rangle + \langle \gamma \epsilon^s \rangle + \langle \epsilon^s \gamma \rangle + \langle \epsilon^s \epsilon^s \rangle \tag{1.7}$$

$$= GG + GI + IG + II. \tag{1.8}$$

If we assume that galaxies are not intrinsically oriented towards one another, then the only correlations in the shape and orientation of observed galaxies are due to the gravitational lensing effect of the intervening mass distribution between the sources and the observer, $\langle \gamma \gamma \rangle$.

Here the only nonzero term is the GG (shear-shear) auto correlation. In the case of a non-negligible intrinsic alignment of galaxies, the second term is also nonzero, i.e. part of the correlation between the shape and orientation of galaxies is *intrinsic*. If the same gravitational forces that shear the light emitted from a galaxy also tidally influence the intrinsic shape of other galaxies, then this will produce a nonzero cross-correlation between shear and intrinsic shape (GI). The term IG is zero since a foreground galaxy cannot be lensed by the same structure that is tidally influencing a background galaxy, unless their respective positions along the line of sight is confused due to large errors in the redshift measurements. In **Chapter 4** we study the *intrinsic* alignment of galaxies in hydrodynamical cosmological simulations.

Most intrinsic alignment theories predict the alignment of dark matter structures. On the other hand, the alignment of galaxies is measured using the light of galaxies, and so it is often assumed that the light traces dark matter, or alternatively that the misalignment between haloes and galaxies is known. The misalignment between galaxies and their haloes is studied in **Chapter 3** using hydrodynamical cosmological simulations.

Galaxy-galaxy lensing

The shear is measured with respect to the projected separation vector of the source-lens pair: the tangential shear, γ_t, defined as

$$\gamma_t = \gamma \cos(2\Phi), \tag{1.9}$$

where Φ is the position angle.

One of the advantages of measuring the tangential shear is that it is directly related to the excess surface mass density (ESD)

$$\gamma_t(r_p) = \frac{\Delta\Sigma(r_p)}{\Sigma_{cr}}, \tag{1.10}$$

that in turn relates to the surface density as

$$\Delta\Sigma(r_p) \equiv \bar{\Sigma}(< r_p) - \Sigma(r_p), \tag{1.11}$$

where r_p is the projected distance from the centre of the halo, $\bar{\Sigma}(< r_p)$ is the surface density within r_p and $\Sigma(r_p)$ is the surface density at r_p. By stacking the shear signal in concentric rings, the radial profile of the ESD can be studied. Galaxy-galaxy weak lensing (GGL) is the study of the ESD profile of galaxies measured through staking the shear signal of background galaxies around the position of lenses. Since the shear of a single lens system is usually too small to be detected, the properties of individual galaxies can not be constrained with GGL. Instead, if the lenses are chosen according to a given property, for example the stellar mass, the statistical property of a selected population of lens galaxies can be studied. In recent years, galaxy-galaxy weak lensing has become a viable way to statistically constrain the mass, density profiles and ellipticity of the dark matter haloes. In **Chapter 5** we compare the ESD from simulations with the galaxy-galaxy lensing signal observed using background galaxies imaged by the KiDS survey around spectroscopically confirmed foreground galaxies from the GAMA survey.

Another application of galaxy-galaxy lensing is the measure of the average shape of haloes. For this application, in the stacking process the lens galaxies are re-oriented according to the direction of the major axis of their light distribution. In this way, for a triaxial halo the value of the tangential shear along the major and minor axes of the galaxy is expected to be different. It is possible to translate this difference into a constraint on the halo shape of galaxies. The halo shapes represent an interesting line of study because they are expected to be triaxial in the ΛCDM framework and spherical in some alternative cosmology theories. Of

course, the main assumption of these studies is that the major axis of light distribution aligns with the major axis of the halo. Therefore, the shape measurement can be affected by the presence of a misalignment between the galaxy and the dark matter halo.

1.3   This Thesis

Our first step in exploring the connection between baryons and their host haloes is to study how the inclusion of baryonic physics can alter the properties of haloes in a cosmological simulation. Along these lines, in **Chapter 2** we study the effect of baryon physics on the masses and profiles of haloes and consequently on their abundance in a unit volume (known as the halo mass function). We find that gas expulsion and the associated dark matter (DM) expansion induced by supernova-driven winds are important for haloes with masses $M_{200} \leq 10^{13} \, M_\odot$, lowering their masses by up to 20% relative to a DM-only model. AGN feedback has a significant impact on halo masses up to cluster scales ($M_{200} \sim 10^{15} \, M_\odot$). Baryon physics changes the total mass profiles of haloes out to several times the virial radius, a modification that can only capture by changing the the functional forms used to commonly fit halo profiles. The decrease in the total halo mass causes a decrease in the halo mass function of about 20%. The analysis presented in **Chapter 2** indicates that baryonic processes can significantly alter the properties of dark matter haloes. Therefore, their inclusion is essential in simulations aiming to give theoretical support to observations that are probing the total mass of haloes.

Gravitational lensing offers a way to detect dark matter haloes in observations. In this context, hydrodynamical cosmological simulations serve as a tool to study and mitigate possible shortcomings of gravitational lensing. In **Chapter 3** we make use of four hydrodynamical simulations, run in increasingly larger volumes and covering over four orders of magnitude in mass ($11 \leq \log_{10}(M_{200}/[\, h^{-1} \, M_\odot]) \leq 15$), to study the misalignment between galaxies and their host haloes as a function of radius and redshift. The galaxy-halo misalignment has profound implications for gravitational lensing studies that aim to constrain the shape of dark matter haloes, since lens galaxies are stacked according to the direction of their major axis under the assumption that it aligns with the major axis of the halo. The shape of the halo is constrained by examining the different shear signals along the major and the minor axes of the stacked galaxies. Therefore, the shape measurement can be affected by the presence of a misalignment between the galaxy and the dark matter halo for which the shape is measured. Moreover, the misalignment between galaxies and their haloes also represents a relevant field of study for the intrinsic alignment. In fact, most intrinsic alignment theories predict the alignment of dark matter structures and thus, in order to be tied to the observations, the misalignment between haloes and galaxies must be known. In our study, we found that galaxies align well with the local distribution of the total (but mostly dark) matter, however, the stellar distributions on galactic scales exhibit a significant misalignment with respect to their host haloes. This misalignment is reduced in the most massive haloes. This implies that the orientation of galaxies is a good tracer of the dark matter, only on comparable scales. On larger scales, a significant misalignment exists between galaxies and their haloes, which must be accounted for in weak lensing studies.

Another potential contaminant of weak gravitational lensing measurements is the alignment of galaxies between themselves, known as intrinsic galaxy alignment. In fact, in the limit of a very weak lensing signal, the distortion induced via gravitational forces (giving rise to an *intrinsic* alignment) can be a non-negligible fraction of the distortion due to the pure gravitational lensing effect (often termed *apparent* alignment). In **Chapter 4** we report results

for the *intrinsic* alignment of galaxies in hydrodynamical cosmological simulations. Specifically, we focus on the orientation-direction, which is the angle between the major axis and the direction of a nearby galaxy, and the orientation-orientation alignment, which is the angle between the major axes of nearby galaxies. We find that while the strength of the alignment is a strongly decreasing function of the distance between galaxies, it can remain significant up to ~ 100 Mpc, with more massive haloes demonstrating stronger alignment. We find a significant decrease in the alignment of galaxies when orientations are computed using only stars within the typical observable extent of a galaxy rather than using all stars associated with the subhalo. This difference may account for the common findings reported in the literature of galaxy alignments being systematically stronger in simulations than in observations. Particular care must be taken in the sample of stars that are considered, since this can ultimately result in very different alignment strengths. The orientation-orientation alignment is always weaker than the orientation-direction alignment, which suggests that galaxy alignment is driven by the position of nearby haloes, whereas the mutual orientations of nearby galaxies is a consequence of this effect.

Through galaxy-galaxy lensing, it is possible to directly measure the mass and mass profiles of haloes as well as the connection between the stellar and halo masses of galaxies. In **Chapter 5** we report the study of the galaxy-galaxy lensing signal for galaxies in the EAGLE simulations divided into six stellar mass bins. We compare the results from the simulations to the observed signal, which was measured using background galaxies imaged by the KiDS survey around spectroscopically confirmed foreground galaxies from the GAMA survey. The GAMA group catalogue offers us the possibility to compare the central and satellite contribution to the total signal separately. Overall, the predicted lensing signal is in broad agreement with the observations, as expected due to the fact that the EAGLE simulation has been calibrated to reproduce the observed $z \sim 0$ galaxy stellar mass function. We find good agreement between the data and predictions from EAGLE for both central and satellite galaxies. When satellite and central galaxies are analyzed jointly, the agreement worsens. This stems from the fact that the total GGL profile is a linear combination of central and satellite profiles with the satellite fraction as the linear coefficient, and the satellite fraction in the EAGLE simulation is always lower than that of the GAMA group catalogue.

Bibliography

Bartelmann, M., & Schneider, P. 2001, Phys. Rep., 340, 291

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

Frenk, C. S., & White, S. D. M. 2012, Annalen der Physik, 524, 507

Hoekstra, H., & Jain, B. 2008, Annual Review of Nuclear and Particle Science, 58, 99

Kaiser, N. 1992, ApJ, 388, 272

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, MNRAS, 441, 1270

McCarthy, I. G., Le Brun, A. M. C., Schaye, J., & Holder, G. P. 2014, MNRAS, 440, 3645

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schneider, P. 2005, ArXiv Astrophysics e-prints, astro-ph/0509252

Springel, V. 2010, ARA&A, 48, 391

2

The impact of galaxy formation on the total mass, mass profile and abundance of haloes

We use cosmological hydrodynamical simulations to investigate how the inclusion of physical processes relevant to galaxy formation (star formation, metal-line cooling, stellar winds, supernovae and feedback from Active Galactic Nuclei, AGN) change the properties of haloes, over four orders of magnitude in mass. We find that gas expulsion and the associated dark matter (DM) expansion induced by supernova-driven winds are important for haloes with masses $M_{200} < 10^{13} \, \mathrm{M_\odot}$, lowering their masses by up to 20% relative to a DM-only model. AGN feedback, which is required to prevent overcooling, has a significant impact on halo masses all the way up to cluster scales ($M_{200} \sim 10^{15} \, \mathrm{M_\odot}$). Baryon physics changes the total mass profiles of haloes out to several times the virial radius, a modification that cannot be captured by a change in the halo concentration. The decrease in the total halo mass causes a decrease in the halo mass function of about 20%. This effect can have important consequences for the abundance matching technique as well as for most semi-analytic models of galaxy formation. We provide analytic fitting formulae, derived from simulations that reproduce the observed baryon fractions, to correct halo masses and mass functions from DM-only simulations. The effect of baryon physics (AGN feedback in particular) on cluster number counts is about as large as changing the cosmology from WMAP7 to Planck, even when a moderately high mass limit of $M_{500} \approx 10^{14} \, \mathrm{M_\odot}$ is adopted. Thus, for precision cosmology the effects of baryons must be accounted for.

2.1 Introduction

N-body simulations using only gravitationally interacting dark matter (DM) particles have
played an important role in the development of astrophysical cosmology (e.g. Frenk & White
2012). DM simulations have for example been used to predict the large-scale distribution
of matter, the halo mass function, and the density profiles of haloes. The results from such
simulations form the basis for halo-based models and abundance matching techniques (e.g.
Seljak 2000; Cooray & Sheth 2002). DM simulations are also used as the starting point for
semi-analytic models that associate galaxies to DM haloes in post-processing, and then follow
the evolution of these galaxies according to different prescriptions that are calibrated such that
the model reproduces a limited set of observables (e.g. Baugh 2006).

Neglecting the hydrodynamics and the feedback processes that affect the gas compo-
nent also allows one to perform simulations with a dynamic range that would otherwise not
be achievable due to the higher computational cost associated with the inclusion of baryon
physics. Explicitly accounting for baryons in simulations means computing hydrodynamic
forces and including processes like radiative cooling, photo-heating, star formation, metal en-
richment, and also energetic feedback processes such as winds driven by supernovae and AGN
that are able to generate outflows and eject baryons from (the inner parts of) DM haloes.

For many purposes it is a reasonable approximation to assume that the effect of baryon
physics on the matter distribution is small, such as in the intergalactic medium (e.g., Theuns
et al. 2002; Viel et al. 2013) and on the outskirts of galaxy clusters where the gas has a long
cooling time and approximately traces the dark matter (e.g., Le Brun et al. 2014). However,
it is clear that on small scales and in lower mass haloes, where cooling can allow the gas
to condense to high densities, baryonic processes such as galactic winds can have important
effects on the matter distribution. Moreover, the distribution of the DM will itself adjust to the
resulting change in the gravitational potential. Indeed, it appears that the observed rotation
curves of dwarf galaxies cannot be reproduced by simulations that assume the standard cold
dark matter paradigm unless they include the effect of outflows (e.g. Governato et al. 2010).

In recent years hydrodynamical simulations have for example been used to quantify the
effect of baryons on the DM halo density profiles (e.g. Gnedin et al. 2004; Gustafsson et al.
2006; Duffy et al. 2010; Tissera et al. 2010), spins (e.g. Bett et al. 2010; Bryan et al. 2013),
shapes (e.g. Abadi et al. 2010; Kazantzidis et al. 2004; Read et al. 2008; Bryan et al. 2012,
2013), and substructure (e.g. Dolag et al. 2009; Romano-Díaz et al. 2009) of dark haloes,
as well as on the matter power spectrum (e.g. Jing et al. 2006; Rudd et al. 2008; Guillet
et al. 2010; van Daalen et al. 2011; Casarini et al. 2011) and the clustering of subhaloes (van
Daalen et al. 2014). Because the physics of galaxy formation is uncertain, it is important to
vary the parameters of the model. In particular, it has recently become clear that the efficient
feedback that is required to reproduce observations, and which is thought to be driven by star
formation and by AGN at low and high halo masses, respectively, leads to results that differ
qualitatively from the predictions of earlier simulations that suffered from overcooling. For
example, efficient feedback reduces, or even reverses, adiabatic contraction in the inner parts
of massive haloes (e.g. Duffy et al. 2010; Mead et al. 2010; Teyssier et al. 2011; Killedar et al.
2012; Martizzi et al. 2013). The ejection of baryons by outflows reduces the matter power
spectrum on remarkably large scales (van Daalen et al. 2011) with dramatic consequences for
future cosmological weak lensing studies (Semboloni et al. 2011, 2013; Zentner et al. 2013).

One of the most important quantities characterising the distribution of matter is the halo
mass function (HMF hereafter), i.e. the number density of haloes as a function of their mass.
The evolution of the HMF, in particular its massive end, is for example a powerful tool for
constraining cosmological parameters such as the dark energy equation of state, using future
large surveys such as eROSITA (Pillepich et al. 2012), XMM-XXL (Pierre et al. 2011), and

XCS (Mehrtens et al. 2012) in X-ray; Planck[1] using the Sunyaev-Zel'dovich effect; and DES[2], Euclid (Laureijs et al. 2011) and LSST[3] using weak lensing. For a given cosmology, the HMF is usually predicted using DM-only simulations (e.g. Jenkins et al. 2001; Reed et al. 2003; Warren et al. 2006; Lukić et al. 2007; Tinker et al. 2008). To exploit the capacity of upcoming surveys, the theoretical HMF needs to be calibrated at the per cent level (Wu et al. 2010). Therefore, even if the impact of baryon physics is only of the order of a per cent or more, it should be taken into account when computing the theoretical HMF.

The effect of baryon physics on the HMF has recently been studied using hydrodynamical simulations by, among others, Sawala et al. (2013) and Khandai et al. (2014) at low halo masses and by Stanek et al. (2009), Cui et al. (2012), Martizzi et al. (2014) and Cusworth et al. (2014) at high masses. Although the studies differ in their detailed findings (and sometimes on the sign of the effect), there is nevertheless a growing consensus that baryon physics will significantly affect the HMF.

Here we will use the suite of cosmological simulations from the OverWhelmingly Large Simulations project (OWLS; Schaye et al. 2010) to study the effect of galaxy formation on the mass function and internal structure of haloes more massive than $M_{200}^{\mathrm{crit}} = 10^{11.5} \, h^{-1} \, \mathrm{M_\odot}$. For this purpose we will also make use of new larger volume, lower resolution versions of a subset of the OWLS models (an extension of OWLS, called cosmo-OWLS; see Le Brun et al. 2014). OWLS is well-suited to study baryonic effects as it consists of a wide range of models that were run from identical initial conditions, but employing a wide variety of recipes for the uncertain baryonic processes. OWLS also includes a DM-only model as well as a model with AGN feedback that reproduces both optical and X-ray observations of groups and clusters of galaxies (McCarthy et al. 2010; Le Brun et al. 2014). These last two models are therefore particularly well-suited to our needs and we will employ them to provide fitting formulas that can be used to correct the HMFs predicted by DM-only models for the effect of baryons.

This paper is organised as follows. In § 2.2 we describe our simulations and explain the methods used to find haloes and to match them between different simulations. In § 2.3 we show how baryon physics, such as radiative cooling and feedback from supernovae and AGN alter the masses of haloes. In § 2.4 we provide analytic fitting formulae to correct the masses of haloes in the DM-only simulation for the effect of baryon physics. In § 2.5 we show the impact of baryon physics on the halo mass function and discuss the implications for cluster number count studies. In § 2.6 we compare our findings to those of previous studies. Finally, in § 2.7 we summarize and conclude.

---

[1]http://www.rssd.esa.int/index.php?project=Planck

[2]http://www.darkenergysurvey.org/

[3]http://www.lsst.org/lsst/

Simulation	L (h^{-1} Mpc)	$N_{particle}$	Cosmology	m_b (h^{-1} M$_\odot$)	m_{dm} (h^{-1} M$_\odot$)	Description
DMONLY	100	512^3	WMAP3	–	4.9×10^8	Only gravitationally interacting particles
NOSN_NOZCOOL	100	2×512^3	WMAP3	8.7×10^7	4.1×10^8	No SN feedback, primordial cooling
NOZCOOL	100	2×512^3	WMAP3	8.7×10^7	4.1×10^8	Cooling assumes primordial abundances
REF	100	2×512^3	WMAP3	8.7×10^7	4.1×10^8	SN feedback, metal line cooling, no AGN
WDENS	100	2×512^3	WMAP3	8.7×10^7	4.1×10^8	Wind mass loading and velocity depend on ρ_{gas}
AGN 8.0	100	2×512^3	WMAP3	8.7×10^7	4.1×10^8	Includes AGN
DMONLY L050N512	50	512^3	WMAP3	–	6.2×10^7	High-res version of DMONLY, smaller box
REF L050N512	50	2×512^3	WMAP3	1.1×10^7	5.1×10^7	High res version of REF, smaller box
DMONLY W7	100	512^3	WMAP7	–	5.6×10^8	Different cosmology w.r.t. DMONLY
REF W7	100	2×512^3	WMAP7	9.4×10^7	4.7×10^8	Different cosmology w.r.t. REF
AGN 8.0 W7	100	2×512^3	WMAP7	9.4×10^7	4.7×10^8	Different cosmology w.r.t. AGN 8.0
AGN 8.5 W7	100	2×512^3	WMAP7	9.4×10^7	4.7×10^8	Different heating temperature w.r.t. AGN 8.0 W7
DMONLY L400 W7	400	1024^3	WMAP7	–	4.5×10^9	Larger box, lower res w.r.t. DMONLY W7
REF L400 W7	400	2×1024^3	WMAP7	7.5×10^8	3.7×10^9	Larger box, lower res w.r.t. REF W7
AGN 8.0 L400 W7	400	2×1024^3	WMAP7	7.5×10^8	3.7×10^9	Larger box, lower res w.r.t. AGN 8.0 W7
AGN 8.5 L400 W7	400	2×1024^3	WMAP7	7.5×10^8	3.7×10^9	Larger box, lower res w.r.t. AGN 8.5 W7

Table 2.1: List of the simulations used in this work. Most simulations use a box of $100\,h^{-1}$ Mpc , with 2×512^3 particles. We carry out resolution tests using simulations with 8 times higher and lower mass resolution. We also use simulations with a different cosmology, WMAP7 instead of WMAP3, to see if our analysis is cosmology dependent. Finally, we take advantage of $400\,h^{-1}$ Mpc , 2×1024^3 version of the OWLS models to extend our analysis to higher halo masses.

2.2 Simulations

The analysis carried out in this paper is based on simulations that are part of the OverWhelmingly Large Simulations project (OWLS Schaye et al. 2010) which includes over 50 large, cosmological, hydrodynamical simulations run with a modified version of the smoothed particle hydrodynamics (SPH) code GADGET3 (last described in Springel 2005). The aim of the OWLS project is to explore the sensitivity of the theoretical predictions to both resolvable and 'subgrid' physics thought to be important for galaxy formation (such as supernova (SN) feedback, stellar mass loss, radiative cooling processes and AGN feedback) in fully self-consistent cosmological hydrodynamical simulations. In this section we will give a brief description of the simulations used in this paper and the physical processes implemented in each of them.

The simulations used in this work were run with either a WMAP3 cosmology (Spergel et al. 2007) $\{\Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s, h\} = \{0.238, 0.0418, 0.762, 0.74, 0.951, 0.73\}$ or a WMAP7 cosmology (Komatsu et al. 2011) $\{\Omega_m, \Omega_b, \Omega_\Lambda, \sigma_8, n_s, h\} = \{0.272, 0.0455, 0.728, 0.81, 0.967, 0.704\}$. Most simulations used in this work were run in periodic boxes of $100\,h^{-1}$ Mpc ($400\,h^{-1}$ Mpc) comoving, and each of the runs uses 512^3 (1024^3) dark matter and equally many baryonic particles (representing collisionless star or collisional gas particles). The particle masses in the 2×512^3 particle $100\ h^{-1}$ Mpc WMAP3 (1024^3 particle $400\,h^{-1}$ Mpc WMAP7) simulations are $4.06 \times 10^8\ h^{-1}\,\mathrm{M_\odot}$ ($3.75 \times 10^9\ h^{-1}\,\mathrm{M_\odot}$) for dark matter and $8.66 \times 10^7\ h^{-1}\,\mathrm{M_\odot}$ ($7.53 \times 10^8\ h^{-1}\,\mathrm{M_\odot}$) for baryons. Note, however, that baryonic particle masses change during the course of the simulation due to mass transfer from star to gas particles.

Comoving gravitational softenings were set to $1/25$ of the initial mean inter-particle spacing but were limited to a maximum physical scale of $2\,h^{-1}$ kpc ($4\,h^{-1}$ kpc) for the $100\,h^{-1}$ Mpc ($400\,h^{-1}$ Mpc) simulations. The switch from a fixed comoving to a fixed proper softening happens at $z = 2.91$ in all simulations. We used $N_{\mathrm{ngb}} = 48$ neighbours for the SPH interpolation.

The physical models considered here are (following the naming convention of Schaye et al. 2010):

- DMONLY: a dark matter only simulation, intended to simulate a set of particles that interact only gravitationally. Such simulations are commonly used to compute the HMF that forms the input of semi-analytic models and abundance matching studies. We use this simulation as a base and evaluate differences with respect to this model when baryon physics is added. Recall that a particle in this simulation becomes two particles in a baryonic simulation: namely a DM particle of mass $\frac{\Omega_m - \Omega_b}{\Omega_m} \times m^{\mathrm{dmonly}}$ and a gas particle of mass $\frac{\Omega_b}{\Omega_m} \times m^{\mathrm{dmonly}}$.

- REF: this is the reference model for the OWLS suite, but is not intended to be the 'best' model. This model includes most of the mechanisms that are thought to be important for the star formation history (see Schaye et al. 2010 for a detailed discussion), but not AGN feedback. The implementation of radiative cooling, star formation, supernova driven winds, and stellar evolution and mass loss have been described in Wiersma et al. (2009a), Schaye & Dalla Vecchia (2008), Dalla Vecchia & Schaye (2008), and Wiersma et al. (2009b), respectively. This simulation represents a standard scenario assumed in cosmological hydrodynamic simulations. The SN feedback is kinetic and is performed by kicking the particles stochastically in random directions. The parameters that regulate the feedback process are the mass loading $\eta = 2$, which represents the average number of particles kicked per star particle in the case of equal mass particles, and the initial wind velocity $v_w = 600$ km s^{-1}. In this simulation the wind parameters are kept fixed and correspond to an injection of energy that is 40% of the available energy of the SN explosion.

- AGN 8.0: this model is identical to REF with the exception that it also includes a prescription for black hole (BH) growth and AGN feedback, following Booth & Schaye (2009). In this approach, which is a modified version of the one introduced by Springel et al. (2005), the accretion of gas on to the BH follows the Bondi-Hoyle accretion formula only if the gas is expected to be warm (i.e. $\gtrsim 10^4 K$). However, if the pressure is sufficiently high that a cold interstellar phase is expected to form, but which is unresolved by our simulations, then the accretion is regulated by a parameter that depends on the density of the gas, multiplied by the Bondi-Hoyle accretion rate. A certain fraction of the rest mass energy of the accreted gas, ϵ, is stored until it is able to heat up a number of randomly selected neighbouring gas particles, n_{heat}, by raising their temperatures by an amount $\Delta T_{\text{heat}} = 10^8$ K. In this way, the heated gas particles do not radiate away their thermal energy immediately but instead they drive supersonic outflows that are able to displace a large amount of gas far from the AGN themselves. A value of $\epsilon = 0.015$ yields a good match to the $z = 0$ relations between BH mass and stellar mass and velocity dispersion and the $z = 0$ cosmic BH density. On the scales of groups and clusters this is the most realistic simulation because it reproduces many observational data sets, such as the relations between X-ray luminosity, temperature, gas mass fraction, and SZ flux, as was shown by McCarthy et al. (2010, 2011) and Le Brun et al. (2014).

- AGN 8.5: this model is identical to AGN 8.0 but with an increased AGN heating temperature of $\Delta T_{\text{heat}} = 10^{8.5}$ K. As per feedback event the same mass of gas is being heated in this model as in the fiducial AGN 8.0 model, more time is required for the BH to accrete sufficient energy to heat the gas by the higher temperature. In practice, therefore, the duty cycles differ between the two models with the AGN 8.5 model having longer quiescent periods but a more energetic release of thermal energy in the surrounding medium for a given event. We note that ΔT_{heat} can not be increased to arbitrarily high temperatures since this would lead to unrealistically long time periods between feedback episodes and would prevent self-regulation of the AGN feedback (see Booth & Schaye 2009). In a WMAP7 cosmology, Le Brun et al. (2014) find that the AGN 8.0 and AGN 8.5 models effectively bracket the observed baryon fractions of local groups and clusters (see also Fig. 2.5). We refer to Le Brun et al. (2014) for the analysis of the BH population properties and for the BHs scaling relations showing that the BH formed in the simulations used in this paper are consistent with observational results and theoretical models.

- NOSN_NOZCOOL: in this simulation the SN feedback is removed and the gas cooling assumes primordial abundances. No AGN feedback is included.

- NOZCOOL: SN feedback is included but the gas cooling still assumes primordial abundances. No AGN feedback is included.

- WDENS: the SN feedback parameters depend on the local gas density of the star-forming particles from which the star particles that produce the SNe are formed. The initial wind velocity scales with density as $v_{\text{w}} = 600$ km s$^{-1}(n_{\text{H}}/10^{-1}\text{cm}^{-3})^{\frac{1}{6}}$, which implies that v_{w} scales with the sound speed of the equation of state that we impose on the unresolved multiphase interstellar medium (Schaye & Dalla Vecchia 2008), and the mass loading factor as $\eta = 2(n_{\text{H}}/10^{-1}\text{cm}^{-3})$. In this way the total amount of feedback energy per unit stellar mass is kept fixed. The higher wind velocity in dense gas results in a more efficient feedback in massive galaxies (Haas et al. 2013). No AGN feedback is included.
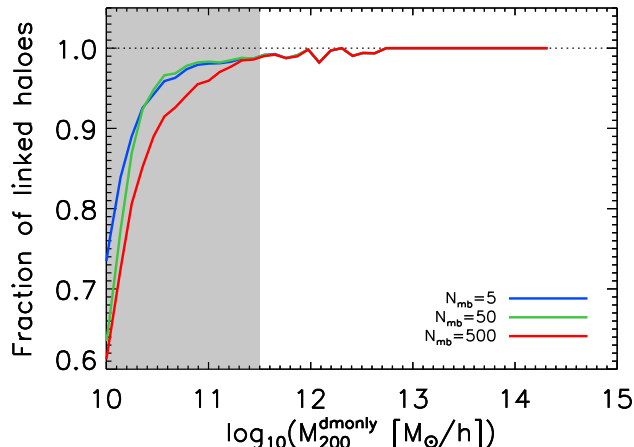
Figure 2.1: The fraction of haloes that are successfully linked as a function of dark matter halo mass for FoF groups in DMONLY and REF. Each line shows what fraction of the FoF groups in DMONLY are linked to a FoF group in REF. Different colours are used for different values of the number of most-bound particles used to match haloes, the fiducial value being $N_{mb} = 50$. The grey shaded region is below the resolution limit. For haloes above this limit the fraction linked is very close to unity.

A complete list of simulations used in this paper, with detailed information on the box size and resolution, is reported in Table 2.1.

2.2.1   Finding and matching haloes between simulations

Haloes are identified in our simulations using the Friends-of-Friends algorithm combined with a spherical over-density algorithm centred on the minimum of the gravitational potential as implemented in SUBFIND (Springel et al. 2001; Dolag et al. 2009). As every simulation from the OWLS project has identical initial conditions for a fixed box size, it is in principle possible to identify the same haloes in each simulation as these should contain mostly the same DM particles, which can be identified using their unique particle IDs. By linking haloes between simulations we can investigate how changes in physics influence the properties of a fixed sample of haloes. Specifically, we are able to examine how the halo mass changes from model to model.

The haloes linking procedure works as follows: for every halo in simulation A we flag the N_{mb} most-bound particles, meaning the particles with the highest absolute binding energy. Next, we locate these particles in the other simulations. If we find a halo in simulation B that contains at least 50% of these flagged particles, a first link is made. The link is confirmed only if, by repeating the process starting from simulation B, the previous halo in simulation A is found.

Fig. 2.1 shows the fraction of friends-of-friends (FoF) groups at $z = 0$ that are successfully linked between DMONLY and REF as a function of dark matter halo mass. Different colours are used for different values of the number of most-bound particles used to match haloes, the fiducial value being $N_{mb} = 50$. For haloes above the resolution limit that we use for this work (see Appendix 2.A), shown by the grey shaded region, the linked fraction is very close to unity. While this fraction insensitive to the value of N_{mb}, using only a few particles to match haloes between simulations may lead to spurious matches, and, more importantly, increases the sensitivity to baryonic cooling and feedback. On the other hand, using values

of N_{mb} that are too high means that even the most loosely bound particles of a halo are used up to relatively high masses, which leads to an even greater sensitivity to baryonic processes. The results for other pairs of simulations are similar, though the curves shift to the left when comparing simulations with baryons to each other, as haloes in these simulations are identified at lower masses. Matched haloes are not considered for our analysis if their mass M_{200} in the DMONLY simulation is less than the mass resolution limit (see Appendix 2.A).

2.3 How baryons alter the masses of haloes

2.3.1 Change in total mass between different realisations of the same halo

In this section we compare the masses of haloes that have been matched as described in § 2.2.1. Haloes are binned according to their mass in the DMONLY simulations (M_{200}^{dmonly}). For each mass bin we plot the median value of the relative difference in mass with respect to the DMONLY realisation. We limit our analysis to haloes with $M_{200}^{\text{dmonly}} > 10^{11.5} \, h^{-1} \, \text{M}_\odot$ (which corresponds to a halo with about 600 particles; see Appendix 2.A), use mass bins of width $\delta \log_{10}(M_{200}) = 0.25$, and include all bins with at least 10 haloes. All the simulations assuming the WMAP7 cosmology (REF, AGN 8.0 and AGN 8.5) are extended to higher masses by the use of cosmo-OWLS (performed in a 400 h^{-1} Mpc box). The transition between the two box sizes happens when in a mass bin the number of haloes in the smaller box falls below a chosen value. Throughout the paper we adopt a value of 50, except for M_{2500} for which we adopt 10.

It is important to note that due to the spherical overdensity definition of masses M_Δ used in this work (and commonly adopted in the literature), a change in the total mass of haloes can be caused by a change in mass within the haloes as well as by change in their density profile. If the density profile changes, then so does the radius at which the value of the mean internal density reaches $\rho(r) = 200\rho_{\text{crit}}$. In this way a redistribution of matter inside the halo can change its total mass just by how M_{200}^{crit} is defined. In order to isolate these two effects, we also show the variation of the total mass inside a given radius that is chosen to be the same in every realisation of the same halo among different simulations.

In Fig. 2.2 we show the relative difference in mass for haloes in different simulations with respect to the DMONLY simulation. There are two different curves for every analysed simulation. The continuous curve represents the relative difference in the M_{200}^{crit} mass between the simulation with baryon physics and the simulation without, namely $(M_{200} - M_{200}^{\text{dmonly}})/M_{200}^{\text{dmonly}}$. In order to isolate the contraction or the ejection of baryons within a common radius, we also plot the quantity $[M(r < R_{200}^{\text{dmonly}}) - M_{200}^{\text{dmonly}}]/M_{200}^{\text{dmonly}}$ (dashed lines). These curves give us insight into the different mass content of each halo realisation, because any differences are due only to the different amount of mass inside a common radius.

We focus first on the top left panel of Fig. 2.2.

For the run without SN feedback, NOSN_NOZCOOL (dark blue lines), it is clear from the low-mass end of Fig. 2.2 that the haloes are more dense than when they are simulated using only gravitationally interacting particles (DMONLY). This is due to the absence of a mechanism that is able to heat the gas and prevent it from overcooling. However, the mass difference becomes smaller with increasing halo mass, due to the increasing importance of gravitational shock heating which limits the cooling rate of the gas (e.g., White & Rees 1978).

Next, we consider the addition of SN feedback with the simulation NOZCOOL (purple lines). Note that the cooling rate of the gas is still computed assuming primordial abundances. For halo masses $M_{200} \lesssim 10^{12} \, h^{-1} \, \text{M}_\odot$, SN feedback leads to a $\approx -20\%$ change in the mass with
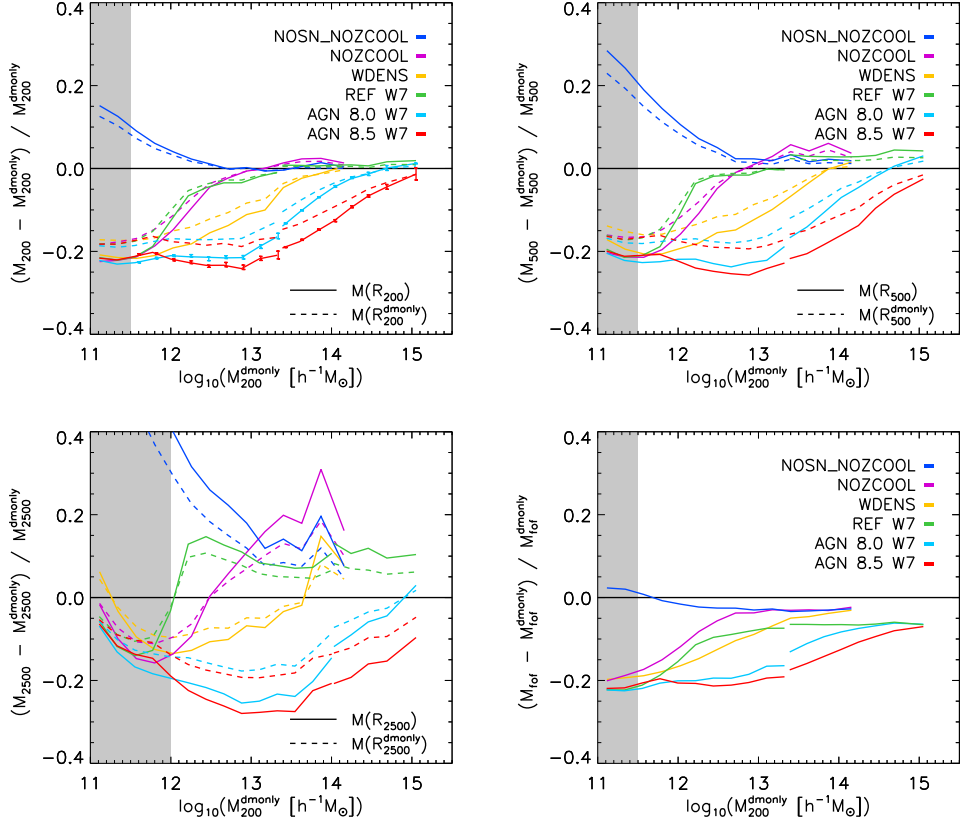
Figure 2.2: The median relative differences in mass (M_{200}, M_{500}, M_{2500} and friends-of-friends mass M_{fof}) between matched haloes in the same mass bin at $z = 0$ for all the different simulations. The dashed lines in the first three panels show the variation of the total mass inside a radius ($R_\Delta^{\mathrm{dmonly}}$) that is the same for any given different realisation of the same halo. Dark blue lines show the results for simulations when gas physics is introduced but not SN feedback or metal-line cooling (NOSN_NOZCOOL). The effect of introducing SN feedback is shown by the purple lines (NOZCOOL). Green lines show the simulations with metal-line cooling and SN feedback (REF). The effect of increasing the efficiency of the SN feedback for massive haloes is shown by the yellow line (WDENS). Cyan lines represent the simulations that include SN feedback, metal-line cooling and AGN feedback with a heating temperature of $\Delta T_{\mathrm{heat}} = 10^{8.0}$ (AGN 8.0). The red lines show the version of the AGN simulation with higher heating temperature $\Delta T_{\mathrm{heat}} = 10^{8.5}$ (AGN 8.5), that produces a more efficient energy release on a longer duty cycle. All the simulations assuming the WMAP7 cosmology (REF, AGN 8.0 and AGN 8.5) are extended to higher masses by the use of cosmo-OWLS (done in a 400 h^{-1} Mpc box). The transition between the two box sizes happens when in a mass bin the number of haloes in the smaller box falls below 50 (10 for M_{2500}). The shaded grey region is below the resolution limit found in Appendix 2.A. In addition, for the upper left panel, 2-sigma errors computed via bootstrapping are shown for the two AGN simulations.

respect to the DMONLY simulation. However, at higher masses the effects of SN feedback are minor and the difference with respect to the DMONLY simulation tends towards zero. In other words, we find that SNe are incapable of ejecting gas from galaxy groups and clusters, consistent with previous studies (e.g., Kravtsov et al. 2005; Ettori et al. 2006).

If metal-line cooling is switched on (while retaining the SN feedback) as in the REF run (green lines), the predictions start to converge to the DMONLY simulation at slightly smaller halo masses. This is because the feedback is less effective due to the increased cooling rate of the gas. Note that for this simulation we use a WMAP7 cosmology in order to extend the dynamic range probed by our analysis by adding the cosmo-OWLS version of the REF simulation done in a 400 h^{-1} Mpc box. We will explore the sensitivity of the results to changes in cosmology in § 2.3.6 (they are minimal).

In the WDENS simulation (yellow lines), the parameters that regulate the wind scale with the local gas density, such that the wind velocity increases with density while the mass-loading decreases. This has the net effect of increasing the effectiveness of the feedback in denser environments and high-mass haloes relative to the fixed-wind REF model. Indeed, the WDENS curve has a similar shape to the REF curve, but the mass range over which the winds are effective is much more extended.

When AGN feedback is turned on (red and cyan lines) the picture changes. The decrease in halo mass extends to much higher masses than in models with SN feedback alone. While the two AGN feedback models show the same qualitative behaviour, the model that invokes higher heating temperature (AGN 8.5) is able to extend the relative change in mass to higher masses due to the more effective energy release. However, even AGN feedback is insufficient to significantly alter the total masses of the most massive galaxy clusters (with $M \sim 10^{15}\,h^{-1}\,\mathrm{M}_\odot$).

Interestingly, in the highest mass bin in the AGN 8.0 simulation we obtain a slight *increase* in the mass of the halo simulated with AGN 8.0 with respect to the DMONLY case (the REF model also displays this behaviour). This could be explained by the fact that the inclusion of baryons leads to more spherical haloes compared to the DMONLY simulation (e.g., Bryan et al. 2013), and so the haloes in baryonic simulations tend to have more mass within a common radius due to this geometric effect.

So far we have discussed the top left panel in Fig. 2.2. The top right (M_{500}) and the bottom left (M_{2500}) panels show the same analysis using higher over densities that probe the inner part of the haloes. We note that for the simulations that include cooling but no feedback from AGN (NOSN_NOZCOOL, NOZCOOL, REF, WDENS), the trend resembles the one seen in the M_{200} panel, but with all the curves shifted upwards. This is because in the central region the overcooling dominates over the SN feedback (if any). In the case of central feedback (AGN) the difference with respect to the DMONLY simulation increases compared to the top left panel, since the feedback is more efficient in removing gas from the centre of the halo.

For completeness, in the bottom right panel in Fig. 2.2 we show the relative difference in the FoF mass (M_{fof}). This mass represents the mass of all the particles that are linked together using the FoF algorithm and therefore within a DM isodensity contour. The linking of the particles in the FoF algorithm is done considering only DM particles (with linking length 0.2). Every baryonic particle is associated with its nearest DM particle. If this DM particle is in a FOF group, then the baryonic particle is assigned to the same FOF group. Note that the relation will in general not converge to zero if the baryons do not have the same 3D spatial distribution as the DM. However, we have tested that the DM component in the baryonic simulations converges to zero mass difference for high masses when the DMONLY simulation is rescaled to take into account the universal baryon fraction.

In order to evaluate the significance of the trends shown in Fig. 2.2, we computed the errors on the median using the bootstrapping technique. These errors are shown in Fig. 2.2 for the
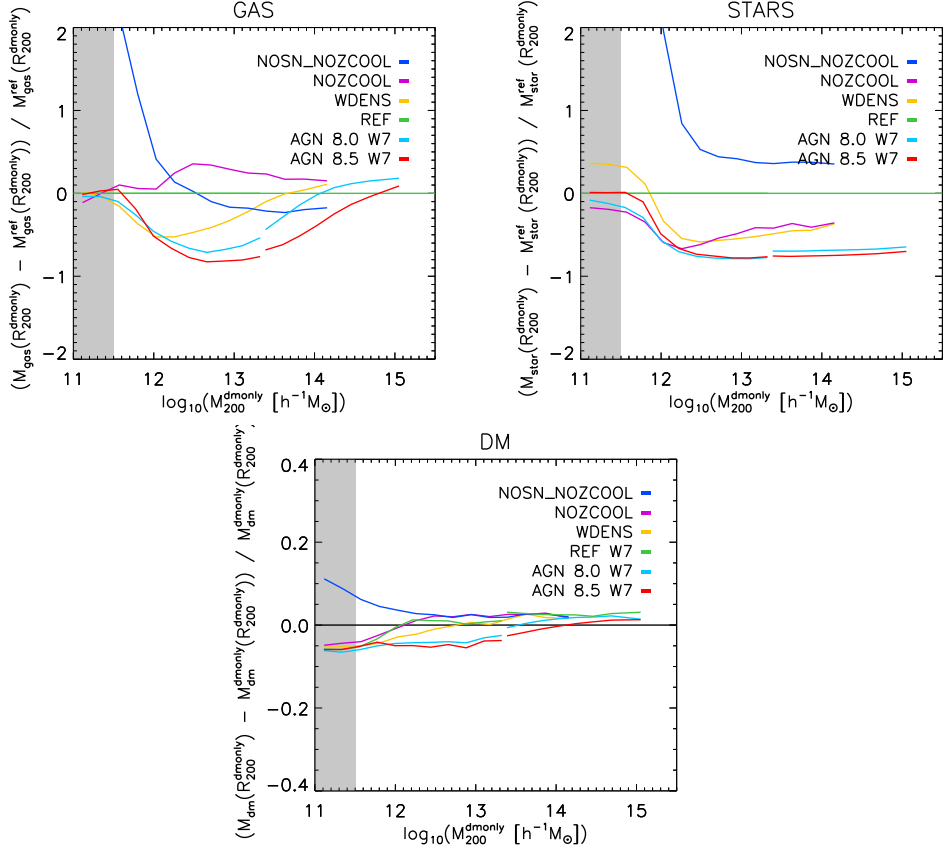
Figure 2.3: Top left panel: the relative difference in the median gas mass enclosed in $R_{200}^{\mathrm{dmonly}}$ with respect to the REF model in each simulation binned in $M_{200}^{\mathrm{dmonly}}$ mass. Top right panel: the relative difference in stellar mass in $R_{200}^{\mathrm{dmonly}}$ with respect to the REF model. Bottom panel: the relative difference in dark matter mass in $R_{200}^{\mathrm{dmonly}}$ with respect to the DMONLY model, notice the change in scale.

two AGN models, but the amplitude is similar for the other mass definition and models. The errors are computed by taking the standard deviation of the distributions of the medians drawn from 1000 bootstrap realizations of the data in every mass bin. These errors are very small, which indicates that median values quoted in our analysis are robust. For the most massive bin the errors suggest that the median values are consistent with no change with respect to the DMONLY model.

The above analysis shows that important subgrid physics, particularly AGN feedback, can substantially alter the *total* masses of haloes, by up to $\approx$ 20% within R_{200} (and by larger amounts within smaller radii). This is suggestively close to (though slightly larger than) the universal baryon fraction, but as we will show later this does not imply that all of the baryons have been removed from the haloes (for an analysis of the gas and baryon fractions in these haloes we refer to § 2.3.3).

In what follows immediately below, we quantify the effects of these subgrid processes on the stellar, gas, and dark matter masses separately.

2.3.2 Change in baryon mass and back-reaction on dark matter

In the top left and top right panels of Fig. 2.3 we show the relative difference in the gas mass and stellar mass, respectively, within $R_{200}^{\mathrm{dmonly}}$ with respect to the REF model. The dark blue lines show the results for simulations that include gas physics but without SN feedback and metal-line cooling (NOSN_NOZCOOL). For this model it is clearly visible that the mass in stars is much higher than the REF model (middle panel). Again, this is due to the absence of a mechanism that prevents the gas from overcooling. In terms of gas (top panel), at the low-mass end uninhibited cooling boosts the accretion of gas within $R_{200}^{\mathrm{dmonly}}$ compared to REF. By contrast, at the high-mass end, where the SNe are not able to affect the total mass enclosed in REF, the amount of gas is less in the NOSN_NOZCOOL model, since a larger fraction of the gas was converted into stars in REF (due to the increased cooling rate from metal lines).

When SN feedback is included but metal-line cooling remains off (NOZCOOL), a lower mass of stars is formed due to the increased effectiveness of SN feedback in the absence of metal-line cooling. This leads to an increase in the mass of gas within $R_{200}^{\mathrm{dmonly}}$ with respect to the REF simulation.

Inclusion of AGN feedback (AGN 8.0 and AGN 8.5) reduces the mass of stars and gas with respect to the REF model, the latter being due to ejection from the haloes. The same qualitative effect is obtained at the high-mass end by varying the efficiency of SN feedback with the local gas density (WDENS). On the other hand, at the low-mass end the SN feedback is less efficient at suppressing the star formation, increasing the amount of mass in stars with respect to the REF model.

