WHAT DO SIMULATIONS PREDICT FOR THE GALAXY STELLAR MASS FUNCTION AND ITS EVOLUTION IN DIFFERENT ENVIRONMENTS?

Benedetta Vulcani1, Gabriella de Lucia2, Bianca M. Poggianti3, Kevin Bundy4, Surhud More4, and Rosa Calvi4

1 Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, The University of Tokyo, Kashiwa 277-8582, Japan; benedetta.vulcani@impu.jp
2 INAF-Astronomical Observatory of Trieste, I-34143 Trieste, Italy
3 INAF-Astronomical Observatory of Padova, I-35122 Padova, Italy
4 Astronomical Department, Padova University, I-35122 Padova, Italy

1 INTRODUCTION

The Λ cold dark matter (CDM) hierarchical paradigm describes well the formation of large-scale structures in the universe. It is, however, difficult to explain all of the observed trends for galaxy populations in the context of this paradigm. This is, at least in part, due to the difficulties in treating relevant physical processes (e.g., feedback from supernovae and active galactic nuclei (AGNs), stellar winds, etc.) whose understanding is far from complete.

From an observational point of view, many studies have tried to quantify the role of the environment in shaping the physical properties of galaxies at different cosmic epochs. In general, denser environments host larger fractions of early-type galaxies in the field (meant as a wide portion of the sky, including all environments), in the local universe (z ∼ 0.06), and at intermediate redshift (z ∼ 0.6), with the aim to shed light on the processes which regulate the mass distribution in different environments. While the mass functions in the field and in its finer environments (groups, binary, and single systems) are well matched in the local universe down to the completeness limit of the observational sample, the model overpredicts the number of low-mass galaxies in the field at z ∼ 0.6 and in clusters at both redshifts. Above $M_* = 10^{10.25} M_\odot$, it reproduces the observed similarity of the cluster and field mass functions but not the observed evolution. Our results point out two shortcomings of the model: an incorrect treatment of cluster-specific environmental effects and an overefficient galaxy formation at early times (as already found by, e.g., Weinmann et al.).

Next, we consider only simulations. Also using the Guo et al. model, we find that the high-mass end of the mass functions depends on halo mass: only very massive halos host massive galaxies, with the result that their mass function is flatter. Above $M_* = 10^{9.4} M_\odot$, simulations show an evolution in the number of the most massive galaxies in all environments. Mass functions obtained from the two prescriptions are different, however, results are qualitatively similar, indicating that the adopted methods to model the evolution of central and satellite galaxies still have to be better implemented in semi-analytic models.

Key words: galaxies: evolution – galaxies: formation – galaxies: general – galaxies: luminosity function, mass function

Online-only material: color figures

1. INTRODUCTION

The Λ cold dark matter (CDM) hierarchical paradigm describes well the formation of large-scale structures in the universe. It is, however, difficult to explain all of the observed trends for galaxy populations in the context of this paradigm. This is, at least in part, due to the difficulties in treating relevant physical processes (e.g., feedback from supernovae and active galactic nuclei (AGNs), stellar winds, etc.) whose understanding is far from complete.

From an observational point of view, many studies have tried to quantify the role of the environment in shaping the physical properties of galaxies at different cosmic epochs. In general, denser environments host larger fractions of early-type galaxies (Hubble & Humason 1931), which are typically more massive, redder, more concentrated, less gas-rich, and show lower star formation rates (SFRs) than late-type galaxies (e.g., Kauffmann et al. 2004; Baldry et al. 2006; Weinmann et al. 2006a). These trends may be driven by the environment and be the result of physical processes coming into play only after galaxies have become part of a structure, like a group or cluster, or they might be mainly driven by intrinsic properties closely related to galaxy-intrinsic conditions (i.e., stellar mass) and be established beforehand, due to the fact that galaxy formation occurs at an accelerated rate in overdense regions. As discussed in De Lucia et al. (2012), separating the two scenarios and differentiating their role in driving galaxy evolution is hard, since they are strongly and physically connected.

In recent years, much attention has been focused on quantifying the environmental trends at fixed stellar mass (but see caveats pointed out in De Lucia et al. 2012). In this paper, we will focus on one specific property of galaxy populations, that is, the galaxy stellar mass function. Since it results from a combination of hierarchical mass assembly of dark matter (DM) halos and different physical processes driving galaxy evolution (e.g., Dekel & Silk 1986; Benson et al. 2002; Wang et al. 2008), comparing observational data at different cosmic epochs with theoretical predictions can provide important constraints on the entire galaxy formation processes.

The first studies focused mainly on field galaxies (e.g., Fontana et al. 2004, 2006; Drory et al. 2005; Gwyn & Hartwick 2005; Borch et al. 2006; Bundy et al. 2006; Pozzetti et al. 2007, 2010; Ilbert et al. 2010, 2013; Baldry et al. 2012; Muzzin et al. 2013). They showed that the number density of galaxies with $M_* \geq 10^{11} M_\odot$ exhibits relatively modest evolution from $z = 1$ to $z = 0$. This implies that the assembly of relatively massive objects is essentially complete by $z \sim 1$. On the other hand, the mass function of less massive galaxies evolves more strongly than that of massive ones, displaying a rapid increase of their number density from $z \sim 1$ (but see Drory et al. 2009).

As pointed out by Marchesini et al. (2009), however, a detailed analysis of random and (in particular) systematic uncertainties significantly weakens any claim on the mass-dependent evolution, particularly for the massive end.

Recently, a few studies have started to investigate the mass function in galaxy clusters. Vulcani et al. (2011) found a quite strong evolution with redshift, consistent with field evolution. Calvi et al. (2013, hereafter C13) and Vulcani et