In the bottom panel of Fig. 2.3 we show the relative difference in the dark matter mass inside $R_{200}^{\mathrm{dmonly}}$ with respect to the DMONLY simulation due to baryon physics (notice the change in scale on the plot). At the low-mass end the effect of the overcooling and the consequent adiabatic contraction of the DM haloes is clearly visible for the simulation without SN feedback. When SN feedback is introduced, the removal of gas expands the DM mass distribution, reducing the amount of DM mass that is present in the haloes. At the high-mass end gravity again becomes dominant and the adiabatic contraction of the DM distribution can only be offset by introducing AGN feedback. However, even AGN feedback cannot prevent some slight contraction of the dark matter component on the scale of R_{200} for the most massive haloes (clusters).
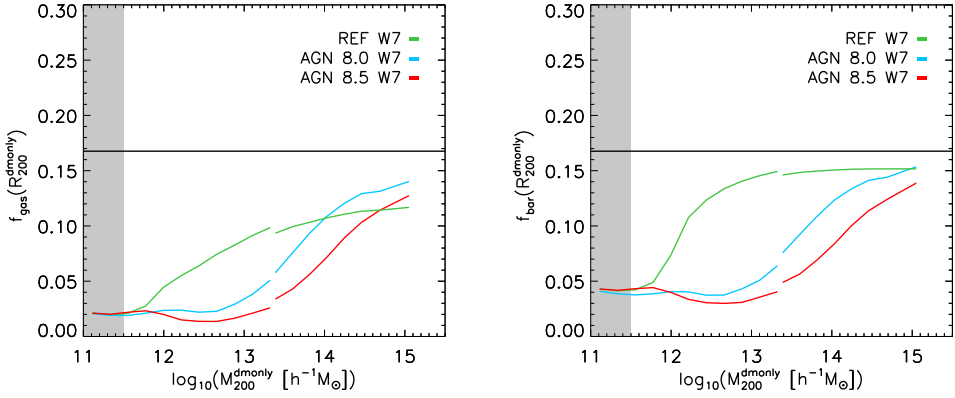
Figure 2.4: The gas fraction (left panel) and baryon fraction (right panel) inside a sphere with a radius corresponding to the R_{200} of the halo in the DMONLY simulation. The results are for the REF simulation (green lines) and the AGN simulations (cyan and red line). The continuous black line represent the universal baryon fraction for the WMAP7 cosmology.
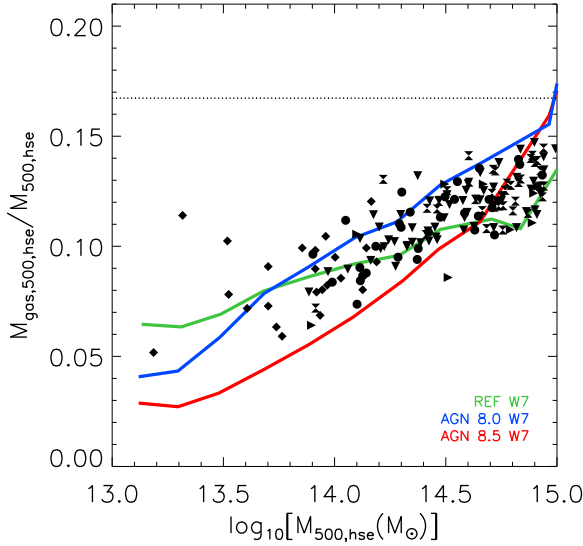

Figure 2.5: The gas mass fraction within $r_{500,hse}$ as a function of $M_{500,hse}$ at $z = 0$. The filled black circles (clusters), right-facing triangles (clusters), downward triangles (clusters), hourglass (clusters) and diamonds (groups) represent the observational data of Pratt et al. (2009), Vikhlinin et al. (2006), Lin et al. (2012), Maughan et al. (2008) and Sun et al. (2009), respectively. The coloured solid curves represent the median gas mass fraction–$M_{500,hse}$ relations in bins of $M_{500,hse}$ for the different simulations. The observed trend is approximately bracketed by the AGN 8.0 and the AGN 8.5 models for a WMAP7 cosmology.

2.3.3   Baryon fractions

In § 2.3.1 we argued that the relative change in the total halo mass due to baryon physics is produced mainly by the ejection of baryons from the haloes, at least for models with efficient feedback. The maximum magnitude of the effect is similar to the universal baryon fraction, which might naively suggest that most of the baryons have been ejected from the haloes. Here we examine the baryon and gas mass fractions of haloes in the different simulations.

In Fig. 2.4 we show the mass fraction in gas (left panel) and gas+stars (right panel) inside $R_{200}^{\mathrm{dmonly}}$ radius corresponding to R_{200} of the same halo in the DMONLY simulation. We show the results for the REF simulation (green lines) and the AGN simulations (cyan and red lines). The continuous black line represents the universal baryon fraction for the WMAP7 cosmology. It is clear that the amount of gas in the haloes is always less in the AGN simulations, with a minimum value that is $\approx 15\%$ of the universal baryon fraction. Interestingly, for the highest mass bin in these simulations REF and the AGN have nearly the same gas mass but for very different reasons: in the REF simulation the gas is removed because it has been locked up in stars, while in the AGN simulations feedback from supermassive black holes expels much of the gas that would otherwise have been turned into stars.

Adding the stars to the gas we obtain the baryon fraction (bottom panel) and in this case the difference between the REF and the AGN models is striking. The REF simulation produces haloes with baryon fractions close to the universal mean all the way down to low halo masses $M_{200}^{\mathrm{dmonly}} \sim 10^{12.5}\,h^{-1}\,\mathrm{M_\odot}$, in strong disagreement with observations of groups and clusters (e.g., Budzynski et al. 2014). The baryon fraction trend in the AGN simulations, by contrast, is very similar to that of the gas fraction, leading to a much better agreement with the observations, as shown by Le Brun et al. (2014).

From the above we see that even in the AGN models the haloes are not devoid of baryons. This raises the question of how the total masses can be altered by up to $\approx 20\%$ (for haloes with $M_{200}^{\mathrm{dmonly}} \lesssim 10^{13.5}\,h^{-1}\,\mathrm{M_\odot}$). The explanation is simply that the ejection of a large fraction of the gas expands the dark matter as well, as shown in the bottom panel of Fig. 2.3. It is the combination of gas ejection and dark matter expansion that causes a total mass change that is comparable to (slightly larger than) the universal baryon fraction.

It is important to compare our results for the simulated gas fraction with observations. Hence, in Fig. 2.5, we plot the gas mass fraction–$M_{500,\mathrm{hse}}$ relation at $z = 0$ for the various simulations and compare to observations of individual X-ray-selected systems. The gas mass fraction is measured within $r_{500,\mathrm{hse}}$. For the simulations, we use the results from the synthetic X-ray analysis of Le Brun et al. (2014) to 'measure' the halo mass and gas mass fraction of the simulated systems, thus the 'hse' subscript in the masses indicates that they have been derived by a synthetic hydrostatic X-ray analysis. As can be seen in Fig. 2.5, the observed trend is approximately bracketed by AGN 8.0 and AGN 8.5, at least up to masses of $10^{14.7}\,h^{-1}\,\mathrm{M_\odot}$ in the WMAP7 cosmology (see Le Brun et al. 2014 for discussion). The REF model, which neglects AGN feedback, also yields reasonable gas mass fractions but they are achieved by overly efficient star formation as already shown in Fig. 2.4. Note that for $M_{500,\mathrm{hse}} < 10^{13.5}\,\mathrm{M_\odot}$ the observational samples are likely biased. We note also that the scatter about the median gas fraction trends (not shown here) is comparable to that in the observed relation (see Le Brun et al. 2014 for further discussion).

2.3.4   Enclosed mass profiles

In this section we explore the variation of the total cumulative mass profile for haloes in different mass bins. This is instructive for explaining the trends in the previous plots. The reader who is interested only in the net effect on the halo mass function, may wish to skip to

the next section.

For this analysis we use mass bins of size $\delta\log_{10}M_{200}^{\text{dmonly}} = 0.25$, again selecting haloes by their mass in the DMONLY simulation. For each simulation we take all the haloes in the mass bin and produce a median stacked total enclosed mass profile over the radial range $-2.5 \leq \log_{10}(R_{\text{min}}/R_{200}^{\text{dmonly}}) \leq 1.0$. We use 47 bins over this radial range but plot only those bins which exceed the softening length of the simulation (below three softening lengths we use dotted lines). The results are shown in Fig. 2.6. (Note that the variation of the mass enclosed at the radius equal to R_{200}^{dmonly} is what is shown by the dashed line in Fig. 2.2 for a given mass bin.) The dot-dashed, short-dashed, and long-dashed vertical lines represent the median values of R_{2500}^{dmonly}, R_{500}^{dmonly}, and R_{200}^{dmonly}, respectively. Three different mass bins are shown for the $100\,h^{-1}$ Mpc simulation box, and the last panel (bottom right) is taken from the $400\,h^{-1}$ Mpc simulation.

In the top left panel, we show the first mass bin, $10^{11.50} < M_{200}^{\text{dmonly}}/[\,h^{-1}\,\text{M}_\odot] < 10^{11.75}$. In this mass range it is clearly visible that in the inner regions (the central $\sim 10\%$ of R_{200}^{dmonly}) the baryonic component dominates in all the simulations. However, at larger radii ($R \gtrsim R_{2500}^{\text{dmonly}}$) the mass enclosed becomes less than that in the case of DMONLY for simulations which include SN feedback (AGN, REF, WDENS, NOZCOOL). Remarkably, at this mass scale convergence to the DMONLY result only occurs at very large radii of $R \gtrsim 5R_{200}^{\text{dmonly}}$.

In the top right panel we consider the mass range $10^{12.50} < M_{200}^{\text{dmonly}}/[\,h^{-1}\,\text{M}_\odot] < 10^{12.75}$. The trends are qualitatively similar to those in the top left panel, except that there is a much larger spread in the predictions of the models at small radii, due to the ineffectivness of SN feedback and the increasing importance of AGN feedback at these high masses. Furthermore, models with no feedback or SN feedback alone converge to the DMONLY result at smaller radii ($R \sim 1-2R_{200}^{\text{dmonly}}$) than models which also include AGN feedback ($R \gtrsim 5R_{200}^{\text{dmonly}}$), as qualitatively expected.

The two bottom panels consider higher halo masses still: $10^{13.50} < M_{200}^{\text{dmonly}}/[\,h^{-1}\,\text{M}_\odot] < 10^{13.75}$ (bottom left) and $10^{14.50} < M_{200}^{\text{dmonly}}/[\,h^{-1}\,\text{M}_\odot] < 10^{14.75}$ (bottom right). Continuing the trends discussed above, the differences between the models are largest at small radii due to excessive overcooling in models without AGN feedback compared to those with it. Note that even in massive galaxy clusters AGN feedback can noticeably alter the total mass distribution all the way out to $\sim R_{500}^{\text{dmonly}}$.

We point out that the large-radii variation with respect to the DMONLY simulation (due to SN feedback at low masses and AGN feedback at high masses) does not necessarily mean that the baryons are ejected out to several times the virial radius, since much of the mass surrounding a given halo is in infalling galaxies which are also driving outflows and influencing their local environments.

2.3.5 Evolution with redshift

For completeness we explore the evolution of these trends with redshift, which is important since it is the evolution of the HMF that is the focus of upcoming cosmological studies. In particular, in Fig. 2.7 we compare the effects of SN and AGN feedback on the total mass of haloes at two different redshifts ($z = 0, 1$) using the three cosmo-OWLS simulations performed using a WMAP7 cosmology. We show only these cases since they are representative of the evolution of the mass difference for all other simulations.

In all the simulations it is clear that the absolute difference in the mass of haloes increases with time. This is expected, since haloes are denser with increasing redshift and the the binding energy of a halo of fixed spherical overdensity mass therefore increases within increasing redshift. Thus, more energy is required to alter the mass distribution of haloes (at fixed mass)
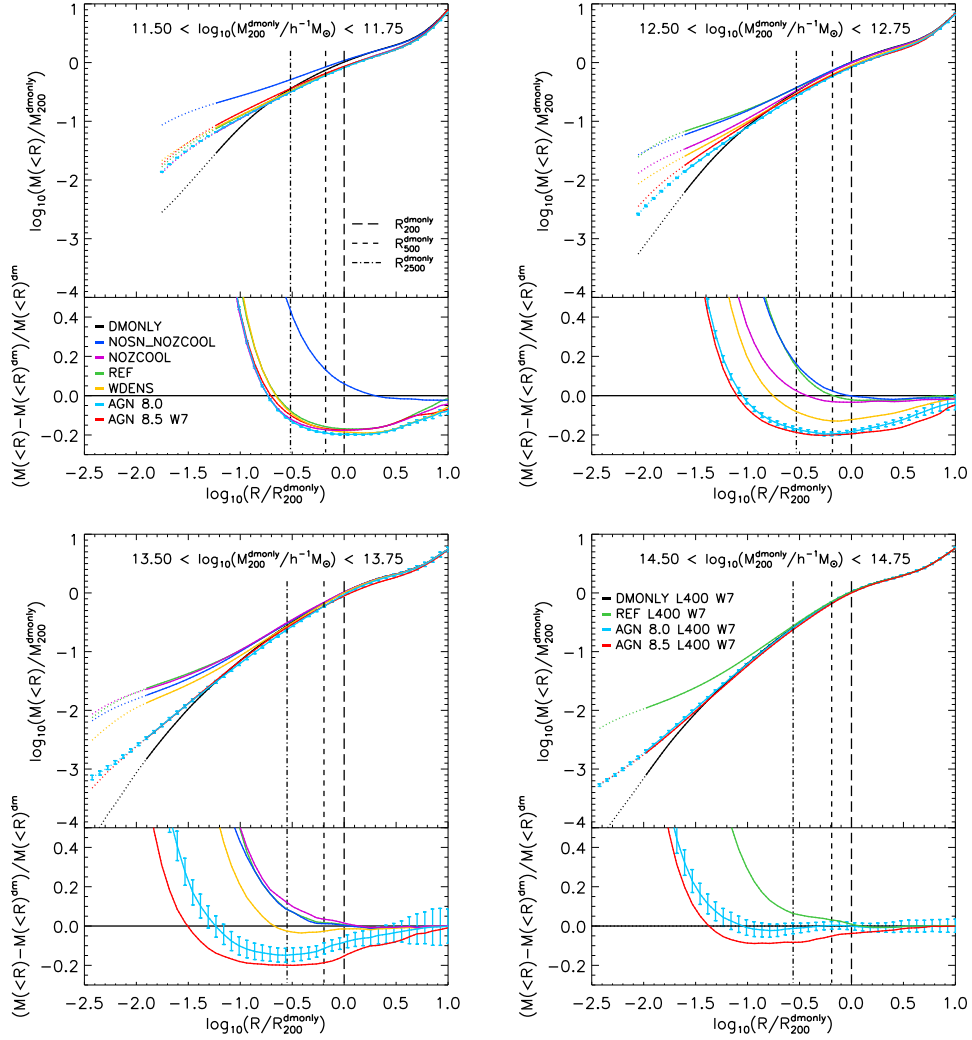
Figure 2.6: Median stacked profiles of the total enclosed mass as a function of radius in units of $R_{200}^{\mathrm{dmonly}}$. Each panel shows a different mass interval in $M_{200}^{\mathrm{dmonly}}$. The long dashed vertical lines represent $R_{200}^{\mathrm{dmonly}}$ the radius of the halo in the DMONLY simulation, the dashed vertical lines represent the $R_{500}^{\mathrm{dmonly}}$, and the dashed-dotted lines represent the $R_{2500}^{\mathrm{dmonly}}$. In the bottom part of every panel we show the relative difference between the curves of the simulations with baryons and the DMONLY simulations. 2-sigma errors from bootstrapping are shown for the AGN model. Every curve is dotted below three softening lengths and stopped at the softening length.
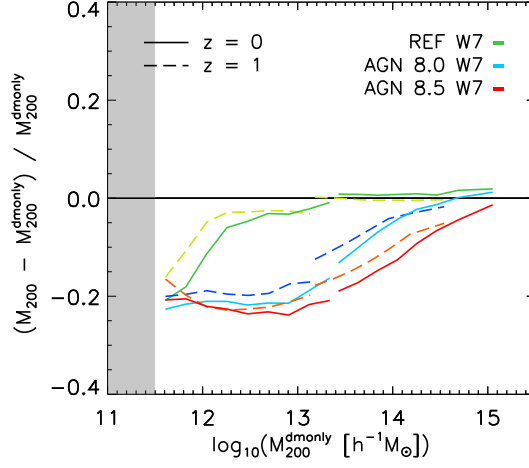
Figure 2.7: Evolution with redshift of the relative difference in M_{200} mass between simulations with baryon physics and DMONLY. The results for redshift zero are represented by continuous lines, while the results for redshift one by long dashed lines.
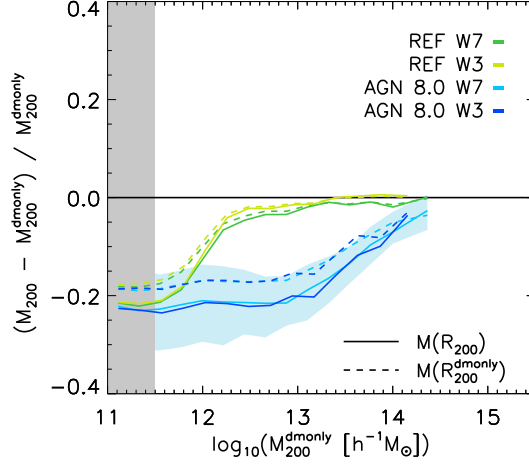

Figure 2.8: Comparison of the effect of baryons on the mass of haloes for simulations with different cosmological parameters. The cyan curve shows the AGN 8.0 simulation in $100 \, h^{-1}$ Mpc box and using a WMAP3 cosmology. The blue line shows the same simulation done with the same box size but with a WMAP7 cosmology. For the REF simulations the green lines represents the WMAP7 variation and the light green the WMAP3 version. It is clear that the changes in the cosmological parameters between the WMAP3 and WMAP7 cosmologies do not lead to a significant change in the relative difference in halo mass. The light cyan band represent the area between the upper 84th and the lower 16th percentile for the AGN 8.0 W7 model.

at early times. Overall, however, the difference in the trends between $z = 0$ and $z = 1$ for a given model is relatively minor.

We do not explore the difference in the relation for $z > 1$ since our box is too small to provide a statistical sample of high-mass haloes at higher redshift.

2.3.6 Effect of cosmology

We now test the sensitivity of our results to changes in the cosmological parameter values adopted in the simulations.

In Fig. 2.8 we show the relative difference in halo mass due to baryonic effects using the AGN 8.0 model in a WMAP7 cosmology (cyan line) and compare it with the same model in WMAP3 cosmology (blue line). The two simulations were run with the same resolution and with the same box size ($100\,h^{-1}$ Mpc). The two lines are fully consistent with each other.[4] The same test is presented for the REF simulations where the green line is the WMAP7 version and the light green represents the WMAP3 realization. The convergence is excellent for both the total mass and the total mass enclosed in $R_{200}^{\mathrm{dmonly}}$ in this case. (The slight difference between the two is plausibly due to the increased universal baryon fraction in the WMAP7 cosmology, which leads to slightly more gas cooling and slightly less efficient SN feedback.) We also show the upper 84th and the lower 16th percentile around the median for the AGN 8.0 W7 model that is also representative for the scatter of the other simulations (for a more detailed description of the scatter see § 2.4).

We conclude that the analysis done on the variation of the mass of the halo due to baryon physics is largely independent of small variations in the input cosmological parameter values used.

2.4 Analytic fitting formula for the change in halo mass

Combining the results obtained from the simulations in the $100\,h^{-1}$ Mpc box and the larger $400\,h^{-1}$ Mpc box, we are able to provide a fitting function that reproduces the median change in the mass of haloes due to baryon physics over four orders of magnitude in mass. We adopt the following functional form:

$$\log_{10}\left(\frac{M_\Delta}{M_\Delta^{\mathrm{dmonly}}}\right) = A + \frac{B}{1 + \exp\left(-\frac{\log_{10}(M_\Delta^{\mathrm{dmonly}})+C}{D}\right)}, \tag{2.1}$$

This equation reaches the constant value $A + B$ (A) in the high-mass limit and the constant value A ($A + B$) in the low mass limit when the parameter $D > 0$ ($D < 0$).

We also provide a linear fitting function for the standard deviation of the change in halo mass at a given mass. The scatter is Gaussian when the logarithm of the difference is considered and is well approximated by the following fitting function:

$$\sigma(\log_{10}(M_\Delta^{\mathrm{dmonly}})) = E + F\log_{10}(M_\Delta^{\mathrm{dmonly}}) \tag{2.2}$$

The scatter is always a decreasing function of mass. While the scatter is physical in origin, it is not strongly correlated with basic properties of the haloes like concentration. We leave an examination of the origin of the scatter for future work.

---

[4]The WMAP7 simulation line extends to slightly higher masses due to the larger value of σ_8 which yields a slightly larger number of high-mass haloes.

When fitting Eqs. 2.1 and 2.2, we assign an error to each mass bin that is equal to the standard deviation of the distribution obtained from 1000 bootstrapped re-samplings.

For the REF model we make use of three different simulations: a high resolution 50 h^{-1} Mpc version done using the WMAP3 cosmology that enables us to push the resolution limit down to $M_{200}^{\mathrm{dmonly}}/ = 10^{10.5}\ h^{-1}\ \mathrm{M_\odot}$, the standard 100 h^{-1} Mpc WMAP7 version and the 400 h^{-1} Mpc box, also WMAP7, for the high-mass bins where the number of haloes in the smaller box falls below 50. For the two simulations with AGN feedback the high-resolution version is not available, so we only combine the 100 and the 400 h^{-1} Mpc box both run with a WMAP7 cosmology.

In Table 2.2 we report the parameters of the fitting function (Eq. 2.1) and the scatter (Eq. 2.2), for different masses (M_{200}, M_{500}) and for the REF and the two AGN simulations. For other masses and redshifts the complete list of fitting parameters as well as the mean relations are available at `http://www.strw.leidenuniv.nl/MV14/`.

We stress that this function is not meant to be extrapolated to masses lower than the mass resolution presented in this work, especially for the two simulations with AGN feedback. In fact, it is clear from the resolution tests that the relative difference in mass continues to increase in amplitude when the lowest mass regime is explored using simulations with higher resolution (Appendix 2.14). At the high-mass end the fitting parameters suggest that there is a constant offset value in the relative change in mass. However, we expect that the difference in mass converges towards no difference when the halo mass becomes sufficiently large. With this in mind, we also provide fitting functions that are constrained to asymptote to zero at high masses (given by the *zero asint* tag in Table 2.2).

In Fig. 2.9 we show the fitting function (constrained to asymptote to zero at high masses) for the differences in M_{500} when the feedback processes are included. For the 100 h^{-1} Mpc simulation the unresolved regime is represented by a gray shaded region. It is clear that the fitting functions reproduce the trend of the simulations well. Moreover, it is clearly visible that for the REF simulation the relative change in mass increases in amplitude towards the low-mass end when the simulation with higher resolution is used for the fitting. This result is in agreement with the work of Sawala et al. (2013) at the dwarf mass scale.

2.5   Effects of baryons on the halo mass function

The halo mass function (HMF) gives the average number of haloes in a given mass range per unit volume. Usually the HMF is defined as:

$$f \equiv \frac{\mathrm{d}n}{\mathrm{d}(\log_{10}M)}, \tag{2.3}$$

where n is the number density of haloes and M is the mass of haloes.

To calculate this function, we use the total halo mass M_{500}^{mean} enclosed within the radius R_{500}^{mean}, defined as the radius within which the mean internal density reaches a value of $500 \times \rho_{\mathrm{mean}}$, where ρ_{mean} is the mean density of the Universe at that time. We switch to using spherical overdensity masses defined with respect to the mean density as opposed to the critical density, since the former is more commonly used in mass function work (e.g., Tinker et al. 2008). Note that $M_{500}^{\mathrm{mean}} \approx 1.5 M_{500}^{\mathrm{crit}}$.

In Fig. 2.10, we show the relative difference between the HMF from the simulations with baryons and the HMF obtained from the simulation with only gravitationally interacting particles (DMONLY). The general trends in the relative difference in the HMFs are very similar to those in the relative change in mass (see Fig. 2.2). This suggests that the major role in altering the HMF is played by the change in the masses of haloes rather than by a change in the
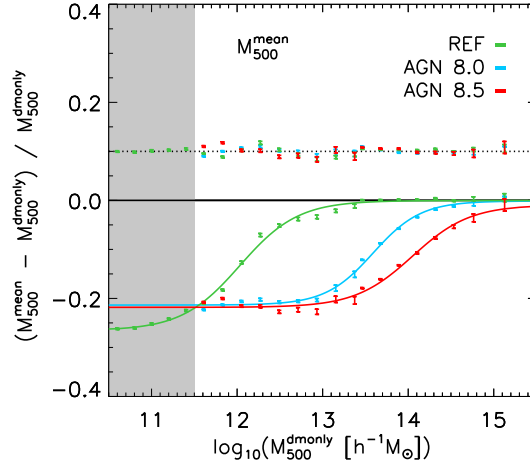
Figure 2.9: Fitting function, constrained to asymptote to zero at high masses for the differences in M_{500}^{mean} due to baryonic processes. The error bars show the $2 - \sigma$ bootstrapped confidence interval. The lines show the best-fitting model using the function in Eq. 2.1. The shaded region shows the resolution limit for the $100\,h^{-1}$ Mpc box, for the REF simulations the points below the resolution limit are taken from a high resolution version of the REF simulation. The error bars in the top half of the plot indicate the residuals between the fit and the points, shifted up by 0.1 for clarity.
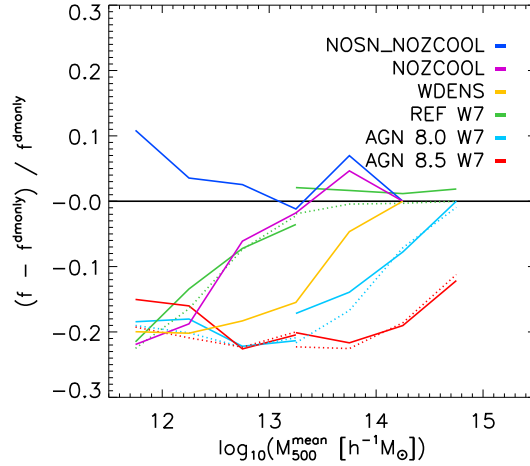

Figure 2.10: Relative difference in the HMF from simulations with baryons with respect to the DMONLY simulation. For the models for which the $400\,h^{-1}$ Mpc version is available we switch to the bigger box when the number of haloes per bin in the smaller box falls below 50. The dotted lines show the HMF computed by correcting the mass of every halo in the DMONLY simulation for the effects of baryon physics, using the fitting function provided in § 2.4. By applying this mass correction to the DMONLY haloes, we are able to reproduce the general trend of the change in HMF from simulations that explicitly include baryon physics.

Sim	z	Mass	A (E)	B (F)	C	D
AGN 8.0	0	M_{200}	-0.1080	0.1100	-13.5861	0.2920
zero asint	0	M_{200}	-0.1077	0.1077	-13.5715	0.2786
		σ_{200}	0.1294	-0.0082		
AGN 8.0	0	M_{500}	-0.1141	0.1232	-13.6581	0.3114
zero asint	0	M_{500}	-0.1133	0.1133	-13.5947	0.2678
		σ_{500}	0.1166	-0.0069		
AGN 8.5	0	M_{200}	-0.0038	-0.1069	-14.0424	-0.3398
zero asint	0	M_{200}	-0.1109	0.1109	-14.0745	0.3579
		σ_{200}	0.1104	-0.0064		
AGN 8.5	0	M_{500}	-0.0035	-0.1151	-14.0871	-0.3333
zero asint	0	M_{500}	-0.1187	0.1187	-14.1132	0.3461
		σ_{500}	0.1073	-0.0060		
REF	0	M_{200}	-0.1385	0.1412	-11.9307	0.3423
zero asint	0	M_{200}	-0.1367	0.1367	-11.9234	0.3148
		σ_{200}	0.1024	-0.0065		
REF	0	M_{500}	-0.1415	0.1517	-11.7863	0.2791
zero asint	0	M_{500}	-0.1366	0.1366	-11.7623	0.2135
		σ_{500}	0.0910	-0.0053		

Table 2.2: Fitting function parameters of Eq. 2.1 and for the scatter in Eq. 2.2, for different simulations and different masses M_Δ^{crit}. The tag *zero asint* gives the fitting function constrained to asymptote to zero at high masses

abundance of haloes. We also tested that the baryon physics does not change the abundance of haloes and found that the number of haloes that are matched between the simulations varies by less than one percent in every mass bin among the different simulations.

We now test our ability to reconstruct the HMF of the simulations with baryon physics by starting from the DMONLY simulations and applying the change in halo mass. We first apply the mass change fitting functions presented in §2.4 to every halo in the DMONLY simulation and then recompute the HMFs assuming that the scatter in the mass change does not play an important role. We use the fitting functions that are constrained to asymptote to zero for high masses, since even a small constant change in halo mass at very high masses can produce a non-converging result in the HMF due to its steepness at the high-mass end.

The results are shown by the dotted lines in Fig. 2.10. It is immediately apparent that the dotted lines reproduce the general trend of the change in the HMF correctly. The main difference with the HMF of the REF simulation at the high-mass end is due to adopting the fitting function that goes to zero for higher masses when it is clear that in the REF simulation there is a constant positive offset in the halo masses with respect to the DMONLY case. Both the AGN cases are reproduced quite well by this change in mass, especially at high masses.

Thus, the change in the mass of haloes is responsible for the differences in the HMFs introduced by baryon physics, and we have shown by applying the fitting function for the change in halo masses, we are able to reproduce the HMFs of simulations with baryons starting from a simulation with only dark matter. Below we will apply the change in mass due to baryon physics to a generic HMF fitting formula obtained from N-body simulations that can be applied to different cosmologies.

Sim	z	Mass	A	B	C	D
REF	0	M_{200}^{mean}	-0.1155	0.1155	-12.0603	0.4230
REF	0	M_{500}^{mean}	-0.1203	0.1203	-11.9864	0.3487
AGN 8.0	0	M_{200}^{mean}	-0.0872	0.0872	-13.6339	0.3509
AGN 8.0	0	M_{500}^{mean}	-0.0942	0.0942	-13.7063	0.2717
AGN 8.5	0	M_{200}^{mean}	-0.0881	0.0881	-14.4100	0.4280
AGN 8.5	0	M_{500}^{mean}	-0.0976	0.0976	-14.4808	0.3795
REF	1	M_{200}^{mean}	-0.1232	0.1232	-11.7513	0.3970
REF	1	M_{500}^{mean}	-0.1174	0.1174	-11.5540	0.1876
AGN 8.0	1	M_{200}^{mean}	-0.0903	0.0903	-13.8505	0.2978
AGN 8.0	1	M_{500}^{mean}	-0.0993	0.0993	-14.0034	0.2979
AGN 8.5	1	M_{200}^{mean}	-0.0995	0.0995	-14.7619	0.3603
AGN 8.5	1	M_{500}^{mean}	-0.1149	0.1149	-14.9524	0.3177

Table 2.3: Fitting formula parameters of Eq. 2.9 calculated using the Planck cosmology, for different simulations and different mass definitions.

2.5.1   Analytic fitting formula for the halo mass function

Assuming that the change in the total halo mass is insensitive to small changes in the cosmological parameters, as the analysis in § 2.3.6 suggests, we can apply the mass correction (Eq. 2.1) to a theoretical prescription for the HMF.

We use the formalism of Tinker et al. (2008) for the theoretical mass function. In order to obtain the linear variance over a certain mass scale, $\sigma(M)$, we assume a linear power spectrum, we apply the transfer function as presented in Eisenstein & Hu (1998) and we assume a top-hat window function in real space. We use the fitting parameters, calibrated using DM-only simulations, for the normalized version of the fitting function $g(\sigma)$ as presented in the appendix of Tinker et al. (2008) and used in Tinker et al. (2010). Here we just summarize the equations for calculating the halo mass function:

$$\frac{dn}{dM} = g(\sigma)\frac{\bar{\rho}_{\mathrm{m}}}{M}\frac{d\ln\sigma^{-1}}{dM}. \tag{2.4}$$

Here, the function $g(\sigma)$ is expected to be universal to the changes in redshift and cosmology and is parametrised as

$$g(\sigma) = B\left[\left(\frac{\sigma}{e}\right)^{-d} + \sigma^{-f}\right]e^{-g/\sigma^2} \tag{2.5}$$

and normalized as follows:

$$\int g(\sigma)\,d\ln\sigma^{-1} = 1. \tag{2.6}$$

The expression for σ is

$$\sigma^2 = \frac{1}{2\pi}\int P(k)\hat{W}^2(kR)k^2dk, \tag{2.7}$$

where $P(k)$ is the linear matter power spectrum as a function of wavenumber k, and $\hat{W}$ is the Fourier transform of the real-space top-hat window function of radius R.

Since the scatter (in the mass change due to baryons) does not play a major role in shaping the HMF, we apply only the median change in mass relation presented in § 2.4 in order to get

the HMF with the effects of AGN feedback included, according to:

$$\frac{dn}{dM}(M_\Delta^{\text{agn}}) = \left(\frac{dn}{dM}\right)^{\text{dmonly}}(M_\Delta^{\text{dmonly}}(M_\Delta^{\text{agn}})). \tag{2.8}$$

Moreover, we can fit the relative difference in the halo mass function using the functional form already used in the previous section, providing in this way an easy to use correction function. The fitting function becomes:

$$\log_{10}\left(\frac{f_{\text{agn}}}{f_{\text{dmonly}}}\right) = A + \frac{B}{1 + \exp\left(-\frac{\log_{10}(M_\Delta^{\text{dmonly}})+C}{D}\right)}, \tag{2.9}$$

where f is defined in Eq. 2.3. The parameters of the fitting function in Eq. 2.9 are presented in Table 2.3.

As an example, we show in Fig. 2.11 the $z = 0$ HMF for M_{200}^{mean} (continuous line) and M_{500}^{mean} (dash-dotted line). The black lines correspond to the uncorrected DMONLY HMF using the fitting formula from Tinker et al. (2008) with Planck best-fitting cosmological parameters (Planck+WP+highL+BAO; Planck Collaboration et al. 2013). The cyan lines show the mass-corrected HMF according to the results from one AGN simulation (AGN 8.0), using the fitting functions that are asymptote to zero at high masses. In the bottom panel we show the relative difference between the curves also, adding in red the results from the AGN simulation with a higher heating temperature (AGN 8.5).

In Fig. 2.12 we show the relative difference in the halo mass function for two different redshifts. Interestingly, the relative difference in the HMF at $z = 1$ is larger than at $z = 0$, a trend that is opposite with respect to the trend in the relative mass change (see Fig. 2.7). This is due to the rapid evolution of the HMF between these two redshifts. The HMF at $z = 1$ is steeper than it is at $z = 0$ and, even though the relative change in halo mass is smaller, this results in a larger change in the HMF at $z = 1$.

2.5.2 Implications for cluster number counts

As discussed in § 2.1, the number density of high-mass haloes and its evolution with redshift are sensitive to a number of fundamental cosmological parameters that control the growth rate of structure. There are numerous ongoing and planned surveys whose main aim is to constrain these parameters by counting the number of high-mass systems on the sky. As we have shown, however, the mass function is also sensitive to the (subgrid) physics of galaxy formation. Here we propagate these effects to show the impact on the predicted number of massive haloes.

We define a cluster to have a mass of $M_{500}^{\text{mean}} \geq 10^{14}\, h^{-1}\, \text{M}_\odot$ and compute the number of haloes above this mass limit at a given time for a comoving volume element.

More specifically, we calculate the function:

$$\mathcal{N}(z) = \frac{dV}{dz} \int_{M_1}^{M_2} n(M, z)dM, \tag{2.10}$$

where $n(M, z)$ represents the HMF and dV/dz is the comoving volume element, which in a flat universe takes the form:

$$\frac{dV}{dz} = 4\pi r^2(z)\frac{dr}{dz}(z), \tag{2.11}$$

with $r(z)$ denoting the comoving radial distance out to redshift z:

$$r(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')}. \tag{2.12}$$
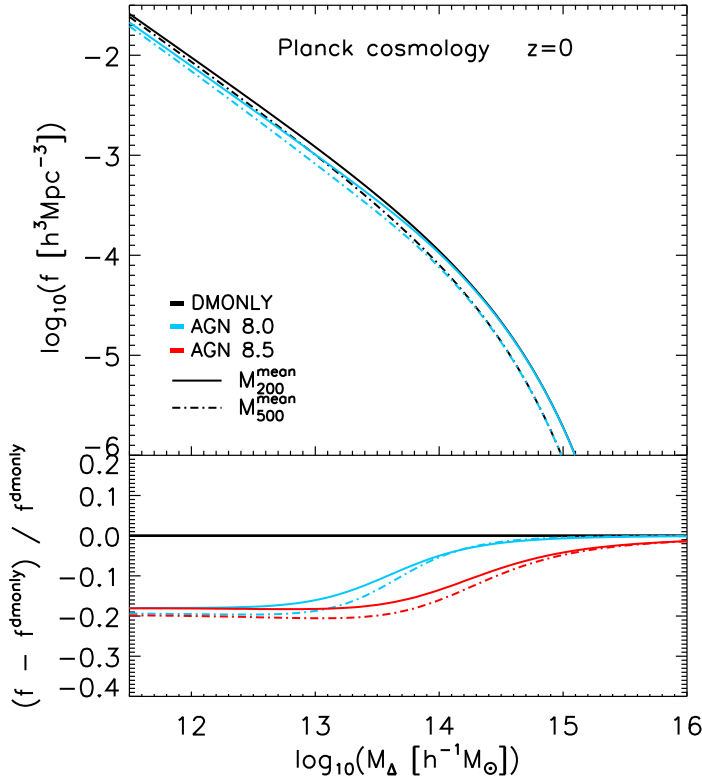
Figure 2.11: In the top panel the black lines show the halo mass function computed using the fitting formula from Tinker et al. (2008) fot Planck cosmological parameters (Planck Collaboration et al. 2013). The cyan curves show the Tinker HMF but corrected for the change in mass calibrated on the AGN 8.0 simulations. In the bottom panel we show the relative difference with respect to the uncorrected Tinker HMF, and we also add the relative change when AGN 8.5 is used (red). The different lines are: M_{200}^{mean} (continuous line) and M_{500}^{mean} (dash-dotted line).
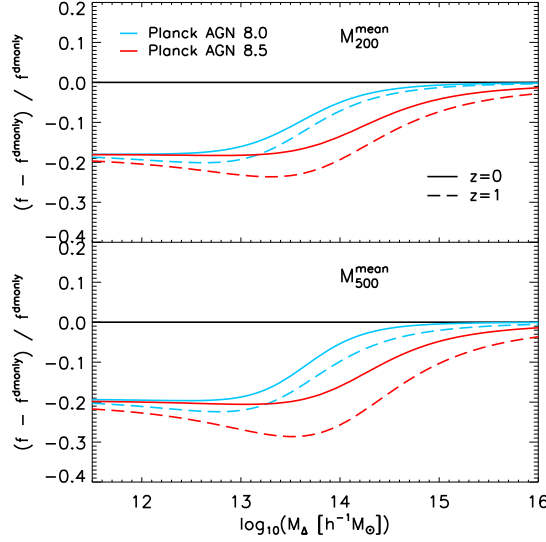
Figure 2.12: Relative differences of the halo mass function when the correction for the change in mass is applied. In the top panel are shown the results for M_{200}^{mean} at redshifts $z = 0$ (continuous line) and at $z = 1$ (dashed line). Results in the bottom panel refer to M_{500}^{mean}.

We account for the effect of baryon physics by using the HMF modified to include the change in the mass of haloes as described in the previous section. We examine only the AGN models since we know that SN feedback alone is insufficient to change the masses of haloes in this mass range (and also leads to significant overcooling in disagreement with observations). For every redshift we integrate the halo mass function at that redshift corrected by the effect on the total mass at redshift zero, in this way we assume that the relative change in mass does not vary with redshift. A more consistent way would be to interpolate the parameters of the available fitting functions for every given redshift where we calculate the number counts. Since the difference in the fitting functions at the three redshifts for which we compute them, z=0, 0.5, 1, are small, and because we have many more haloes at z=0 than at higher redshift, we assume that the change in mass does not vary with redshift. Moreover, we neglect issues having to do with survey completeness and selection effects that can alter the cluster number counts. These issues clearly need to be properly addressed when comparing to a specific survey.

In Fig. 2.13 we show the comoving number density of haloes more massive than $M_{500}^{\mathrm{mean}} = 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ as a function of redshift for the WMAP7 (dashed black line) and the Planck cosmology (continuous black line) predicted by Tinker et al. (2008), i.e. for a DM only universe. We also show the effect of using the HMF corrected for the change in mass of the haloes calibrated on the AGN 8.0 (continuous cyan line) and AGN 8.5 (continuous red line) simulations. Here we see that the inclusion of baryon physics (AGN feedback in particular) can lead to an effect that is of the same order as a change between the best-fit WMAP7 and Planck cosmologies. Thus, for precision cosmological work it is clear that the effects of baryon physics on the HMF must be modelled.

Finally, we tested what the effect is of assuming that the change in mass does not vary with redshift by adding in Fig. 2.13 two points, one for each model, that indicate the cluster number count at $z = 1$ using the mass correction predicted by the simulations for the same redshift.

The difference between the points and the lines represents the error introduced by assuming a change in the mass relation that does not vary with redshift. This difference is indeed very small, thus validating our initial assumption. It is important to note that the fitting function at $z = 0$ was based on 982 haloes with masses $M_{500}^{\mathrm{mean}} > 10^{14}\,h^{-1}\,\mathrm{M_\odot}$ for the AGN 8.0 simulation (862 for the AGN 8.5), while for the fitting function at z=1 only 43 haloes are considered for the same mass range. This means that the fitting function at z=0 is better constrained than the fitting function at $z = 1$.

It is important to note, however, that the magnitude of the effect is quite sensitive to the mass limit used to define a cluster. Here we have adopted a mass limit of $M_{500}^{\mathrm{mean}} = 10^{14}\,h^{-1}\,\mathrm{M_\odot}$, which is roughly comparable to that of surveys such as XMM-XXL, XCS, and GAMA. Surveys such as REFLEX II and Planck, which have mass limits that are a factor of several higher than this, will be considerably less sensitive to the effects of baryons on the HMF. When a higher mass limit of $M_{500}^{\mathrm{mean}} = 10^{15}\,h^{-1}\,\mathrm{M_\odot}$ is used, correcting the masses of the cluster according to the results for AGN 8.5 reduces the cluster counts only by 10% at $z = 1$ (1% for AGN 8.0). Instead, for the same mass limit and redshift, the change in the cosmological parameters from Planck to WMAP7 has a much bigger impact, reducing the cluster counts by 50%. Note however that we only have 6 haloes with masses $M_{500}^{\mathrm{mean}} > 10^{15}\,h^{-1}\,\mathrm{M_\odot}$ at z=0 for AGN 8.5 (7 for AGN 8.0) to constrain the behaviour, in the high-mass regime, of the fitting function that we apply to derive the corrected cluster number counts. Because of the poor statistics for very high mass haloes in our simulation box the results for this higher mass limit are less robust than the correction to the cluster number counts with a limit mass of $M_{500}^{\mathrm{mean}} = 10^{14}\,h^{-1}\,\mathrm{M_\odot}$. Nonetheless, this result suggests that the impact of baryons on the cluster number counts becomes less severe when a higher mass limit is adopted.

2.6   Comparison with previous studies

There have been several recent works examining the inclusion of baryons on the HMF. In this section we compare our findings to those of previous studies.

Cui et al. (2012) explored the HMF in simulations with radiative cooling, star formation, and supernova feedback but no AGN feedback[5]. For massive haloes, they concluded that the HMF is affected at only the few per cent level. This is generally consistent with the results of our REF model, which has same physics but with somewhat different subgrid parametrisations. However, as shown by many previous authors, models that neglect AGN feedback lead to a significant overcooling problem at high masses resulting in groups and clusters with unrealistic properties.

Balaguera-Antolínez & Porciani (2013) study the effect of baryon physics on the HMF by combining the HMF from dark matter only simulations with the observed trend in baryon fraction with halo mass of local groups and clusters. They obtain a (negative) difference in the cluster mass function of 10-15%, depending on which observational data set they use. This is similar to what we find in our AGN models, which we have shown to reproduce the observations (see Fig. 2.5 and Le Brun et al. (2014)). An important caveat of this simple method is that by relying on observations this limits the applicability of this method to relatively low redshifts, where the baryon fractions of clusters can be reasonably well measured. However, even at low redshifts care must be taken to assess the importance of selection effects and mass estimation biases (see Le Brun et al. 2014 for further discussion).

---

[5]In the final stages of preparing this paper, Cui et al. (2014) posted a paper to the arXiv exploring the effects of AGN feedback on the HMF. Consistent with our results, they find a shift of $\approx -20\%$ compared to a DM only simulation.
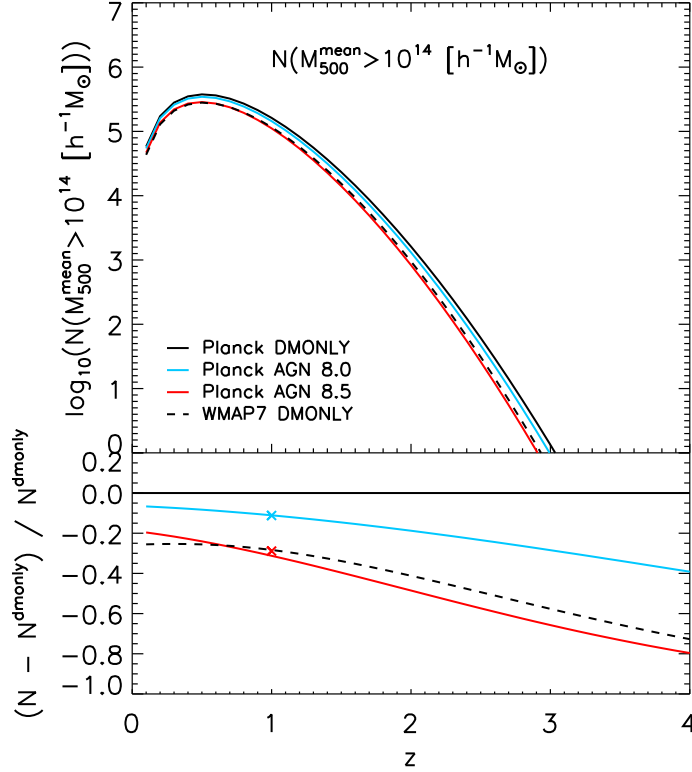
Figure 2.13: The comoving number density of haloes more massive than $M_{500}^{\mathrm{mean}} = 10^{14}\,h^{-1}\,\mathrm{M}_\odot$ as a function of redshift for the WMAP7 (dashed black line) and the Planck cosmology (continuous black line) predicted by Tinker et al. (2008), i.e. for a DM only universe. We also show the effect of using the HMF corrected for the change in mass of the haloes calibrated on the AGN 8.0 simulation (continuous cyan line) and for AGN 8.5 simulation (continuous red line). For this analysis we assume that the relative change in mass does not vary with redshift. In the bottom panel we show the relative difference of the functions with respect to the Planck DMONLY case. The two crosses at $z = 1$ represent the values obtained by applying the change in mass fitting function at the same redshift. It is clear from this figure that the effect on the masses of the haloes due to baryon physics can produce a difference of the same order as the one produced by interesting variations of the cosmological parameters.

Martizzi et al. (2014) use and extend the formalism in Balaguera-Antolínez & Porciani (2013) by allowing for the associated expansion/contraction of the dark matter component. They calibrate their models using a sample of 51 zoom simulations of clusters, as opposed to using observed baryon fractions. Surprisingly, they find that even with the inclusion of AGN feedback the obtained baryon fraction is very close to universal, in contradiction with recent observations. The net effect is that they obtain a small (5%) *positive* variation in the HMF for the runs with AGN feedback, in stark contrast with our results. We hypothesise that if their simulations had simultaneously matched the stellar and gas mass fractions of observed groups and clusters, they would have found a similar negative offset in the HMF.

Cusworth et al. (2014) (see also Stanek et al. 2009) use the Millennium Gas simulations, which include a run with 'pre-heating' and cooling (PC) as well as a hybrid simulation (FO, for feedback only) that combines a semi-analytic model of galaxy formation (with AGN feedback) with a non-radiative cosmological hydrodynamical simulation. Note that both models have been tuned to some degree to match the properties of local groups and clusters. Both result in a shift in the local HMF of around -15%, comparable to what we find in our self-consistent AGN models. Given the relatively large differences in the subgrid implementations of the PC and FO models and our own AGN models, it is plausible that there will be much larger differences in the predictions at higher redshifts.

Thus far we have focused on the high-mass end and the implications for cluster number counts. At the low-mass end we compare our results with the work of Sawala et al. (2013). We find good agreement when we use the high-resolution version of the REF simulation in order to be able to resolve smaller haloes. As already mentioned, our findings suggest an increase in the relative change in mass towards smaller halo masses, although our work suggests a slightly smaller effect due to the fact that the GIMIC simulation used in Sawala et al. (2013) has somewhat more efficient SN feedback (leading to haloes with slightly lower-than-observed stellar mass fractions; see McCarthy et al. 2012).

2.7 Summary and Conclusions

In this paper we have explored the effects of the introduction of important baryon physics associated with galaxy formation on the total mass of haloes, the mass profile up to large radii ($10R_{200}^{\mathrm{dmonly}}$) and the halo mass function (HMF). In order to isolate the effects of baryon physics, we used several simulations from the OWLS project with identical initial conditions, box size and resolution, starting from only gravitationally interacting particles (DMONLY). On top of that we added gas hydrodynamics, star formation and primordial cooling in the NOSN_NOZCOOL simulation, an implementation of kinetic SN feedback in the NOZCOOL simulation, introduction of metal-line cooling in the REF simulation and finally the seeding and growth of black holes and AGN feedback in the AGN 8.0 and AGN 8.5 simulations. We also explored a different prescription for SN feedback in which the mass loading and the wind velocity depend on the local star-forming gas density in the WDENS simulation. By comparing the results of different simulations, we were able to isolate the importance of different galaxy formation physics on the HMF.

An important aspect of this work was to compare exactly the same set of haloes when different physical processes are introduced in order to isolate the effects of baryons without introducing a bias. We therefore applied a halo finding and matching algorithm that takes advantage of the fact that the dark matter particles have unique IDs and the simulations all used identical initial conditions.

Using the matched haloes, we compared the relative difference in the total mass of haloes between the DMONLY and the baryonic simulations. Our results span nearly four orders of magnitude in halo mass, $10^{11.5} < M_{200}^{\mathrm{crit}}/[\,h^{-1}\,\mathrm{M}_\odot] < 10^{15.2}$. We found that at the low-mass end, SN feedback produces haloes that are 20% less massive, at $z = 0$, with respect to their DMONLY counterparts, due to the ejection of baryons from the haloes as well as some expansion of the dark matter itself. This difference decreases with increasing halo mass, as the escape velocity gradually becomes too high for the gas to escape, reaching no difference for masses $M_{200}^{\mathrm{crit}} > 10^{12.5}\,h^{-1}\,\mathrm{M}_\odot$. The mass range over which SN feedback can alter the HMF can be extended somewhat if the wind velocity increases with local gas density (as in WDENS). However, only AGN feedback can produce a substantial alteration of the HMF of galaxy groups and clusters with halo masses up to $M_{200}^{\mathrm{crit}} = 10^{14.8}\,h^{-1}\,\mathrm{M}_\odot$.

A direct effect of the change in the total mass of the haloes is the modification of the HMF. Similarly to the total mass variation, we found that supernova feedback is particularly important in shaping the HMF in the mass range $10^{11.5} < M_{200}^{\mathrm{crit}} < 10^{12.5}\,h^{-1}\,\mathrm{M}_\odot$, with a decrease of 20% with respect to the DM-only simulation. Including only stellar feedback does not produce a significant effect for haloes more massive than $M_{200}^{\mathrm{crit}} = 10^{13}\,h^{-1}\,\mathrm{M}_\odot$. In the higher mass range, AGN feedback can induce a similar 20% decrease with respect to the DM-only scenario.

Baryon physics is able to significantly change the total mass profiles of haloes out to several times the virial radius. This means that also the environment in which the haloes reside has significantly different properties with respect to the simulations with only gravitationally interacting particles. This effect could be very important for gravitational lensing measurements, which are sensitive to the mass profile of haloes (e.g. Mead et al. 2010; Semboloni et al. 2011; Killedar et al. 2012; van Daalen et al. 2014).

We have provided a set of analytic functions that can be used to correct the masses of DMONLY simulated haloes for the presence of baryons (for several mass definitions). We have shown that the change in mass of the haloes due to baryon physics does not depend on small changes in the values of the cosmological parameters. We also used the analytic fitting formulas to correct the Tinker et al. universal HMF for the effects of baryons. In this way we are able to predict the abundance of haloes in different cosmologies. In particular, we showed

that the shift in the HMF is about 20%, which has important implication for cluster number counts (e.g., the effect of baryon physics is of the same order as switching the cosmological parameters between WMAP7 and Planck). To help alleviate this problem, we advocate using only the highest-mass clusters for number counts test, for example $M_{500}^{\mathrm{mean}} > 10^{15}\, h^{-1}\, \mathrm{M}_\odot$, where the effect of baryon physics on the mass of the haloes is far less pronounced.

In conclusion we have shown that the masses of haloes inferred from DM-only simulations are not reliable, and when baryon physics is included this can lead to a difference up to 20% in the mass of the halo and a similar shift in the HMF. The magnitude of the effect *far* exceeds the percentage precision requirement on the HMF (Wu et al. 2010) for future surveys that aim to constrain the dark energy equation of state, such as XMM-XXL, eROSITA, Planck, DES, Euclid and LSST. Thus, it is beyond question that baryons must be properly modelled for future precision cosmological studies, as well as for any other theoretical studies that require halo masses to be known with better than $\sim 20\%$ accuracy.

Acknowledgements

Bibliography

Abadi, M. G., Navarro, J. F., Fardal, M., Babul, A., & Steinmetz, M. 2010, MNRAS, 407, 435

Balaguera-Antolínez, A., & Porciani, C. 2013, JCAP, 4, 22

Baugh, C. M. 2006, Reports on Progress in Physics, 69, 3101

Bett, P., Eke, V., Frenk, C. S., Jenkins, A., & Okamoto, T. 2010, MNRAS, 404, 1137

Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53

Bryan, S. E., Kay, S. T., Duffy, A. R., et al. 2013, MNRAS, 429, 3316

Bryan, S. E., Mao, S., Kay, S. T., et al. 2012, MNRAS, 422, 1863

Budzynski, J. M., Koposov, S. E., McCarthy, I. G., & Belokurov, V. 2014, MNRAS, 437, 1362

Casarini, L., Macciò, A. V., Bonometto, S. A., & Stinson, G. S. 2011, MNRAS, 412, 911

Cooray, A., & Sheth, R. 2002, Phys. Rep., 372, 1

Cui, W., Borgani, S., Dolag, K., Murante, G., & Tornatore, L. 2012, MNRAS, 423, 2279

Cui, W., Borgani, S., & Murante, G. 2014, preprint (arXiv:1402.1493), arXiv:1402.1493

Cusworth, S. J., Kay, S. T., Battye, R. A., & Thomas, P. A. 2014, MNRAS, 439, 2485

Dalla Vecchia, C., & Schaye, J. 2008, MNRAS, 387, 1431

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Duffy, A. R., Schaye, J., Kay, S. T., et al. 2010, MNRAS, 405, 2161

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605

Ettori, S., Dolag, K., Borgani, S., & Murante, G. 2006, MNRAS, 365, 1021

Frenk, C. S., & White, S. D. M. 2012, Annalen der Physik, 524, 507

Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16

Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203

Guillet, T., Teyssier, R., & Colombi, S. 2010, MNRAS, 405, 525

Gustafsson, M., Fairbairn, M., & Sommer-Larsen, J. 2006, Phys. Rev. D, 74, 123522

Haas, M. R., Schaye, J., Booth, C. M., et al. 2013, MNRAS, 435, 2931

Jenkins, A., Frenk, C. S., White, S. D. M., et al. 2001, MNRAS, 321, 372

Jing, Y. P., Zhang, P., Lin, W. P., Gao, L., & Springel, V. 2006, ApJ, 640, L119

Kazantzidis, S., Kravtsov, A. V., Zentner, A. R., et al. 2004, ApJ, 611, L73

Khandai, N., Di Matteo, T., Croft, R., et al. 2014, preprint (arXiv:1402.0888), arXiv:1402.0888

Killedar, M., Borgani, S., Meneghetti, M., et al. 2012, MNRAS, 427, 533

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Kravtsov, A. V., Nagai, D., & Vikhlinin, A. A. 2005, ApJ, 625, 588

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, preprint (arXiv:1110.3193), arXiv:1110.3193

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, preprint (arXiv:1312.5462), arXiv:1312.5462

Lin, Y.-T., Stanford, S. A., Eisenhardt, P. R. M., et al. 2012, ApJ, 745, L3

Lukić, Z., Heitmann, K., Habib, S., Bashinsky, S., & Ricker, P. M. 2007, ApJ, 671, 1160

Martizzi, D., Mohammed, I., Teyssier, R., & Moore, B. 2014, MNRAS, 440, 2290

Martizzi, D., Teyssier, R., & Moore, B. 2013, MNRAS, 432, 1947

Maughan, B. J., Jones, C., Forman, W., & Van Speybroeck, L. 2008, ApJS, 174, 117

McCarthy, I. G., Schaye, J., Bower, R. G., et al. 2011, MNRAS, 412, 1965

McCarthy, I. G., Schaye, J., Font, A. S., et al. 2012, MNRAS, 427, 379

McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, MNRAS, 406, 822

Mead, J. M. G., King, L. J., Sijacki, D., et al. 2010, MNRAS, 406, 434

Mehrtens, N., Romer, A. K., Hilton, M., et al. 2012, MNRAS, 423, 1024

Pierre, M., Pacaud, F., Juin, J. B., et al. 2011, MNRAS, 414, 1732

Pillepich, A., Porciani, C., & Reiprich, T. H. 2012, MNRAS, 422, 44

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, preprint (arXiv:1303.5076),
    arXiv:1303.5076

Pratt, G. W., Croston, J. H., Arnaud, M., & Böhringer, H. 2009, A&A, 498, 361

Read, J. I., Lake, G., Agertz, O., & Debattista, V. P. 2008, MNRAS, 389, 1041

Reed, D., Gardner, J., Quinn, T., et al. 2003, MNRAS, 346, 565

Romano-Díaz, E., Shlosman, I., Heller, C., & Hoffman, Y. 2009, ApJ, 702, 1250

Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19

Sawala, T., Frenk, C. S., Crain, R. A., et al. 2013, MNRAS, 431, 1366

Schaye, J., & Dalla Vecchia, C. 2008, MNRAS, 383, 1210

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536

Seljak, U. 2000, MNRAS, 318, 203

Semboloni, E., Hoekstra, H., & Schaye, J. 2013, MNRAS, 434, 148

Semboloni, E., Hoekstra, H., Schaye, J., van Daalen, M. P., & McCarthy, I. G. 2011, MNRAS,
    417, 2020

Spergel, D. N., Bean, R., Doré, O., et al. 2007, ApJS, 170, 377

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Stanek, R., Rudd, D., & Evrard, A. E. 2009, MNRAS, 394, L11

Sun, M., Voit, G. M., Donahue, M., et al. 2009, ApJ, 693, 1142

Teyssier, R., Moore, B., Martizzi, D., Dubois, Y., & Mayer, L. 2011, MNRAS, 414, 195

Theuns, T., Viel, M., Kay, S., et al. 2002, ApJ, 578, L5

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

Tissera, P. B., White, S. D. M., Pedrosa, S., & Scannapieco, C. 2010, MNRAS, 406, 922

van Daalen, M. P., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 3649

van Daalen, M. P., Schaye, J., McCarthy, I. G., Booth, C. M., & Vecchia, C. D. 2014, MNRAS, 440, 2997

Viel, M., Schaye, J., & Booth, C. M. 2013, MNRAS, 429, 1734

Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, ApJ, 640, 691

Warren, M. S., Abazajian, K., Holz, D. E., & Teodoro, L. 2006, ApJ, 646, 881

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS, 399, 574

Wu, H.-Y., Zentner, A. R., & Wechsler, R. H. 2010, ApJ, 713, 856

Zentner, A. R., Semboloni, E., Dodelson, S., et al. 2013, Phys. Rev. D, 87, 043509

2.A   Resolution test

In this section we perform resolution tests for the analysis done in § 2.3.1. In Fig. 2.14 we repeat the same analysis performed on the standard resolution (REF green lines) using eight times better mass resolution (REF L050N512) and eight times worse mass resolution with a larger box size (REF L400N1024). We do not have the higher resolution version of the AGN models (blue lines), but we argue that at the low-mass end the behaviour is similar to the REF model since the AGN feedback is not efficient for those low-mass haloes. Instead, we show the effect of a low-resolution version (AGN L400N1024). The high-resolution run and the standard run agree reasonably well for masses $M_{200}^{\mathrm{dmonly}} > 10^{11.5}\,h^{-1}\,\mathrm{M_\odot}$, and this is the reason we choose this mass as the lower limit in our analysis. The upturn that is visible at low masses in both simulations is a resolution artefact since it is present at the low-mass end of every simulation but shifted by a factor of eight in mass, i.e. the difference in mass resolution between the two simulations. The vertical arrows show the chosen resolution limits that approximately correspond to 600 DM particles in the DMONLY simulation with standard resolution. In the other panels we show resolution tests for the other mass definitions used. We find that $M_{200}^{\mathrm{dmonly}} = 10^{11.5}\,h^{-1}\,\mathrm{M_\odot}$ is also a good choice for M_{500} and M_{fof}. However, for M_{2500} a better choice for the resolution mass limit is $M_{200}^{\mathrm{dmonly}} = 10^{12}\,h^{-1}\,\mathrm{M_\odot}$.

We also show in Fig. 2.15 the same resolution test for M_{500}^{mean} and for M_{fof} when they are plotted as functions of DMONLY mass with the same mass definition. These relations are used to produce the fitting functions and here we show that the mass resolution limit is the same as in the previous figure where all the quantities are plotted as a function of $M_{200}^{\mathrm{dmonly}}$. The same applies to the other mass definitions used in the fitting function.
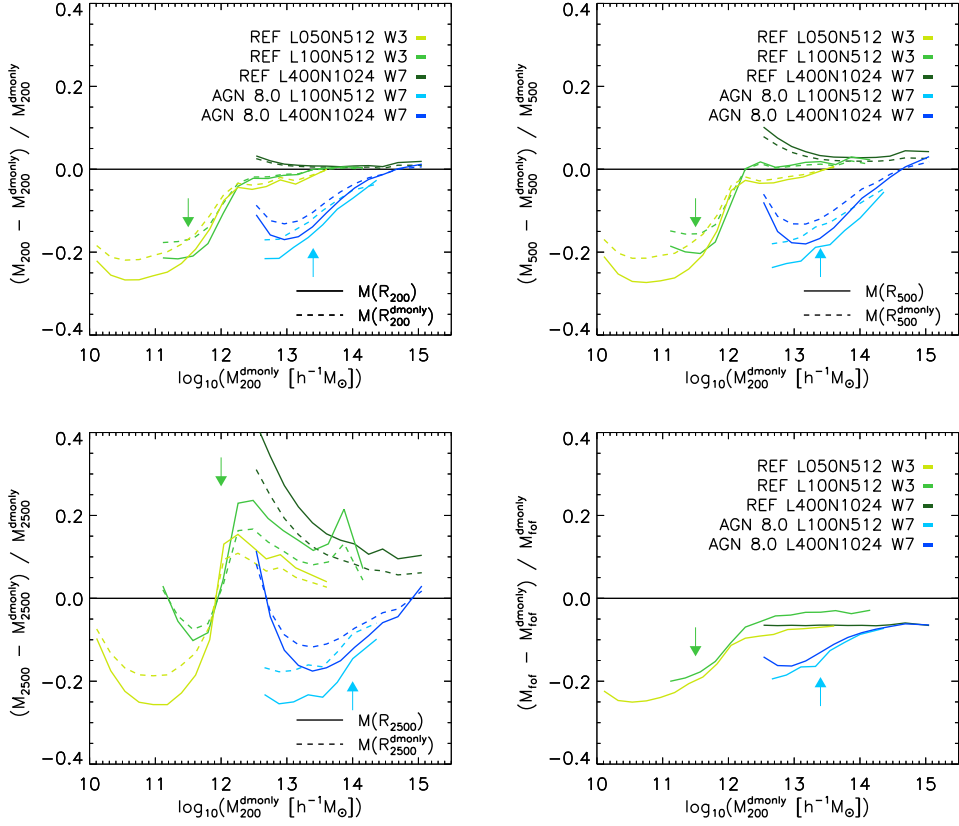
Figure 2.14: Resolution test for the analysis done in § 2.3.1. We show the same panels as in Fig. 2.2 with the difference that we include only the results from the REF simulations and AGN 8.0, and for every simulation we show the effect of changing the resolution. The arrows show the resolution limits adopted in this work, the arrows that are pointing downward refer to the resolution limit for the simulations done in the 100 h^{-1} Mpc box, while the upwards pointing arrows show where we switch to the larger box size.
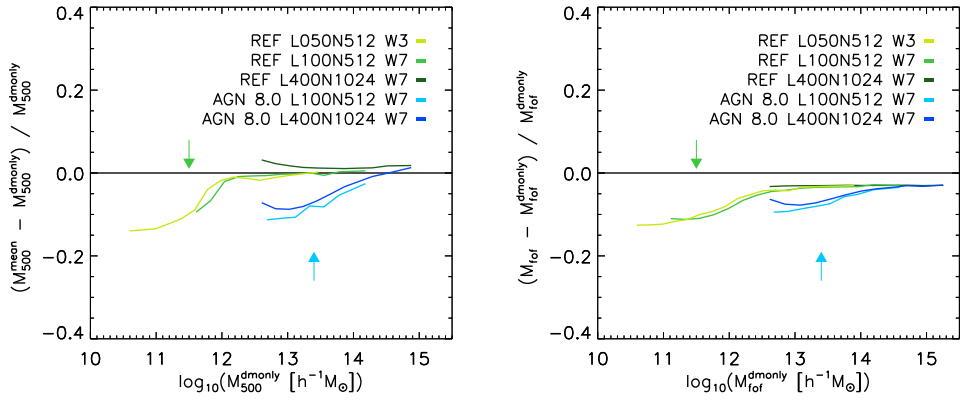
Figure 2.15: Resolution test for the values used in the fitting functions. We explicitly show the resolution test for M_{500}^{mean} and for M_{fof}. The resolution limits are the same as the one used when the haloes are binned in M_{200}^{crit}.

3

The alignment and shape of dark matter, stellar, and hot gas distributions in the EAGLE and cosmo-OWLS simulations

We report the alignment and shape of dark matter, stellar, and hot gas distributions in the EAGLE and cosmo-OWLS simulations. The combination of these state-of-the-art hydrodynamical cosmological simulations enables us to span four orders of magnitude in halo mass ($11 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) \leq 15$), a wide radial range ($-2.3 \leq \log_{10}(r/[\,h^{-1}\,\mathrm{Mpc}]) \leq 1.3$) and redshifts $0 \leq z \leq 1$. The shape parameters of the dark matter, stellar and hot gas distributions follow qualitatively similar trends: they become more aspherical (and triaxial) with increasing halo mass, radius and redshift. We measure the misalignment of the baryonic components (hot gas and stars) of galaxies with their host halo as a function of halo mass, radius, redshift, and galaxy type (centrals *vs* satellites and early- *vs* late-type). Overall, galaxies align well with the local distribution of the total (mostly dark) matter. However, the stellar distributions on galactic scales exhibit a median misalignment of about 45-50 degrees with respect to their host haloes. This misalignment is reduced to 25-30 degrees in the most massive haloes ($13 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) \leq 15$). Half of the *disc* galaxies in the EAGLE simulations have a misalignment angle with respect to their host haloes larger than 40 degrees. We present fitting functions and tabulated values for the probability distribution of galaxy-halo misalignment to enable a straightforward inclusion of our results into models of galaxy formations based on purely collisionless N-body simulations.

3.1   Introduction

The topology of the matter distribution in the Universe is well described as a web-like structure comprising voids, sheets, filaments and haloes. This so-called *cosmic web* arises naturally from the gravitational growth of small initial perturbations in the density field of an expanding cold dark matter dominated (ΛCDM) Universe. The evolution of the properties of the large-scale cosmic web is governed by the dominant components, i.e. dark energy and dark matter, while baryons are expected to trace the distribution of the latter. Specifically, galaxies reside in dark matter haloes and trace them in terms of their positions and, to first order, in terms of their shapes and mutual alignment, albeit in a biased fashion due to the dissipative processes they experience during galaxy formation. Theoretical studies of this *galaxy bias* have been ongoing for several decades (e.g. Kaiser 1984; Davis et al. 1985).

It has become apparent that when galaxies are used to infer the properties of the underlying dark matter distribution, it is convenient to bisect this investigation into two steps: the relation between galaxies and haloes and the relation between haloes and the underlying density field. The latter can be studied directly via cosmological N-body simulations, whereas the former is a far more complicated relation that is potentially affected by virtually all the physical processes associated with galaxy formation. For instance, while the triaxial shape of dark matter haloes is understood in terms of the collisionless nature of dark matter coupled with ellipsoidal collapse, galaxies manifest themselves in a plethora of morphologies ranging from thin to bulge-dominated discs and to ellipsoidals and this is undoubtedly related to the redistribution of angular momentum occurring during galaxy formation and evolution which, in turn, depends on the physical processes in operation. Thus, the characterization of the way galaxy shapes relate to their host haloes holds the potential to unveil the relevant physical mechanisms behind such a rich manifestation of galaxy types.

Numerical simulations have been used to study the mutual alignment of galaxies with their own host haloes. For instance, van den Bosch et al. (2002), Chen et al. (2003), Sharma & Steinmetz (2005), Bett et al. (2010) and Sales et al. (2012) have shown that the angular momentum distributions of gas and dark matter components are partially aligned, with a typical misalignment angle of $\sim 30°$, although this might predominantly apply to disc galaxies. On the other hand, central ellipticals are expected to be aligned with their host haloes if they are formed by mergers (Dubinski 1998; Naab et al. 2006; Boylan-Kolchin et al. 2006), because the orientations of the central ellipticals and of the host dark matter haloes are determined by respectively the orbital angular momenta of their (correlated) progenitor galaxies and haloes. Observationally, there exist different indications of the presence of a misalignment between galaxies and their host haloes. However, different studies have reached somewhat conflicting conclusions about the typical values of this misalignment angle (see e.g. Heymans et al. 2004; Kang et al. 2007; Wang et al. 2008; Okumura et al. 2009).

Beyond its theoretical relevance, the misalignment of a galaxy with its own host halo can be a source of systematics for those studies that aim to: infer the shape of dark matter haloes or constrain cosmological parameters via the measurement of the galaxy shape correlation function. Several current and forthcoming weak lensing surveys (e.g. KiDS, DES, LSST, and Euclid[1]) will achieve the statistical power to probe, observationally, halo shapes and to obtain exquisite measurements of the apparent alignment of galaxy shapes –*cosmic shear*– due to the gravitational lensing effect caused by the underlying (dark) matter distribution. It is

---

[1]KIDS: KIlo-Degree Survey,
http://www.astro-wise.org/projects/KIDS/;
DES: Dark Energy Survey, https://www.darkenergysurvey.org;
LSST: Large Synoptic Survey Telescope, http://www.lsst.org;
Euclid: http://www.euclid-ec.org

therefore of great importance to guide the interpretation of the measured signal with numerical simulations. For instance, the link between the shape of the visible, baryonic matter and the structure of the underlying dark matter distribution, as well as their mutual orientation can be examined. To this end it is necessary to complement the expectations derived from cosmological N-body simulations with the properties of galaxies as inferred from small-scale, high-resolution hydrodynamic simulations and/or semi-analytical models (e.g. Joachimi et al. 2013; van den Bosch et al. 2002; Croft et al. 2009; Hahn et al. 2010; Bett et al. 2010; Bett 2012).

In this paper, we extend previous work by exploiting the wealth of information encoded in hydro-cosmological simulations in which the main physical processes responsible for galaxy formation and evolution are simultaneously at play, thus leading to a more realistic realization of the galaxy-dark matter connection. We use the OverWhelmingly Large Simulations (cosmo-OWLS, Schaye et al. 2010; Le Brun et al. 2014; McCarthy et al. 2014) and the Evolution and Assembly of GaLaxies and their Environments (EAGLE, Schaye et al. 2015; Crain et al. 2015) project. This approach has the advantage that the processes that lead to galaxy formation are self-consistently incorporated in the simulations and are therefore accounted for in the resulting galaxy and halo shapes, as well as in their correlation. During the late phase of this project, a study adopting a similar methodology was Tenneti et al. (2014), hereafter Ten14, which has many aspects in common with our analysis. Throughout the paper, we will therefore compare mutual findings.

Our study is, however, unique as a consequence of several key features of our simulations and analysis. As detailed in § 3.2.1, the use of cosmo-OWLS and EAGLE provides us sufficient cosmological volume and resolution, both of which are crucial for the reliability and the applicability of our results. Specifically, we span four orders of magnitude in halo mass ($11 \leq \log(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$) and over six orders of magnitude in subhalo mass M_{sub}, enabling us to investigate spatial variations of the shape of galaxies and haloes from galactic to cosmological scales. Furthermore, the combination of EAGLE and cosmo-OWLS forms a set of simulations that reproduces the observed abundance of galaxies as a function of stellar mass (the galaxy stellar mass function) at both low ($\log(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 13$) and high ($13 \leq \log([M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$) halo masses. Moreover, it has been shown that the cosmo-OWLS simulations reproduce various (X-ray and optical) observed properties of galaxy groups (Crain et al. 2010; McCarthy et al. 2010; Le Brun et al. 2014) as well as the observed galaxy mass function for haloes more massive than $\log(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) = 13$. Finally, the galaxy size distribution in EAGLE reproduces the observed one (Schaye et al. 2015).

This paper is organized as follows. We summarize the properties of the simulations in § 3.2, where we also introduce the technical definitions used throughout the paper. We highlight some caveats to the shape and angle estimates related to the feedback implementation in § 3.3. In § 3.4 we present the results concerning the sphericity and triaxiality of dark matter haloes, as well as those of the stellar and the hot X-ray emitting gas distribution. The (mis)alignment of the baryonic components with their host haloes is addressed in § 3.5. We summarize and comment on our results in § 3.6.

Throughout the paper, we assume a flat ΛCDM cosmology with massless neutrinos. Such a cosmological model is characterized by five[2] parameters: $\{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, \sigma_8, n_{\mathrm{s}}, h\}$. The simulations used in this paper were run with two slightly different sets of values for these parameters. Specifically, we will refer to PLANCK as the set of cosmological values suggested by the Planck mission $\{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, \sigma_8, n_{\mathrm{s}}, h\} = \{0.307, 0.04825, 0.8288, 0.9611, 0.6777\}$ (Table 9; Planck Collaboration et al. 2014), whereas WMAP7 refers to the cosmological parameters $\{\Omega_{\mathrm{m}}, \Omega_{\mathrm{b}}, \sigma_8, n_{\mathrm{s}}, h\} = \{0.272, 0.0455, 0.728, 0.81, 0.967, 0.704\}$ suggested by the 7th-year

---

[2]Flatness implies that $\Omega_\Lambda = 1 - \Omega_{\mathrm{m}}$.

data release (Komatsu et al. 2011) of the WMAP mission.

3.2 Simulations and Technical Definitions

3.2.1 Simulations

Throughout the paper, we employ the outputs of four cosmological volumes simulated within the context of two distinct projects: EAGLE (Schaye et al. 2015; Crain et al. 2015) and cosmo-OWLS (Le Brun et al. 2014; McCarthy et al. 2014). We use the former to investigate (well-resolved) smaller halo masses in relatively small volumes; whereas the latter is used to study more massive haloes in larger volumes. Table 3.1 lists all relevant specifics of these simulations.

Both EAGLE and cosmo-OWLS were run using a modified version of the *N*-Body Tree-PM smoothed particle hydrodynamics (SPH) code GADGET 3, which was last described in Springel (2005). The main modifications are the formulation of the hydrodynamics, the time stepping and, most importantly, the subgrid physics. All the simulations used in this work include element-by-element radiative cooling (for 11 elements; Wiersma et al. 2009a), star formation (Schaye & Dalla Vecchia 2008), stellar mass loss (Wiersma et al. 2009b), energy feedback from star formation (Dalla Vecchia & Schaye 2008, 2012), gas accretion onto and mergers of supermassive black holes (BHs; Booth & Schaye 2009; Rosas-Guevara et al. 2013), and AGN feedback (Booth & Schaye 2009; Schaye et al. 2015).

The subgrid physics used in EAGLE builds upon that of OWLS (Schaye et al. 2010), GIMIC (Crain et al. 2009a) and cosmo-OWLS (Le Brun et al. 2014; McCarthy et al. 2014). Furthermore, the EAGLE project brings a number of changes with respect to cosmo-OWLS regarding the implementations of energy feedback from star formation (which is now thermal rather than kinetic), the accretion of gas onto BHs (which now accounts for angular momentum), and the star formation law (which now depends on metallicity). More information regarding technical implementation of hydro-dynamical aspects as well as subgrid physics can be found in Schaye et al. (2015).

Arguably, the most important feature of the EAGLE simulation is the calibration of the subgrid physics parameters to reproduce the observed galaxy mass function and galaxy sizes at redshift zero. One of the key feature of the cosmo-OWLS simulations is that they reproduce optical and X-ray scaling relations of groups and clusters of galaxies. In this work we exploit both these unique features by splitting our range of halo masses into four mass bins and by using a different simulation for each one of them. Specifically, for halo masses below the 'knee' of the galaxy stellar mass function we use EAGLE in order to ensure galaxies form with the 'correct' efficiency and size, whereas for haloes above the 'knee' we use cosmo-OWLS. In practice, we create a composite sample of haloes spanning four orders of magnitude in mass ($11 \leq \log(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) \leq 15$).

Simulation	L	N_{particle}	Cosmology	m_{b} $[\,h^{-1}\,\text{M}_\odot]$	m_{dm} $[\,h^{-1}\,\text{M}_\odot]$	ϵ_{prop} $(\,h^{-1}\,\text{kpc})$	Colour	tag
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
EAGLE Recal	25 (Mpc)	2×752^3	PLANCK	1.5×10^5	8.2×10^5	0.5	purple	EA L025
EAGLE Ref	100 (Mpc)	2×1504^3	PLANCK	1.2×10^6	6.6×10^6	0.2	orange[a]	EA L100
cosmo-OWLS AGN 8.0	200 (h^{-1} Mpc)	2×1024^3	WMAP7	8.7×10^7	4.1×10^8	2.0	blue	CO L200
cosmo-OWLS AGN 8.0	400 (h^{-1} Mpc)	2×1024^3	WMAP7	7.5×10^8	3.7×10^9	4.0	green	CO L400

Table 3.1: Simulations used throughout the paper and their relevant properties. Description of the columns: (1) descriptive simulation name; (2) comoving size of the simulation volume; (3) total number of particles; (4) cosmological parameters used in the simulation; (5) initial mass of baryonic particles; (6) mass of dark matter particles; (7) maximum softening length; (8) colour used for the simulation; (9) simulation name tag.

Simulation tag	mass bin *	M_{200}^{crit} *	M_{star} *	$\sigma_{\log_{10}M_{\text{star}}}$ *	r_{200}^{crit} **	$r_{\text{half}}^{\text{dm}}$ **	$r_{\text{half}}^{\text{star}}$ **	N_{halo}	N_{sat}
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EA L025	$[11 - 12]$	11.31	9.50	0.45	96.0	39.8	2.7	156	24
EA L100	$[12 - 13]$	12.27	10.58	0.26	199.3	93.4	4.9	1008	104
CO L200	$[13 - 14]$	13.16	11.21	0.25	396.4	241.8	53.4	2190	137
CO L400	$[14 - 15]$	14.09	12.06	0.19	805.9	505.1	106.7	1152	26

Table 3.2: Values at $z = 0$ of various quantities of interest in each mass bin. Description of the columns: (1) simulation tag; (2) mass range $\log_{10}(M_{200}/[\,h^{-1}\,\text{M}_\odot])$ of the haloes selected from the simulation; (3) median value of the halo mass $\log_{10}(M_{200}^{\text{crit}})$; (4) median value of the stellar mass ($\log_{10}(M_{\text{star}}/[\,h^{-1}\,\text{M}_\odot])$) considering all the star particles that belong to the halo; (5) standard deviation of the stellar mass distribution $\sigma_{\log_{10}M_{\text{star}}}$; (6) median value of halo radius r_{200}^{crit}; (7) median radius within which half of the mass in dark matter is enclosed; (8) median radius within which half of the mass in stars is enclosed; (9) number of haloes; (10) number of *satellite* haloes.

[a] Cyan is used for Figs. 3.11 and 3.12 where the EA L100 simulation is used in order to improve the statistics for the least massive bin.

* $\log_{10}(M/[\,h^{-1}\,\text{M}_\odot])$

** $R/[\,h^{-1}\,\text{kpc}]$

3.2.2   Halo and subhalo definition

Groups of particles are identified in our simulations by applying the Friends-of-Friends algorithm with linking length 0.2 to the dark matter particles (Davis et al. 1985). The mass M_{200}^{crit} and the radius r_{200}^{crit} of the groups are assigned using a spherical over-density algorithm centred on the minimum of the gravitational potential, as implemented in SUBFIND (Springel et al. 2001; Dolag et al. 2009). From each group, dynamically un-bounded particles are discarded. Thus, subhaloes are identified as a collection of bound particles that reside in a local minimum of the gravitational potential computed using all particle types. The most massive subhalo is the *central* subhalo of a given FoF group and all other subhaloes are *satellites*. Particles that are bound to a subhalo belong exclusively to that subhalo. Correspondingly, central subhaloes do not contain particles that reside in other local minima of the potential, even if those particles are within the subhalo boundary. We define the centre of a subhalo as the position of the particle with the minimal gravitational potential. The subhalo radius can be calculated for each component separately. A commonly used estimate is the radius within which half of the mass in dark matter is included, $r_{\text{half}}^{\text{dm}}$. The mass of a subhalo is the sum of the masses of all the particles that constitute it. For the rest of the paper we will use the term 'halo' to refer both to central and satellite subhaloes, unless otherwise specified.

The masses of subhaloes for both centrals and satellites (according to SUBFIND classifications), are indicated with M_{sub}. However, whenever a distinction is required, we shall use M_{200} and r_{200} to characterize the properties of central haloes.

3.2.3   Shape parameter definitions

A fundamental quantity that describes how matter is spatially distributed is the three-dimensional mass distribution tensor (e.g. Davis et al. 1985; Cole & Lacey 1996),

$$M_{ij} = \sum_{p=1}^{N_{\text{P}}} m_p x_{pi} x_{pj}, \tag{3.1}$$

where N_{P} is the number of all particles that belong to the structure of interest, x_{pi} denotes the element i (with $i, j = 1, 2, 3$ for a 3D particle distribution) of the position vector of particle p, and m_p is the mass of the pth particle. This mass distribution tensor is often referred as the inertia tensor, since the two tensors share the same eigenvectors (see Zemp et al. 2011, for a discussion) and, for most astrophysical purposes, those eigenvectors encode the information of interest. Throughout this paper we will refer to the mass distribution tensor as the inertia tensor to conform to the jargon used in the literature.

The eigenvalues of the inertia tensor will be denoted as λ_i (with $i = 1, 2, 3$ for a 3D particle distribution as in our case). Given a particle distribution inertia tensor, the modulus of the major, intermediate, and minor axes of the corresponding ellipsoid can be written in terms of these eigenvalues as $a = \sqrt{\lambda_1}$, $b = \sqrt{\lambda_2}$, and $c = \sqrt{\lambda_3}$, respectively. We interpret this ellipsoid as an approximation to the shape of the halo. Specifically, the sphericity and triaxiality parameters, S and T, are defined as

$$S = \frac{c}{a}, \quad \text{and } T = \frac{a^2 - b^2}{a^2 - c^2}. \tag{3.2}$$

A purely spherical halo will have $S = 1$ with T being undefined. Low values of T (i.e. $T \to 0$) correspond to oblate haloes while high values (i.e. $T \to 1$) correspond to prolate haloes.

We note that the computation of shape parameters in a spherical region biases the shape towards higher sphericity. When computing shapes of dark matter haloes in spherical regions
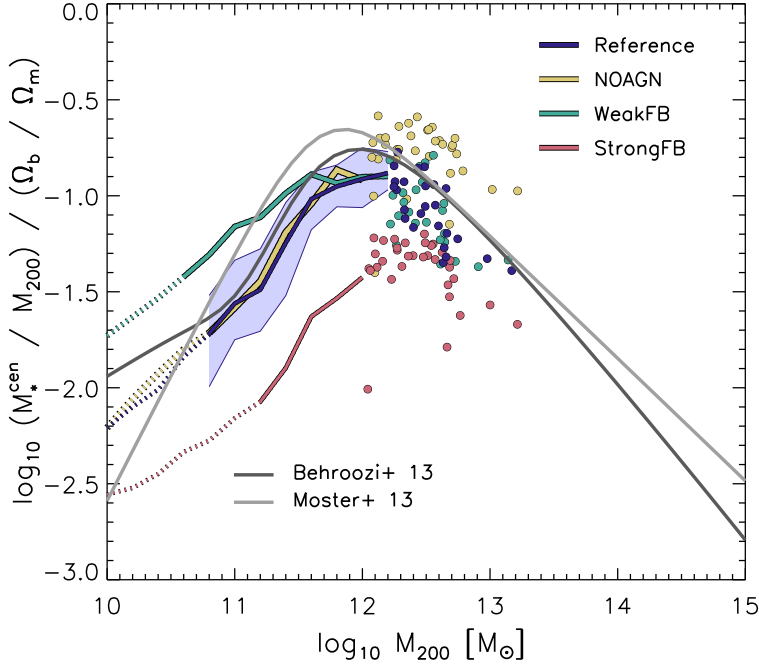
Figure 3.1: The stellar mass to halo mass ratio of central galaxies as a function of halo mass, normalised by the cosmic baryon fraction, for the four feedback variations used in §3.3. The curves are dotted where there are fewer than 100 star particles per galaxy and individual galaxies are showed for bins that contain fewer than 10 galaxies. The 1-sigma scatter about the median of Reference is shown as a shaded region. Dark and light grey lines represent the abundance matching relations of Behroozi et al. (2013) and Moster et al. (2013).

it is possible to correct for this effect applying the simple empirical re-scaling: $S_{true} = S^{\sqrt{3}}$ as suggested in Bailin & Steinmetz (2005). This correction is not implemented in the results presented here since a similar correction is not available for the other quantities that we present.

3.2.4 Axes and misalignment angle definition

The eigenvectors of the inertia tensor, in Eq.3.1, are denoted as $\mathbf{e}_x^i$, with $i = 1, 2, 3$ in the case of a 3D distribution of particles and $x =$ halo, star, gas to indicate total matter[3], stars, or gas, respectively. We relate the ordered eigenvectors $\mathbf{e}_x^1$, $\mathbf{e}_x^2$, and $\mathbf{e}_x^3$ of the inertia tensor to the direction of the major, intermediate and minor axis of the corresponding ellipsoid. We further indicate the radial dependence of the major axis as $e_x^1(r)$, which, unless stated otherwise, has been computed using the volume enclosed by the entire structure as defined by SUBFIND (see §3.2.2). We shall quantify the alignment of different matter components via the scalar product of two major axes, i.e. the misalignment angle θ (Θ in case of projected quantities). Specifically, we will use $\cos\theta$ as the principal quantity of interest and only comment on the actual value of θ when relevant. We stress here that the major axis is a spin-2 quantity, i.e.

---

[3]We do not deal with the specific case of only dark matter because on the scales of interest it almost exactly coincides with the total matter in a halo.
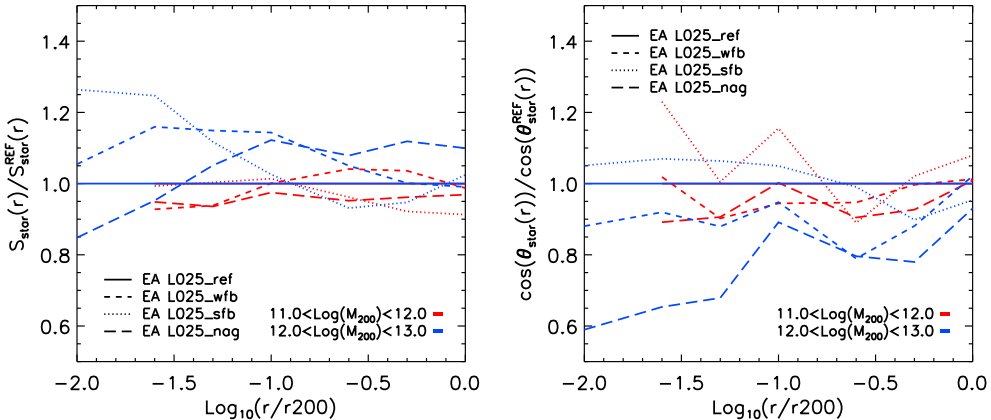
Figure 3.2: Ratios of the average sphericity (left panel) and average misalignment angles (right panel) of the stellar distribution with respect to the average values of the REFERENCE EAGLE simulation (see §3.3). Different colours indicate different mass bins while different line styles refer to different simulations which differ only by the implementation of feedback. The disagreement (up to 20% for sphericity and up to 40% for misalignment angle) stems from the different efficiency of galaxy formation (see discussion in § 3.3). The misalignment angle is more sensitive to galaxy formation efficiencies than the sphericity. The differences always increase towards the centre of the halo.

it is invariant under rotation of 180 degrees. This means that θ only varies between 0 and 90 degrees and, correspondingly, $\cos\theta$ can only assume values between zero and unity.

3.3 The effect of galaxy formation efficiency

A major asset of our composite sample of simulated haloes is that it reproduces the observed stellar-to-halo mass ratio as a function of halo mass. Specifically, EAGLE has been calibrated to reproduce the stellar mass function at redshift zero and cosmo-OWLS has proven successful in reproducing many observable properties of groups and clusters (McCarthy et al. 2010; Le Brun et al. 2014). Moreover, in the halo mass range where cosmo-OWLS haloes are used, their galaxy formation efficiency is consistent with the results of Moster et al. (2013); Behroozi et al. (2013) from abundance matching techniques. This feature is particularly important in the context of our investigation, as one might expect that if a simulation produces either too many or too few stars, then their distribution and consequently, the galaxy shape parameters would also be affected. Note however that, as shown by Crain et al. (2015), this criteria is insufficient to guarantee that the spatial distribution of barionic matter is realistic.

Before showing the main results of our analysis we investigate how different feedback implementations results in different predictions for the shape and orientations of galaxies with respect to their host haloes. To quantify this effect, we make use of a set of feedback variations on the Reference model of the EAGLE simulations that, unlike the Reference model itself, do not reproduce the observed galaxy stellar mass function (i.e. in those simulations haloes do not form stars with the observed efficiency). A detailed description of these simulations can be found in Crain et al. (2015). Here, we only briefly summarize their properties. All simulations adopt the PLANCK cosmology. The simulation boxes have comoving volumes of 25^3Mpc^3, with 2×376^3 particles. We consider four variations:

L025_ref A simulation run in a smaller volume with respect to the main run using the Reference EAGLE implementations, namely: thermal energy feedback from star formation, BH gas accretion that takes into account the gas angular momentum and a star formation law which depends on gas pressure and metallicity. In the thermal feedback implementation the amount of energy injected per feedback event is fixed but there is freedom in the amount of energy that can be injected per unit of stellar mass. This freedom is incorporated in the parameter f_{th} that is the expectation value of the amount of energy injected per unit stellar mass formed, in units of the energy available from core collapse supernova for our IMF. The average number of neighbouring particles heated by a feedback event is $< N_{heat} >\approx 1.3 f_{th} \left(\frac{\Delta T}{10^{7.5} K}\right)^{-1}$ whereas the temperature jump for the single particle is fixed to $\Delta T = 10^{7.5} K$. If the value of f_{th} is constant, then both the energy injected per single event of feedback and the energy per unit of stellar mass are fixed. By varying the parameter f_{th}, it is possible to control the efficiency of the feedback and so to account for the unresolved radiative losses that depend on the physical state of the ISM, or to compensate for numerical losses (see Schaye et al. 2015 and Crain et al. 2015 for a discussion). The value of f_{th} depends on the local physical conditions (density and metallicity) of the gas according to:

$$f_{th} = f_{th,min} + \frac{f_{th,max} - f_{th,min}}{1 + \left(\frac{Z}{0.1 Z_\odot}\right)^{n_Z} \left(\frac{n_{H,birth}}{n_{H,0}}\right)^{-n_n}}, \qquad (3.3)$$

where $n_{H,birth}$ is the density of the parent gas particle at the time it was converted into a stellar particle and Z is the gas metallicity. The value of $n_{H,0} = 0.67$ cm^{-3} was chosen to reproduce the observed present-day GSMF and galaxy sizes, whereas $n_Z = n_n = 2/\ln 10$. We use the asymptotic values $f_{th,max} = 3$ and $f_{th,min} = 0.3$, where the high asymptote $f_{th,max}$ is reached at low metallicity and high density.

L025_wfb Weaker stellar feedback than for the Reference model. In this case the function in Eq. 3.3 is scaled by a factor of 0.5.

L025_sfb Stronger stellar feedback than for the Reference model. In this case the function in Eq. 3.3 is scaled by a factor of 2.

L025_nag Same as Reference but without AGN feedback.

Fig. 3.1 shows the stellar mass to halo mass ratio of central galaxies as a function of halo mass, normalized by the cosmic baryon fraction, for the four aforementioned feedback variations. The galaxy stellar mass function and the galaxy sizes as obtained from these different feedback variations can be seen in Fig. 10, panels a and c, of Crain et al. (2015). Those models produce stellar mass functions with differences of the order of half a dex above (L025_wfb) and below (L025_sfb) the Reference one. The case without AGN feedback differs from the Reference case only for the most massive galaxies. Dark and light grey lines represent the abundance matching relations of Behroozi et al. (2013) and Moster et al. (2013), respectively. The Reference simulation shows good agreement with the abundance matching models.

Fig. 3.2 shows the changes in the main quantities of interest in our analysis for the aforementioned feedback implementations. The left panel displays the ratio of the sphericity of the stellar component of haloes as a function of the distance from the halo centre for each simulation with respect to L025_ref. Different line styles refer to different simulations and we report the results for two halo mass bins. The difference are of the order of 10%. For the triaxiality parameter (not shown here) the differences range from 15% to 50 %. The right panel displays the ratio of the cosine of $\theta(r)$ (the angle between the halo's first eigenvector

and the first eigenvector of the stars inside a given radius) of each simulation with respect to L025_ref. This quantity shows 10% differences at r_{200}^{crit}, while differences as large as 40% for the case without AGN (and 20% in the case of weak SN feedback) are present on scales representative of typical galaxy sizes. We report that the differences between the sphericity of haloes in the different sub-grid implementations (not shown) are smaller than 5% at all radii. This analysis underlines the importance of the calibration of feedback, especially for the shape and alignment of the innermost parts of haloes where most of the stars reside.

A priori, there is no guarantee that reproducing the galaxy stellar mass function is a suffi-cient condition to predict realistic shape parameters. For instance, one may envision a scenario in which the size of galaxies, at the same mass, will also influence their shapes. Crain et al. (2015) have reported four different simulations in which the galaxy stellar mass function is equally well reproduced but the predictions for galaxy sizes are widely different. We com-puted the shape parameters and star-halo misalignment for the same simulations employed in Crain et al. (2015). Although in rough agreement, the relative variance from model to model is $\sim 10\% - 15\%$ for both the sphericity and the misalignment angle (not shown). Clearly, beyond the effect of the 'galaxy formation efficiency', galaxy sizes also play a role in the accuracy of the retrieved shape parameters.

In this section and in the rest of this paper we will not focus on the origin of the different shapes and misalignment of the different populations of haloes. Investigating the physical origin of shapes and misalignments represents an interesting line of inquiry that has been addressed using zoom-in simulations and by following the evolution of galaxies and haloes in time (e.g. Romano-Díaz et al. 2009; Scannapieco et al. 2009; Cen 2014). In this work we focus on exploiting the large dynamical range available to give statistical trends with halo mass and radius and postpone a detailed investigation on their physical origin to future work.

Sim tag (1)	mass bin * (2)	M_{sub} * (3)	M_{star} * (4)	S (5)	T (6)	E_{2D} (7)	θ_{halo}^{star} Deg (8)	Θ_{halo}^{star} Deg (9)	θ_{mass}^{star} Deg (10)	Θ_{mass}^{star} Deg (11)
EA L025	$[11-12]$	11.33	9.50	$0.61_{-0.10}^{+0.17}$	$0.22_{-0.17}^{+0.39}$	$0.82_{-0.07}^{+0.07}$	$47.90_{-24.75}^{+29.60}$	$32.44_{-17.02}^{+18.13}$	$8.21_{-6.19}^{+36.35}$	$4.95_{-3.42}^{+7.67}$
EA L100	$[12-13]$	12.28	10.58	$0.58_{-0.12}^{+0.11}$	$0.31_{-0.23}^{+0.43}$	$0.79_{-0.07}^{+0.06}$	$46.59_{-27.43}^{+29.75}$	$32.34_{-15.88}^{+18.72}$	$3.86_{-2.67}^{+14.37}$	$3.17_{-1.66}^{+4.99}$
CO L200	$[13-14]$	13.25	11.21	$0.65_{-0.08}^{+0.09}$	$0.71_{-0.30}^{+0.16}$	$0.80_{-0.07}^{+0.06}$	$31.04_{-18.69}^{+33.77}$	$24.95_{-14.10}^{+17.82}$	$5.70_{-3.32}^{+8.65}$	$5.62_{-2.96}^{+7.09}$
CO L400	$[14-15]$	14.18	12.06	$0.63_{-0.07}^{+0.08}$	$0.74_{-0.21}^{+0.14}$	$0.77_{-0.07}^{+0.06}$	$24.80_{-14.99}^{+31.20}$	$20.46_{-11.60}^{+18.07}$	$5.61_{-3.08}^{+6.63}$	$5.66_{-2.95}^{+6.20}$

Table 3.3: Values of main quantities of interest for each halo mass bin. Values refer to $z = 0$ and are measured at the half-mass radius in star, r_{half}^{star}, for all subhaloes. Angle θ refers to 3D quantities, whereas Θ refers to the 2D projected equivalent. Description of the columns:(1) simulation tag; (2) mass range of the haloes, $\log_{10}M_{200}$; (3) Median value of the subhalo mass, $\log_{10}M_{sub}$, considering the sum of all the masses of the particles belonging to the subhalo; (4) median value of the stellar mass considering all the star particles belonging to the halo; (5) median value of the sphericity computed at the stellar half-mass radius; (6) median value of the triaxiality computed at the stellar half mass radius; (7) median value of the projected ellipticity (averaged over the three axis projections x, y and z); (8) median angle between the first eigenvector of the stellar component enclosed in r_{half}^{star} and the first eigenvector of the total matter distribution in the halo; (9) same as (8) but for the projected haloes averaged over the three projection axes; (10) median angle between the first eigenvector of the stellar distribution and the total matter distribution, both evaluated at r_{half}^{star}; (11) same as (10) but for the projected haloes averaged over the three projection axes.

* $\log_{10}(M/[\, h^{-1}\, \mathrm{M_{\odot}}])$

3.4 Shape of the different components of haloes

Armed with the simulations described in § 3.2.1 and with the technical definitions introduced in §2.2-2.4, we present here a systematic study of the shape parameters. Specifically, we will present the shape parameters of the entire matter distribution in haloes (in §3.4.1), and of different halo components (stars in §3.4.2 and hot gas in §3.4.3) as well as their mass, spatial, and redshift dependence. In Table 3.3 we summarize the values and the scatter of the shape parameters and the misalignment angles for the stars within $r_{\mathrm{half}}^{\mathrm{star}}$.

It is well known that the reliability of shape estimates of particle distributions depends on the number of particles used to trace those distributions (e.g. Ten14). Motivated by the results presented in Appendix 3.A.2, we measure shapes of structures with at least 300 particles. The resolution criterion is applied separately to the different halo components. Therefore, for a reliable shape measurement of the stellar component we require galaxies containing at least 300 stellar particles. Our tests performed using synthetic NFW haloes show that this choice ensures a precision of 3% and an accuracy better than 10% in the sphericity and triaxiality parameters, see Appendix 3.A.2 for more details. We note that Ten14 performed a similar convergence test according to which using 300 particles leads to $\sim -10\%$ bias in the sphericity of a particle distribution. Our choice ensures relatively high precision while still allowing us to have a large number of haloes for which shape measurements can be performed.

3.4.1 The shape of haloes

Fig. 3.3 displays the sphericity (left panel) and triaxiality (right panel), S and T respectively, for halo masses in the range $9 \leq \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{Mpc}]) \leq 15$. Different colours indicate different simulations and different line styles represent different redshifts (see legend). Notably, despite their difference in resolution, the results agree in the overlapping mass intervals probed via different simulations. The common qualitative result is very simple: haloes become less spherical and more triaxial (prolate) with increasing mass. Sphericity (triaxiality) decreases (increases) from $z = 0$ (solid lines) to $z = 1$ (dotted lines). Haloes thus become more spherical/oblate as cosmic time progresses. This effect is not due to baryon physics since it was also found in dark matter only simulations (e.g. Bryan et al. 2013; Tenneti et al. 2014). For comparison, we also plot the halo sphericity reported by Tenneti et al. (2014) using a dashed line for $z = 0$ and a long dashed line for $z = 1$. Despite the differences in box size, resolution and implementation of baryon physics, the overall agreement with our composite set of simulations is good at both redshifts. The shape of the haloes when all particles are considered is dominated by the dark matter component. In fact, the shape of the dark matter component is nearly identical to the that of the total mass distribution.

Our composite sample suggests that, over a wider range in halo masses, the relation deviates from linear showing a steepening from low to high masses.

3.4.2 Shape of the stellar component of haloes

Fig. 3.4 displays the halo mass dependence of the shape parameters of the stellar distributions. Sphericity is on the left, triaxiality is on the right. As in Fig. 3.3 different colours indicate different simulations according to Table 3.1. We remind the reader that we use a minimum of 300 particles to determine the shape of particle distributions. This inevitably leads to a relatively small halo mass range for each simulation. However, the composite sample of our simulations covers the halo mass range $11 \leq \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$. Note that we have indicated with grey lines the values of the shape parameters obtained when considering haloes comprising fewer than 300 particles. Interestingly, in the overlapping halo mass range, the
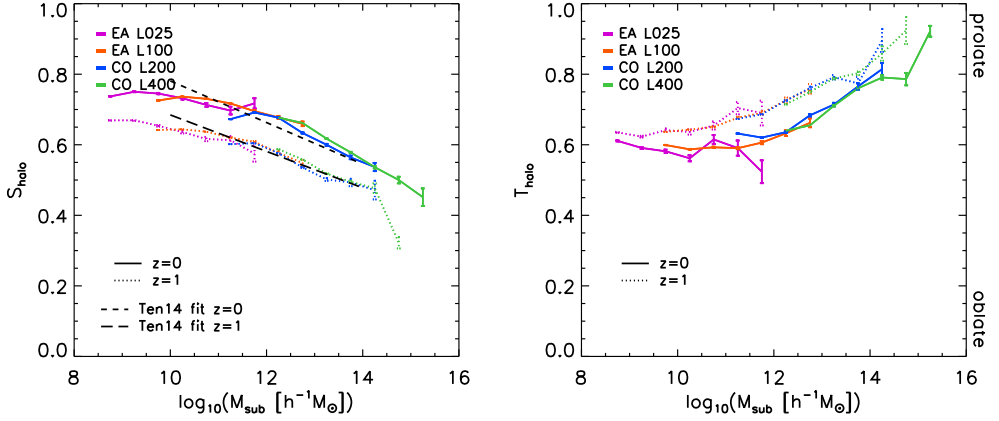
Figure 3.3: Halo sphericity (left) and triaxiality (right) as a function of halo mass. Both central and satellite haloes are considered, hence the choice of M_{sub} (the sum of the masses all the particles belonging to the subhalo) as identifier of the halo mass. Both shape parameters are computed using all the particles in the subhaloes (gas, stars, and dark matter). Using only dark matter would give virtually identical results. Different colours indicate different simulations, whereas solid (dotted) lines refer to $z = 0$ ($z = 1$). The error bars represent the one sigma bootstrap error on the median. Dashed black lines are the values obtained using the fitting functions from Ten14.
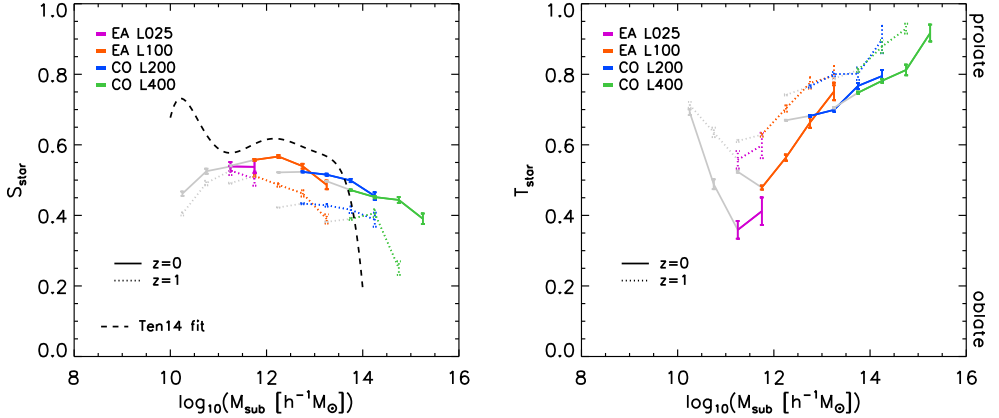
Figure 3.4: Stellar shape parameters (sphericity on the left, triaxiality on the right) as a function of halo mass. Different colours indicate different simulations, whereas different line styles refer to different redshifts. The error bars represent the one sigma bootstrap error on the median. Grey lines are show the results for mass bins containing haloes with less than 300 stellar particles. The dashed black line indicates the sphericity obtained from the fitting function of Ten14. The upturn and the downturn in this fitting function are likely due to selection effects (see discussion in § 3.4.2).
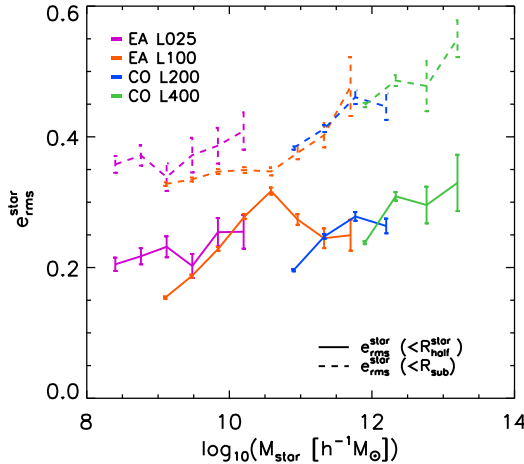
Figure 3.5: Projected rms stellar ellipticity as a function of halo mass. Different colours indicate different simulations (and therefore halo masses), whereas different line styles indicate the region within which the stellar distribution is considered. Specifically, dashed lines indicate the case in which only star particles within the entire halo are considered, whereas solid lines indicate the case in which star particles within the stellar half-mass radius are considered. Both centrals and satellites are considered this analysis.

sphericity parameters derived from simulations with different resolutions agree remarkably well. The general trend seems to suggest that sphericity is a decreasing function of halo mass for $\log_{10}(M_{sub}/[h^{-1} M_\odot]) > 12$ at $z = 0$ and for $\log_{10}(M_{sub}/[h^{-1} M_\odot]) > 11$ at $z = 1$.

We compare our results in Fig. 3.4 with the recent work of Tenneti et al. (2014) by showing their fitting function to the sphericity of the stellar component of haloes (black dashed line). The most prominent feature of their fitting function, namely the sharp upturn at masses $\log_{10}(M_{sub}/[h^{-1} M_\odot]) < 11$, is most likely due to a selection bias. In their work they only compute shapes for subhaloes with more than 1000 stellar particles. This choice imposes a strict limit in stellar mass but not in subhalo mass. This approach only results in an unbiased selection if the minimum stellar mass of all haloes in a given mass bin is higher than $> 1000 m_{star}$ where m_{star} is the mass of a stellar particle. If we impose the same strict limit of 1000 star particles without also limiting the halo masses accordingly, we obtain a similar upturn in the stellar sphericity. Moreover, this upturn occurs at a different mass for different simulations since a fixed number of particles translates into different mass depending on the resolution used.

The triaxiality parameter (right panel of Fig. 3.4) is an increasing function of halo mass at both $z = 0$ and 1. As discussed in Appendix 3.A.2, the accuracy of the triaxiality estimate is more sensitive to the minimum number of particles used to compute it. This manifests itself in the fact that the grey lines in this plot do not continue a monotonic trend beyond the well-resolved mass interval, thus reinforcing the importance of imposing a minimum number of particles used when attempting to recover the triaxiality of a distribution of particles.

The projected stellar rms ellipticity

Under the assumption that galaxies are randomly oriented, averaging the observed projected ellipticity of galaxies gives a measurement of the gravitational lensing effect that, in turn, gives constraints on the matter distribution along the line of sight. The S/N of those measurements depend on the second moment of the distribution of galaxy intrinsic ellipticity, termed $e_{\rm rms}$. Many observational studies have measured the value of the $e_{\rm rms}$ for populations of galaxies. Early results were reported in Hoekstra et al. 2000, and more statistically robust results were obtained using SDSS data (Reyes et al. 2012), COSMOS (Joachimi et al. 2013; Mandelbaum et al. 2014) and the CFHTLenS survey (Heymans et al. 2013; Miller et al. 2013). Unfortunately, despite the tremendous progress in the statistical power of the galaxy surveys employed in these studies, obtaining an accurate estimate of $e_{\rm rms}$ remains challenging, especially because of the fact that the quantity that is accessible observationally always has a (often non-negligible) noise contribution (see e.g. Viola et al. 2014).

For our composite sample of haloes, $e_{\rm rms}$ is defined as:

$$e_{\rm rms}^2 = \frac{1}{N} \sum_i \left(\frac{1 - q_i'^2}{1 + q_i'^2} \right)^2 \tag{3.4}$$

where q_i' is the projected ellipticity of the ith halo $q' = b'/a'$ where a' and b' are the values of the major and minor axis of the projected stellar distribution, and N is the total number of haloes considered.

We use our composite sample of haloes to compute the stellar $e_{\rm rms}$ in bins of halo mass of width 0.5 dex, as a function of halo mass in Fig 3.5. We make use of all star particles that belong to the subhaloes (dashed lines) or only stellar particles within the stellar half-mass radius (solid lines). Both centrals and satellites are considered for this analysis. When all star particles are considered the value of the $e_{\rm rms}$ increases with mass from 0.35 to 0.55. Systematically lower values are found if only stellar particles within the half-mass radius are considered.

The values of the $e_{\rm rms}$ predicted by our composite sample, when all stars are considered, are in broad agreement with the observed noise-corrected values that are of the order of $\approx$ 0.5-0.6 depending on luminosity and galaxy type (e.g. Joachimi et al. 2013). Unfortunately, a direct comparison of our results with those obtained from observational studies is far from trivial. In fact, it would be crucial to mimic all steps in the observational methodology. For instance, $e_{\rm rms}$ measurements are usually only available for a given sub-population of galaxies, those galaxies are further binned in absolute magnitude, and the axis ratio is computed starting from (noisy) images for which flux isophotes need to be identified. In the context of this investigation, we find the current level of agreement satisfactory and ideal as a starting point for future explorations.

Variation of the shape of the stellar component of haloes with the distance from the halo centre

Fig. 3.6 shows the sphericity (left panel) and the triaxiality (right panel) of the stellar component of haloes as a function of the distance from the centre of the halo.

We divide our sample into mass bins that are drawn from different simulations according to Table 3.1. We then compute the inertia tensor for increasingly larger spheres around the centre of each halo. For every sphere we show the median values of the shape parameters of the mass inside the sphere. Radii are given in units of $r_{200}^{\rm crit}$ to allow for a comparison of haloes of different masses. Only particles that are bound to the halo are considered for this analysis. Curves are drawn only on scales where at least 300 particles can be used.
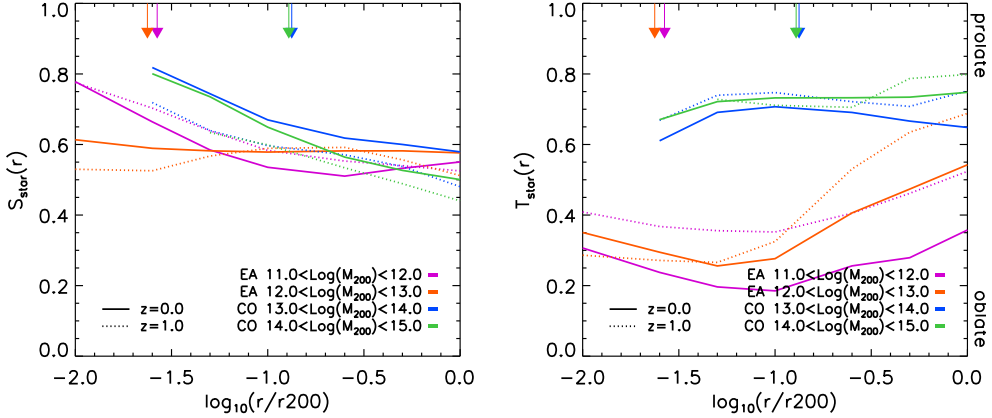
Figure 3.6: Stellar shape parameters (sphericity on the left, triaxiality on the right) as a function of distance from the centre of the halo for haloes in different halo mass bins (see legend). Distances have been rescaled to the mean halo radius in each mass bin, r_{200}^{crit} to ease the comparison of the results for different masses. Different colours indicate different simulations, whereas different line styles refer to different redshifts. The distribution becomes less spherical and more prolate with increasing distance from the halo centre. Vertical arrows indicate the median values of the half-mass radii in stars, $r_{\mathrm{half}}^{\mathrm{star}}$, which can be considered a proxy for the typical extent of a galaxy. The blue arrow lies beneath the green one.

The stellar component of haloes tends to be more spherical near the centre. The triaxiality value shows significant evolution for masses below $M_{\mathrm{sub}} < 10^{12} \, h^{-1} \, \mathrm{M}_\odot$. These trends are qualitatively the same as those found for the dark matter component (not shown) with the exception that the radial profile of the stellar distribution is steeper than that of the dark matter distribution.

The right panel of Fig. 3.6 shows a large difference between the triaxiality values of sub-haloes in the mass bins $12 < \log_{10}(M_{\mathrm{sub}}/[\, h^{-1} \, \mathrm{M}_\odot]) < 13$ (orange curves, EA L100) and $13 < \log_{10}(M_{\mathrm{sub}}/[\, h^{-1} \, \mathrm{M}_\odot]) < 14$ (blue curves, CO L200). This feature might be caused by the different resolution, volume, and/or baryon physics of the two sets of simulations (although the latter is relatively small). To test whether that is the case, we compute the triaxiality parameter of subhaloes with mass $13 < \log_{10}(M_{\mathrm{sub}}/[\, h^{-1} \, \mathrm{M}_\odot]) < 14$ using the EAGLE L100 simulation (not shown). We find the corresponding results to agree with the results obtained using the same mass bin from the cosmo-OWLS L200 simulation. Thus, we interpret the differences between the triaxiality of subhaloes in the mass bins $12 < \log_{10}(M_{\mathrm{sub}}/[\, h^{-1} \, \mathrm{M}_\odot]) < 13$ and $13 < \log_{10}(M_{\mathrm{sub}}/[\, h^{-1} \, \mathrm{M}_\odot]) < 14$ as having a physical origin rather than being due to the resolution, the volume, or the (small) differences in the baryon physics of the two sets of simulations.

3.4.3   Shape of the hot gas component of haloes

In this section we repeat, for the hot gaseous component of haloes, the analysis performed for the total and stellar matter in § 3.4.1 and § 3.4.2, respectively. We present the shape parameters for a subsample of temperature-selected diffuse gas ($T > 10^6 \mathrm{K}$). The selection is quite insensitive on the exact temperature cut, since most of the hot gas in groups and cluster has a temperature that is a factor of two greater than the virial temperature. This temperature
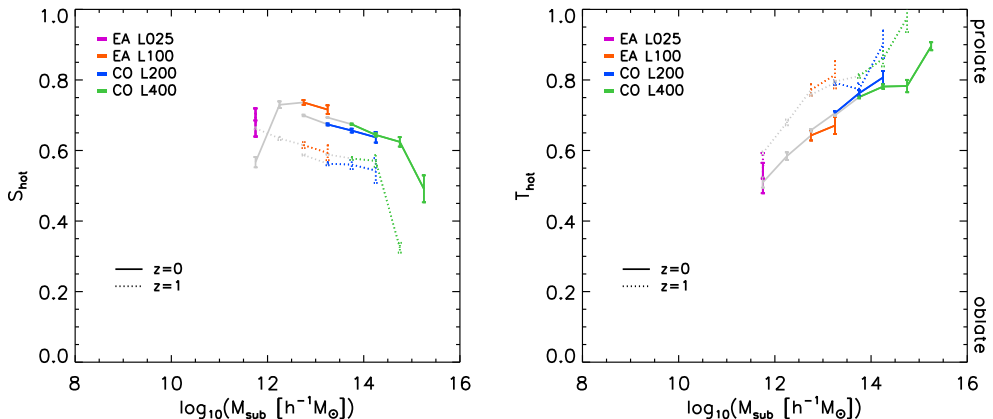
Figure 3.7: Shape parameters (sphericity on the left, triaxiality on the right) of the gas distribution for the hot ($T > 10^6$K) component. Different colours indicate different simulations, whereas different line styles refer to different redshifts. The error bars represent one sigma bootstrap error on the median. Grey lines show the results for mass bins containing haloes with less than 300 hot gas particles.

selection is used as a rough proxy for the hot X-ray emitting gas. A proper selection of X-ray emitting gas is beyond the scope of this paper, as this would require an accurate computation of the X-ray luminosity of the gas particles. A luminosity weighted scheme for the shape of the hot gas would result in the inner regions dominating the shape resulting in more spherical shapes (Crain et al. 2013). On the other hand, a mass weight scheme, as adopted in this work, would be closer to the shape that a Sunyaev-Zeldovich (SZ) experiment would measure since SZ flux is proportional to the gas mass and the temperature, making it potentially testable with a combined SZ-lensing analyses.

Fig. 3.7 presents the variation of the shape parameters, sphericity on the left and triaxiality on the right, of the temperature-selected hot gas particle. The convergence of the sphericity parameter between the different simulations is poorer in this case than for other components shown earlier. By imposing a strict limit on the number of particles needed for measuring the shape we limit our results to only few points for the EAGLE simulations. For instance, is no longer possible to connect the results from L025 and L100. Nonetheless by relaxing the constraint on the number of particles (grey points), it is possible to identify a trend in the shapes that suggests an increasing triaxiality and decreasing sphericity of the hot gas component with host halo mass.

We have also studied the radial dependence of the shape parameters for the hot gas component of haloes (not shown). Given the limit on the minimum number of particles, only three mass bins could be investigated ($M_{200}^{crit} > 10^{12} \, h^{-1} \, M_\odot$) and only down to radius of $r/r_{200}^{crit} = 0.3$, for which no significant radial trend was found.

3.5 Misalignment of Galaxies with their own host haloes

In this section, we show the relation between the orientation of haloes and that of their stellar and hot gas component. Specifically, we will show how the orientation of the major axis of the stellar distribution (§3.5.1) and of the hot gas distribution (§3.5.2) compare to that of the host halo. Similarly to the case of the shape parameters, we will investigate the mass, radial, and
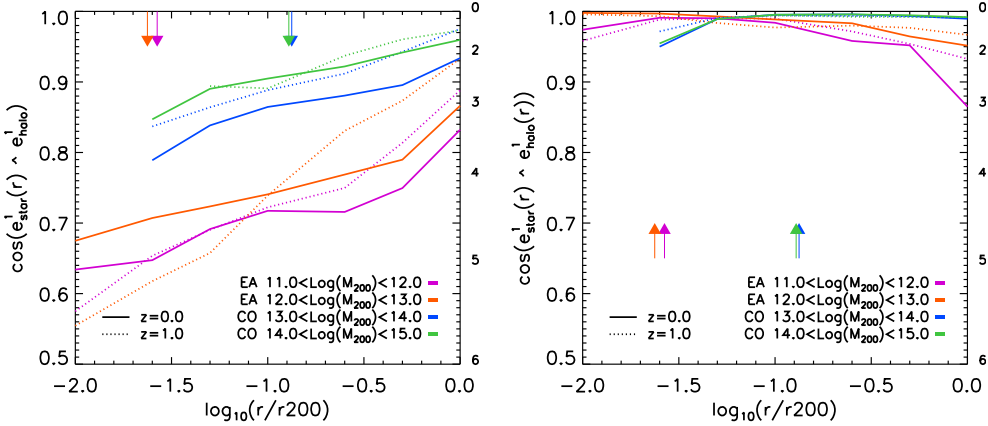
Figure 3.8: Spatial variation of the median cosine of the misalignment angle between the major axis of stars and the underlying (mostly dark) matter distribution. Different colours indicate different halo mass bins, whereas different line styles indicate different redshifts. Radial coordinates are normalized by the mean halo radius, r_{200}^{crit}, of each mass bin to ease the comparison. Only central haloes are used. Vertical arrows represent the median value of r_{half}^{star} in units of r_{crit}^{200} in different mass bins. *Left Panel.* Median value of the cosine of the angle between the major axes of the stellar component and that of the entire halo. Here the direction of the halo is determined using all particles belonging to the halo. *Right Panel.* Median value of the cosine of the angle between the major axes of the stellar component and that of the halo. The misalignment between the stars and halo is caused, to first order, by the misalignment of the inner dark matter halo with the total matter distribution in the halo.

redshift dependence of this relation. Note that we focus our study mainly on central haloes. We remind the reader that a formal definition of the axes of particle distribution and their relative misalignment angles is provided in §3.2.4.

3.5.1 Misalignment of stars with their host haloes

The left panel of Fig. 3.8 shows the median misalignment of stars in spheres of increasingly larger radii with the direction of the total matter distribution within the virial radius for different bins in halo mass and for radii expressed in units of r_{200}^{crit}. Perhaps not surprisingly, the alignment of stars within the total halo increases from the inner to the outer part of the halo. The gradient is relatively steep, with the misalignment angle between the stars and their host haloes decreasing from about 30 degrees (at $r \sim 0.03 r_{200}^{crit}$) to a few degrees (at $r \sim r_{200}^{crit}$) in the case of the most massive haloes. In less massive haloes, the misalignment is larger at all scales. Similar trends hold at $z = 1$ (dotted lines). The right panel of Fig. 3.8 shows the misalignment of stars with the direction of the total (mostly dark) matter, where both are now enclosed in spheres of increasingly larger radii. At each radius, the misalignment is small. Stars are aligned with the total mass to within a few degrees in the most massive haloes, whereas the alignment deteriorates to about 20-30 (10-20) degrees for the least massive haloes at $z = 0$ ($z = 1$).

The misalignment of stars with their host halo can vary substantially depending on the radius and the mass of a halo. The arrows in the plot represent the values, in units of r_{200}^{crit}, of the half mass radius in stars, which is a good indicator of the physical extent of a galaxy. At this radius the orientation of the galaxies is clearly a biased proxy of the orientation of the

halo. Galaxies are, however, much better aligned with the *local* distribution of matter. This indicates that the the stellar orientations follow that of the dark matter, which is the dominant component in mass, and the dark matter itself changes orientation from the inner to the outer halo. This causes the stars to be well aligned with the local mass distribution but misaligned with the orientation of the entire halo.

Probability distribution function of misalignment angles

In the previous section we presented the median value of misalignment between the halo and the stellar component. The upper panel of Fig. 3.9 shows the probability distribution function of the cosine of the misalignment angle between the stars and the entire host halo for central galaxies. Here the stars are taken to be inside $r_{\rm half}^{\rm star}$. Each panel shows a different mass bin and therefore a different simulation. The colour histograms show the misalignment distribution for haloes in that specific mass bin, whereas the black histograms show the probability distribution functions for all haloes that are above the halo mass resolution limit (300 stellar particles inside $r_{\rm half}^{\rm star}$) in the corresponding simulation . The vertical lines show the median values for the distributions and the dashed red curves are analytic fits (see Appendix 3.B). The lower panel of Fig. 3.9 shows the cumulative probability of the cosine of the misalignment angle for early- (dotted curves) and late-type galaxies[4] (dashed curves) as well as for the whole sample of haloes (continuous curves).

The distribution of the cosine of the misalignment angle has a long tail towards low values (i.e. strong misalignment) with a floor value that decreases with increasing halo mass. The misalignment angle distribution of resolved haloes is quite similar in shape for the different simulations. Using the fitting functions provided in Appendix 3.B and the median values of the misalignment angle shown in the previous plots, it is possible to populate dark matter haloes with galaxies oriented such that these misalignment distribution are reproduced.

Bett et al. (2010) quantified the misalignment angle between the stellar and total matter distribution in a sample of (about 90) *disc* galaxies selected from a hydrodynamic simulation in a cubic volume of $35\,h^{-1}$ Mpc by side. They found that half of these galaxies have a misalignment angle larger than 45 degrees. Using the GIMIC simulations (Crain et al. 2009b), Deason et al. (2011) reported that 30% of disc galaxies with average halo mass of $\log_{10}(M_{\rm sub}/[\,{\rm M_\odot}]) = 12.1$ have a misalignment angle of more than 45 degrees. Both these studies are in broad agreement with our findings for similar halo masses. Specifically, in the EAGLE simulations, we find that half of the disc galaxies have misalignment angles larger than 50 (40) degrees in L025 (L100) and 30% of the galaxies in L100 (for which the typical halo mass is close to that in Deason et al. (2011) have misalignment angles larger than 60 degrees.

Fig. 3.10 shows the probability distribution function of the misalignment between the major axes of the *projected* halo and the projected stellar mass component. For comparison, we report with black (mean) and grey (one sigma deviation) dashed lines the results from Okumura et al. (2009) who found that, by assuming a Gaussian misalignment distribution between LRGs and dark matter haloes, they were able to account for the discrepancy between the measured orientation correlation of LRGs and the one predicted by N-body simulations. Furthermore, we overplot analytic fits to our discrete distributions using a double Gaussian (red dashed curves, see Appendix 3.B). Notably, none of our probability distributions resembles a Gaussian function. It is obvious that a single Gaussian function cannot be used as a fair description of the probability functions measured from our simulations.
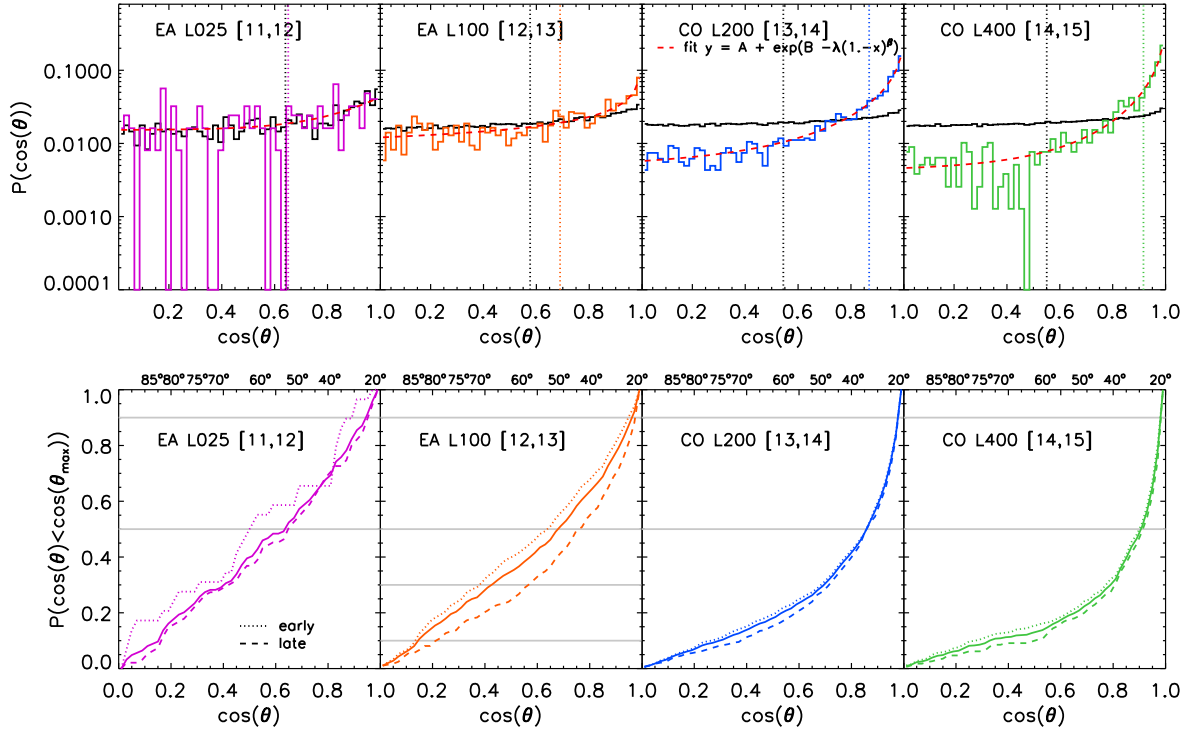
---

[4]See definition of disc galaxies in §3.5.1.

Figure 3.9: Upper panel: probability distribution function of the cosine of the misalignment angle between the major axis of the distribution of stars inside r_{half}^{star}, and the major axis of the entire halo for four halo mass bins. The black histograms indicate the probability distributions for the total sample of haloes that satisfies the resolution criteria, whereas coloured histograms refer only to the subsample of haloes whose mass is indicated in the legend. Vertical lines indicate the median values of the misalignment angle (same colour convention as for the histograms). Red dashed curves represent the analytic fit discussed in Appendix 3.B. Lower panel: cumulative version of the probability function for early- and late-type galaxies (dotted and dashed curves respectively).
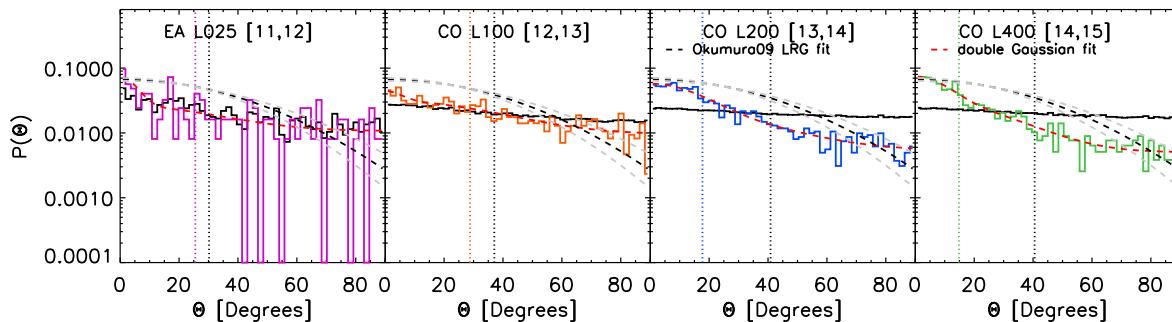
Figure 3.10: Probability distribution function of the 2D misalignment angle between the major axes of the *projected* distribution of stars (inside r_{half}^{star}) and the major axes of the *projected* total matter distribution for four halo mass bins. The black histograms indicate the probability distributions for the total sample of haloes that satisfies the resolution criteria, whereas coloured histograms refer only to the subsample of haloes whose mass is indicated in the legend. Vertical lines indicate the median values of the misalignment angle (same colour convention as for the histograms). Red dashed curves represent the analytic fit discussed in Appendix 3.B, whereas black and grey curves are obtained with analytic functional forms that have been employed in the literature (see text).

Misalignment for early- and late-type galaxies

In this section, we study the alignment between stars and their host haloes in early- and late-type galaxies. Given the galaxy stellar velocity dispersion, σ_{star}, and the halo maximum circular velocity, $V_{\text{circ}}^{\text{max}}$, one can define the ratio $\eta = \sigma_{\text{star}}/V_{\text{circ}}^{\text{max}}$ to quantify whether a galaxy is supported either by ordered (rotational) motion or by the velocity dispersion. We adopt the convention that $\eta \leq 0.5$ indicates a rotationally-supported galaxy (late type), whereas $\eta > 0.5$ indicates a dispersion-supported galaxy (early type).

The left panel of Fig. 3.11 shows the median misalignment of the direction of the entire host halo with that of stars in spheres of increasingly larger radii for early- (dotted lines) and late-type (dashed lines) galaxies. As for the entire galaxy population, the misalignment of stars with their host halo decreases from the inner to the outer part of the halo. The misalignment decreases with mass and is lower for late- than for early-type galaxies. The misalignment of early-type galaxies in low-mass haloes[5] ($11 < \log_{10}(M_{200}/[\,h^{-1}\,\text{M}_\odot]) < 12$ and $12 < \log_{10}(M_{200}/[\,h^{-1}\,\text{M}_\odot] < 13)$) is especially large at all radii and its radial dependence is significantly steeper than in all other cases.

The right panel of Fig. 3.11 shows the ratio, $\cos\theta_{\text{star}}^{\text{late}}/\cos\theta_{\text{star}}^{\text{early}}$, of the cosine of the misalignment angle between the stars of early- and late-type galaxies and the entire halo. At all radii of interest here, early-type galaxies are more misaligned than late-type galaxies. The misalignment angle of late-type galaxies is smaller by about 10-20% at $r \sim 0.03 r_{200}^{\text{crit}}$ approximately the expected physical extent of the galaxy.

A more detailed investigation of the galaxy-halo misalignment as a function of galaxy type is beyond the scope of this paper. We do acknowledge that this is certainly an interesting direction to be further explored, especially in view of the fact that many (current and forthcoming) lensing studies for which the misalignment angle hampers the interpretation of the signal use early-type galaxies such as LRGs. This exploratory work suggests that late-type galaxies are instead less misaligned with their host halo and therefore, in this respect, to be preferred to early-type galaxies.

Misalignment for central and satellites galaxies

In this section, we characterize the alignment of stars with their host halo for centrals and satellites separately. Note that we have only considered centrals in the preceding sections. The left panel of Fig. 3.12 shows the median misalignment of the direction of the entire halo with that of the stars in central and satellite galaxies, whereas the right panel of Fig. 3.12 shows the ratio of the cosine of the misalignment angle between the entire halo and the stars for central and satellite galaxies. As for Fig. 3.11, we employ the EAGLE L100 simulation for the mass bins ($11 < \log_{10}(M_{\text{sub}}/[\,h^{-1}\,\text{M}_\odot]) < 12$ and $12 < \log_{10}(M_{\text{sub}}/[\,h^{-1}\,\text{M}_\odot]) < 13$) to improve the otherwise poor statistics of the EAGLE L025 simulations. Furthermore, we adopt here the dark matter half-mass radius, $r_{\text{half}}^{\text{dm}}$, as a definition of the extent of a halo, as this is properly defined for both centrals and satellites whereas an overdensity with respect to a background/critical value is an ill-defined concept for subhaloes that host satellite galaxies. At all radii, the misalignment angle between the entire halo and the stars in central and satellite galaxies is the same to within 10%. The radial trend is in qualitative agreement with those of the whole sample shown in Fig. 3.8 (i.e. the misalignment decreases from the inner to the outer halo).

The consistently lower misalignment in the outer parts of satellites could be due to the tidal stripping removing the outer (and more misaligned) part of the halo. Instead, in the inner

---

[5]In Fig. 3.11, we use the EAGLE L100 simulation also for the least massive bin (cyan lines) to improve the otherwise poor statistics of the EAGLE L025 simulation.
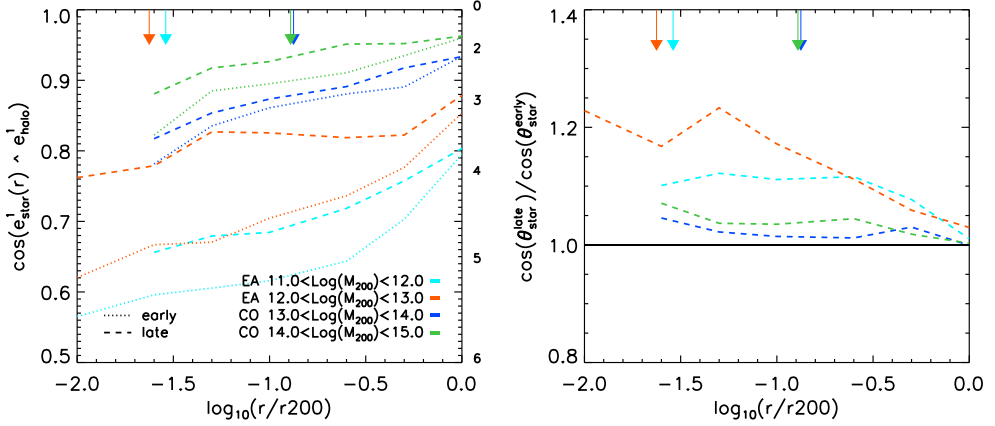
Figure 3.11: *Left Panel.* Spatial variation of the cosine of the median misalignment angle between stars and the underlying (mostly dark) matter distribution. Different colours refer to different halo mass bins, whereas different line styles refer to different galaxy types. Radii are rescaled to the mean halo radius, r_{200}^{crit}, to ease the comparison of the results corresponding to different halo mass bins. Only central haloes are used in order to remove effects that can alter mostly the alignment of satellites. Here the direction of the halo is determined using all particles belonging to the structure. A kinematic classification has been employed (see text) to divide galaxies into early- and late-type. *Right Panel.* Spatial variation of the ratio between the alignment of late and early type galaxies. As in the left panel, the alignment is expressed in terms of the cosine of the angle between galaxy and halo major axes. In both panels the least massive bin is taken from the L100 simulation to improve on the otherwise poor statistics of L025. The vertical arrows represent the median values of r_{half}^{star} in units of r_{crit}^{200} in different mass bins.
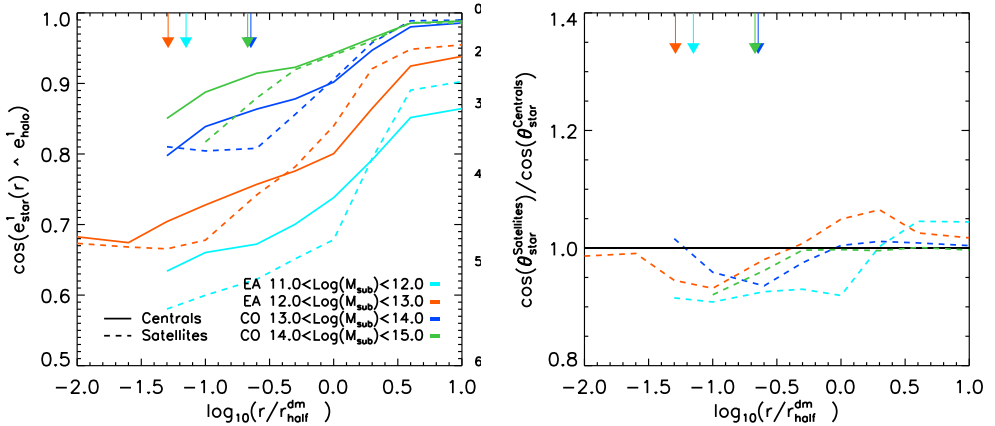
Figure 3.12: Same as Fig. 3.11 but for central and satellite galaxies and subhaloes (see text). To ease the comparison for results of different halo types, radii have been rescaled to the half mass radius for the dark matter mass, r_{half}^{dm}. In both panels the least massive bin is taken from the L100 simulation in order to improve the statistic. The vertical arrows represent the median value of r_{half}^{star} in units of r_{half}^{dm} in the different mass bins.
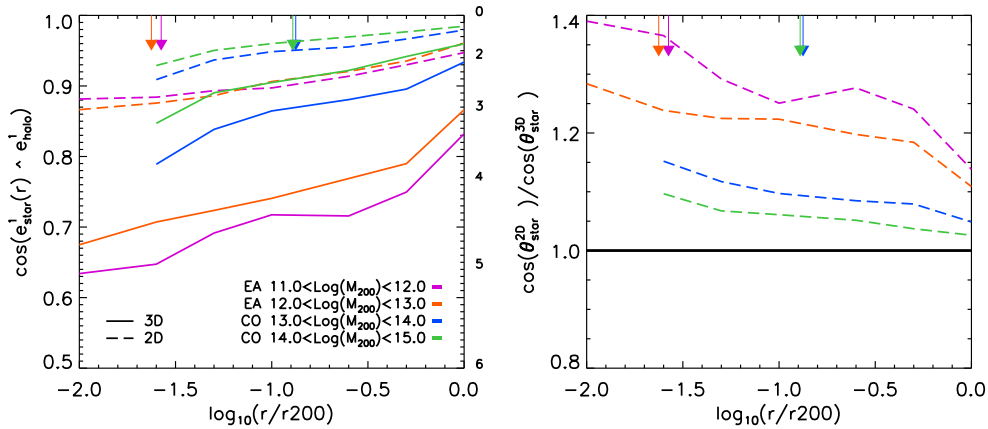
Figure 3.13: Comparison between the spatial variation of the 3D (continuous lines) and 2D (dashed lines) of the cosine of the median misalignment angles of stars with the underlying (mostly dark) matter distribution. Different colours refer to different halo mass bins. Only central haloes are used to exclude effects that can alter mostly the alignment of satellites. As discussed in the main body of the paper, the projection, by reducing the degrees of freedom of the system, increases the alignment.

part the resulting reduction of r_{half}^{dm} (for which the radii are normalized) would produce a shift of the whole relation to the right, effectively increasing the misalignment. The competition between these two processes could explain the transition between a more misaligned inner part to a less misaligned outer part of satellites with respect to central subhaloes.

This result indicates that satellite-specific physical processes (e.g. dynamical friction, tidal stripping) generally do not have a strong impact on the misalignment between the stellar and (mostly dark) matter component.

The effect of projection on the misalignment angle

Observationally one only has access to quantities projected onto the plane of the sky. Therefore, it is of interest to compute the misalignment in a (random) two-dimensional (2D) plane onto which all particles of the simulations have been projected. Correspondingly, one has 2D inertia tensors that describe the matter distribution of each component. In this 2D application, the misalignment angle between the stars and the halo is measured as the angle between the main eigenvectors of the inertia tensor of stars and (mostly dark) matter.

The left panel of Fig. 3.13 shows the radial- and mass-dependence of the median (cosine of the) misalignment angle for the 3D (solid) and the 2D (dashed) case. Clearly, the net effect of projecting the 3D distribution onto a 2D plane is an increase in the alignment at all radii and all halo masses. The right panel of Fig. 3.13 shows the ratio between the cosine of the misalignment angle in 2D and 3D. The ratio decreases with both mass and radius but is always greater than unity. It reaches values of about 1.25-1.35 for the low-mass bins at the radii that are representative of the physical extent of a galaxy. A similar result was reported in Ten14.

3.5.2   Misalignment of hot gas with its host halo

Fig. 3.14 shows the radial and mass dependence of the alignment of the hot component of the gas ($T > 10^6$K) with its host halo. The results are only shown for three mass bins,
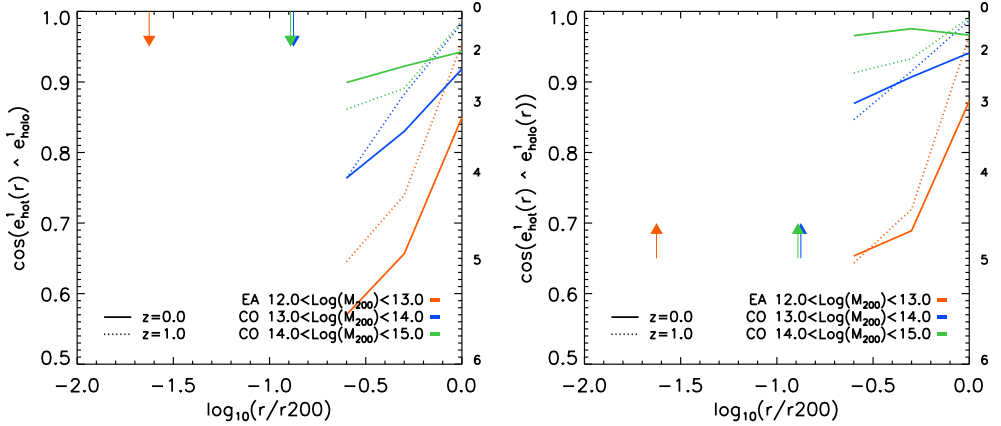
Figure 3.14: Same as Fig. 3.8 but for the hot ($T > 10^6$K) gas component of haloes. The alignment of the hot component increases with radius and mass. Except for the highest-mass haloes, the gas does not follow the dark matter distribution as well as was the case for the stars (c.f. Fig. 3.8).

because the mass bin $11 < \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{Mpc}]) < 12$ does not contain enough hot gas particles to retrieve reliable estimates for the orientation. For the highest halo mass bins (right panel) the spatial variation of the misalignment angle between the hot component and the entire halo (left panel) is similar to that of the stars in the same halo mass bins. On the other hand, the misalignment between the hot gas and the local matter distribution differs from the case of stars: the hot gas component is significantly misaligned with respect to the local matter distribution. Specifically, for haloes in the mass range $12 < \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) <$ 13 the misalignment angle of the hot gas is as large as 50 degrees at $r \sim 0.3\, r_{200}^{\mathrm{crit}}$ and it is ~ 30 (~ 10) degrees for the halo mass range $13 < \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) < 14$ ($14 < \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) < 15$). Results for redshift $z = 1$ (dotted lines) have similar radial and mass dependence as for redshift $z = 0$.

Because it is observable out to larger radii than the stellar distribution of the central galaxy, hot gas represents a valuable tracer of the gravitational potential of massive clusters. Unfortunately, the fact that the hot gas tends to be largely misaligned with the local matter distribution makes it a poor tracer of the shape of the halo, unless $\log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M_\odot}]) > 14$.

3.6 Summary and Conclusions

This paper reports the results of a systematic study of halo and galaxy shapes and their relative alignment in the EAGLE (Schaye et al. 2015; Crain et al. 2015) and cosmo-OWLS (Le Brun et al. 2014; McCarthy et al. 2014) hydro-cosmological simulations. Several aspects of these simulations make them an ideal tool for this investigation. First, the combination of these simulations allows us to apply our study to four orders of magnitude in halo masses with sufficient resolution and statistics. Second, the EAGLE simulations have been calibrated to be in agreement with the observed present-day galaxy stellar mass function and the observed size-mass relation (Schaye et al. 2015). Third, it has been shown that cosmo-OWLS simulations reproduce key (X-ray and optical) observed properties of galaxy groups as well as the observed galaxy mass function for haloes more massive than $\log(M/[\,h^{-1}\,\mathrm{M_\odot}]) = 13$.

We have studied the shapes of the distributions of dark matter, stars and hot gas in haloes

with masses $11 < \log_{10}(M_{200}^{\mathrm{crit}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 15$ and their evolution in the redshift range $0 \leq z \leq 1$. We find that the matter distribution in haloes is more aspherical (and triaxial) at higher halo mass and higher redshift (see Fig. 3.3). The same qualitative trends hold for the star and the hot gas distribution in haloes (see Figs. 3.4 and 3.7). We report (in Fig. 3.6) the spatial variation of the median of the shape parameters of the stellar distribution from $\sim 0.02 r_{200}$ (i.e. a few to tens of kpc) to r_{200} (i.e. up to a few Mpc). We note that at fixed radius and halo mass, stellar distributions are generally less spherical than dark matter haloes. We have measured the r.m.s. of the projected stellar ellipticity as a function of halo mass. We find a modest mass dependence, with r.m.s. stellar ellipticity increasing by 50 % as halo mass increases by four orders of magnitude. We note that the values of the r.m.s. stellar ellipticity vary from ~ 0.2 to ~ 0.35 when one considers only stars within the star half-mass radius. However, the same quantity varies from ~ 0.35 to ~ 0.55 when all stars within the halo are considered (see Fig. 3.5).

Tenneti et al. (2014) recently used the Massive Black II simulation to study the mass dependence and evolution of the stellar and dark matter components of haloes and subhaloes. Their findings are, for the most part, in qualitative agreement with ours. However we find a few differences as reported in the corresponding sections (see e.g. § 3.4 and the discussion of Fig. 3.4). Specifically, we highlighted sources of potential biases in their analysis. As detailed in § 3.3, those biases mostly stem from the use of a hydro-simulation that does not reproduce the observed stellar-halo mass relation and by imposing an artificial cut-off in the minimum stellar mass for which the shape is calculated.

We have measured the misalignment of the baryonic components (stars and hot gas) of galaxies with their own host haloes. We find that stars align well with the underlying (mostly dark) matter distribution, especially when all stars inside the halo are considered (see Fig. 3.8). However, the stellar distributions in the inner parts of the host haloes do exhibit a median misalignment of about 45-50 degrees. The misalignment is smaller in more massive haloes ($13 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$), late-type galaxies (see Fig. 3.11), and central galaxies (see Fig. 3.12). The hot gas distribution can only be traced with a sufficient number of particles only in the outer part ($\geq 0.3 r_{200}$) of massive ($12 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$) haloes. In this range we find that the alignment of the hot gas with the *entire* halo is similar to that of the stellar distribution. However, the hot gas does not align well with the *local* matter distribution, exhibiting misalignment angles larger than 20 (typically 30 to 50) degrees in haloes with masses $13 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$ (see Fig. 3.14).

We have quantified the effect of projection on the median misalignment angles between the stellar distribution and the halo (see Fig. 3.13). Projection reduces the degrees of freedom of the system, increasing the alignment. Finally, we provided the probability distribution of the misalignment angle between the major axis of the stellar distribution inside the stellar half-mass radius and the major axis of the entire halo for the three- and two-dimensional case (see Figs. 3.9 and 3.10, respectively).

We have encapsulated our results in fitting functions (see Appendix B) and tables that allow interested practitioners to straightforwardly include our results into halo catalogues extracted from N-body simulations. The complete list of fitting parameters as well as tabulated values are available at http://www.strw.leidenuniv.nl/MV15a/.

A natural extension of this work is the study of the correlation functions of galaxy shapes. We will present such an investigation in a future publication.

Acknowledgements

Bibliography

Bailin, J., & Steinmetz, M. 2005, ApJ, 627, 647

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57

Bett, P. 2012, MNRAS, 420, 3303

Bett, P., Eke, V., Frenk, C. S., Jenkins, A., & Okamoto, T. 2010, MNRAS, 404, 1137

Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53

Boylan-Kolchin, M., Ma, C.-P., & Quataert, E. 2006, MNRAS, 369, 1081

Bryan, S. E., Kay, S. T., Duffy, A. R., et al. 2013, MNRAS, 429, 3316

Cen, R. 2014, ApJ, 785, L15

Chen, D. N., Jing, Y. P., & Yoshikaw, K. 2003, ApJ, 597, 35

Cole, S., & Lacey, C. 1996, MNRAS, 281, 716

Crain, R. A., McCarthy, I. G., Schaye, J., Frenk, C. S., & Theuns, T. 2010, ArXiv e-prints (arXiv:1011.1906), arXiv:1011.1906

Crain, R. A., McCarthy, I. G., Schaye, J., Theuns, T., & Frenk, C. S. 2013, MNRAS, 432, 3005

Crain, R. A., Theuns, T., Dalla Vecchia, C., et al. 2009a, MNRAS, 399, 1773

—. 2009b, MNRAS, 399, 1773

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

Croft, R. A. C., Di Matteo, T., Springel, V., & Hernquist, L. 2009, MNRAS, 400, 43

Dalla Vecchia, C., & Schaye, J. 2008, MNRAS, 387, 1431

—. 2012, MNRAS, 426, 140

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

Deason, A. J., McCarthy, I. G., Font, A. S., et al. 2011, MNRAS, 415, 2607

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Dubinski, J. 1998, ApJ, 502, 141

Hahn, O., Teyssier, R., & Carollo, C. M. 2010, MNRAS, 405, 274

Heymans, C., Brown, M., Heavens, A., et al. 2004, MNRAS, 347, 895

Heymans, C., Grocutt, E., Heavens, A., et al. 2013, MNRAS, 432, 2433

Hoekstra, H., Franx, M., & Kuijken, K. 2000, ApJ, 532, 88

Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538

Joachimi, B., Semboloni, E., Bett, P. E., et al. 2013, MNRAS, 431, 477

Kaiser, N. 1984, ApJ, 284, L9

Kang, X., van den Bosch, F. C., Yang, X., et al. 2007, MNRAS, 378, 1531

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, MNRAS, 441, 1270

Mandelbaum, R., Rowe, B., Bosch, J., et al. 2014, ApJS, 212, 5

McCarthy, I. G., Le Brun, A. M. C., Schaye, J., & Holder, G. P. 2014, MNRAS, 440, 3645

McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, MNRAS, 406, 822

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121

Naab, T., Khochfar, S., & Burkert, A. 2006, ApJ, 636, L81

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Okumura, T., Jing, Y. P., & Li, C. 2009, ApJ, 694, 214

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16

Reyes, R., Mandelbaum, R., Gunn, J. E., et al. 2012, MNRAS, 425, 2610

Romano-Díaz, E., Shlosman, I., Heller, C., & Hoffman, Y. 2009, ApJ, 702, 1250

Rosas-Guevara, Y. M., Bower, R. G., Schaye, J., et al. 2013, ArXiv e-prints (arXiv:1312.0598), arXiv:1312.0598

Sales, L. V., Navarro, J. F., Theuns, T., et al. 2012, MNRAS, 423, 1544

Scannapieco, C., White, S. D. M., Springel, V., & Tissera, P. B. 2009, MNRAS, 396, 696

Schaye, J., & Dalla Vecchia, C. 2008, MNRAS, 383, 1210

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Sharma, S., & Steinmetz, M. 2005, ApJ, 628, 21

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Tenneti, A., Mandelbaum, R., Di Matteo, T., Feng, Y., & Khandai, N. 2014, MNRAS, 441, 470

Tenneti, A., Singh, S., Mandelbaum, R., et al. 2015, MNRAS, 448, 3522

van den Bosch, F. C., Abel, T., Croft, R. A. C., Hernquist, L., & White, S. D. M. 2002, ApJ, 576, 21

Viola, M., Kitching, T. D., & Joachimi, B. 2014, MNRAS, 439, 1909

Wang, Y., Yang, X., Mo, H. J., et al. 2008, MNRAS, 385, 1511

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS, 399, 574

Zemp, M., Gnedin, O. Y., Gnedin, N. Y., & Kravtsov, A. V. 2011, ApJS, 197, 30

3.A Caveats in shape parameter estimation

3.A.1 The choice of inertia tensor

There exists a plethora of methods designed to characterize the shape of a given three-dimensional particle distribution (say dark matter, star, gas) in the context of cosmological structure formation simulations (see Zemp et al. 2011 and references therein). All those methods are based on the idea that structures can be well described by an ellipsoidal shape. However, the actual algorithms used to retrieve this shape can differ substantially and unfortunately the corresponding results do not often agree (see Zemp et al. 2011 for an analysis of this problem under controlled conditions with known shapes). Most notably, results on the shape of a particle distribution may vary if one adopts the inertia tensor rather than the reduced inertia tensor, or some iterative form of the two (see discussions in Zemp et al. 2011; Tenneti et al. 2015). The differences between the inertia tensor and the reduced inertia tensor are driven by the fact that in the reduced inertia tensor calculation particles are not weighted by their distance from the centre. The net effect is that if the reduced inertia tensor is used, the shape is less dominated by the particles in the outer part of haloes, meaning that the retrieved shape tends to be more spherical as particles in the inner parts of haloes are typically more spherically distributed. We repeated our analysis using the reduced inertia tensor and found that there is

little information content in exploring the radial variation using this method since its properties have almost no variation with radius. We also note that the misalignment angle, which is the quantity of primary interest here, is less affected than shape by the choice of the algorithm that defines the shape parameters. This is especially true when the alignment is calculated between particles distributions at the same distance from the centre.

Another possible variation in the shape calculation is to use an iterative method for the inertia tensor calculation but this method was proven to give very similar results when the inertia tensor is used, as shown in Tenneti et al. (2015).

Throughout the paper we adopted the definition presented in eq. (3.1) and the corresponding shape parameters defined in eq. (3.2). While our adopted method may be considered somewhat arbitrary (e.g. see the discussions in Jing & Suto 2002; Zemp et al. 2011), we used this approach as it is adequate for the comparison we presented (see e.g Bett 2012) and because it allowed us to compare our results with most of the other results in the literature.

3.A.2   The effect of sampling

An important technical aspect regarding shape measurements is to find the minimum number of particles required to obtain a reliable estimate. To this aim, we simulate a three-dimensional halo with a given axis ratio and a Navarro, Frenk, & White (hereafter NFW) density profile (Navarro et al. 1997). More specifically, we choose values for the three-dimensional halo axis a, b and c, and use an analyitical NFW profile with $c = 5$ and $r_{\mathrm{vir}} = a$ (the largest axis). We then generate N_{part} spherical coordinates (r, ϕ, θ) using the NFW profile as a selection function, redrawing any coordinates that fall outside the ellipsoid defined by a, b, and c. Specifically, we use $1 \leq N_{\mathrm{part}} \leq 3000$. For each value of N_{part}, we repeat the sampling 10^5 times so as to obtain a median and a standard deviation.

It is worth noting that the number of particles needed for an unbiased shape determination depend on the intrinsic shape of the halo. Many more particles are needed to retrieve a quasi-spherical shape than for example a disky structure. For our test, the intrinsic shape of the halo was chosen to have sphericity $S = 0.6$ and triaxiality $T = 0.7$, which is representative of the average shape parameters of our halo sample (see e.g. results in §3.4.2 and Fig. 3.6).

In Fig. 3.15 we show the relative error on the retrieved shape parameters, S (green lines) and T (red lines), as a function of N_{part}. Solid lines refer to the median, whereas dashed lines refer to the 16th and 84th percentiles. The retrieved sphericity shows a monotonic trend with the number of test particles. The sphericity increases towards the real value as the number of test particles is increased. This means that any resolution effect will lead to an underestimating of the true sphericity of haloes. For this particular halo shape using 300 particles will lead to an average $\sim 2\%$ error in the determination of the sphericity with an accuracy of $\sim 10\%$. The triaxiality is typically underestimated but converges faster to the true value, with the systematic error dropping below 3% for 30 particles. On the other hand the scatter around the median converging slowly and is still 20% for 300 particles. Triaxiality thus requires more particles than sphericity in order to reduce the random error below a specific value. Throughout the paper, we thus employ $N_{\mathrm{parts}} \geq 300$ as the limit for shape parameter determination. This assures very good estimate of the median value of the shape parameters with a systematic error below 3% and a random error of 10% in the sphericity and 20% in triaxiality. In this work we did not show these systematic errors in the shape measurement but only the statistical errors evaluated using the bootstrapping technique.
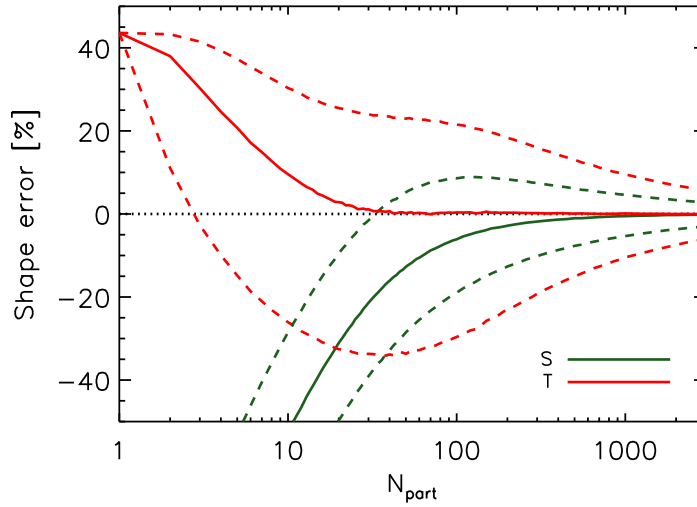
Figure 3.15: Convergence test for shape parameter retrieval. Relative error on the retrieved shape parameter of a synthetic NFW halo as a function of the number of particles used to sample the underlying distribution. The relative errors on sphericity and triaxiality are indicated by the green and red lines, respectively. The test is performed using a typical sphericity value for the synthetic halo, $S = 0.6$ and $T = 0.7$. For each number of particles, distributions are drawn 10^5 times and we report the 50th (continous lines), 16th and 84th percentiles (dashed lines). Retrieving both shape parameters with a systematic error smaller than a few percent requires at least 300 particles.

Simulation	mass bin	A	B	λ	β
EA L025	$[11-12]$	$1.52E-02$	-3.58	5.92	1.01
EA L100	$[12-13]$	$6.43E-03$	-0.05	5.13	0.15
CO L200	$[13-14]$	$4.46E-03$	-1.04	5.64	0.41
CO L400	$[14-15]$	$4.13E-03$	-0.53	7.13	0.42

Table 3.4: Fit parameters for Eq. 3.5 that describes the misalignment angle distribution between the direction of the stellar component inside $r_{\text{half}}^{\text{star}}$ and that of the entire halo.

3.B   Analytic fits for the misalignment angle distributions

In this section we provide fitting functions[6] for the distribution of the cosine of the 3D misalignment angle θ, as well as for the 2D misalignment angle, Θ. We note that the choice of using the cosine as the variable of the fitting function stems from the notion that the distribution of the cosine of the alignment angle of a random set of 3D vectors is flat, whereas the distribution of the angle itself is not, as it is skewed towards large alignments.

We employ the following functional form:

$$\mathcal{M}^{\text{3D}}(x) = A + \exp\left[B - \lambda(1-x)^{\beta}\right], \qquad (3.5)$$

where $x = \cos(\theta)$ and $0 < \theta < \pi/2$. This functional form has four free parameters: A, B, λ, β. We find this number of parameters necessary to adequately reproduce the main features of the results obtained from the simulations. In the main body of the paper (see Fig. 3.9), we have employed this fitting function to describe the misalignment angle between the stellar component and its host halo in four halo mass bins and for the typical extent of a galaxy, the half mass radius $r_{\text{half}}^{\text{star}}$. The corresponding fitting parameters are given in Table 3.4. Parameters that refer to other components, radius definitions, and halo mass bins, as well as tabulated median values, can be found at http://www.strw.leidenuniv.nl/MV15a/.

We analytically describe the probability function of the cosine of the 2D misalignment angle with the following functional form:

$$\mathcal{M}^{\text{2D}}(x) = C\exp\left(-\frac{x^2}{2\sigma_1^2}\right) + D\exp\left(-\frac{x^2}{2\sigma_2^2}\right) + E, \qquad (3.6)$$

where $C, \sigma_1, D, \sigma_2, E$ are the 5 free parameters required to describe a double Gaussian plus a 'floor'. The level of complexity of this functional form is motivated by the results obtained from the simulations. In the main body of the text (see especially Fig.3.10), we describe the probability distribution of the 2D misalignment angle between stars and their host haloes in four halo mass bins and accounting only for stars within the typical extent of a galaxy, the half mass radius $r_{\text{half}}^{\text{star}}$. The corresponding fitting parameters are given in Table 3.5. Parameters that refer to other components, radius definitions, and halo mass bins, as well as tabulated median values, can be found at http://www.strw.leidenuniv.nl/MV15a/. It is instructive to compare these 2D misalignment angle distributions to the commonly assumed single-Gaussian distribution (see e.g. Okumura et al. 2009). None of the distributions found in this study resembles a single-Gaussian and we therefore caution interested practitioners against adopting this assumption.

---

[6]The analytic fits provided in this section reproduce the median of the distributions obtained from the simulations with an accuracy better than 1%.

mass bin	σ_1	σ_2	C	D	E
$[11-12]$	5.00	28.17	$4.69E-02$	$4.69E-02$	$4.69E-02$
$[12-13]$	5.00	31.65	$1.31E-02$	$1.31E-02$	$1.31E-02$
$[13-14]$	14.70	32.52	$3.61E-02$	$3.61E-02$	$3.61E-02$
$[14-15]$	9.42	25.71	$4.28E-02$	$4.28E-02$	$4.28E-02$

Table 3.5: Fit parameters for the double Gaussian fitting function Eq. 3.6 that describes the misalignment angle distribution between the direction of the projected stellar component inside r_{half}^{star} and that of the entire (projected) halo.

3.C   Resolution test

In this section we make use of our different simulations to test the influence of resolution on our results. For this test we make use of the fact that the L025 simulation is the high-resolution version of L100 simulated using a smaller box size. The same is true for L200 and L400. We do not compare results from simulations that were not run with the same code.

In Fig. 3.16 we show in the upper panels the variation of the sphericity of the stellar component. In the left panel we show the two mass bins for which is it possible to obtain results for both the L025 and L100 EAGLE simulations. On the right we do the same for the L200 and L400 cosmo-OWLS simulations. Different colours refer to different mass bins where as different line styles refer to different simulations. In the lower panels we show in the same manner the misalignment between the stellar component and the whole halo.

The convergence is generally good, especially at larger radii, even though the box size, and hence the halo samples, also change between the different simulations. The only case that shows a relatively poor convergence is the misalignment for the least massive bin of the cosmo-OWLS simulations (blue lines) for which the shape of the curves are similar but the values are shifted between the two simulations.
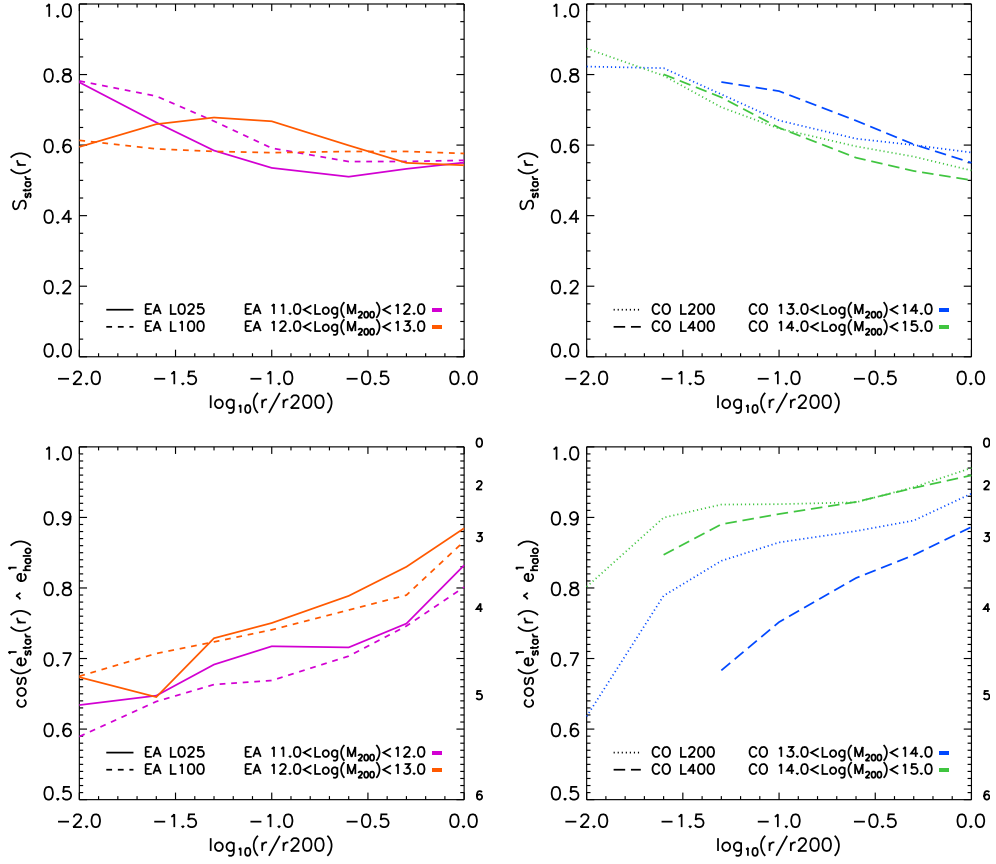
Figure 3.16: Resolution test for the variation of the sphericity (upper panels) and misalignment with the halo (lower panels) of the stellar component as a function of radius. We show separately the results for EAGLE (on the left) and cosmo-OWLS (on the right). For each set of simulations we show the results in two distinct mass bins.

4

Intrinsic alignments of galaxies in the EAGLE and cosmo-OWLS simulations

---

We report results for the alignments of galaxies in the EAGLE and cosmo-OWLS hydro-cosmological simulations as a function of galaxy separation ($-1 \leq \log_{10}(r/[\,h^{-1}\,\mathrm{Mpc}]) \leq 2$) and halo mass ($10.7 \leq \log_{10}(M_{200}/[\,h^{-1}\,\mathrm{M}_\odot]) \leq 15$). We focus on two classes of alignments: the orientations of galaxies with respect to either the *directions* to, or the *orientations* of, surrounding galaxies. We find that the strength of the alignment is a strongly decreasing function of the distance between galaxies. For galaxies hosted by the most massive haloes in our simulations the alignment can remain significant up to ~ 100 Mpc. Galaxies hosted by more massive haloes show stronger alignment. At a fixed halo mass, more aspherical or prolate galaxies exhibit stronger alignments. The spatial distribution of satellites is anisotropic and significantly aligned with the major axis of the main host halo. The major axes of satellite galaxies, when all stars are considered, are preferentially aligned towards the centre of the main host halo. The predicted projected direction-orientation alignment, $\epsilon_{\mathrm{g+}}(r_{\mathrm{p}})$, is in broad agreement with recent observations. We find that the orientation-orientation alignment is weaker than the orientation-direction alignment on all scales. Overall, the strength of galaxy alignments depends strongly on the subset of stars that are used to measure the orientations of galaxies and it is always weaker than the alignment of dark matter haloes. Thus, alignment models that use halo orientation as a direct proxy for galaxy orientation overestimate the impact of intrinsic galaxy alignments.

4.1   Introduction

Tidal gravitational fields generated by the formation and evolution of large-scale structures tend to align galaxies due to correlations of tidal torques in random gaussian fields (e.g. Heavens & Peacock 1988). Analytic theories have been developed to describe these large-scale alignments (linear alignment theory; Catelan et al. 2001), but these are only applicable to low matter density contrasts (the linear regime of structure formation) and do not account for drastic events such as mergers of structures, which may erase initial correlations.

To overcome these limitations, galaxy alignments have been studied via N-body simulations (see e.g. West et al. 1991; Tormen 1997; Croft & Metzler 2000; Heavens et al. 2000; Jing 2002; Lee et al. 2008; Bett 2012). The most common ansatz in such studies is that galaxies are perfectly aligned with their dark matter haloes and that one can therefore translate the alignments of haloes directly into those of the galaxies that they host. However, the observed light from galaxies is emitted by the baryonic component of haloes and hydro-dynamical simulations of galaxy formation have revealed a misalignment between the baryonic and dark matter components of haloes (Deason et al. 2011; Tenneti et al. 2014; Velliscig et al. 2015). On spatial scales characteristic of a galaxy, baryon processes (radiative cooling, supernova explosions and AGN feedback) play an important role in shaping the spatial distribution of the stars that constitute a galaxy. Specifically, the ratio between cooling and heating determines the way baryons lose angular momentum and consequently the way they settle inside their dark matter haloes. Furthermore, feedback from star formation and AGN can heat and displace large quantities of gas and inhibit star formation (Springel et al. 2005; Di Matteo et al. 2005, 2008; Booth & Schaye 2009; McCarthy et al. 2010). These processes, which determine when and where stars form, may influence the observed morphology of galaxies and in turn their observed orientations. In hydrodynamical simulations of galaxy formation in a cosmological volume, such processes are modelled simultaneously, leading to a potentially more realistic realization of galaxy alignments. The study of such models can unveil patterns that encode important information concerning both the initial conditions that gave rise to the large-scale structure, and the evolution of highly non-linear structures like groups and clusters of galaxies.

Beyond their relevance to galaxy formation theory, galaxy alignments are a potential contaminant of weak gravitational lensing measurements. Although this contamination is relatively mild, it is a significant concern for large-area cosmic shear surveys (Joachimi et al. 2015; Kirk et al. 2015; Kiessling et al. 2015, and references therein). Requirements for the precision and accuracy of such surveys are very challenging, as their main goal is to constrain the dark energy equation-of-state parameters at the sub-percent level. Weak lensing surveys are used to measure the effect of the bending of light paths of photons emitted from distant galaxies due to intervening matter density contrasts along the line of sight. The distortion and magnification of galaxy images is so weak that it can only be characterized by correlating the shapes and orientations of large numbers of background galaxies. In a pure weak gravitational lensing setting, the observed ellipticity of a galaxy, ϵ, is the sum of the intrinsic shape of the galaxy, ϵ^s, and the shear distortion that the light of the galaxy experiences due to gravitational lensing, γ,

$$\epsilon = \epsilon^s + \gamma. \tag{4.1}$$

If galaxies are randomly oriented, the average ellipticity of a sample of galaxies, $\langle \epsilon^s \rangle$, vanishes. Therefore, any detection of a nonzero $\langle \epsilon \rangle$ is interpreted as a measurement of gravitational shear γ. However, in the limit of a very weak lensing signal, the distortion induced via gravitational forces (giving rise to an *intrinsic* alignment) can be a non-negligible fraction of the distortion due to the pure gravitational lensing effect (often termed *apparent* alignment, see Crittenden et al. 2001 and Crittenden et al. 2002 for a statistical description of this effect).

Cosmic shear measurements are obtained in the form of projected 2-point correlation functions (or their equivalent angular power spectra) between shapes of galaxies. Following Eq. 4.1:

$$\langle \epsilon\epsilon \rangle = \langle \gamma\gamma \rangle + \langle \gamma\epsilon^s \rangle + \langle \epsilon^s\gamma \rangle + \langle \epsilon^s\epsilon^s \rangle \,, \tag{4.2}$$

$$= GG + GI + IG + II \,. \tag{4.3}$$

If we assume that galaxies are not intrinsically oriented towards one another, then the only correlations in the shape and orientation of observed galaxies is due to the gravitational lensing effect of the intervening mass distribution between the sources and the observer, $\langle \gamma\gamma \rangle$. In this case the only nonzero term is the GG (shear-shear) auto correlation. In the case of a non negligible intrinsic alignment of galaxies, the II term is also nonzero, i.e. part of the correlation between the shape and orientation of galaxies is *intrinsic*. If the same gravitational forces that shear the light emitted from a galaxy also tidally influence the intrinsic shape of other galaxies, then this will produce a nonzero cross correlation between shear and intrinsic shape (GI). The term IG is zero since a foreground galaxy cannot be lensed by the same structure that is tidally influencing a background galaxy, unless their respective position along the line of sight is confused due to large errors in the redshift measurements.

In this paper we report results for the *intrinsic* alignment of galaxies in hydro-cosmological simulations. Specifically, we focus on the orientation-direction and orientation-orientation galaxy alignments. To this aim, we define as galaxy orientation the major eigenvector of the inertia tensor of the distribution of stars in the subhalo. We then compute the mean values of the angle between the galaxy orientation and the separation vector of other galaxies, as a function of their distance. In the case of orientation-orientation alignment we compute the mean value of the angle between the major axes of the galaxy pairs, as a function of their distance. While the orientation-orientation alignment can be interpreted straightforwardly as the II term in Eq. (4.2), the orientation-direction is related to the GI term in a less direct way (see Joachimi et al. 2011, for a derivation of the GI power spectrum from the ellipticity correlation function).

In this paper we make use of four complementary simulations to explore the dependence of the orientation-direction alignment over four orders of magnitude in subhalo mass, and spanning physical separations of hundreds of Mpc. The use of four simulations of different cosmological volumes offer both resolution and statistics, whilst also incorporating baryon physics. The EAGLE simulations used in this work have been calibrated to reproduce the observed present-day galaxy stellar mass function and the observed size-mass relation of disc galaxies (Schaye et al. 2015), whereas the cosmo-OWLS (Le Brun et al. 2014; McCarthy et al. 2014) simulations reproduce key (X-ray and optical) observed properties of galaxy groups and clusters, in addition to the observed galaxy mass function for haloes more massive than $\log(M/[\,h^{-1}\,\mathrm{M}_\odot]) = 13$. In Velliscig et al. (2015) we used the same set of simulations to study the shape and relative alignment of the distributions of stars, dark matter, and hot gas within their own host haloes. One of the conclusions was that although galaxies align relatively well with the local distribution of the total (mostly dark) matter, they exhibit much larger misalignments with respect to the orientiation of their complete host haloes.

After the submission of this manuscript, a paper by Chisari et al. (2015) appeared on the arXiv. They study the alignment of galaxies at $z = 0.5$ in the cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014) run with the adaptive-mesh-refinement code RAMSES (Teyssier 2002). The HORIZON-AGN simulation is run in a $(100\,h^{-1}\,\mathrm{Mpc})^3$ volume with a dark matter particle mass resolution of $m_{\mathrm{dm}} = 8 \times 10^7\,\mathrm{M}_\odot$. They focus on a galaxy stellar mass range of $9 < \log_{10}(M_{\mathrm{star}}/[\,\mathrm{M}_\odot]) < 12.36$ and separations up to $25\,h^{-1}\,\mathrm{Mpc}$. Their analysis differs in various technical, as well as conceptual, aspects from the study presented

here. However, they also report that the strength of galaxy alignments depends strongly on the subset of stars that are used to measure the orientations of galaxies, as found in our investigation.

Throughout the paper, we assume a flat ΛCDM cosmology with massless neutrinos. Such a cosmological model is characterized by five parameters: $\{\Omega_m, \Omega_b, \sigma_8, n_s, h\}$. The EAGLE and cosmo-OWLS simulations were run with two slightly different sets of values for these parameters. Specifically, EAGLE was run using the set of cosmological values suggested by the Planck mission $\{\Omega_m, \Omega_b, \sigma_8, n_s, h\} = \{0.307, 0.04825, 0.8288, 0.9611, 0.6777\}$ (Table 9; Planck Collaboration et al. 2014), whereas cosmo-OWLS was run using the cosmological parameters suggested by the 7th-year data release (Komatsu et al. 2011) of the WMAP mission $\{\Omega_m, \Omega_b, \sigma_8, n_s, h\} = \{0.272, 0.0455, 0.728, 0.81, 0.967, 0.704\}$.

This paper is organized as follows. In Section 4.2 we summarize the properties of the simulations employed in this study (§ 4.2.1) and we introduce the technical definitions used throughout the paper (§ 4.2.2 and § 4.2.3). In Section 4.3 we report the dependence of the orientation-direction alignment of galaxies on subhalo mass (§ 4.3.1), matter components (§ 4.3.2), galaxy morphology (§ 4.3.3) and subhalo type (§ 4.3.4). In Section 4.4 we compare our results with observations of the orientation-direction alignment. In section 4.5 we report results for the orientation-orientation alignment of galaxies. We summarize our findings and conclude in Section 4.6.

4.2   Simulations and Technical Definitions

4.2.1   Simulations

In this work we employ two different sets of hydrodynamical cosmological simulations, EAGLE (Schaye et al. 2015; Crain et al. 2015) and cosmo-OWLS (Schaye et al. 2010; Le Brun et al. 2014; McCarthy et al. 2014). Specifically, from the EAGLE project we make use of the simulations run in domains of boxsize $L = 25$ and 100 comoving Mpc in order to study with sufficient resolution central and satellite galaxies hosted by subhaloes with mass from $\log_{10}(M_{sub}/[h^{-1} M_\odot]) = 10.7$ up to $\log_{10}(M_{sub}/[h^{-1} M_\odot]) = 12.6$, whereas from cosmo-OWLS we select the simulations run in domains of boxsize $L = 200$ and 400 comoving h^{-1} Mpc which enable us to extend our analysis to $\log_{10}(M_{sub}/[h^{-1} M_\odot]) = 15$. For each simulation the minimum value of subhalo mass is chosen to be the subhalo mass above which all haloes have at least 300 stellar particles. Using 300 particles ensures a reliable estimation of the subhalo shape (Velliscig et al. 2015). Table 4.1 lists relevant specifics of these simulations. A relevant feature of our composite sample of haloes, taken from four different simulations, is that it reproduces the stellar mass halo mass relation inferred from abundance matching techniques studies (Schaye et al. 2015), which ensures that galaxies in our simulations reside in subhaloes of the right mass.

EAGLE and cosmo-OWLS were both run using modified versions of the N-Body Tree-PM smoothed particle hydrodynamics (SPH) code GADGET 3 (Springel 2005). The simulations employed in this work make use of element-by-element radiative cooling (Wiersma et al. 2009a), star formation (Schaye & Dalla Vecchia 2008), stellar mass losses (Wiersma et al. 2009b), stellar feedback (Dalla Vecchia & Schaye 2008, 2012), Black Hole (BH) growth through gas accretion and mergers (Booth & Schaye 2009; Rosas-Guevara et al. 2013), and thermal AGN feedback (Booth & Schaye 2009; Schaye et al. 2015).

The subgrid physics used in cosmo-OWLS is identical to that used in the OWLS run "AGN" (Schaye et al. 2010). EAGLE includes a series of developments with respect to cosmo-OWLS in the subgrid physics, namely the use of thermal (Dalla Vecchia & Schaye 2012),

instead of kinetic, energy feedback from star formation, BH accretion that depends on the gas angular momentum (Rosas-Guevara et al. 2013) and a metallicity dependent star formation law. More information regarding the technical implementation of EAGLE's hydro-dynamical aspects, as well as the subgrid physics, can be found in Schaye et al. (2015).

4.2.2 Halo and subhalo definition

Haloes are identified by first applying the Friends-of-Friends (FoF) algorithm to the dark matter particles, with linking length 0.2 (Davis et al. 1985). Baryonic particles are associated to their closest dark matter particle and they inherit their group classification. Subhaloes are identified as groups of particles in local minima of the gravitational potential. The gravitational potential is calculated for the different particle types separately and then added in order to avoid biases due to different particle masses. Local minima are identified by locating saddle points in the gravitational potential. All particles bound to a given local minimum constitute a subhalo. The most massive subhalo in a given halo is the *central* subhalo, whereas the others are *satellite* subhaloes. Minima of the gravitational potential are used to identify the centers of subhaloes. The subhalo mass M_{sub} is the sum of the masses of all the particles belonging to the subhalo. For every subhalo we define the radius $r_{\text{half}}^{\text{dm}}$ within which half the mass in dark matter is found. Similarly, but using stellar particles, we define $r_{\text{half}}^{\text{star}}$ (usually around one order of magnitude smaller than $r_{\text{half}}^{\text{dm}}$), which represents a proxy for the typical observable extent of a galaxy within a subhalo. The r_{200}^{crit} is the radius of the sphere, centered on the central subhalo, that encompasses a mean density that is 200 times the critical density of the Universe. The mass within r_{200}^{crit} is the halo mass M_{200}^{crit}. The aforementioned quantities are computed using SUBFIND (Springel et al. 2001; Dolag et al. 2009).

In Table 4.2 we summarize the $z = 0$ values of various quantities of interest for the halo mass bins analysed here.

Simulation	L	N_{particle}	Cosmology	m_{b} $[\,h^{-1}\,\text{M}_\odot]$	m_{dm} $[\,h^{-1}\,\text{M}_\odot]$	ϵ_{prop} $[\,h^{-1}\,\text{kpc}]$	tag
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EAGLE Recal	25 [Mpc]	2×752^3	PLANCK	1.5×10^5	8.2×10^5	0.2	EA L025
EAGLE Ref	100 [Mpc]	2×1504^3	PLANCK	1.2×10^6	6.6×10^6	0.5	EA L100
cosmo-OWLS AGN 8.0	200 $[\,h^{-1}\,\text{Mpc}]$	2×1024^3	WMAP7	8.7×10^7	4.1×10^8	2.0	CO L200
cosmo-OWLS AGN 8.0	400 $[\,h^{-1}\,\text{Mpc}]$	2×1024^3	WMAP7	7.5×10^8	3.7×10^9	4.0	CO L400

Table 4.1: List of the simulations used and their relevant properties. Description of the columns: (1) descriptive simulation name; (2) comoving size of the simulation box; (3) total number of particles; (4) cosmological parameters; (5) initial mass of baryonic particles; (6) mass of dark matter particles; (7) maximum proper softening length; (8) simulation name tag.

Simulation tag	mass bin *	M_{200}^{crit} *	M_{star} *	$\sigma_{\log_{10} M_{\text{star}}}$ *	r_{200}^{crit} **	$r_{\text{half}}^{\text{dm}}$ **	$r_{\text{half}}^{\text{star}}$ **	N_{halo}	N_{sat}	Color
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
EA L025	$[10.70 - 11.30]$	10.87	8.72	0.46	68.1	28.0	2.3	234	43	black
EA L100	$[11.30 - 12.60]$	11.59	9.92	0.45	118.4	50.7	3.2	4530	745	red
CO L200	$[12.60 - 13.70]$	12.78	10.88	0.27	295.6	175.7	31.1	5745	450	green
CO L400	$[13.70 - 15.00]$	13.82	11.85	0.22	656.3	416.4	73.5	3014	94	blue

Table 4.2: Values at $z = 0$ of various quantities of interest for our four subhalo mass bins. Description of the columns: (1) simulation tag; (2) subhalo mass range $\log_{10}(M_{\text{sub}}/(\,h^{-1}\,\text{M}_\odot))$; (3) median value of the halo mass $\log_{10}(M_{200}^{\text{crit}})$ for centrals; (4) median value of the stellar mass ($\log_{10}(M_{\text{star}}/(\,h^{-1}\,\text{M}_\odot))$); (5) standard deviation of the stellar mass distribution $\sigma_{\log_{10} M_{\text{star}}}$; (6) median value of halo virial radius r_{200}^{crit} for centrals; (7) median radius within which half of the mass in dark matter is enclosed; (8) median radius within which half of the mass in stars is enclosed; (9) number of haloes; (10) number of *satellite* haloes (11) color used throughout the paper for this particular mass bin and, with different shades, for the simulation from which the mass bin is drawn from.

* $\log_{10}[M/(\,h^{-1}\,\text{M}_\odot)]$
** $[\,h^{-1}\,\text{kpc}]$

4.2.3 Shape parameter definitions

To describe the morphology and orientation of a subhalo we make use of the three-dimensional mass distribution tensor, also referred to as the inertia tensor (e.g. Cole & Lacey 1996),

$$M_{ij} = \sum_{p=1}^{N_{\text{part}}} m_p x_{pi} x_{pj}, \tag{4.4}$$

where N_{part} is the number of particles that belong to the structure of interest, x_{pi} denotes the element i (with $i, j = 1, 2, 3$ for a 3D particle distribution) of the position vector of particle p, and m_p is its mass.

The eigenvalues of the inertia tensor are λ_i (with $i = 1, 2, 3$ and $\lambda_1 > \lambda_2 > \lambda_3$, for a 3D particle distribution as in our case). The moduli of the major, intermediate, and minor axes of the ellipsoid that have the same mass distribution as the structure of interest, can be written in terms of these eigenvalues as $a = \sqrt{\lambda_1}$, $b = \sqrt{\lambda_2}$, and $c = \sqrt{\lambda_3}$. Specific ratios of the moduli of the axes are used to define the sphericity, $S = c/a$, and triaxiality, $T = (a^2 - b^2)/(a^2 - c^2)$, parameters (see Velliscig et al. 2015). The eigenvectors $\hat{e}_i$, associated with the eigenvalues λ_i, define the orientation of the ellipsoid and are a proxy for the orientation of the structure itself. We interpret this ellipsoid as an approximation to the shape of the halo and the axis represented by the major eigenvector as the orientation of the halo in a 3D space.

4.3 Orientation-direction alignment

In this section we present results concerning the alignment between the orientations of the stellar distributions in subhaloes, defined as the major eigenvector of the inertia tensor, $\hat{e}_1$, and the normalized separation vector, $\hat{d}$, of a galaxy at distance r. Note that all quantities are defined in a 3D space. We define ϕ as:

$$\phi(r) = \arccos(|\hat{e}_1 \cdot \hat{d}(r)|), \tag{4.5}$$

where $\hat{e}_1$ is the major eigenvector of a galaxy in the orientation sample, and $\hat{d}$ is the separation vector pointing towards the position of a galaxy in the position sample (see Fig. 4.1). Note that, following Eq.4.5, $0 < \phi < \pi/2$. The value of $\langle \cos(\phi) \rangle$ is then computed as an average over pairs of galaxies from the orientation and position samples. Values of $\langle \cos(\phi) \rangle$ close to unity indicate that on average galaxies are preferentially oriented towards the direction of neighbouring subhaloes. We remind the reader that we use the term subhalo to refer to the ensemble of particles bound to a local minimum in the gravitational potential. Central galaxies are hosted by the most massive subhalo in a FoF group (see § 4.2.2). Throughout the text and in the figures we use (+) to refer to properties of galaxies in the orientation sample, whereas we use (g) for galaxies in the position sample.

Observations typically measure the product of the cosine of the angle ϕ and the ellipticity of the galaxy in the orientation sample. We opt to begin our analysis by presenting results only for the angle ϕ since it has a clearer interpretation that is independent on the shape determination of the galaxy. We present results for observationally accessible proxies in Section 4.4.

4.3.1 Dependence on subhalo mass and separation

The left panel of Fig. 4.2 shows $\langle \cos(\phi) \rangle$ for pairs of galaxies (both centrals and satellites) binned in subhalo mass and as a function of 3D separation r. Subhaloes in the orientation
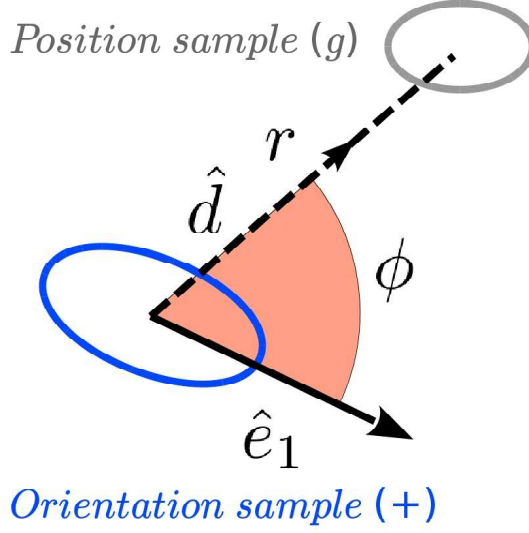
Figure 4.1: Diagram of the angle ϕ between the major eigenvector $\hat{e}_1$ of a subhalo in the orientation sample (+), and the separation vector $\hat{d}$ pointing towards the direction of a subhalo in the position sample (g). Note that all quantities are defined in 3D space.
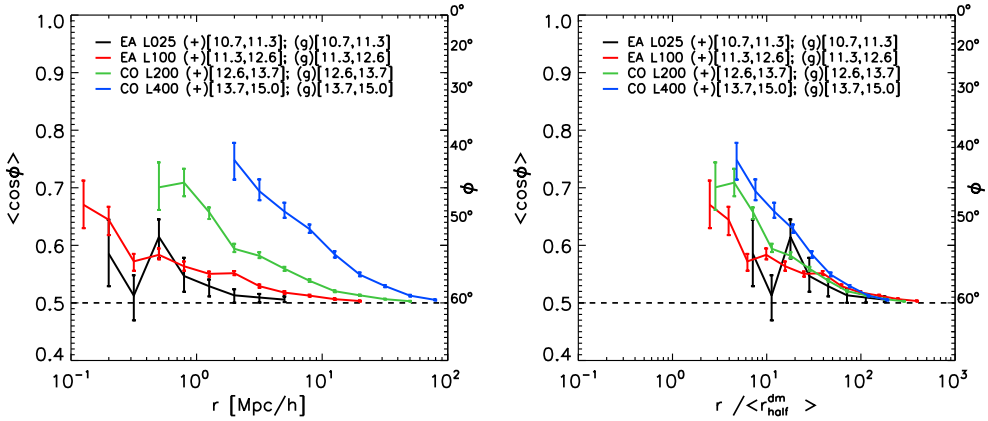

Figure 4.2: *Left*: Mean value of the cosine of the angle ϕ between the major eigenvector of the stellar distribution and the directions towards subhaloes with comparable masses as a function of 3D galaxy separation. Every mass bin is taken from a different simulation. The simulation identifiers used in the legends refer to column (8) of Table 4.1. The minimum subhalo mass in every bin ensures that only haloes with more than 300 stellar particles are selected. The curves are not shown for 3D separations larger than approximately 1/3 of the simulation volume. *Right*: Same as left panel but with physical distances rescaled by the r_{half}^{dm} of the subhaloes. In both panels the error bars represent one sigma bootstrap errors. The horizontal dashed line indicates the expectation value for random orientations. The orientation-direction alignment decreases with distance and increases with mass. The mass dependence is greatly reduced when the distances are normalized by r_{half}^{dm}.
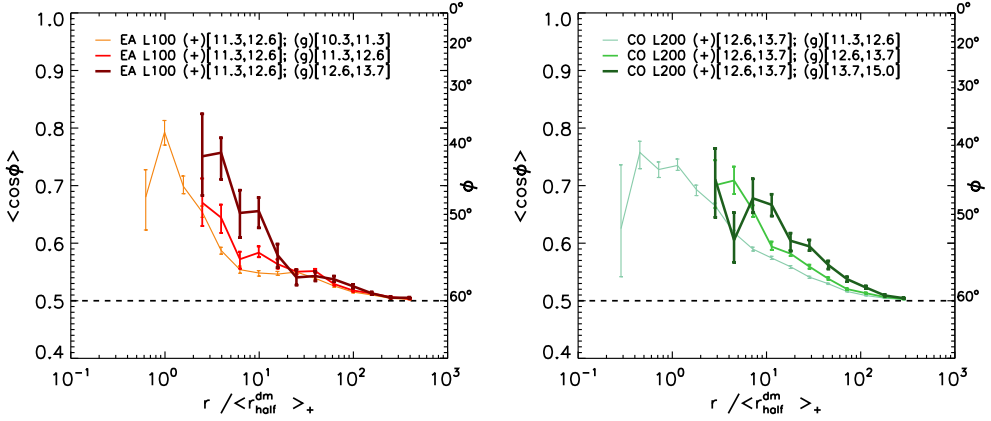
Figure 4.3: As for the right panel of Fig. 4.2, but in this case the masses of the subhaloes in the orientation sample (+) are kept fixed whereas subhaloes in the position (g) sample are selected from mass bins above, below or equal to the mass bin of the orientation sample. Physical distances are rescaled by the $r_{\mathrm{half}}^{\mathrm{dm}}$ of the subhaloes in the orientation sample. In the left panel the subhaloes are taken from the EAGLE L100 simulation and in the right panel they are taken from the cosmo-OWLS L200 simulation. In both panels the error bars represent one sigma bootstrap errors. Thicker lines indicate more massive subhaloes for the position sample. The orientation-direction alignment is stronger for more massive subhaloes in the position subsample.

sample (+) are chosen to have the same mass limits as the subhaloes in the position sample (g). Values are shown for four different choices of subhalo masses, where every mass bin is taken from a different simulation (see legend). Errors are estimated via the bootstrap technique. Specifically, we use the 16th and the 84th percentiles of 100 realizations to estimate the lower and upper limits of the error bars. The cosine of the angle between the orientation of galaxies and the direction of neighbouring galaxies is a decreasing function of distance and it increases with mass. For large separations the angle tends to the mean value for a randomly distributed galaxy orientation, i.e. $\langle \cos(\phi) \rangle = 0.5$. The physical scale at which this asymptotic behaviour is reached increases with increasing subhalo mass.

The right panel of Fig. 4.2 shows $\langle \cos(\phi) \rangle$ as a function of the physical separation rescaled by the average size of subhaloes, $\left\langle r_{\mathrm{half}}^{\mathrm{dm}} \right\rangle$, in that mass bin. This rescaling removes most, but not all, of the offset between different halo mass bins. On average, subhalo pairs separated by more than $100 r_{\mathrm{half}}^{\mathrm{dm}}$ show only weak alignment ($\langle \cos(\phi) \rangle \leq 0.52$ at $100 r_{\mathrm{half}}^{\mathrm{dm}}$).

Fig. 4.3 shows $\langle \cos(\phi) \rangle$ as a function of the separation rescaled by the average size of the subhaloes in the orientation (+) sample. In this case the masses of the subhaloes for which we measure the orientation of the stellar distribution are kept fixed whereas haloes in the position (g) sample are selected from mass bins above, below or equal to the mass bin of the orientation sample. Results are shown for two of the four simulations: in the left panel for EAGLE L100 and in the right panel for cosmo-OWLS L200. The line thickness is proportional to the subhalo mass of the position sample. The orientation of subhaloes of a given mass tends to be more aligned with the position of higher-mass subhaloes.

We note that the two suites of simulations employed here, cosmo-OWLS and EAGLE, differ in resolution, volume, cosmology and subgrid physics. Testing how each of these differences impacts our mean results is beyond the scope of this study (we would need as many simulations as differences that we wish to test), therefore we examine the overall convergence
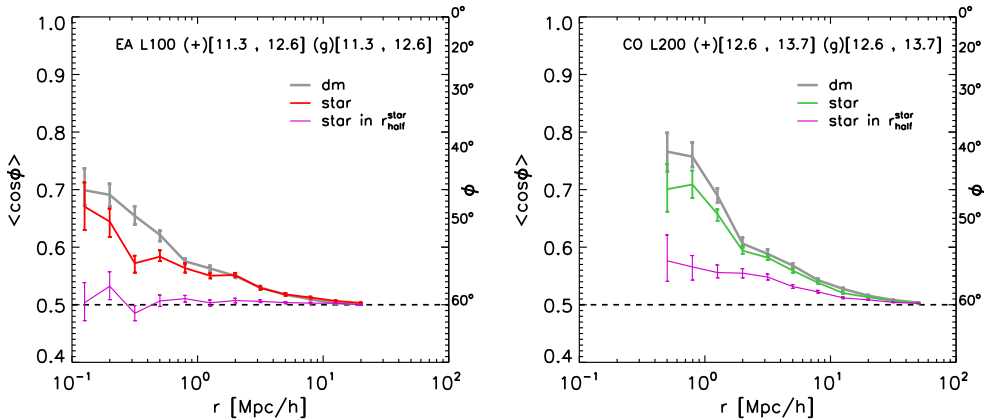
Figure 4.4: Mean value of the cosine of the angle ϕ between the major eigenvectors of the distributions of stars (red curve in the left panel and green curve in right panel as in Fig. 4.2), dark matter (gray curves), or stars within $r_{\mathrm{half}}^{\mathrm{star}}$ (purple curves) and the direction towards subhaloes with comparable masses as a function of 3D galaxy separation. The subhaloes used for the left panel are taken from the EAGLE L100 $(11.3 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 12.6)$ simulation while in the right panel they are taken from the cosmo-OWLS L200 simulation $(12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 13.7)$. Thicker lines indicate components with stronger alignment. In both panels the error bars represent one sigma bootstrap errors. The orientation of the dark matter component is most strongly aligned with the directions of nearby subhaloes, whereas the orientation of stars inside $r_{\mathrm{half}}^{\mathrm{star}}$ shows the weakest alignment.

of the two simulations by selecting a subhalo mass bin $12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 13.1$ that yields an orientation sample of galaxies that is numerous enough in the EAGLE L100 simulation, as well as resolved in the cosmo-OWLS L200 simulation. We find that, in this specific case, the results are consistent within the bootstrapped errors, both for stars and stars within $r_{\mathrm{half}}^{\mathrm{star}}$ (not shown).

Subhalo mass plays an important role in the strength of the orientation-direction alignment of subhaloes. The dependence on the subhalo mass weakens with distance but only becomes negligible for separation $\gg 100$ times the subhalo radius.

4.3.2   Dependence on the choice of matter component

In this section we report the orientation-direction alignment for the case in which the orientation of the subhalo is calculated using, respectively, dark matter, stars (as in the previous section) and stars within the half-mass radius $r_{\mathrm{half}}^{\mathrm{star}}$. An alternative choice of a proxy for the typical extent of a galaxy would be to consider only stars within a fixed 3D aperture of 30Kpc that gives similar galaxy properties as the 2-D Petrosian apertures often used in observational studies (Schaye et al. 2015). Note that the two definitions coincide for the subhalo mass bin $12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 13.7$ (CO L200). We note that $r_{\mathrm{half}}^{\mathrm{star}}$ varies among the four mass bins used in this work (see Table 4.2 column (8)).

Fig. 4.4 shows the cosine of the angle ϕ between the direction of nearby subhaloes and the orientation of the distribution of dark matter, stars (as shown in Fig. 4.2) and stars within the half-mass radius of subhaloes in the same mass bin. The left panel displays the results for the subhalo mass bin $11.3 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 12.6$ (from the EAGLE L100 simulation), whereas the right panel refers to the subhalo mass bin $12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M}_\odot]) < 13.7$

(from the cosmo-OWLS L200 simulation).

Irrespective of the subhalo mass and separation, the orientation of the dark matter component shows the strongest alignment with the directions of nearby haloes, whereas the orientation of stars inside $r_{\text{half}}^{\text{star}}$ shows the weakest alignment.

These results are suggestive of a scenario in which the alignment between subhaloes and the surrounding density field is imprinted mostly on the dark matter distribution. Therefore, when the orientation of the subhalo is computed using all stars or the stars within $r_{\text{half}}^{\text{star}}$, the signal is weakened according to the internal misalignment angle between the specified component and the total dark matter distribution. The trend shown by Fig. 4.4 therefore follows naturally from the results of Velliscig et al. (2015): stars within $r_{\text{half}}^{\text{star}}$ exhibit a weaker alignment with the total dark matter distribution than all stars in the subhalo.

The difference between the orientation-direction alignment obtained using the dark matter, all the stars or the stars within the typical extent of the galaxy, could account for the common finding reported in the literature of galaxy alignment, that such alignments are systematically stronger in simulations than when measured in observational data (see the recent reviews of Kiessling et al. 2015 and Kirk et al. 2015 for a detailed comparison between observational and computational studies). Observations are limited to the shape and orientation of the region of a galaxy above a limit surface brightness (often within surface brightness isophotes), whereas simulations need to rely on proxies for the extent of those regions (e.g. using baryonic over-density thresholds Hahn et al. 2010; Codis et al. 2015; Welker et al. 2014; Dubois et al. 2014) or to employ weighting schemes to the sample of star particles that constitute a galaxy (see e.g. use of the *reduced* inertia tensor in Tenneti et al. 2015).

4.3.3 Dependence on galaxy morphology

Theory predicts that the alignment of early-type galaxies and late-type galaxies arises from different physical processes (e.g. Catelan et al. 2001). It is of interest then to study the alignment as a function of galaxy morphologies.

In this section we report the orientation-direction alignment of galaxies with different sphericities in order to explore the effect of the shape of galaxies on the orientation-direction alignment. We divide our sample of subhaloes according to the sphericity of their whole stellar distribution, defined as $S = c/a$ where a and c are the squareroot of the major and minor eigenvalues of the inertia tensor respectively (see §4.2.3). We choose a threshold value for the sphericity of 0.5 that yields a similar numbers of galaxies in the two subsamples, as the median sphericity of the total sample is 0.55. This galaxy selection by sphericity represent a simple proxy for galaxy morphology.

Fig. 4.5 shows the mean values of the cosine of the angle ϕ for galaxies of sphericity above and below the threshold, as well as for the total sample. The left panel displays the results for the subhalo mass bin $11.3 < \log_{10}(M_{\text{sub}}/[\,h^{-1}\,\text{M}_\odot]) < 12.6$ (from the EAGLE L100 simulation), whereas the right panel refers to the subhalo mass bin $12.6 < \log_{10}(M_{\text{sub}}/[\,h^{-1}\,\text{M}_\odot]) < 13.7$ (from the cosmo-OWLS L200 simulation).

More spherical galaxies (thinner lines) show a weaker orientation-direction alignment. The differences between the two shape selected samples of haloes are within the errors for scales larger than $1\,h^{-1}$ Mpc, suggesting that the effect of shape is dominated by subhaloes of the same hosts. A similar trend (not shown) is found using triaxiality, see §4.2.3, as the indicator of galaxy shape. Prolate ($T > 0.5$) stellar distributions show the strongest orientation-direction alignment, whereas oblate ($T < 0.5$) ones show the weakest. The better alignment of prolate or aspherical galaxies is probably due to the fact that these galaxies align better with their underlying dark matter distributions (not shown), which in turn produces a stronger orientation-direction alignment (see Fig. 4.4).
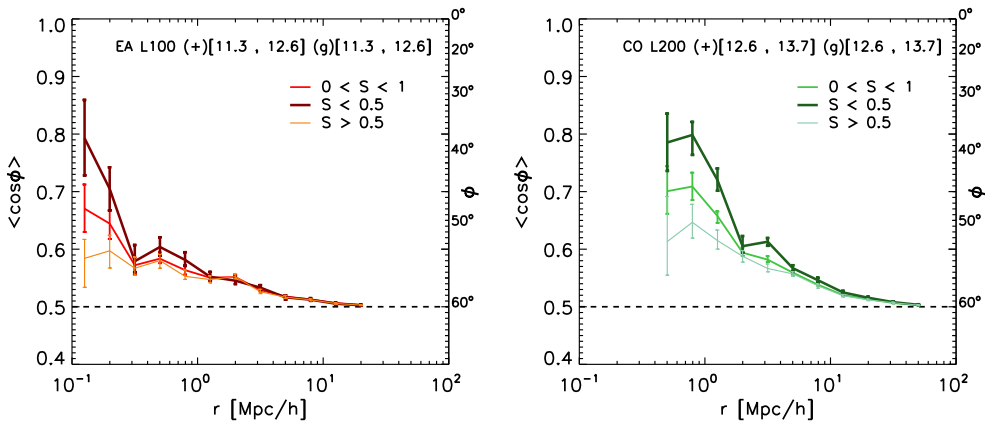
Figure 4.5: Mean value of the cosine of the angle ϕ between the major eigenvector of the stellar distribution and the direction towards neighbouring subhaloes as a function of 3D galaxy separation, for galaxies in the orientation sample selected based on their shape. The selection is based on the sphericity of the whole stellar distribution defined as $S = c/a$ where a and c are the square root of the major and minor eigenvalues of the inertia tensor respectively. We choose a threshold value for the sphericity of 0.5. The subhaloes used for the left panel are taken from the EAGLE L100 ($11.3 < \log_{10}(M_{\mathrm{sub}}/[h^{-1}\,M_{\odot}]) < 12.6$) simulation while in the right panel they are taken from the cosmo-OWLS L200 simulation ($12.6 < \log_{10}(M_{\mathrm{sub}}/[h^{-1}\,M_{\odot}]) < 13.7$). Thicker lines indicate components with stronger alignment. In both panels the error bars represent one sigma bootstrap errors. More spherical galaxies show a weaker orientation-direction alignment.

We note that the orientation of a perfectly spherical distribution ($S = 1$) of stars is ill defined. Although this can potentially affect our measurements, less than 2% of galaxies in our sample have a sphericity higher than 0.8. We also note that more massive haloes, for which the orientation-direction alignment is strongest, tend to be less spherical and more triaxial (see Velliscig et al. 2015). Therefore, selecting haloes by shape biases the sample towards systematically different masses: however, the mass difference in the two shape-selected samples is about 4%, which is too small to explain the differences in alignment of haloes with different shapes.

Observations indicate that ellipsoidal galaxies show stronger intrinsic alignment than blue disk galaxies (Hirata et al. 2007; Singh et al. 2015). However, we caution the reader that there are still many complications to take into account before one can compare the trends discussed above with these observational results. First, one would need to select galaxies based on their colors, which requires stellar population synthesis models. Second, the sphericity of the stellar component is a simplistic proxy for selecting disc galaxies. Selecting galaxies according to their morphology, in a similar way as done observationally, would require a stellar light decomposition in bulge and disc component.

4.3.4   Alignment of satellite and central galaxies

The increased probability of finding satellites along the major axis of the central galaxy

In the previous sections we studied the orientation-direction alignment of galaxies irrespective of their classification as centrals or satellites. In this subsection we report the alignment
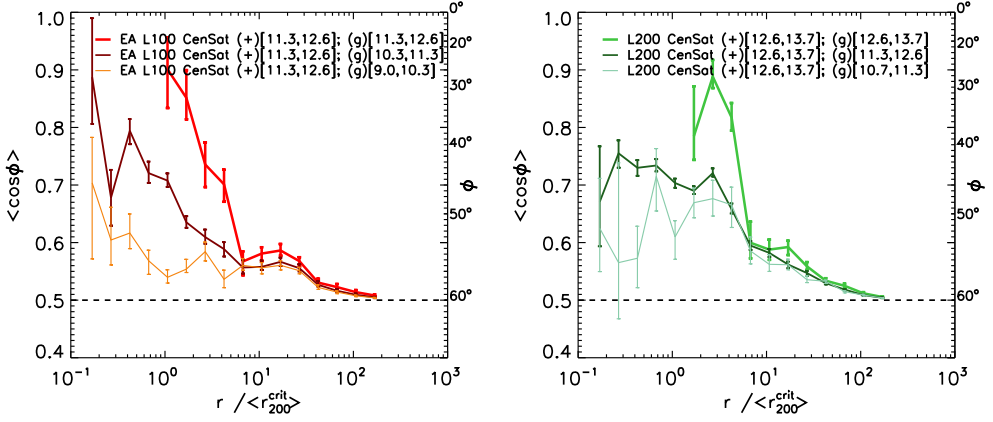
Figure 4.6: Mean value of the cosine of the angle between the orientation of the stars in the central galaxy and the direction of satellite galaxies as a function of the 3D galaxy separation, rescaled by the host halo r_{200}^{crit}. The central galaxies used for the left panel are taken from the EAGLE L100 ($11.3 < \log_{10}(M_{sub}/[h^{-1}M_\odot]) < 12.6$) while in the right panel they are taken from the cosmo-OWLS L200 simulation ($12.6 < \log_{10}(M_{sub}/[h^{-1}M_\odot]) < 13.7$). In both panels the error bars represent one sigma bootstrap errors. Thicker lines indicate higher mass. The satellite distribution is aligned with the central galaxy out to $\sim 100 r_{200}^{crit}$. For $r < 10 r_{200}^{crit}$ the alignment is substantially stronger for higher-mass satellites.
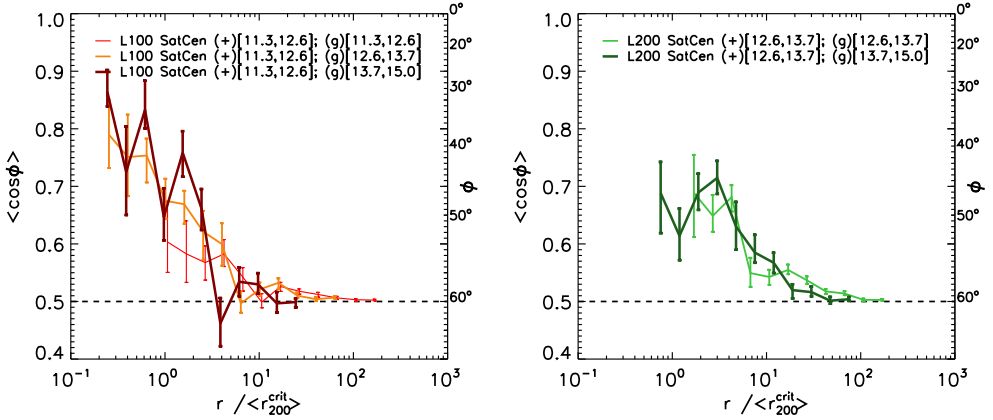
Figure 4.7: As for Fig. 4.6 but in this case the orientation is computed for the stellar distribution in satellite galaxies and the angle is measured with respect to the directions of central galaxies hosted by subhaloes of different masses. The alignment between satellites and the directions of centrals decreases with distance but is insensitive to the mass of the host halo.

between the orientations of central galaxies (g) and the directions of satellite[1] galaxies (+) and, in turn, the probability of finding satellite galaxies distributed along the major axis of the central galaxy. This effect has been studied both theoretically, making use of N-Body (e.g. Faltenbacher et al. 2008; Agustsson & Brainerd 2010; Wang et al. 2014a) and hydrodynamical simulations (Libeskind et al. 2007; Deason et al. 2011), and observationally (Sales & Lambas 2004; Brainerd 2005; Yang et al. 2006; Wang et al. 2008; Nierenberg et al. 2012; Wang et al. 2014b; Dong et al. 2014). Those studies report that the distribution of satellites around central galaxies is anisotropic, with an excess of satellites aligned with the major axis of the central galaxy.

Fig. 4.6 shows the average angle between the orientation of the stellar distribution of central subhaloes and the position of satellite galaxies. Values of $\langle \cos \phi \rangle$ that are significantly greater than 0.5 indicate that the positions of satellites are preferentially aligned with the major axis of the central galaxy. We use two different mass bins taken from two simulations: $11.3 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M_\odot}]) < 12.6$ from EAGLE L100 (left) and $12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M_\odot}]) < 13.7$ from cosmo-OWLS L200 (right). The line thickness is proportional to the subhalo mass of the position (g) sample. In both panels the physical separations between the pairs are normalized by the $\left\langle r_{200}^{\mathrm{crit}} \right\rangle$ of the haloes hosting the central galaxies.

For separations up to $100 \left\langle r_{200}^{\mathrm{crit}} \right\rangle$, the positions of satellite galaxies are significantly aligned with the orientation of central galaxies (not necessarily in the same host halo), with more massive satellites showing a stronger alignment. The same qualitative behaviour is found for both mass bins, but the effect is stronger for the more massive central subhaloes. On scales larger than $\sim \left\langle 10\, r_{200}^{\mathrm{crit}} \right\rangle$ the alignment depends only weakly on the mass of the satellite subhaloes. We speculate that the alignment of satellites with central galaxies of different host haloes is likely driven by the correlation between the orientation of the central galaxies and the surrounding large-scale structure, which in turn influences the positions of satellite galaxies.

The radial alignment of satellite galaxies with the direction of the host galaxy

Here we investigate the radial alignment of the orientations of satellites (+) with the direction of the central galaxy (g), whereas in the previous section report the results for alignment between the orientations of the central galaxy and the direction of satellites. The orientation of satellite subhaloes is computed using all the stars bounded to the subhalo. Theoretical studies using N-body simulations (Kuhlen et al. 2007; Pereira et al. 2008; Faltenbacher et al. 2008) and hydrodynamic simulations (Knebe et al. 2010) found that on average the orientation of satellite galaxies is aligned with the direction of the centre of their host halo.

Fig. 4.7 shows the average value of the cosine of the angle between the orientation of the satellite and the direction of the centrals as a function of the separation rescaled by the average virial radius $(r_{200}^{\mathrm{crit}})$. The mass of the subhaloes in the orientation sample (+) is kept fixed whereas the masses of the central haloes (g) are chosen to have similar or higher masses. Values of $\langle \cos \phi \rangle$ that are significantly greater than 0.5 indicate that the orientation of satellites galaxies are preferentially aligned towards the direction of central galaxies. As for the previous subsection, we use two different mass bins taken from two simulations: $11.3 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M_\odot}]) < 12.6$ from EAGLE L100 (left) and $12.6 < \log_{10}(M_{\mathrm{sub}}/[\,h^{-1}\,\mathrm{M_\odot}]) < 13.7$ from cosmo-OWLS L200 (right). The line thickness is proportional to the subhalo mass of the position (g) sample. In both panels the physical separations between the pairs are normalised by the $\left\langle r_{200}^{\mathrm{crit}} \right\rangle$ of the haloes hosting the central galaxies.

---

[1]In this subsection, satellite galaxies do not necessarily belong to the same haloes that host the paired central galaxies.
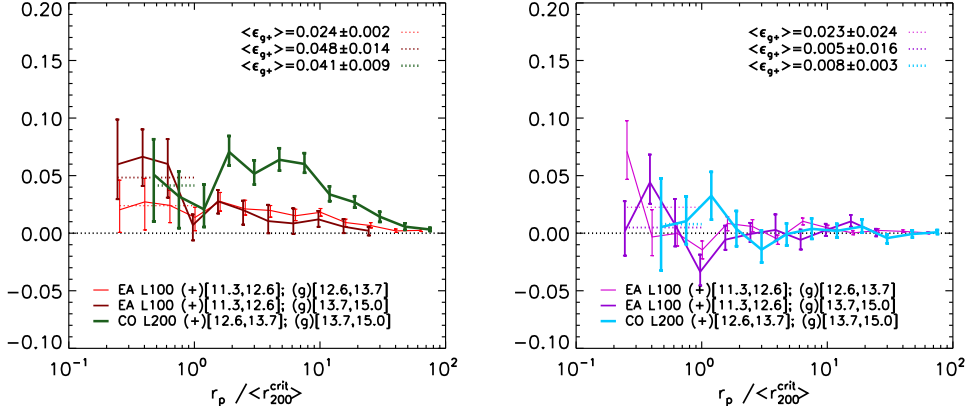
Figure 4.8: Values of the observationally accessible proxy for orientation-direction alignment, ϵ_{g+} (see Eq. 4.9), as a function of the projected separation r_p. Only pairs that are separated by less than $2.5\,h^{-1}$ Mpc along the projected axis are considered for this analysis. The error bars indicate one sigma bootstrap errors. In the left panel the orientation of the subhalo is measured using all the stellar particles whereas in the right panel only stars within the r_{half}^{star} are used, which greatly reduces the alignment. Thicker lines indicates higher masses. The coloured dotted lines show the average value of ϵ_{g+} within the virial radius of the central galaxy. Observational measurements from Sifón et al. (2015) constrained the average ellipticity to be $\epsilon_{g+} = -0.0037 \pm 0.0027$.

The major axes of satellite galaxies, when all stars are considered, are significantly aligned towards the direction of the centrals within their virial radius. The strength of the alignment declines very rapidly with radius and is very small outside the virial radius. There is only a weak dependence on the central subhalo mass.

We note that by considering only stars in r_{half}^{star} the trends shown in Fig. 4.6 and in Fig. 4.7 are weakened (not shown). This results in a less significant alignment for galaxies hosted by subhaloes with masses $11.3 < \log_{10}(M_{sub}/[\,h^{-1}\,M_\odot]) < 12.6$ from the EAGLE L100 simulation, whereas a still significant alignment is found for galaxies with $12.6 < \log_{10}(M_{sub}/[\,h^{-1}\,M_\odot]) < 13.7$ from the cosmo-OWLS L200 simulation.

4.4 Towards observations of orientation-direction galaxy alignment

In this subsection we report results for observationally accessible proxies for the orientation-direction alignment, which depend on the shape of galaxies as well as on their orientation, making them tightly connected to cosmic shear studies. All the relevant quantities for the following analysis are defined in a 2D space.

Observationally, the ellipticity is decomposed into the projected tangential (ϵ_+) and transverse ($\epsilon_\times$) components with respect to the projected separation vector of the galaxy pair:

$$\epsilon_+ = |\epsilon| \cos(2\Phi) \tag{4.6}$$

$$\epsilon_\times = |\epsilon| \sin(2\Phi) \tag{4.7}$$

$$|\epsilon| = \frac{1 - b/a}{1 + b/a}, \tag{4.8}$$

where Φ is the position angle[2] between the projected orientation of the galaxy and the direction of a galaxy at projected distance r_p and b/a is the axis ratio of the projected galaxy.

Then the function ϵ_{g+} is defined as:

$$\epsilon_{g+}(r_p) = \sum_{i \neq j | r_p} \frac{\epsilon_+(j \mid i)}{N_{pairs}}, \tag{4.9}$$

where the index i represents a galaxy in the shape sample, whereas the index j represents a galaxy in the position sample. The function $\epsilon_{g+}(r_p)$ is the average value of ϵ_+ at the projected separation r_p.

Groups and clusters of galaxies, where strong tidal torques are expected to align satellite galaxies toward the centre of the host's gravitational potential, are ideal environments to study orientation-direction alignment. However, the task of measuring this alignment has proven to be very challenging (see Kirk et al. 2015, and references therein). In group and cluster environments, the measured quantity, ϵ_{g+} (see Eq. 4.9), is the mean value of the angle between the projected orientation of the *satellite* galaxy and the direction of the *host*, multiplied by the projected ellipticity of the satellite. Typical values of the root mean square of galaxy shape parameter, $e = (1-(b/a)^2)/(1+(b/a)^2)$, in the set of simulations employed in this study, can be found in Fig. 3.5 of Chapter 3. Those values are in broad agreement with the observed noise-corrected values (about 0.5-0.6 depending on luminosity and galaxy type, (e.g. Joachimi et al. 2013)) when all stars in subhaloes are considered. However, when only stars within r_{half}^{star} are considered, Velliscig et al. (2015) found typical values for e_{rms} of ≈ 0.2-0.3, that is a factor of 2 lower than the observed value. This suggests that galaxy shapes computed using stars within r_{half}^{star} are rounder than the observed shapes, potentially leading to an underestimate of the ϵ_{g+}. To quantify this effect, we would need to analyse synthetic galaxy images from simulations with the shape estimator algorithms used in weak lensing measurements. We defer such an investigation to future works.

Recent observational studies of the orientation-direction alignment in galaxy groups and clusters reported signals consistent with zero alignment (Chisari et al. 2014; Sifón et al. 2015). Specifically, Sifón et al. (2015) used a sample of $\approx 14,000$ spectroscopically confirmed galaxy members of 90 galaxy clusters with median mass of $\log_{10}(M_{200}/[M_\odot]) = 14.8$ and median redshift of $z = 0.14$, selected as part of MENeaCS (Multi-Epoch Nearby Cluster Survey; Sand et al. 2012) and CCCP (Canadian Cluster Comparison Project; Hoekstra et al. 2012). They constrained the average ellipticity, within the host virial radius, to be $\epsilon_{g+} = (-3.7 \pm 2.7) \times 10^{-3}$ or $\epsilon_{g+} = (0.4 \times \pm 3.1) \times 10^{-3}$ depending on the shape estimation method employed. Chisari et al. (2014) measured galaxy alignments in 3099 photometrically-selected galaxy groups in the redshift range between $z = 0.1$ and $z = 0.4$ of masses $\log_{10}(M_{200}/[M_\odot]) = 13$ in SDSS Stripe 82 and constrained the alignments to similar values as Sifón et al. (2015).

The left panel of Fig. 4.8 shows the value of ϵ_{g+} calculated for the simulations using all the stellar particles of subhaloes for host masses and satellite masses that are roughly comparable to the range of masses explored in Chisari et al. (2014) and Sifón et al. (2015). We only consider pairs separated by less than $2.5\,h^{-1}$ Mpc along the projection axis to confine the measurement to the typical extent of massive bound structures. Within the virial radii of groups or clusters the statistical uncertainties are large. The average values of ϵ_{g+} for distances smaller than the host virial radii are $\approx 2 - 4 \times 10^{-2}$ with errors of $\approx 0.1 - 2 \times 10^{-2}$, indicating positive alignment. We repeat the same analysis using only stars within r_{half}^{star} (see right panel of Fig. 4.8). In this case the average value of ϵ_{g+} for distances that are smaller than the host virial radius is consistent with zero, in agreement with the observations of Chisari et al.

---

[2]The symbol Φ is used to indicate an angle between vector in 2D, whereas the symbol ϕ (see Eq. 4.5) indicates an angle between vectors in 3D.
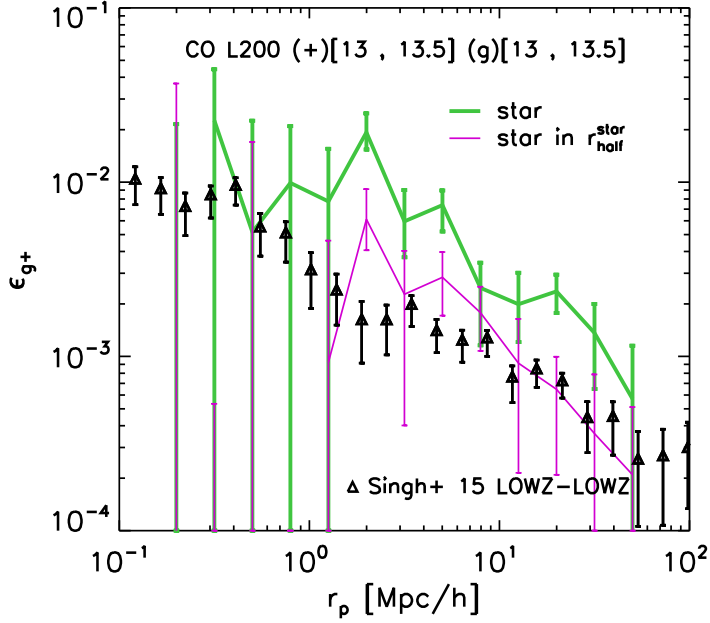
Figure 4.9: Values of the observationally accessible proxy for orientation-direction alignment, ϵ_{g+} (see Eq. 4.9), as a function of the projected separation, r_p, form simulations considering all stars bound to the subhalo (green curves) and only stars within r_{half}^{star} (magenta curve). The data points are the observational results from Singh et al. (2015) for the LOWZ sample of LRG galaxies (their Fig. 19, note that in their work ϵ_{g+} is denoted $\langle \gamma \rangle$ for its direct connection with the shear). We consider pairs of galaxies that have a separation along the projected axis smaller than $100\,h^{-1}$ Mpc. The error bars on the curves indicate one sigma bootstrap errors. When we consider all the stars, the predicted alignment is stronger than observed. However, when we only use stars in the part of the galaxy that might typically be observed, we find good agreement with the data.

(2014) and Sifón et al. (2015). Using deeper observations, in order to probe the lower surface brightness parts of satellite galaxies, could represent a way to reveal the alignment that is seen in observations when all stars bounded to subhaloes are considered.

Recently, Singh et al. (2015) measured the relative alignment of SDSS-III BOSS DR11 LOWZ Luminous Red Galaxies (LRGs) in the redshift range $0.16 < z < 0.36$ observed spectroscopically in the BOSS survey (Dawson et al. 2013). As opposed to the case of galaxy groups and clusters, these measurements are obtained by integrating along the line of sight between ±100 Mpc. Furthermore, Singh et al. (2015) reported the average halo masses of those galaxies, as obtained from galaxy-galaxy lensing analysis. We perform the same measurements as in Singh et al. (2015) on our simulations. Given the observed halo mass ($\log_{10}$ ($M_{180}^{\mathrm{mean}}[h^{-1}\,\mathrm{M_\odot}]) = 13.2$) and the line of sight integration limits, we employ the cosmo-OWLS L200 in this analysis.

Fig. 4.9 shows the values of $\epsilon_{\mathrm{g+}}(r_{\mathrm{p}})$ from our simulation together with the measurements from Singh et al. (2015). Note that we have used a halo mass bin ($13 < \log_{10}(M_{\mathrm{sub}}^{\mathrm{crit}}/[h^{-1}\,\mathrm{M_\odot}])$ < 13.5) half a magnitude wide to obtain statistically robust measurements ($N_{\mathrm{haloes}} = 1677$). As for the case of satellite galaxies in clusters, the agreement with observational results depends strongly on the subset of stars used to compute the galaxy orientations. When one considers all stars bound to subhaloes, the values obtained for $\epsilon_{\mathrm{g+}}(r_{\mathrm{p}})$ are systematically higher than the values in observations, whereas broad agreement is found when using only stars inside $r_{\mathrm{half}}^{\mathrm{star}}$.

As noted before, when only stars within $r_{\mathrm{half}}^{\mathrm{star}}$ are considered, simulated galaxies exhibit rounder shapes than observed. Therefore, the results presented here may underestimate the values of $\epsilon_{\mathrm{g+}}(r_{\mathrm{p}})$. Thus, when more observationally motivated algorithms would be employed to analyse the simulations, it is not guaranteed that the agreement found in Fig. 4.9 would still hold.

4.5 Orientation-orientation alignment

In this section, we present the results for the II (intrinsic-intrinsic) term of the intrinsic alignment that is given by the angle between the orientations of different haloes. We define ψ as

$$\psi(|\vec{r}|) = \arccos(|\hat{e}_1(\vec{x}) \cdot \hat{e}_1(\vec{x} + \vec{r})|). \tag{4.10}$$

where $\hat{e}_1$ are the major eigenvectors of the 3D stellar distributions of a pair of galaxies separated by a 3D distance $r = |\vec{r}|$ (see Fig. 4.10).

Fig. 4.11 shows the average value of the cosine of the angle ψ for pairs of subhaloes with similar masses at a given 3D separation r (in h^{-1} Mpc). Values are shown for four different choices of subhalo mass, where each mass bin is taken from a different simulation (see legend). To estimate the errors, we bootstrap the shape sample 100 times and take as 1-sigma error bars the 16th and the 84th percentile of the bootstrap distribution. Values of $\langle\cos\psi\rangle$ equal to 0.5 indicate a random distribution of galaxy orientations, whereas values of $\langle\cos\psi\rangle$ higher than 0.5 indicate that on average galaxies are preferentially oriented in the same direction.

The alignment between the orientation of the stellar distribution decreases with distance and increases with subhalo mass. Comparing with Fig. 4.2, the orientation-orientation alignment is systematically lower than the orientation-direction angle alignment. Beyond $50\,h^{-1}$ Mpc the alignment is consistent with a random distribution, whereas in the orientation-direction case a positive alignment was found for scales up to $100\,h^{-1}$ Mpc. This is suggestive of the *direction* of nearby galaxies as being the main driver of the orientation-orientation alignment, as a weaker orientation-orientation alignment naturally stems from the dilution of the orientation-direction alignment.
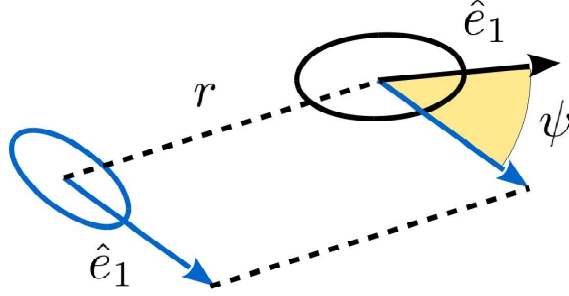
Figure 4.10: Diagram of the angle ψ formed between $\hat{e}_1$ of galaxy pairs at a distance r.

Similarly to ϵ_{g+} (in Eq. 4.9), we can define the projected orientation-orientation ϵ_{++} as:

$$\epsilon_{++}(r_p) = \sum_{i \neq j | r_p} \frac{\epsilon_+^i \epsilon_+^j(j \mid i)}{N_{pairs}}, \tag{4.11}$$

where ϵ_+ is defined in Eq. 4.6. Galaxies are selected to have at least 300 star particles.

Fig. 4.12 shows the projected orientation-orientation alignment, ϵ_{++}, for the same halo mass bin and integration limits as employed in Fig. 4.9. Green and magenta curves refer to the cases where one uses all stellar particles in subhaloes and only stellar particles confined within r_{half}^{star}, respectively. For comparison, $\epsilon_{g+}(r_p)$ is overplotted in grey. As expected, the $\epsilon_{++}(r_p)$ profile has an overall lower normalization. Interestingly, $\epsilon_{++}(r_p)$ is steeper than $\epsilon_{g+}(r_p)$, although the significance of this trend is diminished by the noisy behaviour of the $\epsilon_{++}(r_p)$ profile.

The presence of a non-vanishing $\epsilon_{++}(r_p)$ profile reveals a net alignment of galaxies with the orientations of nearby galaxies, thus suggesting a potential II term in cosmic shear measurements for galaxies residing in haloes with masses $13 < \log_{10}(M_{sub}^{crit}/[h^{-1} M_\odot]) < 13.5$.

4.6 Conclusions

This paper reports the results of a systematic study of the orientation-direction and orientation-orientation alignment of galaxies in the EAGLE (Schaye et al. 2015; Crain et al. 2015) and cosmo-OWLS (Le Brun et al. 2014; McCarthy et al. 2014) hydro-cosmological simulations. The combination of these state-of-the-art hydro-cosmological simulations enables us to span four orders of magnitude in subhalo mass ($10.7 \leq \log_{10}(M_{sub}/[h^{-1} M_\odot]) \leq 15$) and a wide range of galaxy separations ($-1 \leq \log_{10}(r/[h^{-1} Mpc]) \leq 2$). For the orientation-direction alignment we define the galaxy orientation to be the major eigenvector of the inertia tensor of the distribution of stars in the subhalo, $\hat{e}_1$. We then compute the mean values of the angle ϕ between $\hat{e}_1$ and the normalized separation vector, $\hat{d}$, towards a neighbouring galaxy at the distance r, for galaxies in different subhalo mass bins. In the case of orientation-orientation alignment, we compute the mean value of ψ, the angle between the major axes $\hat{e}_1$ of galaxy pairs separated by a distance r.
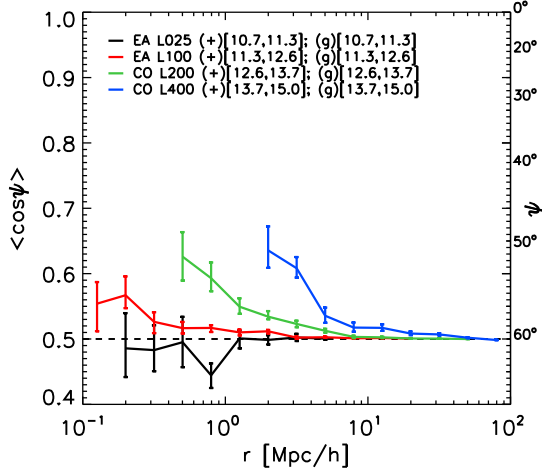
Our key findings are:

Figure 4.11: Mean value of the cosine of the angle ψ between the major axes of the stellar distributions of subhaloes as a function of their 3D separation. Each mass bin is taken from a different simulation. The minimum subhalo mass in every bin ensures that only haloes with more than 300 stellar particles are selected. Orientations are computed using all stars bound to the subhaloes. The orientation-orientation alignment decreases with distances and increases with mass. It is weaker than the orientation-direction alignment (cf. left panel of Fig. 4.2).
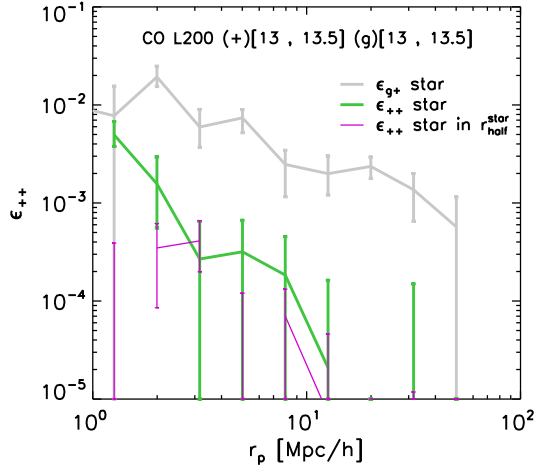
Figure 4.12: Dependence of ϵ_{++} (Eq. 4.11), a measure of orientation-orientation alignment, obtained from the simulations using an integration limit of $100\,h^{-1}$ Mpc. Both the values for the whole stellar distribution (in green) and for the stars within $r_{\text{half}}^{\text{star}}$ are shown (in purple). The error bars indicates one sigma bootstrap errors. The results for ϵ_{g+} (grey curve) are shown for comparison.

- Subhalo mass affects the strength of the orientation-direction alignment of galaxies for separations up to tens of Mpc, but for distances greater than approximately ten times the subhalo radius the dependence on mass becomes insignificant. The strength of the signal is consistent with no orientation-direction alignment for separations $\gg 100$ times the subhalo radius (Figs. 4.2-4.3).

- The difference between the orientation-direction alignment obtained using the dark matter, all the stars or the stars within $r_{\text{half}}^{\text{star}}$ to define galaxy orientations, could account for the common findings reported in the literature of galaxy alignment being systematically stronger in simulations than reported by observational studies (Fig. 4.4). Since observations are limited to the shape and orientation of the region of a galaxy above a limit surface brightness, simulations have to employ proxies for the extent of this region.

- At a fixed mass, subhaloes hosting more aspherical or prolate stellar distributions show stronger orientation-direction alignment (Fig. 4.5).

- The distribution of satellites is significantly aligned with the orientation of the central galaxy for separations up to 100 times the virial radius of the host halo (r_{200}^{crit}), within $10\, r_{200}^{\text{crit}}$ higher-mass satellites show substantially stronger alignment (Fig. 4.6).

- Satellites are radially aligned towards the directions of the centrals. The strength of the alignment of satellites decreases with radius but is insensitive to the mass of the host halo (Fig. 4.7).

- Predictions for the radial profile of the projected orientation-direction alignment of galaxies, $\epsilon_{\text{g}+}(r_{\text{p}})$, depend on the subset of stars used to measure galaxy orientations. When only stars within $r_{\text{half}}^{\text{star}}$ are used, we find agreement between results from our simulations and recent observations from Sifón et al. (2015) and Singh et al. (2015)(see Figs. 4.8 and 4.9, respectively).

- Predictions for the radial profile of the orientation-orientation alignment of galaxies, $\epsilon_{++}(r_{\text{p}})$, are systematically lower than those for the orientation-direction alignment, $\epsilon_{\text{g}+}(r_{\text{p}})$, and have a steeper radial dependence (Figs. 4.11 and 4.12). Although low, the non vanishing $\epsilon_{++}(r_{\text{p}})$ profile reveals a net alignment of galaxies with the orientations of nearby galaxies, thus suggesting a potential intrinsic-intrinsic term in cosmic shear measurements for galaxies residing in haloes with masses $13 < \log_{10}(M_{\text{sub}}^{\text{crit}}/[\,h^{-1}\,\text{M}_\odot]) < 13.5$.

For a direct comparison with the observations, in order to validate the models or to explain the observations, particular care has to be taken to compare the same quantities in simulations and observations. A future development of this work will be to extend the comparison with observations further by using the same selection criteria for luminosity, colour, and morphology in the simulations and in the observations.

The strength of galaxy alignments depends strongly on the subset of stars that are used to measure the orientations of galaxies and it is always weaker than the alignment of the dark matter components. Thus, alignment models that use halo orientation as a direct proxy for galaxy orientation will overestimate the impact of intrinsic galaxy alignments on weak lensing analyses.

Acknowledgements

Bibliography

Agustsson, I., & Brainerd, T. G. 2010, ApJ, 709, 1321

Bett, P. 2012, MNRAS, 420, 3303

Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53

Brainerd, T. G. 2005, ApJ, 628, L101

Catelan, P., Kamionkowski, M., & Blandford, R. D. 2001, MNRAS, 320, L7

Chisari, N. E., Mandelbaum, R., Strauss, M. A., Huff, E. M., & Bahcall, N. A. 2014, MNRAS, 445, 726

Chisari, N. E., Codis, S., Laigle, C., et al. 2015, ArXiv e-prints (arXiv:1507.07843), arXiv:1507.07843

Codis, S., Gavazzi, R., Dubois, Y., et al. 2015, MNRAS, 448, 3391

Cole, S., & Lacey, C. 1996, MNRAS, 281, 716

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937

Crittenden, R. G., Natarajan, P., Pen, U.-L., & Theuns, T. 2001, ApJ, 559, 552

—. 2002, ApJ, 568, 20

Croft, R. A. C., & Metzler, C. A. 2000, ApJ, 545, 561

Dalla Vecchia, C., & Schaye, J. 2008, MNRAS, 387, 1431

—. 2012, MNRAS, 426, 140

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

Deason, A. J., McCarthy, I. G., Font, A. S., et al. 2011, MNRAS, 415, 2607

Di Matteo, T., Colberg, J., Springel, V., Hernquist, L., & Sijacki, D. 2008, ApJ, 676, 33

Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Dong, X. C., Lin, W. P., Kang, X., et al. 2014, ApJ, 791, L33

Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453

Faltenbacher, A., Jing, Y. P., Li, C., et al. 2008, ApJ, 675, 146

Hahn, O., Teyssier, R., & Carollo, C. M. 2010, MNRAS, 405, 274

Heavens, A., & Peacock, J. 1988, MNRAS, 232, 339

Heavens, A., Refregier, A., & Heymans, C. 2000, MNRAS, 319, 649

Hirata, C. M., Mandelbaum, R., Ishak, M., et al. 2007, MNRAS, 381, 1197

Hoekstra, H., Mahdavi, A., Babul, A., & Bildfell, C. 2012, MNRAS, 427, 1298

Jing, Y. P. 2002, MNRAS, 335, L89

Joachimi, B., Mandelbaum, R., Abdalla, F. B., & Bridle, S. L. 2011, A&A, 527, A26

Joachimi, B., Semboloni, E., Bett, P. E., et al. 2013, MNRAS, 431, 477

Joachimi, B., Cacciato, M., Kitching, T. D., et al. 2015, Space Sci. Rev., arXiv:1504.05456

Kiessling, A., Cacciato, M., Joachimi, B., et al. 2015, ArXiv e-prints (arXiv:1504.05546),
    arXiv:1504.05546

Kirk, D., Brown, M. L., Hoekstra, H., et al. 2015, ArXiv e-prints (arXiv:1504.05465),
    arXiv:1504.05465

Knebe, A., Libeskind, N. I., Knollmann, S. R., et al. 2010, MNRAS, 405, 1119

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Kuhlen, M., Diemand, J., & Madau, P. 2007, ApJ, 671, 1135

Le Brun, A. M. C., McCarthy, I. G., Schaye, J., & Ponman, T. J. 2014, MNRAS, 441, 1270

Lee, J., Springel, V., Pen, U.-L., & Lemson, G. 2008, MNRAS, 389, 1266

Libeskind, N. I., Cole, S., Frenk, C. S., Okamoto, T., & Jenkins, A. 2007, MNRAS, 374, 16

McCarthy, I. G., Le Brun, A. M. C., Schaye, J., & Holder, G. P. 2014, MNRAS, 440, 3645

McCarthy, I. G., Schaye, J., Ponman, T. J., et al. 2010, MNRAS, 406, 822

Nierenberg, A. M., Auger, M. W., Treu, T., et al. 2012, ApJ, 752, 99

Pereira, M. J., Bryan, G. L., & Gill, S. P. D. 2008, ApJ, 672, 825

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16

Rosas-Guevara, Y. M., Bower, R. G., Schaye, J., et al. 2013, ArXiv e-prints
    (arXiv:1312.0598), arXiv:1312.0598

Sales, L., & Lambas, D. G. 2004, MNRAS, 348, 1236

Sand, D. J., Graham, M. L., Bildfell, C., et al. 2012, ApJ, 746, 163

Schaye, J., & Dalla Vecchia, C. 2008, MNRAS, 383, 1210

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Sifón, C., Hoekstra, H., Cacciato, M., et al. 2015, A&A, 575, A48

Singh, S., Mandelbaum, R., & More, S. 2015, MNRAS, 450, 2195

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Tenneti, A., Mandelbaum, R., Di Matteo, T., Feng, Y., & Khandai, N. 2014, MNRAS, 441, 470

Tenneti, A., Singh, S., Mandelbaum, R., et al. 2015, MNRAS, 448, 3522

Teyssier, R. 2002, A&A, 385, 337

Tormen, G. 1997, MNRAS, 290, 411

Velliscig, M., Cacciato, M., Schaye, J., et al. 2015, MNRAS, 453, 721

Wang, Y., Yang, X., Mo, H. J., et al. 2008, MNRAS, 385, 1511

Wang, Y. O., Lin, W. P., Kang, X., et al. 2014a, ApJ, 786, 8

—. 2014b, ApJ, 786, 8

Welker, C., Devriendt, J., Dubois, Y., Pichon, C., & Peirani, S. 2014, MNRAS, 445, L46

West, M. J., Villumsen, J. V., & Dekel, A. 1991, ApJ, 369, 287

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS, 399, 574

Yang, X., van den Bosch, F. C., Mo, H. J., et al. 2006, MNRAS, 369, 1293

5

Galaxy-Galaxy Lensing in EAGLE: comparison with data from 100 square degrees of the KiDS and GAMA surveys

We present predictions for the galaxy-galaxy lensing (GGL) profile from the EAGLE hydro-dynamical cosmological simulation. These predictions are computed at redshift zero, in the spatial range $-1.7 < \log_{10}[r_p/(h^{-1}\mathrm{Mpc})] < 0.3$, and for 6 logarithmically equispaced stellar mass bins in the range $10 < \log_{10}(M_{star}[\,M_\odot]) < 11.8$. We compare them with the observed signal measured using background galaxies imaged by the KiDS survey around spectroscopically confirmed foreground galaxies from the GAMA survey. Overall, the predicted lensing signal is in broad agreement with observations, as expected from the fact that the EAGLE simulation has been calibrated to reproduce the redshift zero galaxy stellar mass function. Exploiting the GAMA galaxy group catalogue, the GGL profiles of central and satellite galaxies are also computed independently but this split is limited to groups with at least five members to minimize contamination by false identification. We find good agreement between the EAGLE predictions and the observations for both central and satellite galaxies. When central and satellite galaxies in groups with at least five members are analyzed jointly, predictions result in a poorer agreement with observations. This stems from the fact that the total GGL profile is a linear combination of central and satellite profiles with the satellite fraction as the linear coefficient and that the satellite fraction in the EAGLE simulation is always lower than that in the GAMA group catalogue. The discrepancy in the satellite fraction may, at least partially, originate from the comparison of a flux limited sample (GAMA) to a volume limited one (EAGLE). As the precision of the measurements is about 10%, we find it important to explore the effect of possible systematics in the stacking procedure. Specifically, we focus on two possible sources of error in assigning stellar masses to simulated galaxies. We assume a random error of 0.1 dex to mimic observational uncertainties in inferring stellar masses and we restrict the definition of stellar mass to the sum of masses of star particles within 30 kpc to mimic the observational caveat that stars in a galaxy's outskirts do not enter into the estimation of the galaxy flux. The inclusion of random errors has a very small effect on the GGL profile, whereas considering only stars within 30 kpc increases the estimated ESD profile.

Velliscig, Cacciato, Schaye, Hoekstra

5.1   Introduction

The connection between observable galaxy properties and the underlying (mostly dark) matter density field is the result of galaxy formation and evolution in a cosmological context and as such it is extensively studied from various complementary perspectives. With the advent of large and homogeneous galaxy surveys (e.g. 2dFGRS, SDSS, CFHTLS, KiDS[1]), the link between the stellar content of galaxies and their dark matter halos can be addressed statistically.

Numerous methods are available to probe the mass of dark matter haloes within the galaxy formation framework : galaxy clustering (see e.g. Jing et al. 1998; Peacock & Smith 2000; Zehavi et al. 2002; van den Bosch et al. 2003; Anderson et al. 2014), abundance matching (see e.g. Vale & Ostriker 2004; Moster et al. 2013; Behroozi et al. 2013) and stacked satellite kinematics (see e.g. Zaritsky & White 1994; Prada et al. 2003; Conroy et al. 2005; More et al. 2011). These methods assume, in various ways, a prior knowledge of galaxy formation theory. They are, therefore, limited in their capacity to produce a stellar mass versus halo mass relation that can serve as a test for the galaxy formation framework itself.

For single galaxies, direct methods for estimating the halo mass are available (see for a recent review Courteau et al. 2014). The rotation curves of spiral galaxies or the velocity dispersion of ellipticals can give estimates of total galaxy mass, albeit at relatively small scales. Furthermore, the light of a galaxy can be lensed into multiple images by another galaxy along the line of sight (strong galaxy-galaxy lensing), providing a measurement of the total projected mass within the Einstein radii of galaxies (Kochanek 1991; Bolton et al. 2008; Collett 2015, and references therein). The masses of groups or clusters of galaxies can be estimated via the dynamics of their satellite galaxies (see e.g. Prada et al. 2003; Conroy et al. 2005), using strong lensing (see e.g. Fort & Mellier 1994; Massey et al. 2010) or X-ray emission (Ettori et al. 2013, and references therein).

For a population of galaxies, galaxy-galaxy weak lensing (see e.g. Brainerd et al. 1996; Wilson et al. 2001; Hoekstra et al. 2004; Mandelbaum et al. 2006; van Uitert et al. 2011; Velander et al. 2014; Viola et al. 2015; van Uitert et al. 2015) offers the possibility to measure directly the average halo mass and therefore represents a viable alternative to test galaxy formation models.

Galaxy-galaxy lensing measures the distortion and magnification of the light of faint background galaxies (sources) caused by the deflection of light rays by intervening matter along the line of sight (lenses). The effect is independent of the dynamical state of the lens, and the projected mass of the lens is measured without any assumption on the physical state of the matter. The gravitational lensing signal due to a single galaxy is too weak to be detected (it is typically 10 to 100 times smaller than the true ellipticity of galaxies). Therefore the galaxy-galaxy signal must be averaged over many lenses.

The link between haloes and galaxies can be studied theoretically with an *ab-initio* approach using Semi Analytical Models (SAMs) and hydrodynamical cosmological simulations. Simulations in particular aim at the direct modelling of the physical processes that are thought to be important for the formation of galaxies, as well as the energetic feedback from supernovae and AGN that is thought to regulate their growth (see Somerville & Davé 2014, for a recent review). However, many of these processes are happening on scales that are unresolved by simulations and as such they must be treated as 'subgrid' physics. A key test for such studies is to reproduce the observed abundances of galaxies as a function of their stellar mass (galaxy stellar mass function; hereafter GSMF), as this is interpreted as the achievement of

---

[1]2dFGRS: http://www.2dfgrs.net/;
SDSS: http://www.sdss.org/;
KIDS: KIlo-Degree Survey, http://www.astro-wise.org/projects/KIDS/;
CFHTLS: http://www.cfht.hawaii.edu/Science/CFHTLS/

a successful mapping between the stellar mass and the halo mass. However, reproducing the GSMF has proven to be extremely challenging for simulations of galaxy formation. Since most of the radiative losses are due to unresolved physics, a novel and useful approach for hydrodynamical simulations is to calibrate the feedback efficiency to reproduce the present-day GSMF. This is the approach adopted by the EAGLE simulation (Schaye et al. 2015; Crain et al. 2015).

Is a simulation that reproduces the GSMF, and thus presumably also the stellar to halo mass relation, also able to reproduce the galaxy-galaxy signal as a function of galaxy stellar mass? This is not a trivial question to ask since, even if the correct stellar mass is assigned to haloes, there are other possible sources of discrepancy. In fact, the radial dependence of the galaxy-galaxy lensing signal depends also on the radial density profile of haloes, as well as the number and position of satellite galaxies within their host haloes. To answer this question, we compute predictions from the EAGLE hydrodynamical cosmological simulation of the galaxy-galaxy lensing profile at redshift zero. These predictions are compared with the observed signal measured using background galaxies imaged by the KiDS survey around spectroscopically confirmed foreground galaxies from the GAMA survey. KiDS (de Jong et al. 2013) is an optical imaging survey with the OmegaCAM wide-field imager (Kuijken 2011) on the VLT Survey Telescope (Capaccioli & Schipani 2011) with exquisite image quality. KiDS overlap with the Galaxy And Mass Assembly GAMA spectroscopic survey (Driver et al. 2011) which provides reliable redshift estimates and reliable group catalogues (for groups with more than four members with stellar masses above the completeness limit of GAMA). The combination of good image quality for shape determination of the source galaxies from KiDS and the spectroscopic redshift information of the lenses from GAMA provides an ideal set-up for galaxy-galaxy lensing measurements. In the rest of the paper we refer to the combination of KiDS and GAMA data with the term KiDSxGAMA. Recent works used the same observations to study the density profiles and masses of galaxy groups and clusters (Viola et al. 2015), as well as the total subhalo mass of satellite galaxies, the average satellite distance from the host halo and the average satellite-to-host mass ratio (Sifón et al. 2015).

This paper is organized as follows. In Section 5.2 we summarize important aspects of KiDSxGAMA, as well as a summary of the ESD measurements from galaxy shapes(§ 5.2.1). In § 5.2.2 we summarize the properties of the simulations employed in this study, the algorithm used to produce the group catalogue from simulations (§ 5.2.2) and the steps taken to measure the galaxy-galaxy lensing signal (§ 5.2.2). In Section 5.3 we report the results for the galaxy-galaxy lensing signal from simulations and the comparison with KiDSxGAMA data for central (§ 5.3.1) and satellites galaxies (§ 5.3.2). In § 5.3.3 we compare the $\Delta\Sigma$ profile for the whole galaxy population against the KiDSxGAMA observations as well as testing the effect of the stellar mass uncertainties on the simulation-observations comparison (§ 5.3.3). We discuss the limitations and the possible future improvements of this study in Section 5.4, we then summarize our findings and conclude in Section 5.5. In Appendix 5.A we test our results against changes in the simulation volume and resolution.

5.2 Methods

5.2.1 KiDSxGAMA

The data used in this paper are obtained from a cross-analysis of two surveys: KiDS and GAMA. KiDS is an ongoing ESO optical imaging survey (de Jong et al. 2013) with the OmegaCAM wide-field imager on the VLT Survey Telescope. When completed, it will cover a total area of 1500 square degrees in four bands (u, g, r, i). KiDS was designed to have both

good galaxy shape measurements and photometric redshift estimates to identify the location of background galaxies. The mean redshift of the sources is $z = 0.55$.

In this paper, we use the initial galaxy-galaxy lensing measurements based on observations from KiDS covering 100 square degrees in all four optical bands. Their data are part of the first and second 'KiDS-DR1/2' data releases to the ESO community, as described in de Jong et al. (2015).

A key feature of the KiDS survey is the overlap with the GAMA spectroscopic survey (Driver et al. 2011) carried out using the AAOmega multi-object spectrograph on the Anglo-Australian Telescope (AAT). The GAMA survey is 98% complete down to r-band magnitude 19.8, and covers $\sim$ 180 square degrees of sky. The available spectroscopy allows reliable identification of galaxy groups (Robotham et al. 2011), which in turn permits a separation between central and satellite galaxies. This distinction will be used extensively throughout the paper. The redshift distribution of GAMA galaxies (median redshift $z \sim 0.25$) is ideal for measurements of the galaxy-galaxy lensing signal.

GAMA group finder

One of the main products of the GAMA survey is the group catalogue, G3Cv7 (Robotham et al. 2011). The group finder is based on a friends-of-friends (FoF) algorithm, which links galaxies based on their projected and line-of-sight proximity. Groups are therefore identified using spatial and spectroscopic redshift information (Baldry et al. 2014) of all the galaxies targeted by GAMA. The linking length has been calibrated using mock data (Robotham et al. 2011; Merson et al. 2013) from the Millennium simulation (Springel et al. 2005). The calibration also ensures that the grouping algorithm reproduces the basic properties, such as the mass, radius and velocity dispersion of FoF groups found in simulations with linking length of $b = 0.2$ and has thus been used as a base for the mock catalogue. The group catalogue has been tested against the mock data and ensures reliable central-satellite distinction against interlopers for groups with 5 or more members, $N_{\mathrm{FoF}} \geq 5$, above the completeness limit of GAMA of approximately $\log_{10}(M_{\mathrm{star}}[\,\mathrm{M_\odot}]) = 8$ (Robotham et al. 2011). Throughout the paper, unless stated otherwise, the galaxy-galaxy lensing signal is only computed for galaxies in groups with 5 or more members, $N_{\mathrm{FoF}} \geq 5$. This selection leaves 18712 out of an initial sample of 58642 galaxies in the overlapping region with the KiDS DR1 and DR2.

Lensing analysis

The galaxy-galaxy lensing measurements are based on the r-band exposures since these yield the highest image quality in KiDS. The images are then processed with the THELI pipeline (optimized for lensing applications, Erben et al. 2013), and galaxy ellipticities are derived using the LENSFIT code (Miller et al. 2007; Kitching et al. 2008; Miller et al. 2013).

The LENSFIT algorithm gives an estimate of the ellipticity (e_1, e_2) with respect to an equatorial coordinate system for every galaxy. Shape measurements are calibrated against a multiplicative bias that arises from the non-linear transformation of the image pixels for galaxies with low signal-to-noise ratio and small sizes (e.g., Melchior & Viola 2012; Refregier et al. 2012; Miller et al. 2013; Viola et al. 2014) using the same method as in Miller et al. (2013). The biases from non-perfect PSF deconvolution, centroid bias and pixel level detector effects are quantified and corrected for using the residual average ellipticity over the survey area. More details can be found in Kuijken et al. (2015) .

For every source-lens pair the measured ellipticity of the source is projected along the

separation of the lens in a tangential (e_+) and cross ($e_\times$) component as:

$$\begin{pmatrix} e_+ \\ e_\times \end{pmatrix} = \begin{pmatrix} -\cos(2\Phi) & -\sin(2\Phi) \\ \sin(2\Phi) & -\cos(2\Phi) \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}, \tag{5.1}$$

where Φ is the position angle of the source with respect to the lens. Each lens-source pair is then assigned a weight

$$\tilde{w}_{ls} = w_s \tilde{\Sigma}_{crit}^{-2}, \tag{5.2}$$

which is the product of the LENSFIT weight w_s, assigned according to the estimated reliability of the source ellipticity (Miller et al. 2007), and $\tilde{\Sigma}_{crit}$ which assigns a weight that is proportional to the lens-source pair distance, effectively down-weighting pairs that are close together and so less sensitive to lensing. The effective critical surface density, $\tilde{\Sigma}_{crit}$, is:

$$\tilde{\Sigma}_{crit} = \frac{4\pi G}{c^2} \int\limits_{z_l}^{\infty} \frac{D_l(z_l)D_{ls}(z_l z_s)}{D_s(z_s)} p(z_s) dz_s, \tag{5.3}$$

where D_l is the angular diameter distance of the lens calculated using the spectroscopic redshift z_l, D_s is the angular diameter distance of the source, $p(z_s)$ is the redshift distribution of the sources, and D_{ls} is the distance between the lens and the source. The distances of the lenses are known from the GAMA spectroscopy whereas for the sources the distances are computed based on the photometric redshifts derived from the KiDS-ESO-DR1/2 *ugri* images in the ESO data release. The the excess surface density, ESD, is then computed in bins of projected distance r_p:

$$\Delta\Sigma(r_p) = \gamma_t(r_p)\tilde{\Sigma}_{crit} = \left(\frac{\sum_{ls} \tilde{w}_{ls} e_+ \tilde{\Sigma}_{crit}}{\sum_{ls} \tilde{w}_{ls}} \right) \frac{1}{1 + K(r_p)} \tag{5.4}$$

where the sum is over all source-lens pairs in the distance bin, and

$$K(r_p) = \frac{\sum_{ls} \tilde{w}_{ls} m_s}{\sum_{ls} \tilde{w}_{ls}} \tag{5.5}$$

is the correction to the ESD profile that takes into account the multiplicative noise bias m_s in the LENSFIT shape estimates. Typically, the value of the $K(r_p)$ correction is around 0.1, largely independent of r_p.

Galaxy-galaxy lensing offers a direct measure of the $\Delta\Sigma$ profile:

$$\gamma_t(r_p)\tilde{\Sigma}_{crit} = \Delta\Sigma(r_p) \equiv \bar{\Sigma}(< r_p) - \Sigma(r_p), \tag{5.6}$$

where $\Delta\Sigma$ is the difference between the surface density averaged within, and measured at, r_p ($\bar{\Sigma}(< r_p)$ and $\Sigma(r_p)$, respectively). This implies that the $\Delta\Sigma$ calculated from simulations using mass surface densities can be directly compared to the observed $\Delta\Sigma$ from weak lensing analysis.

The error on the ESD measurement is then estimated by:

$$\sigma^2_{\Delta\Sigma} = \sigma^2_{e_+} \left(\frac{\sum_{ls} \tilde{w}_{ls}^2 \tilde{\Sigma}_{cr}^2}{(\sum_{ls} \tilde{w}_{ls})^2} \right), \tag{5.7}$$

where $\sigma^2_{e_+}$ is the variance of all source ellipticities combined. The ESD calculated from Eq. 5.4 can be directly compared to the ESD signal calculated from the simulations (see Eq. 5.6).

5.2.2   Simulations

We compare the observations to the hydrodynamical cosmological simulations from the EAGLE project (Schaye et al. 2015; Crain et al. 2015) with a cubic volume of 100 Mpc per side. EAGLE was run using a modified version of the *N*-Body Tree-PM smoothed particle hydrodynamics (SPH) code GADGET-3, which was last described in Springel (2005). The main modifications with respect to GADGET-3 regard the formulation of the hydrodynamics, the time stepping and the subgrid physics. Dark matter and baryons are represented by 2×1504^3 particles, with an initial particle mass of $m_b = 1.2 \times 10^6 \, h^{-1} \, M_\odot$ and $m_{dm} = 6.6 \times 10^6 \, h^{-1} \, M_\odot$ for baryons and a dark matter, respectively. EAGLE was run using the set of cosmological values suggested by the Planck mission $\{\Omega_m, \, \Omega_b, \sigma_8, \, n_s, \, h\} = \{0.307, 0.04825, 0.8288, 0.9611, 0.6777\}$ (Table 9; Planck Collaboration et al. 2013).

EAGLE includes element-by-element radiative cooling (for 11 elements; Wiersma et al. 2009a), pressure and metallicity-dependent star formation (Schaye 2004; Schaye & Dalla Vecchia 2008), stellar mass loss (Wiersma et al. 2009b), thermal energy feedback from star formation (Dalla Vecchia & Schaye 2012), angular momentum dependent gas accretion onto supermassive black holes (Rosas-Guevara et al. 2013) and AGN feedback (Booth & Schaye 2009; Schaye et al. 2015). The subgrid feedback parameters were calibrated to reproduce the present day observed GSMF as well as the sizes of galaxies (Schaye et al. 2015). More information regarding the technical implementation of hydro-dynamical aspects as well as subgrid physics can be found in Schaye et al. (2015).

Halo catalogue

Groups of connected particles are identified by applying the FoF algorithm (Davis et al. 1985) to the dark matter particles using a linking length of 0.2 times the mean interparticle separation. Baryons are then linked to their closest dark matter particle and they are assigned to the same FoF group, if any. Subhaloes in the FoF group are identified using SUBFIND (Springel et al. 2001; Dolag et al. 2009). SUBFIND identifies local minima in the gravitational potential using saddle points. All particles that are gravitationally bound to a local minimum are grouped into a subhalo. Particles that are bound to a subhalo belong to that subhalo only. We define the subhalo center as the position of the particle for which the gravitational potential is minimum. The mass of a subhalo is the sum of the masses of all the particles that belong to that subhalo. The most massive subhalo is the *central* subhalo of a given FoF group and all other subhaloes are *satellites*.

The mass M_{200}^{crit} and the radius r_{200}^{crit} of the halo are assigned using a spherical over-density algorithm centered on the minimum of the gravitational potential, such that r_{200}^{crit} encompasses a region within which the mean density is 200 times the critical density of the universe.

An important aspect for the analysis is that the group finder of EAGLE links particles in real space whereas the GAMA group finder connects members in redshift space. This difference could be particularly important if a large fraction of interlopers were wrongly assigned to groups for GAMA. However, the GAMA group finder was tested against mock catalogues and found to be robust against interlopers for groups with 5 or more members ($N_{FOF}^{GAMA} \geq 5$) above the completeness limit of GAMA of approximately $\log_{10}(M_{star}[\,M_\odot]) = 8$ (Robotham et al. 2011).

Another caveat is that the galaxy sample in EAGLE can be considered to be volume limited whereas the galaxies in GAMA represent a flux limited sample. This can produce differences for the selection of rich groups between simulations and observations, the impact of which will be studied in future work.

Computation of the galaxy-galaxy lensing signal in EAGLE

In section § 5.2.1 we showed that the galaxy-galaxy lensing signal from observations is a direct proxy for the $\Delta\Sigma$ profile. Therefore, in order to compare with the observations, we need to calculate the $\Delta\Sigma$ profiles from EAGLE.

To calculate the surface density of a subhalo, we project onto the $x - y$ plane all the particles within a sphere with radius $2\,h^{-1}$ Mpc centered on the location of the subhalo. We divide the projected radial range into 150 equally spaced bins. At every projected radius r_p, we calculate the surface density within r_p, $\bar{\Sigma}(< r_p)$, as the sum of the mass of all the particles within the projected radius r_p, $M(< r_p)$, divided by the area $A = \pi r_p^2$. The surface density at r_p, $\Sigma(r_p)$, is the mass enclosed in the annulus with inner radius $(r_p - \delta r_p/2)$ and outer radius $(r_p + \delta r_p/2)$ divided by the area $2\pi r_p\delta r_p$, where $\log_{10}\delta r_p = \log_{10}(2[\,h^{-1}\,\mathrm{Mpc}])/150$. We tested different choices for the shape and extension of the projection volume. In principle the lensing signal is affected by the matter between the source and the lens and not only up to a given radial distance. We verified that projecting a cylindrical section around the center of a subhalo instead of a sphere has a small effect on the ESD profile but a large impact on the computation time. We thus opted for the spherical region. We also tested the impact of a using different radii. Specifically, we found that using spheres of $3\,h^{-1}$ Mpc instead of $2\,h^{-1}$ Mpc has a negligible effect on the signal.

Subhaloes are binned according to their stellar mass, calculated as the sum over all stellar particles that belong to the subhalo. The $\Delta\Sigma$ in a given stellar mass bin is then calculated by averaging the $\Delta\Sigma$ profiles of single subhaloes. The statistical errors are calculated using bootstrapping: galaxies in each mass bin are re-sampled 1000 times and the standard deviation of the resulting distribution of mean values of $\Delta\Sigma$ is reported as a proxy for the 1-sigma error.

5.3 Results

In the following we present the results for the excess surface density $\Delta\Sigma$ computed from the simulations (for details see §5.2.2). Galaxies are divided into 6 stellar mass bins ranging from $\log_{10}(M_{star}[\,\mathrm{M_\odot}]) = 10$ to $\log_{10}(M_{star}[\,\mathrm{M_\odot}]) = 11.8$. In the simulations we consider all stellar mass particles bound to a subhalo for the stellar mass determination. We note however that this choice may overestimate the stellar mass content of a galaxy since in observations stars in the galaxy outskirts are often not detectable. We address this caveat in § 5.3.3. The ESD in a given stellar mass bin is the mean value of the $\Delta\Sigma$ of all galaxies in that mass bin. The galaxy center is defined by the position of the particle belonging to the subhalo hosting the galaxy, for which the gravitational potential is minimal. The $\Delta\Sigma$ profile is computed in the simulations using 150 equally spaced logarithmic radial bins up to $2\,h^{-1}$ Mpc.

We compare each prediction from the simulation to the corresponding data from KiD-SxGAMA. We note that we compare results from the EAGLE simulation at $z = 0$ with the ESD of galaxies that have a mean redshift of $z \approx 0.25$. This is expected to have a minor impact on our results, as discussed in Section 5.4. To ensure a fair and robust comparison between predictions and observations (see discussion in § 5.2.2), throughout the paper we consider only galaxies hosted by haloes with 5 or more members with stellar masses above $\log_{10}(M_{star}[\,\mathrm{M_\odot}]) = 8$ unless otherwise specified (see § 5.3.3).

We first present results for central and satellite galaxies separately (see § 5.3.1 and § 5.3.2) and their comparison with observations (see § 5.3.1 and § 5.3.2). We then present the results for both galaxy types combined (§ 5.3.3). This signal can be interpreted as a linear combination of the signal from satellite and central galaxies, where the relative importance of either of the two terms is modulated by the value of the satellite fraction (§ 5.3.3).

$M_{\rm star}$ *	$M_{200}^{\rm crit}\|_{\rm cen}$ **	$M_{200}^{\rm crit}\|_{\rm sat}$ **	$M_{\rm sub}^{\rm cen}$ **	$M_{\rm sub}^{\rm sat}$ **	$M_{\rm sub}^{\rm sat}/M_{200}^{\rm crit}$	$d_{\rm sat}$ ***	$r_{\rm half}^{\rm dm}\|_{\rm cen}$ ***	$r_{\rm half}^{\rm dm}\|_{\rm sat}$ ***	$N_{\rm gal}$	$f_{\rm sat}^{\rm EAGLE}$	$f_{\rm sat}^{\rm GAMA}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
[10.0 − 10.3]	11.96	13.80	12.01	11.18	0.04	575.9	90.2	17.0	586	0.96	0.99
[10.3 − 10.6]	12.17	13.83	12.19	11.46	0.06	608.7	100.4	22.3	498	0.85	0.97
[10.6 − 10.9]	12.42	13.87	12.46	11.80	0.07	806.5	113.0	30.6	375	0.57	0.93
[10.9 − 11.2]	12.71	13.98	12.75	12.23	0.09	805.3	139.7	47.8	241	0.28	0.78
[11.2 − 11.5]	13.11	14.03	13.13	12.48	0.06	624.7	186.4	51.6	106	0.10	0.46
[11.5 − 11.8]	13.51	14.17	13.53	12.81	0.07	615.1	257.0	67.8	32	0.16	0.22

Table 5.1: Values at $z = 0$ of various quantities of interest for each stellar mass bin. From left to right of the columns list: (1) stellar mass range; (2) average value of the host halo mass, $M_{200}^{\rm crit}$, for central galaxies; (3) same as (2) but for satellite galaxies; (4) mean value of the subhalo mass for central galaxies, considering all the particles bound to the subhalo; (5) same as (4) but for satellite galaxies; (6) average ratio between the mass of the satellite subhalo, $M_{\rm sub}$, and the mass of its host halo $M_{200}^{\rm crit}$; (7) average 3D distance between the satellite galaxy and the center of its host halo; (8) mean radius of central galaxies within which half of the mass in dark matter is enclosed; (9) same as (8) but for satellite galaxies; (10) total number of galaxies in the stellar mass bin; (11) average satellite fraction expressed as the total number of satellites divided by the total number of galaxies in the mass bin; (12) average satellite fraction in GAMA.

* $\log_{10}(M/[\,{\rm M_\odot}])$
** $\log_{10}(M/[\,h^{-1}\,{\rm M_\odot}])$
*** $R/[\,h^{-1}\,{\rm kpc}]$

5.3.1   The galaxy-galaxy lensing signal around central galaxies

The left panel of Fig. 5.1 shows the ESD profile around central galaxies in the EAGLE simulation as a function of the projected distance from the center of the galaxy. In all mass bins $\Delta\Sigma$ is a monotonically decreasing function of the projected radius. This is expected since the matter density peaks at the center of the halo where the central galaxy resides. Fluctuations in the surface density profiles can arise due to the presence of satellites, but these are usually not massive enough to significantly alter the azimuthally averaged surface density. Moreover, since the signal is averaged over many galaxies, any deviation due to a single massive satellite would be averaged out in the stacking process.

The right panel of Fig. 5.1 shows the values of $\Delta\Sigma$ at a separation $r_{\rm p} = 0.05\,h^{-1}\,{\rm Mpc}$ (red curve) as a function of stellar mass, normalized to the value in the stellar mass bin $10 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 10.3$. We also report the mean $M_{\rm sub}$ as a function of stellar mass (black curve), normalized to the subhalo mass of galaxies in $10 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 10.3$. Both $\Delta\Sigma(r_{\rm p} = 0.05\,h^{-1}\,{\rm Mpc})$ and the the mean mass $M_{\rm sub}$ are monotonically increasing functions of the stellar mass. The slopes of the two relations are different: $\Delta\Sigma(r_{\rm p} = 0.05\,h^{-1}\,{\rm Mpc})$ shows a weaker dependence on stellar mass (with a slope of 0.6), whereas $M_{\rm sub}$ increases linearly with the stellar mass. This is in line with the fact that haloes do not have single power law matter density profiles but a double power law with the characteristic radius depending on halo mass. Nonetheless, central galaxies with higher $\Delta\Sigma$ amplitudes are hosted by more massive haloes. Therefore, the amplitude of the $\Delta\Sigma$ profile at small scales is a proxy for the typical mass of the subhaloes hosting central galaxies in a given stellar mass bin.

Comparison with observations

Fig. 5.2 shows the $\Delta\Sigma$ signal in EAGLE (red curves) whereas $\Delta\Sigma$ from the observations is indicated with black diamonds. For stellar masses $10 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 10.9$ the uncertainties in the data are large due to the limited number of low mass galaxies that are centrals in rich groups ($N_{\rm FoF} \geq 5$) and therefore not representative of the entire central galaxy population (Viola et al. 2015). For stellar masses $10.6 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 11.5$ the uncertainties are smaller and the radial dependence of the signal is better constrained.

For the three most massive stellar mass bins the normalization and the radial dependence of the signal from EAGLE are consistent with the KiDSxGAMA data. For $11.2 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 11.5$ the observed $\Delta\Sigma$ seems to favour a shallower excess surface density profile at radii larger than $400\,h^{-1}\,{\rm kpc}$. Note, however, that the error bars in EAGLE are asymmetric and correlated; consequently the estimates of $\Delta\Sigma$ appear more significant than they are.

The agreement between the ESD in EAGLE and KiDS suggest that central galaxies, with masses $10.9 < \log_{10}(M_{\rm star}[\,{\rm M}_\odot]) < 11.8$, are hosted in the simulation by subhaloes of approximately the correct mass and the right density profile. This is perhaps not surprising considering that EAGLE was calibrated to reproduce the GSMF and therefore to assign the correct stellar mass to subhaloes. The typical host halo masses predicted by EAGLE for galaxies in the 6 stellar mass bins shown can be found in Table 5.1, column (3). We have computed analytical $\Delta\Sigma$ profiles corresponding to haloes with NFW matter density profiles for the halo masses reported in Table 5.1. These analytical profiles reproduce the overall normalisation of the signal but poorly match the radial dependence of the numerical profiles. At this stage we speculate that this might be due to having assumed an incorrect concentration-halo mass relation and/or having neglected the contribution of baryonic matter to the $\Delta\Sigma$ profile. We defer a systematic analysis of this mismatch to future investigations.
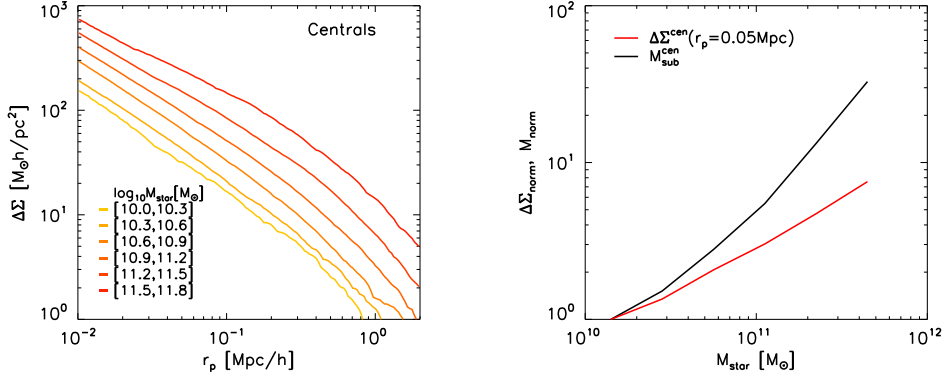
Figure 5.1: *Left panel*: Profiles of the excess surface density, $\Delta\Sigma$, of matter around central galaxies up to projected separations of $2\,h^{-1}$ Mpc from the center of the galaxy. Only galaxies hosted by groups with 5 or more members with stellar mass $\log_{10}(M_{star}[\,M_\odot]) > 8$ are taken into account for this analysis in order to mimic the GAMA selection of galaxies. Central galaxies are divided into 6 stellar mass bins ranging from $\log_{10}(M_{star}[\,M_\odot]) = 10$ to $\log_{10}(M_{star}[\,M_\odot]) = 11.8$. *Right panel*: The values of $\Delta\Sigma$ at a separation $r_p = 0.05\,h^{-1}$ Mpc as a function of stellar mass (red curve), normalized by the value in the stellar mass bin $10 < \log_{10}(M_{star}[\,M_\odot]) < 10.3$. The mean M_{sub} as a function of stellar mass (black curve), normalized by the subhalo mass value for galaxies in $10 < \log_{10}(M_{star}[\,M_\odot]) < 10.3$, is also reported. The slope of the linear function fitting the $\Delta\Sigma$ values at the separation $r_p = 0.05\,h^{-1}$ Mpc is 0.59 whereas for M_{sub} the slope is 1.01.
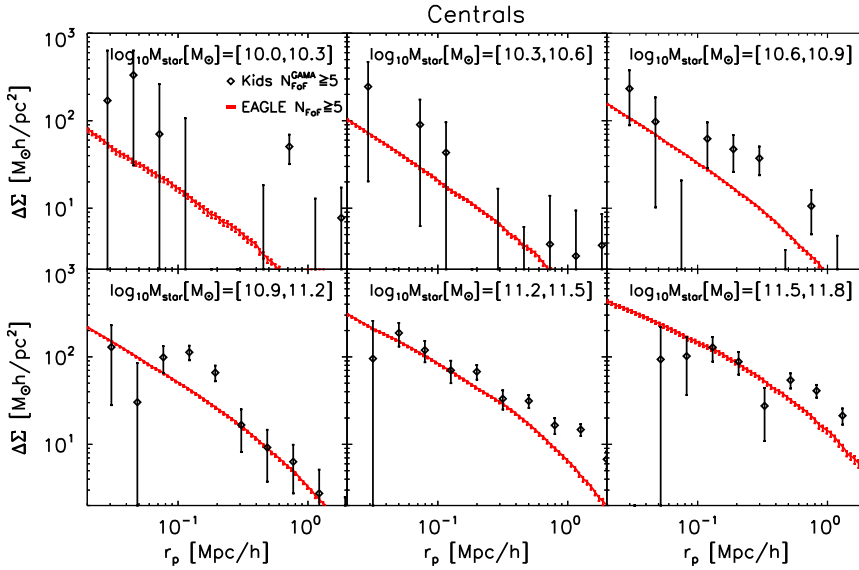
Figure 5.2: Excess surface density profiles in the KiDS survey (black diamonds) and in the EAGLE simulation (red curves) for central galaxies hosted by groups with 5 or more members that each have stellar masses greater than $\log_{10}(M_{star}[\,M_\odot]) = 8$ in order to mimic the GAMA selection of galaxies. Each panel contains a different bin in central galaxy stellar mass. Asymmetric error bars show the 1-σ scatter in each r_p bin.

5.3.2 The galaxy-galaxy lensing signal around satellite galaxies

Unlike central galaxies, the $\Delta\Sigma$ profiles of the satellites galaxies are *not* expected to be monotonically decreasing functions of the separation from the centre. For a single satellite galaxy the profile should become negative at the projected separation from the center of the host halo (Yang et al. 2006). This effect is due to the surface density at the center of the host halo being larger than the mean internal surface density, $\Sigma(r_{\rm p,center}^{\rm halo}) > \bar{\Sigma}(< r_{\rm p,center}^{\rm halo})$. At larger separations than the distance to the host halo, the $\Delta\Sigma$ profile first increases due to the inclusion of the center of the host halo in the term $\Sigma(< r_{\rm p})$, before decreasing again at still larger separations. Stacking the $\Delta\Sigma$ of satellites in a given stellar mass bin averages out the negative parts of the profiles since the distances between each satellite and its host halo are different. On the other hand, the increase in the signal at larger radii is preserved by the stacking; the amplitude of the satellite bump can be used as a proxy for the typical mass of the host haloes in which satellites reside.

The left panel of Fig. 5.3 shows the excess surface density as a function of the projected distance from the center of the satellite galaxy. The small scale ($r_{\rm p} = 0.05 \, h^{-1}$ Mpc) normalization of the $\Delta\Sigma$ profile is an increasing function of the stellar mass of the satellite. The three lowest stellar mass bins show a comparable amplitude of the satellite bump. This similarity can be explained by the fact that the richness cut effectively selects host haloes by mass. In fact, most of the satellites with stellar mass $10 < \log_{10}(M_{\rm star}[\, {\rm M}_\odot]) < 10.9$ reside in host haloes of comparable masses ($\log_{10}(M_{200}^{\rm crit}[\, h^{-1} \, {\rm M}_\odot]) \approx 13.8$; see also Table 5.1, column (3)). The prominence of the satellite bump decreases up to $\log_{10}(M_{\rm star}[\, {\rm M}_\odot]) = 11.2$, a trend that is explained by the fact that the ratio $M_{\rm sub}^{\rm sat}/M_{200}^{\rm crit}$ increases from 0.04 to 0.09 in the considered mass range (see Table 5.1, column 6). For higher stellar mass bins, the relative importance of the satellite bump decreases.

The radius at which the excess surface density profile starts to be dominated by the host halo mass (the satellite bump) increases with stellar mass up to $\log_{10}(M_{\rm star}[\, {\rm M}_\odot]) = 11.2$. This effect is driven by the increasing average distance between satellites and their host haloes raising from $\sim 500 \, h^{-1}$ kpc to $\sim 800 \, h^{-1}$ kpc in the mass range considered (see table 5.1, column (7)).

The right panel of Fig. 5.3 shows $\Delta\Sigma$ for satellite galaxies at a separation $r_{\rm p} = 0.05 \, h^{-1}$ Mpc (blue continuous curve) and at separation $r_{\rm p} = 0.5 \, h^{-1}$ Mpc (blue dashed curve), as a function of stellar mass, normalized by their values in the stellar mass bin $10 < \log_{10}(M_{\rm star}[\, {\rm M}_\odot]) < 10.3$. The mean $M_{\rm sub}$ (black continuous curve) and the host halo mass $M_{200}^{\rm crit}$ (black dashed curve) are also shown as a function of stellar mass.

At smaller separations ($r_{\rm p} = 0.05 \, h^{-1}$ Mpc), the ESD of satellite galaxies increases with stellar mass. The same trend is shared by the average subhalo mass for satellite galaxies since satellites with higher stellar masses are hosted by more massive dark matter subhaloes. As in the case of central galaxies, the similar dependence on the stellar mass suggests that the amplitude of $\Delta\Sigma$ at small separations can be considered a proxy for the subhalo mass hosting the satellite galaxy.

For larger separations ($r_{\rm p} = 0.5 \, h^{-1}$ Mpc), the $\Delta\Sigma$ profile starts to be dominated by the contribution of the halo hosting the satellite galaxy. In this case $\Delta\Sigma$ shares a similar trend with stellar mass as the mean host halo mass for satellite galaxies, $M_{200}^{\rm crit}$. The dependence on stellar mass is remarkably similar for both quantities and shows very little variation up to $\log_{10}(M_{\rm star}[\, {\rm M}_\odot]) = 10.9$ since, as discussed before, the richness cut effectively results in a selection in host halo mass as well. The similar dependence of $\Delta\Sigma$ with halo mass at larger radii highlights the fact that the amplitude of the satellite bump is tightly correlated to the host halo mass. In principle the amplitude of the satellite bump should depend on the satellite's subhalo mass as well as on the host halo mass. In practice the satellite's subhalo mass is
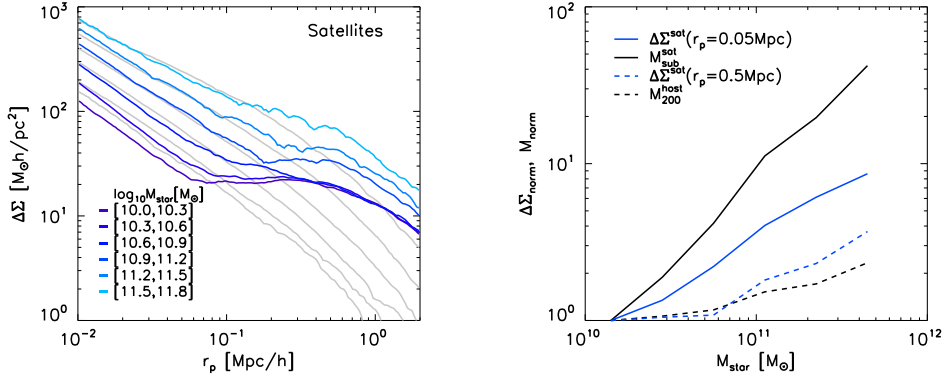
Figure 5.3: As in Fig. 5.1 but for satellite galaxies. To ease the comparison in the left panel, the results for the central galaxies are reported with gray curves. In the right panel the value of $\Delta\Sigma(r_p = 0.5\,h^{-1}\,\mathrm{Mpc})$ is reported as a function of the stellar mass (blue dashed curve) along with the mean host halo mass M_{200}^{crit} (black dashed curve). The slope of a linear fit of the $\Delta\Sigma$ values at a separation $r_p = 0.05\,h^{-1}$ Mpc is 0.65 whereas for M_{sub} the slope is 1.10.

always around $\sim$ 5% of the host halo mass and therefore it plays a minor role in setting the amplitude of the satellite bump. We note that the ratio $M_{\mathrm{sub}}/M_{200}^{\mathrm{crit}}$ is often inferred in N-body simulations or in semi-analytical models of infalling satellite galaxies (e.g. van den Bosch et al. 2005; Jiang & van den Bosch 2014) and it has been recently measured via galaxy-galaxy lensing by Sifón et al. (2015) who report values ranging from 0.005 to 0.015, in agreement with our results.

Comparison with observations

Fig. 5.4 shows the comparison between the observed $\Delta\Sigma$ profile of satellite galaxies in KiD-SxGAMA (black squares) and the corresponding signal in the EAGLE simulations (blue curves) for 6 stellar mass bins. For all stellar masses there is good agreement between simulation and observation in both the predicted normalization of the ESD profile and in the location and the amplitude of the satellite bump.

For stellar masses $10.6 < \log_{10}(M_{\mathrm{star}}[\,\mathrm{M}_\odot]) < 10.9$ the data show a dip in the $\Delta\Sigma$ profile at $r_p \sim 100\,h^{-1}$ kpc that is not present in the simulations. This unreproduced feature can be due to a different radial distribution of satellite galaxies in the simulation and observations. For example, if a large part of the satellites in the observations are located in a very narrow range of distances from their host haloes this can produce a similar dip in the observed signal. On the other hand, it is quite unlikely that satellites of different host haloes are preferentially located at a particular distance from their hosts. The richness is likely to play a role since the dip is less pronounced in the signal computed using the full sample of galaxies without a richness cut (cf. Fig. 5.10). Understanding this feature in the observations would require a comparison of the radial distribution of satellites in both the simulation and observations. We leave this comparison to future work. The amplitude of the signal at small and large radii for this mass bin is well reproduced by the simulation.
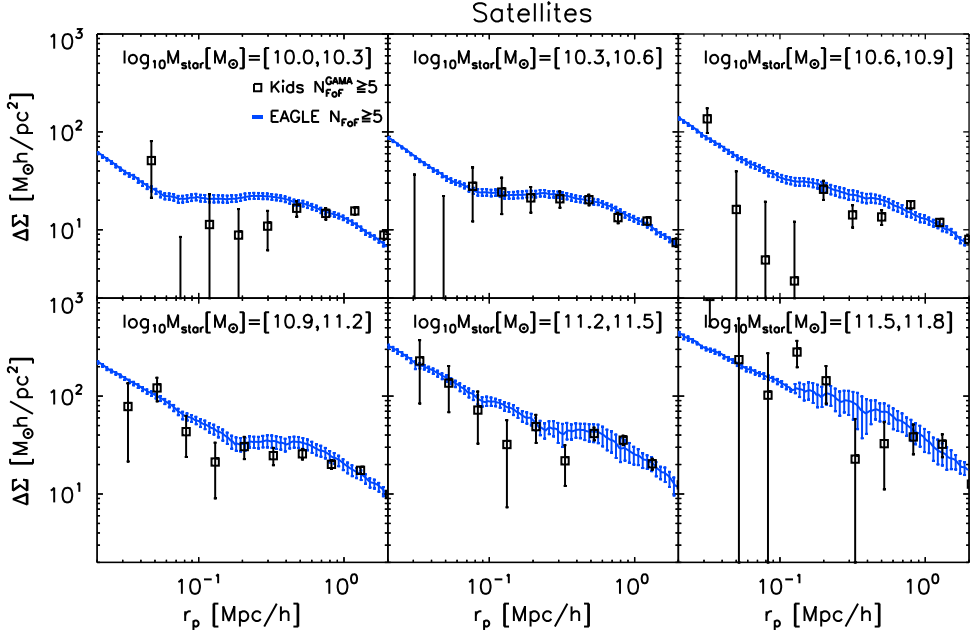
Figure 5.4: As in Fig. 5.2 but for satellite galaxies.

5.3.3   The galaxy-galaxy lensing signal around all galaxies

In this section we present the ESD calculated considering all galaxies without distinguishing between centrals and satellites, while still selecting only rich groups.

The $\Delta\Sigma$ profile of the whole population of galaxies of a given stellar mass is essentially the linear combination of the profiles for satellites, $\Delta\Sigma_{sat}$, and centrals, $\Delta\Sigma_{cen}$,:

$$\Delta\Sigma = f_{sat}\Delta\Sigma_{sat} + (1 - f_{sat})\Delta\Sigma_{cen} \tag{5.8}$$

where f_{sat} is the satellite fraction of galaxies in a given stellar mass bin. The relative importance of either of the two terms is set by the value of f_{sat}.

In order to illustrate more clearly the role of the satellite fraction, in the left panel of Fig. 5.5 we show the ESD profile of galaxies in the stellar mass bin $10.9 < \log_{10}(M_{star}[\,M_\odot]) < 11.2$ for the central (red curve), the satellite (blue curve), and the full galaxy population (black curve).

The $\Delta\Sigma$ profiles for satellites and centrals have similar amplitudes at small radii, which implies that they have comparable central surface densities and subhalo masses. It is clear that the $\Delta\Sigma$ profile of the whole galaxy population lies in between those for satellite and central galaxies. In this mass bin $f_{sat} \sim 28\%$; see Table 5.1 for the tabulated values in the other stellar mass bins.

The right panel of Fig. 5.5 shows the ESD as a function of the projected distance from the center of the galaxy for central and satellite galaxies combined. For stellar masses smaller than $\log_{10}(M_{star}[\,M_\odot]) = 10.9$ the signal is close to that of satellites (compare with Fig. 5.3). This is due to the high satellite fraction ($\geq 60\%$, see Table 5.1). For higher stellar mass bins the relative importance of the satellite component is downweighted by the lower satellite fractions. For $11.2 < \log_{10}(M_{star}[\,M_\odot]) < 11.8$ the satellite fraction is smaller than 20% and
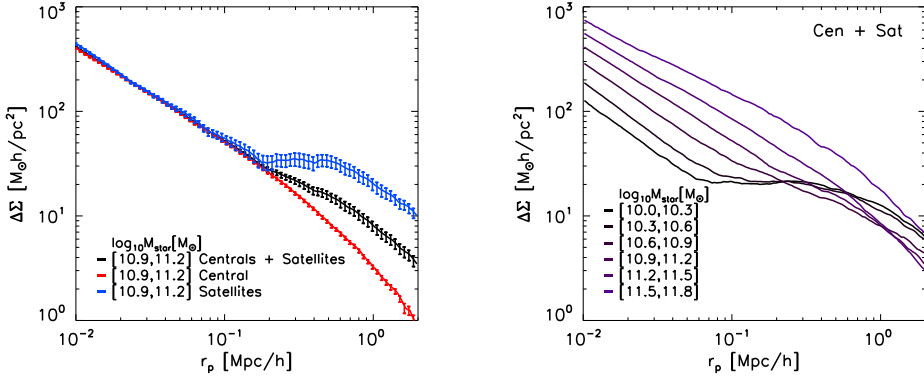
Figure 5.5: *Left panel*: Excess surface density for the central (red curves), the satellite (blue curve), and the full galaxy population (black curve) for a representative stellar mass bin ($10.9 < \log_{10}(M_{\mathrm{star}}[\,M_\odot]) <$ 11.2). Error bars show 1-σ uncertainties calculated with a 1000 bootstrap resampling of the galaxies. *Right panel*: As in the left panel of Fig. 5.1 but for the full galaxy population (i.e. both centrals and satellites).

the total $\Delta\Sigma$ profile is almost completely dominated by the $\Delta\Sigma$ profile of the central galaxies (compare with Fig. 5.1).

Comparison with observations

In the previous section we underlined the importance of the satellite fraction in shaping the $\Delta\Sigma$ signal when both satellites and central galaxies are taken into account. It is of interest to start the comparison with the observational data by considering the satellite fraction in the simulation and in the observations. Fig. 5.6 shows the satellite fraction in EAGLE (black curve) and GAMA (blue filled circles) for groups with 5 or more members, as a function of stellar mass. We remind the reader that, as discussed in § 5.2.2 the group finders used in the simulation and observations are not identical, which could also lead to differences in the satellite-central galaxy classification. Moreover, since the GAMA survey is flux limited, some of the groups that are considered rich in EAGLE could have some of their satellites with stellar mass $\log_{10}(M_{\mathrm{star}}[\,M_\odot]) < 8$, below the GAMA detection limit, thus excluded from the measurements. Due to this effect, groups in GAMA will tend to be systematically richer (especially at the high redshift end), resulting in a higher fraction of satellites with large stellar masses compared to the groups in EAGLE. Hence the satellite fraction at fixed stellar mass is expected to be higher for GAMA, which may potentially account for the discrepancies in the satellite fraction between the simulation and observations (at least partially). We defer a proper investigation of this effect to future work, as one would need to create mock GAMA observations from the EAGLE simulation.

The satellite fraction for stellar masses $\log_{10}(M_{\mathrm{star}}[\,M_\odot]) < 10.9$ is close to one because rich groups are unlikely to have low-mass centrals. With increasing stellar mass the satellite fraction becomes smaller. There is a clear offset between the f_{sat} in simulations and observations. The satellite fraction in GAMA is consistently higher than in EAGLE and the difference is largest for $10.9 < \log_{10}(M_{\mathrm{star}}[\,M_\odot]) < 11.2$. We also tested that the satellite fraction in EAGLE is well converged with resolution and simulation volume (not shown).

Fig. 5.7 compares the $\Delta\Sigma$ profiles obtained from observations using KiDSxGAMA (black
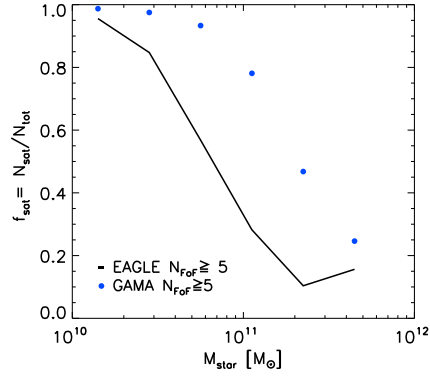
Figure 5.6: Satellite fraction, $f_{\rm sat}$, in EAGLE (black curve) and in KiDS (blue filled circles) for groups with 5 or more members with stellar mass above the stellar mass limit of $\log_{10}(M_{\rm star}[\,{\rm M}_\odot]) = 8$.
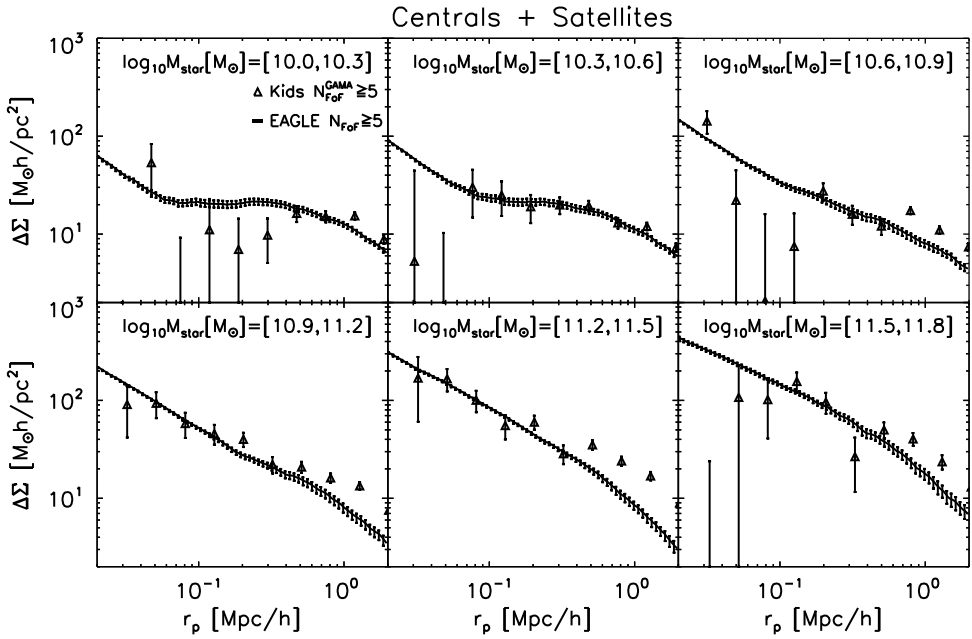

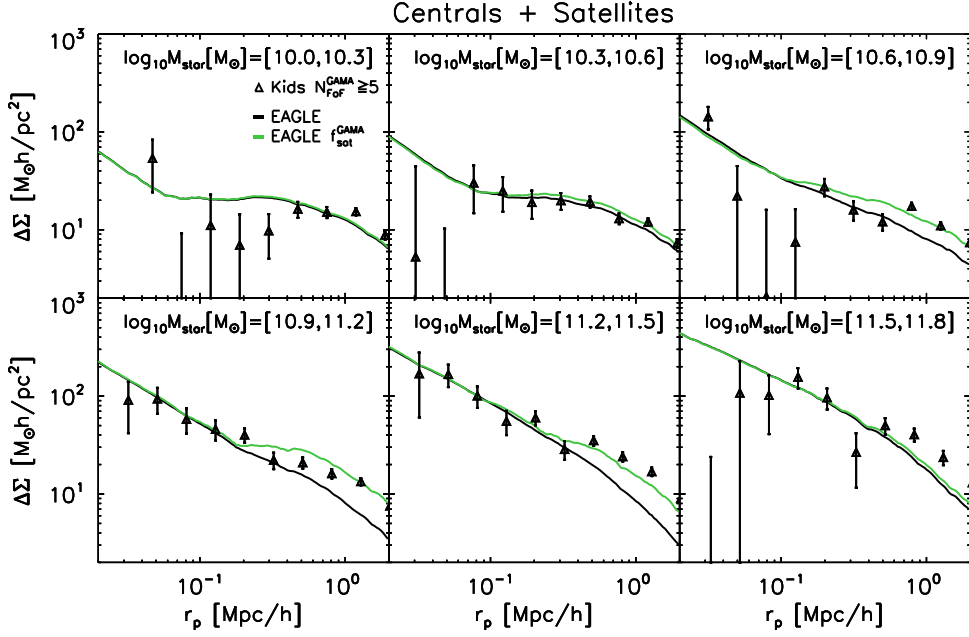Figure 5.7: As in Fig. 5.2 but for the whole galaxy population (i.e. both centrals and satellites).

Figure 5.8: Excess surface density profiles in KidsxGAMA (data points), in EAGLE (black curves) and in EAGLE after rescaling by the satellite fraction from GAMA (green curves).

triangles) and the EAGLE simulations (black curves). Most of the differences between $\Delta\Sigma$ in the simulation and observations are in line with what we expect from our previous results. Specifically, the unreproduced dip at $r_p \sim 100\,h^{-1}$ kpc for the mass bin $10.6 < \log_{10}(M_{star}[\,M_\odot]) < 10.9$ is a consequence of the same feature in the $\Delta\Sigma$ profile of satellites. In the same way, the steepness of the the outer profile for $10.9 < \log_{10}(M_{star}[\,M_\odot]) < 11.8$ arises from the different satellite fraction between EAGLE and GAMA, as $\Delta\Sigma$ for all galaxies is a linear combination of $\Delta\Sigma$ from satellite and central galaxies (see Eq. 5.8). A smaller satellite fraction tends to reduce $\Delta\Sigma$ at large separations (see left panel of Fig. 5.5). Therefore, since the satellite fraction in EAGLE is underestimated, f_{sat} might be responsible for the shallower $\Delta\Sigma$ profiles at large distances ($r_p > 400$ kpc). In order to test this, we rescale $\Delta\Sigma$ from EAGLE by the observed satellite fraction:

$$\Delta\Sigma_{rescaled}^{EAGLE} = f_{sat}^{GAMA}\Delta\Sigma_{sat}^{EAGLE} + (1 - f_{sat}^{GAMA})\Delta\Sigma_{cen}^{EAGLE}, \qquad (5.9)$$

where f_{sat}^{GAMA} is the satellite fraction from GAMA.

Fig. 5.8 shows the rescaled $\Delta\Sigma$ profile for all galaxies (green curves) as well as the original EAGLE profile (black curves) and the observed signal from KiDS (black triangles). The rescaled signal better reproduces the shallow outer radial profile for the stellar mass bin $10.6 < \log_{10}(M_{star}[\,M_\odot]) < 11.5$. On the other hand, for $10.9 < \log_{10}(M_{star}[\,M_\odot]) < 11.2$ the amplitude of the satellite bump is now too high. As expected, the different satellite fractions have no influence on the $\Delta\Sigma$ profile at small radii ($r_p < 100\,h^{-1}$ kpc).

The effect of stellar mass uncertainties in the comparison with observations

In the comparison between simulation and observations an important role is played by stellar mass errors, both random and systematic. We consider here the effect of a random error of ~ 0.1 dex (Behroozi et al. 2013) associated with uncertainties in the stellar mass estimation from broadband photometry. We are not considering here the effect of systematic errors that might arise from different choices in the stellar population synthesis model or in the initial stellar mass function. These errors can be significantly larger ($\sim 0.3 - 0.4$ dex) (Conroy et al. 2009; Behroozi et al. 2010; Pforr et al. 2012; Mitchell et al. 2013) than the random error considered here.

In the case of random errors in the stellar mass estimation, since the number of galaxies decreases with stellar mass, there are always more low-mass galaxies scattered to high masses than vice versa (Furlong et al. 2015). The importance of this effect depends on the steepness of the GSMF. For low mass galaxies $\log_{10}(M_{\rm star}[\,{\rm M_\odot}]) < 10.9$ where the GSMF is reasonably flat, a comparable number of galaxies are scattered towards higher masses than to lower masses. On the other hand, for higher stellar masses where the GSMF is steeper, comparably more low mass galaxies are scattered towards higher masses. Therefore the effect of random errors is expected to be stronger at higher masses.

A potential source of systematic errors is the different definition of stellar mass in the simulation and observations. Throughout the paper we have defined the stellar mass of a galaxy as the sum of the masses of the stars bound to the subhalo hosting the galaxy. However, in observations stars that reside in the outskirts of a galaxy are often undetected and will thus not contribute to the stellar mass. We study this source of systematic errors by redefining stellar masses to include only stars in a given aperture (30 physical kpc, see the discussion in Schaye et al. (2015)). As for the previous case, this effect is expected to be stronger at higher masses, for which the typical extent of a galaxy is greater than 30 kpc.

Fig. 5.9 shows the $\Delta\Sigma$ profile for all galaxies in the case where a random error of 0.3 dex is applied to the stellar masses (purple dotted curves), as well as a case in which only stars within 30 kpc are considered for the computation of the stellar masses (orange dashed curves). For comparison we also show the original EAGLE profiles (black curves) and the observed signal from KiDS (black triangles).

The uncertainties in the stellar mass determinations play a very minor role for all stellar mass bins. The effect of random errors on the $\Delta\Sigma$ profiles is well within the errors on the simulation results (cf. Fig 5.7).

Including only stars within a 30 kpc aperture effectively lowers the stellar mass estimate of galaxies with $10.9 < \log_{10}(M_{\rm star}[\,{\rm M_\odot}]) < 11.8$. This systematic decrease of the stellar mass tends to assign galaxies, previously contributing to higher stellar mass bins, to lower stellar mass bins. Therefore, the average subhalo mass of galaxies in a given stellar mass bin is increased, which in turn increases the amplitude of the average $\Delta\Sigma$ profile in that bin.

We note that, due to the stochasticity of the stellar mass determination, when random errors are included different galaxies contribute to the $\Delta\Sigma$ profile in the different cases.

Comparison with observations without richness cut

So far we have only considered groups with 5 or more members with stellar mass above the stellar mass limit of $\log_{10}(M_{\rm star}[\,{\rm M_\odot}]) = 8$. Galaxies in rich groups are used to ensure a robust classification in central and satellite galaxies in GAMA for an easier interpretation of the signal of both populations combined. It is also of interest to compare the $\Delta\Sigma$ profile for all galaxies not selected by their host group richness.
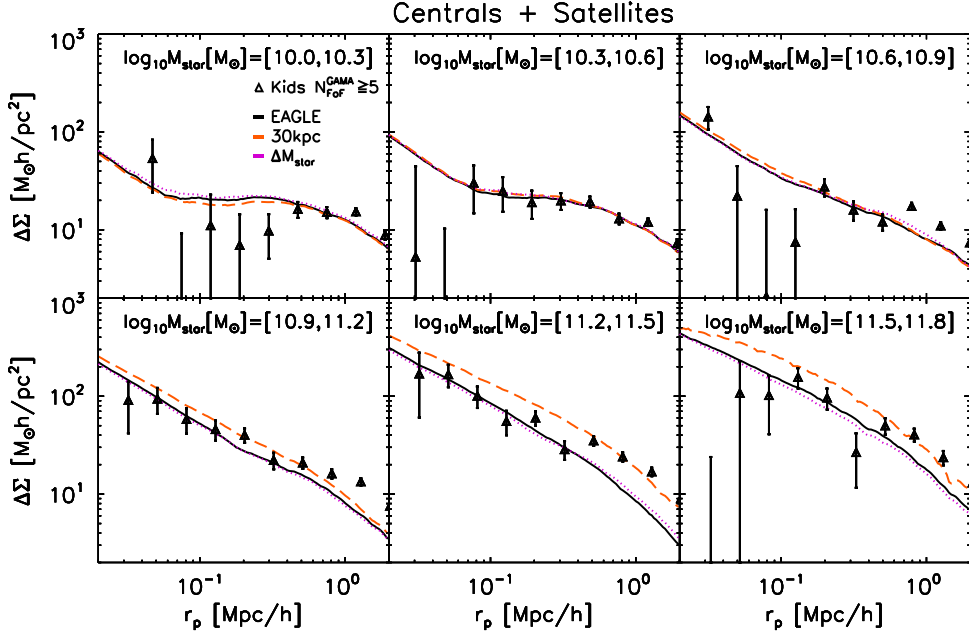
Figure 5.9: As in Fig. 5.7, but showing the $\Delta\Sigma$ profile for galaxies for which a random error in the stellar mass estimation of ~ 0.1 dex is included (purple dotted curves), the $\Delta\Sigma$ profile of galaxies for which only stars within 30 kpc are considered (orange dashed curves), as well as the original EAGLE profile (black curves) and the observed signal from KiDS (black triangles).
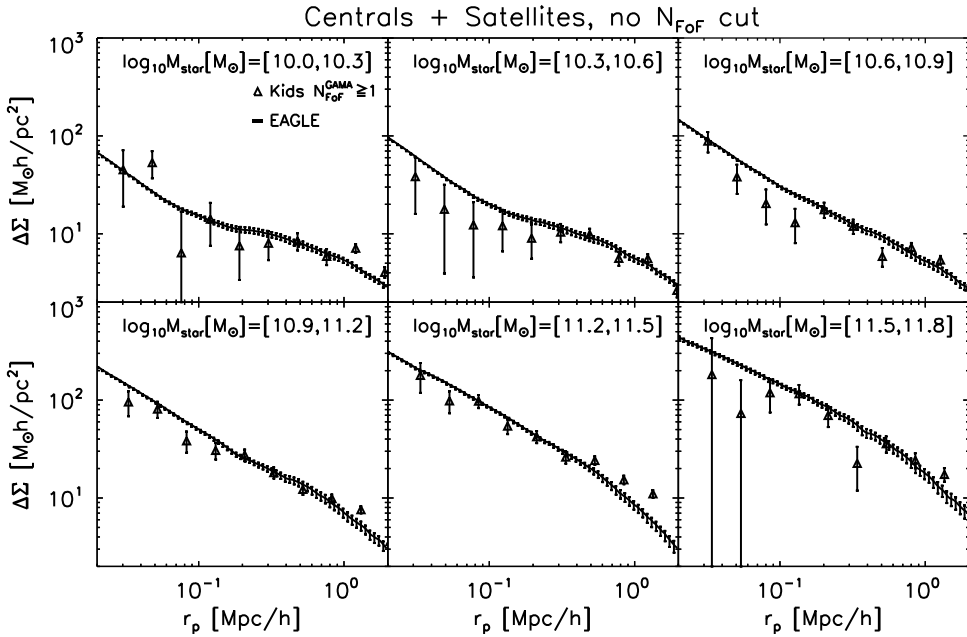
Figure 5.10: As in Fig. 5.7, but without a richness cut in both data and simulations.

Fig. 5.10 shows the $\Delta\Sigma$ profiles in EAGLE and in KiDSxGAMA for all galaxies irrespective of the richness of the group in which they reside. Due to the lower satellite fraction, at lower steller masses the amplitude of the satellite bump is reduced relative to that measured when including the richness cut (see Fig. 5.7). Without the richness cut the differences between EAGLE and the observations are smaller. However, for $10.3 < \log_{10}(M_{star}[\,M_\odot]) < 10.6$ the simulation slightly overestimates the $\Delta\Sigma$ profile, and therefore the masses of the haloes hosting these galaxies, which is consistent with the slight underestimate of the GSMF at these masses shown in Schaye et al. (2015). Also, for $10.6 < \log_{10}(M_{star}[\,M_\odot]) < 10.9$, the dip in the $\Delta\Sigma$ profile is less pronounced, suggesting that the richness cut could play a role in the strength of this feature.

5.4 Discussion

In this section we discuss some of the limitations of our study and we highlight possible future improvements. The main points are:

- The group finder of EAGLE identifies groups in real space whereas the GAMA group finder uses redshift space. This may cause differences in the $\Delta\Sigma$ profile, in particular if interlopers are wrongly assigned to groups, which would artificially increase the richness of the observed group. Therefore, the observed signal would be artificially lowered by the contribution of less massive groups hosting fewer than 5 members. To be fully consistent, the same algorithm should be employed in both simulations and observations.

- The GAMA survey is a flux limited survey whereas the EAGLE simulation can be con-

sidered as a volume limited one. In practice some of the groups that are considered rich in EAGLE could have some of their satellites with stellar mass $\log_{10}(M_{star}[\,M_\odot]) <$ 8, below the GAMA detection limit, thus excluded from the measurements. Due to this effect, groups in GAMA with at least 5 detected members will tend to be systematically richer than EAGLE groups with 5 or more members with stellar mass $\log_{10}(M_{star}[\,M_\odot]) > 8$ (especially at the high redshift end), resulting in a higher fraction of satellites with large stellar masses compared to the groups in EAGLE. Hence the satellite fraction at fixed stellar mass is expected to be higher for GAMA, which may potentially account for the discrepancies in the satellite fraction between the simulation and observations (at least partially). Moreover, this effect could also explain why the agreement with EAGLE improves if no richness cut is applied to the data. We defer a proper investigation of this effect to future work, as one would need to create mock GAMA observations from the EAGLE simulation.

- The centering in observations is done according to the light emitted by the galaxies – the center of a group is defined as location of the Brightest Group Galaxy – whereas in simulations we adopt the minimum of the gravitational potential as the center. Thus any significant misalignment between the center of light and the deepest point in the gravitational potential could cause differences in our results. However Schaller et al. (2015), have shown that in EAGLE the majority of the galaxies (> 95%) have offsets between the center of mass of their stellar and dark matter distribution that are smaller than the simulation's gravitational softening length ($\sim$ 700pc); Therefore, this effect is likely to be unimportant.

- The $\Delta\Sigma$ signal in KiDSxGAMA is calculated around galaxies that are on average at $z \sim 0.25$ whereas we compare the results from the EAGLE simulation at $z = 0$. This discrepancy is mitigated by the fact that from $z = 0.25$ to $z = 0$ there is little evolution in the GSMF (Furlong et al. 2015), but a consistent comparison should use the same redshift for the observations and the model.

- An interesting line of inquiry to better explain some of our results would be to compare the satellite radial distributions in the simulations and observations. This could potentially unveil the source of the unreproduced feature in the $\Delta\Sigma$ profile of satellite galaxies in $10.6 < \log_{10}(M_{star}[\,M_\odot]) < 10.9$.

- In this work we mostly assume that the good agreement between the simulation and observations stems from the ability of EAGLE to reproduce the observed GSMF. A comprehensive study should be made to test how sensitive this agreement is to the level at which the GSMF is reproduced by the simulations. This can be studied by employing the EAGLE models using different subgrid parameters (Crain et al. 2015).

5.5 Conclusions

In this work we compare the excess surface density signal $\Delta\Sigma$ obtained from the state-of-the-art hydrodynamical cosmological EAGLE simulation project to the observed weak galaxy-galaxy lensing signal using sources with accurate measurements from the KiDS survey around spectroscopically confirmed galaxies from GAMA.

We select galaxies in EAGLE that are hosted by groups with more than four members with stellar masses above the completeness limit of GAMA. For this selection the GAMA group catalogue has been tested to be robust against interlopers (Robotham et al. 2011).

Galaxies are divided into six logarithmically equispaced stellar mass bins in the range $10 < \log_{10}(M_{\text{star}}[\,\text{M}_\odot]) < 11.8$.

Thanks to the GAMA group catalogue the observed $\Delta\Sigma$ signal can be calculated independently for satellite and central galaxies. The $\Delta\Sigma$ signal from central galaxies (Fig. 5.1) in EAGLE is composed of a shallower inner part and a steeper outer part. We compare the $\Delta\Sigma$ signal from central galaxies in EAGLE with the observed signal in KiDS. For stellar mass bins for which the uncertainties on the data are small enough that the radial dependence on the signal can be appreciated ($10.6 < \log_{10}(M_{\text{star}}[\,\text{M}_\odot]) < 11.5$), the normalization and the radial dependence of the signal from EAGLE are consistent with the KiDSxGAMA data (Fig. 5.2). This suggests that the average subhalo mass, as well as the density profile of central galaxies in EAGLE, is consistent with observations.

The $\Delta\Sigma$ profiles of satellite galaxies show a deviation from the profiles of central galaxies due to the contribution to the surface density of the mass of the halo hosting the satellite galaxies (Fig. 5.3). Comparing the predicted satellite signal with observations, we find a good agreement with the exception of a notable feature in the $\Delta\Sigma$ profile of satellite galaxies for $10.6 < \log_{10}(M_{\text{star}}[\,\text{M}_\odot]) < 10.9$. The good agreement between data and simulation suggests that the density profile, the subhalo mass and the satellite-host mass ratio of satellite galaxies in EAGLE are consistent with observations.

The $\Delta\Sigma$ signal of the whole population of galaxies is a linear combination of the signals from satellite and central galaxies (left panel of Fig. 5.5) where the multiplicative factor is the fraction of galaxies, f_{sat}, that are satellites. The slope of the signal depends on f_{sat}, on their typical host halo mass and only minimally on the typical subhalo masses of satellites (right panel of Fig. 5.5). This result indicates that galaxy-galaxy lensing has the potential to constrain quantities, such as f_{sat}, that are not strictly dependent on the stellar-halo mass relation.

The differences between observations and the simulation in the steepness of the outer $\Delta\Sigma$ profile (Fig. 5.7), for all galaxies with $10.9 < \log_{10}(M_{\text{star}}[\,\text{M}_\odot]) < 11.8$, originate from the different satellite fractions in EAGLE and GAMA (Fig. 5.6). Indeed, after rescaling the total signal in EAGLE by the observed satellite fraction from GAMA the agreement with the data improves (Fig. 5.8). The discrepancies in the satellite fraction potentially originate from the comparison of a flux-limited sample (GAMA) to a volume-limited one (EAGLE). This effect could explain, at least partially, the differences in the satellite fraction of EAGLE and GAMA.

Including only stars within an aperture of 30 kpc, to account for the caveat that only stars in the inner part of a galaxy are detected in observations, increases the ESD profile (see fig. 5.9) because in this procedure the same stellar masses are obtained for more massive haloes.

Without the richness cut the differences between EAGLE and the observations are somewhat smaller, although some discrepancies are still present (see Fig. 5.10).

We discussed some possible caveats of this study (see Section 5.4), such as the difference between the halo finders of EAGLE and GAMA and the use of $z = 0$ results from EAGLE to compare with $z = 0.25$ median redshift observations of the GAMA galaxies. We argue that these limitations are unlikely to affect our results significantly. We suggest some possible future improvements of this study such as the comparison of the radial distributions of satellites in EAGLE and GAMA as well as a study of the sensitivity of $\Delta\Sigma$ profiles on the level with which the GSMF is reproduced by the simulation.

Bibliography

Anderson, L., Aubourg, É., Bailey, S., et al. 2014, MNRAS, 441, 24

Baldry, I. K., Alpaslan, M., Bauer, A. E., et al. 2014, MNRAS, 441, 2440

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57

Bolton, A. S., Burles, S., Koopmans, L. V. E., et al. 2008, ApJ, 682, 964

Booth, C. M., & Schaye, J. 2009, MNRAS, 398, 53

Brainerd, T. G., Blandford, R. D., & Smail, I. 1996, ApJ, 466, 623

Capaccioli, M., & Schipani, P. 2011, The Messenger, 146, 2

Collett, T. E. 2015, ArXiv e-prints, arXiv:1507.02657

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486

Conroy, C., Newman, J. A., Davis, M., et al. 2005, ApJ, 635, 982

Courteau, S., Cappellari, M., de Jong, R. S., et al. 2014, Reviews of Modern Physics, 86, 47

Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, ArXiv e-prints, arXiv:1501.01311

Dalla Vecchia, C., & Schaye, J. 2012, MNRAS, 426, 140

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371

de Jong, J. T. A., Kuijken, K., Applegate, D., et al. 2013, The Messenger, 154, 44

de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., et al. 2015, ArXiv e-prints, arXiv:1507.00742

Dolag, K., Borgani, S., Murante, G., & Springel, V. 2009, MNRAS, 399, 497

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971

Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545

Ettori, S., Donnarumma, A., Pointecouteau, E., et al. 2013, Space Sci. Rev., 177, 119

Fort, B., & Mellier, Y. 1994, A&A Rev., 5, 239

Furlong, M., Bower, R. G., Theuns, T., et al. 2015, MNRAS, 450, 4486

Hoekstra, H., Yee, H. K. C., & Gladders, M. D. 2004, ApJ, 606, 67

Jiang, F., & van den Bosch, F. C. 2014, ArXiv e-prints, arXiv:1403.6827

Jing, Y. P., Mo, H. J., & Börner, G. 1998, ApJ, 494, 1

Kitching, T. D., Miller, L., Heymans, C. E., van Waerbeke, L., & Heavens, A. F. 2008, MNRAS, 390, 149

Kochanek, C. S. 1991, ApJ, 373, 354

Kuijken, K. 2011, The Messenger, 146, 8

Kuijken, K., Heymans, C., Hildebrandt, H., et al. 2015, ArXiv e-prints, arXiv:1507.00738

Mandelbaum, R., Seljak, U., Kauffmann, G., Hirata, C. M., & Brinkmann, J. 2006, MNRAS, 368, 715

Massey, R., Kitching, T., & Richard, J. 2010, Reports on Progress in Physics, 73, 086901

Melchior, P., & Viola, M. 2012, MNRAS, 424, 2757

Merson, A. I., Baugh, C. M., Helly, J. C., et al. 2013, MNRAS, 429, 556

Miller, L., Kitching, T. D., Heymans, C., Heavens, A. F., & van Waerbeke, L. 2007, MNRAS, 382, 315

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858

Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2013, MNRAS, 435, 87

More, S., van den Bosch, F. C., Cacciato, M., et al. 2011, MNRAS, 410, 210

Moster, B. P., Naab, T., & White, S. D. M. 2013, MNRAS, 428, 3121

Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144

Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, preprint (arXiv:1303.5076), arXiv:1303.5076

Prada, F., Vitvitska, M., Klypin, A., et al. 2003, ApJ, 598, 260

Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951

Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, MNRAS, 416, 2640

Rosas-Guevara, Y. M., Bower, R. G., Schaye, J., et al. 2013, ArXiv e-prints, arXiv:1312.0598

Schaller, M., Robertson, A., Massey, R., Bower, R. G., & Eke, V. R. 2015, ArXiv e-prints, arXiv:1505.05470

Schaye, J. 2004, ApJ, 609, 667

Schaye, J., & Dalla Vecchia, C. 2008, MNRAS, 383, 1210

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Sifón, C., Cacciato, M., Hoekstra, H., et al. 2015, ArXiv e-prints, arXiv:1507.00737

Somerville, R. S., & Davé, R. 2014, ArXiv e-prints, arXiv:1412.2712

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189

van den Bosch, F. C., Tormen, G., & Giocoli, C. 2005, MNRAS, 359, 1029

van den Bosch, F. C., Yang, X., & Mo, H. J. 2003, MNRAS, 340, 771

van Uitert, E., Cacciato, M., Hoekstra, H., & Herbonnet, R. 2015, A&A, 579, A26

van Uitert, E., Hoekstra, H., Velander, M., et al. 2011, A&A, 534, A14

Velander, M., van Uitert, E., Hoekstra, H., et al. 2014, MNRAS, 437, 2111

Viola, M., Kitching, T. D., & Joachimi, B. 2014, MNRAS, 439, 1909

Viola, M., Cacciato, M., Brouwer, M., et al. 2015, ArXiv e-prints, arXiv:1507.00735

Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009a, MNRAS, 393, 99

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009b, MNRAS,
    399, 574

Wilson, G., Kaiser, N., Luppino, G. A., & Cowie, L. L. 2001, ApJ, 555, 572

Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2006, MNRAS, 373, 1159

Zaritsky, D., & White, S. D. M. 1994, ApJ, 435, 599

Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, ApJ, 571, 172

5.A   Convergence tests

In this section we report the effect on the $\Delta\Sigma$ profiles of varying the volume of the simulations keeping the resolution fixed (i.e. the initial particle masses). We make use of a simulation run in a smaller volume with respect to EAGLE, 50^3 Mpc (2×752^3 particles) instead of 100^3 Mpc (2×1504^3 particles). In this way we can isolate the effect of the size of the simulated box from the effect of changing the resolution.

Fig. 5.11 shows the $\Delta\Sigma$ profiles for all galaxies divided into 5 stellar mass bins in the range $10. < \log_{10}(M_{star}[\,M_\odot]) < 11.5$. We change the range relative to that used in the rest of the paper because in the smaller volume simulations there are no galaxies in the stellar mass bin $11.5 < \log_{10}(M_{star}[\,M_\odot]) < 11.8$. The results from the main EAGLE simulation previously employed in this work are presented by black lines whereas the $\Delta\Sigma$ profiles of the smaller volume simulations are shown by green dashed lines.

In all the stellar mass bins the $\Delta\Sigma$ profile to the left of the satellite bump is unaffected by the change in the simulated volume. The amplitude of the satellite bump is always higher in the simulations with the larger volume. This effect is due to the absence of the most massive host haloes of the 100^3 Mpc in the 50^3 Mpc volume which results in a smaller average host halo mass of satellite haloes.

We also tested the effect of varying the volume of the simulations on the satellite fraction and on the stellar mass-halo mass relation, finding very good convergence in both cases. Moreover, we report the results for the effect of resolution on the $\Delta\Sigma$ profile of all galaxies. We make use of the EAGLE simulation run on a smaller volume of 50^3 Mpc (2×752^3 particles) but with the same particle mass of the main EAGLE run, and a simulation run in the same volume 50^3 Mpc (2×376^3 particles) but with a factor of 8 decrease in mass resolution.

Fig. 5.12 shows the $\Delta\Sigma$ signal calculated in 5 stellar mass bins for all galaxies in the 50^3 Mpc volume (green curves) and for the low resolution version of the 50^3 Mpc volume (blue dotted curves). There is no clear trend with resolution apart from an increase of the overall noise of the $\Delta\Sigma$ profiles.
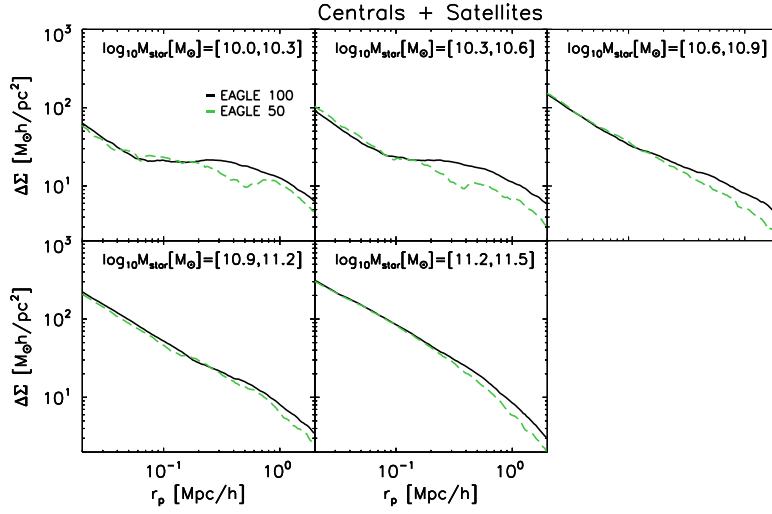
Figure 5.11: ΔΣ profiles for all galaxies in EAGLE (black curves) and in the smaller volume version of EAGLE run in a 50³ Mpc volume with the same particle resolution (dashed green curves). The last stellar mass bin is missing since there are no galaxies that massive in the smaller voloume simulation.
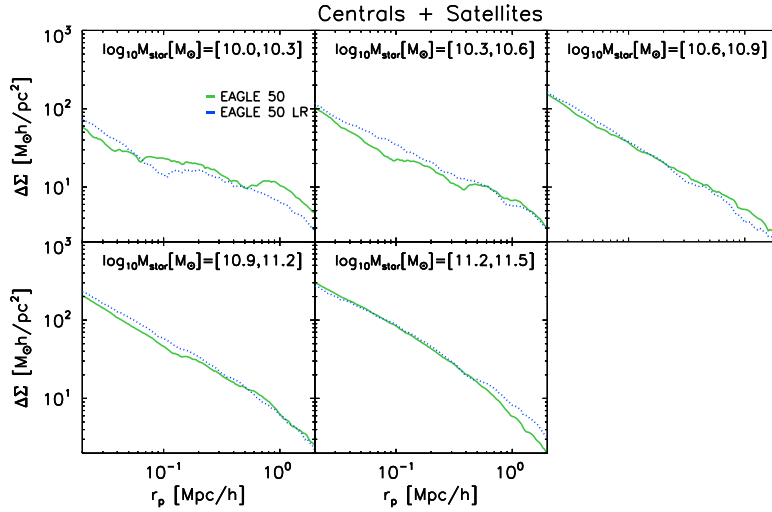

Figure 5.12: ΔΣ profiles for all galaxies in the smaller volume 50³ Mpc version of EAGLE (green curves) and in the low-resolution version of the same simulation (blue dotted curves).

6

Nederlandse Samenvatting

De meeste materie in ons Universum is donker. Deze donkere materie vormt de bouwsteen van de grootschalige, kosmische structuren, waarin sterrenstelsels leven. Door zijn botsingloze natuur is donkere materie namelijk beter in staat structuren te vormen dan normale ("baryonische") materie. Deze structuren bestaan uit vlakken, filamenten en knopen, die samen ook wel het kosmisch web worden genoemd. Sterrenstelsels bewonen de centra van grotere "halo's" van donkere materie. Deze halo's zijn zelf niet zichtbaar en het licht uitgezonden door sterrenstelsels kan ons alleen iets vertellen over het binnendeel van deze halo's.

In dit proefschrift trachten we meer over halo's te weten te komen. Hiertoe maken we gebruik van kosmologische, hydrodynamische simulaties, waarin we niet alleen de donkere maar ook de zichtbare materie meenemen, alsmede alle processen die gedacht worden belangrijk te zijn voor de vorming en groei van sterrenstelsels. Dergelijke simulaties bieden ons de mogelijkheid om het verband tussen zichtbare en donkere materie te verkennen, aangezien beide componenten tegelijk en zelfconsistent worden gesimuleerd.

In waarnemingen kan dit verband onderzocht worden door gebruik te maken van zwaartekrachtlenzen. De werking van dergelijke lenzen is gebaseerd op de afbuiging van fotonen (lichtdeeltjes) wanneer deze door een zwaartekrachtspotentiaal reizen. Zodoende ondervindt licht dat van ver in het heelal naar ons toe reist, onderweg verschillende kleine afbuigingen. Als gevolg hiervan zien wij het beeld van de bron als verplaatst, vergroot en verstoord. Het zwaartekrachtlenseffect kan gebruikt worden om verschillende eigenschappen van (materie in) het Universum te meten, waaronder de totale massa en het massaprofiel van halo's, de vormen van halo's, de efficiëntie van de vorming van sterrenstelsels en uiteindelijk ook de fundamentele kosmologische parameters van ons Universum. Door gebruik te maken van kosmologische, hydrodynamische simulaties kunnen we ook mogelijke effecten onderzoeken die ons ervan weerhouden om zwaartekrachtlenswerking te gebruiken om de fundamentele eigenschappen van de structuren waaruit ons Universum is opgebouwd, te meten.

6.1 Dit Proefschrift

Aangezien de meeste materie donker is en donkeremateriedeeltjes niet botsen, zou het mogelijk moeten zijn om met uitsluitend deeltjes die alleen maar zwaartekracht ondervinden het Universum te simuleren en tegelijkertijd toch ruwweg de juiste eigenschappen van halo's en de grootschalige structuren af te leiden. Dergelijke simulaties noemen we "N-body" simulaties. Voor veel doeleinden is de achterliggende aanname dat andere fysische processen betrokken bij de vorming van sterrenstelsels de verdeling van materie nauwelijks beïnvloeden redelijk. Deze benadering is bijvoorbeeld geldig aan de rand van clusters van sterrenstelsels, waar het aanwezige gas zeer langzaam koelt en ongeveer de verdeling van de donkere materie volgt. Echter, op - kosmisch gezien - kleine schaal en in minder massieve halo's, waar gas hoge dichtheden kan bereiken door koeling, kunnen baryonische processen zoals galac-

tische winden de verdeling van materie wel degelijk beïnvloeden. Dit verandert indirect ook de verdeling van de donkere materie, die zichzelf bijstelt naar de aangepaste zwaartekrachts- potentiaal.

In **Hoofdstuk 2** vergelijken we de eigenschappen van halo's in N-body simulaties met die van dezelfde halo's in hydrodynamische simulaties. We kijken specifiek naar het effect van baryonische processen die nodig zijn voor de vorming en groei van realistische sterrenstelsels, op de massa's en profielen van halo's, en als gevolg daarvan op de hoeveelheid halo's per een- heid volume als functie van massa. We laten zien dat de uitdrijving van gas veroorzaakt door het exploderen van sterren als supernovae, en de daarop volgende uitbreiding van de donkere materie, de massa's van halo's kunnen verlagen, in vergelijking met dezelfde halo's gesimuleerd met alleen zwaartekracht. Dit effect kan oplopen tot 20% voor halo's met sterren- stelsels kleiner dan de Melkweg. Voor halo's met zwaardere sterrenstelsels tot aan clusters van sterrenstelsels begint het effect van actieve kernen (AGN) te domineren. De aangroei van gas in deze kernen en de daarop volgende krachtige uitstoting van energie in de vormen van jets kan grote hoeveelheden gas verplaatsen en voorkomen dat kouder gas aan de rand van het sterrenstelsel naar het centrum koelt. Baryonische fysica verandert ook de totale massa- profielen van halo's tot ver buiten wat doorgaans als de rand van halo's wordt beschouwd. Dit effect kan alleen worden gemodelleerd door de algemeen gebruikte profielen van halo's aan te passen. De afname van de totale massa van de halo veroorzaakt een afname in de relatieve hoeveelheid halo's, voor een gegeven massa, van ongeveer 20%. De analyse die we presen- teren in **Hoofdstuk 2** toont aan dat baryonische processen de eigenschappen van donkere- materiehalo's significant kunnen wijzigen. Het is daarom essentieel om dergelijke processen mee te nemen in simulaties die een theoretische onderbouwing trachten te geven aan waarne- mingen van de totale massa van halo's. Dit resultaat impliceert sterk dat men kosmologische, hydrodynamische simulaties nodig heeft om de eigenschappen van halo's te onderzoeken.

De zwaartekrachtlenswerking biedt ons een manier om donkerematerihalo's te detecteren in waarnemingen. Binnen deze context dienen kosmologische hydrodynamische simulaties als een hulpmiddel om mogelijke beperkingen van deze methode te bestuderen en te ver- helpen. Een van de mogelijke valkuilen bij het meten van de vormen van halo's door middel van het zwaartekrachtlenseffect is de discrepantie tussen de oriëntaties van de halo's en de sterrenstelsels die ze huisvesten. De vorm van de halo van een sterrenstelsel op de voorgrond wordt bepaald door te kijken naar hoe de afbuiging van het licht van sterrenstelsels op de achtergrond verschilt langs de lange of de korte as van het voorgrondstelsel. Omdat dit effect te zwak is om direct te meten voor een enkel sterrenstelsel, analyseert men het signaal voor vele sterrenstelsels tegelijk. Dit kan door alle sterrenstelsels op de voorgrond zo te oriënteren dat hun lange assen uitgelijnd zijn, onder de aanname dat deze overeenkomen met de lange assen van hun halo's. Hierdoor kan een eventuele discrepantie tussen de oriëntaties van halo's en hun sterrenstelsels de meting van de vorm negatief beïnvloeden.

Deze discrepantie is verder interessant om te bestuderen om meer te weten te komen over de intrinsieke correlaties tussen de vormen en oriëntaties van sterrenstelsels, tegenover de extrinsieke correlaties veroorzaakt door het zwaartekrachtlenseffect. De meeste theorieën in dit veld doen voorspellingen omtrent deze intrinsieke correlaties voor donkere materie. Om deze voorspellingen te kunnen vergelijken met waarnemingen moet het verband tussen de oriëntaties van sterrenstelsels en hun halo's dus bekend zijn. De intrinsieke correlaties tussen de vormen en oriëntaties van sterrenstelsels zijn een last voor onderzoeken gebaseerd op het zwaartekrachtlenseffect en moeten achterhaald worden voordat ze in acht genomen kunnen worden.

In **Hoofdstuk 3** bestuderen we de correlaties tussen de oriëntaties van sterrenstelsels en hun halo's. We maken hiervoor gebruik van vier hydrodynamische simulaties met verschil-

lende volumes om zo vier ordes van magnitude in massa te kunnen beslaan, van sterrenstelsels tot aan clusters van sterrenstelsels. Onze bevindingen zijn dat de oriëntaties van sterrenstelsels als geheel goed overeenkomen met de locale materieverdeling, maar dat de oriëntatie van de verdeling van sterren een significante discrepantie vertoont met die van hun halo's. Deze discrepantie is kleiner in de meest massieve halo's. Dit impliceert dat de oriëntatie van sterrenstelsels alleen op vergelijkbare schaal een goede indicator is van de oriëntatie van de donkere materie. Op grotere schaal is er een significante discrepantie tussen de twee oriëntaties, waar onderzoeken die gebruik maken van de (zwakke) zwaartekrachtlenswerking rekening mee moeten houden.

Zoals eerder vermeld kunnen metingen via het zwaartekrachtlenseffect negatief beïnvloed worden door de intrinsieke correlatie tussen de vorm en oriëntatie van sterrenstelsels, wat men de intrinsieke uitlijning van sterrenstelsels noemt. De verstoringen veroorzaakt door getijdenkrachten (die een *intrinsieke* uitlijning teweeg brengen) zijn niet noodzakelijk verwaarloosbaar in vergelijking met de verstoringen door het lenseffect (vaak de *schijnbare* uitlijning genoemd). In **Hoofdstuk 4** doen we verslag van de intrinsieke uitlijning van sterrenstelsels zoals gemeten in hydrodynamische, kosmologische simulaties. We richten ons specifiek op de correlatie tussen de oriëntatie van lange as van een sterrenstelsel en de richting van een nabij sterrenstelsel (oriëntatie-richting) en de correlatie tussen de oriëntaties van de lange assen van beide stelsels (oriëntatie-oriëntatie). We vinden dat deze correlaties statistisch gezien zwakker zijn als de sterrenstelsels verder uit elkaar staan. Er is echter een kleine, maar significante, correlatie als deze afstand 100 maal zo groot is als de typische grootte van de halo's. De correlaties nemen significant af als we de oriëntaties berekenen met waardes voor de groottes van de sterrenstelsels die beter observationeel te verantwoorden zijn. Dit verschil kan verklaren waarom voorgaande onderzoeken vaak sterker gecorreleerde uitlijningen van sterrenstelsels rapporteren in simulaties dan in waarnemingen. Men moet vooral voorzichtig zijn met de definitie die men kiest voor de grootte van de sterrenstelsels in simulaties in vergelijking tot waarnemingen, aangezien dit kan resulteren in sterk uiteenlopende correlaties. De oriëntatie-oriëntatie correlatie is altijd zwakker dan de oriëntatie-richting correlatie, wat impliceert dat de uitlijning van sterrenstelsels voornamelijk bepaald wordt door de posities van nabije halo's, waar de gezamenlijke oriëntaties van nabije sterrenstelsels slechts een indirect gevolg van is.

Door via de zwaartekrachtlenswerking de typische halomassa's van een collectie sterrenstelsels te meten, kunnen de massa's en massaprofielen van halo's aan hun sterinhoud verbonden worden. Dit kan uiteindelijk leiden tot nieuwe inzichten betreffende de efficiëntie van de vorming van sterrenstelsels, wat bepaalt hoeveel sterren vormen in een halo van een gegeven massa. In **Hoofdstuk 5** doen we verslag van ons onderzoek naar de zwaartekrachtlenswerking tussen sterrenstelsels in de hydrodynamische, kosmologische simulatie EAGLE. We vergelijken de resultaten van simulaties met het waargenomen signaal, dat afgeleid is uit data van KiDS en GAMA. De groepscatalogus van GAMA biedt ons de mogelijkheid om onderscheid te maken tussen de aandelen van centrale stelsels en satellietstelsels aan het totale signaal. Het voorspelde lenssignaal van EAGLE komt in het algemeen goed overeen met de waarnemingen. We vinden een goede overeenkomst voor zowel de centrale stelsels als de satellietstelsels. Als we beide groepen sterrenstelsels los van elkaar analyseren, verslechtert deze overeenkomst. Dit komt doordat het totale lenssignaal een lineaire combinatie is van het signaal van de centrale stelsels en dat van de satellietstelsels met het relatieve aantal satellietstelsels als coëfficiënt. Dit relatieve aantal blijkt altijd lager te zijn in de EAGLE simulatie dan in de GAMA groepscatalogus.

Publications

1. *The impact of galaxy formation on the total mass, mass profile and abundance of haloes*
   Velliscig, Marco; van Daalen, Marcel P.; Schaye, Joop; McCarthy, Ian G.; Cacciato, Marcello; Le Brun, Amandine M. C.; Dalla Vecchia, Claudio
   Monthly Notices of the Royal Astronomical Society, Volume 442, Issue 3, p.2641-2658 (Chapter 2)

2. *The alignment and shape of dark matter, stellar, and hot gas distributions in the EAGLE and cosmo-OWLS simulations*
   Velliscig, Marco; Cacciato, Marcello; Schaye, Joop; Crain, Robert A.; Bower, Richard G.; van Daalen, Marcel P.; Dalla Vecchia, Claudio; Frenk, Carlos S.; Furlong, Michelle; McCarthy, I. G.; Schaller, Matthieu; Theuns, Tom
   Monthly Notices of the Royal Astronomical Society, Volume 453, Issue 1, p.721-738 (Chapter 3)

3. *Intrinsic alignments of galaxies in the EAGLE and cosmo-OWLS simulations*
   Velliscig, Marco; Cacciato, Marcello; Schaye, Joop; Hoekstra, Henk; Bower, Richard G.; Crain, Robert A.; van Daalen, Marcel P.; Furlong, Michelle; McCarthy, I. G.; Schaller, Matthieu; Theuns, Tom
   Monthly Notices of the Royal Astronomical Society, in press (Chapter 4)

4. *The clustering of baryonic matter. II: halo model and hydrodynamic simulations*
   Fedeli, Cosimo; Semboloni, Elisabetta; **Velliscig, Marco**; Van Daalen, Marcel; Schaye, Joop; Hoekstra, Henk
   Journal of Cosmology and Astroparticle Physics, Issue 08, article id. 028, pp. (2014)

Curriculum Vitae

I was born on January 11th, 1985, in Udine, a town in the center of the beautiful Italian region of Friuli where the sea and the mountains meet. I attended the Istituto Tecnico Industriale "Arturo Malignani," one of the most well known schools in the region, with a specialisation in Electronics and Telecommunications. Even though interested in technology, I always had a deep curiosity for science, especially astrophysics. This curiosity, which later became a passion and then a profession, was surely sparked by reading many science fiction novels and watching popular Italian TV shows such as "Super Quark" with Piero Angela and Paco Lanciano.

In 2004, after my high school studies and faced with a decision for university, I decided to switch to physics to explore this passion further. I enrolled in the Bachelor of Physics program at the University of Trieste. For my Bachelor thesis, I worked on PAMELA (Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics) data analysis for 3 months during spring-summer 2008 in collaboration with the INFN (Italian National Institute for Nuclear Physics) and Dr. Emiliano Mocchiutti. That was my first hands-on experience with what research was really about, and I enjoyed it.

In 2008, I started my Masters in Astrophysics and Space Science. I concluded my thesis with Prof. Stefano Borgani and Dr. Luca Tornatore on "The role of Diffusion in Lagrangian SPH codes." After obtaining my degree with honours, I was ready to leave the city where I lived for almost 7 years, and so I applied to PhD programs in Germany and in the Netherlands.

In 2011, I was offered a PhD position at Leiden Observatory working with Prof. Joop Schaye on the shape and alignment of galaxies in cosmological hydrodynamical simulations. I presented my work at conferences and workshops in the Netherlands, Italy, the UK and in the US. I also attended the ICTP summer school in cosmology in July 2012 and the 6th Heidelberg Summer School on Computational Astrophysics in September 2012. Besides my research work, during my years at the Sterrewacht I served as a teaching assistant for the course "The formation of galaxies and large-scale structure" taught by Dr. Jarle Brinchmann. I also had the opportunity to organise the Phd Talks in 2014 as well as the Borrel in 2013.

Acknowledgments

Experience is not what happens to a man;
it is what a man does with what happens to him.
(Aldous Huxley)

Even if sometimes during my PhD I felt alone, I really never was, thanks to the people that shared this experience with me.

First, I want to thank the PhD candidates in my year that were a real support to me and with whom I shared both the good and the bad parts of these four years: Alex South for our Borrel organization together. I will always remember the brainstorming about what beer to buy and the headaches caused by squaring the accounts of the Borrel; Heather, no matter what you can always find a way to laugh and to speak your mind; Andra, even if you were always travelling somewhere you made the times when you were here even more special; Monica, sharing an office with you was fun in so many ways and I was lucky to have someone like you to speak to less than half a meter away; Alex North, you were always kind and helpful beyond the call of duty, no matter if it was for a scientific discussion, proof reading, or just borrowing a ladder.

Then I want to thank the "older" PhD students that were already here when I arrived and that made me feel immediately welcome and at home in this country: Matteo, I could always count on you for a serious talk, and even if we were having different experiences, I was sure that you would understand me; Irene, Irenita linda, thank you for your infinite moral support, for the spritz at your place, my new plants, and mostly for just being there; Berenice, thank you for the discussion about science fiction, the good advice on books, and the brownie recipe (even if I think it was from Jesus); Nicola, thank you for your movie advice and your over the top storytelling; Silvia, thank you for making me realise that I have a "better green" in me; Thanja (and Daan), thank you for your vegetarian cuisine and for your meals at the vrijplaats.

I also want to thank my group members that made all the after group meetings special: Marcel, even if only one year ahead in the PhD process, I always looked up to you and you always had time and good advice for me. It was a pleasure and an honour to be your paranymph; Rob, we truly miss you in the group. Thank you for all the discussions and the support. I really appreciated it; Marijke, thank you for the multiple discussions on Dutch culture, and among other things, for being the buienalarm personified and for your help with the Dutch summary; Joki, thank you for asking me "How are you?" as you really mean it and not only as a formality; Chris, thank you for taking care of my plants, for proofreading my last thesis chapter on such a short notice, and for introducing me to smart phones; Lorrie, thank you for the help with the last bit of the proofreading for the thesis. I'm sorry we can not share the office for longer, it would have been fun; Sylvia, thank you for bringing new enthusiasm in the progress report meetings.

Then I want to thank some people that do not fit in the aforementioned categories: Emanuele thank you for your weird stories, your gossiping, your loudness, and in general for bringing something "different" to the corridor; Alessandra and Fabio, thank you for the discussion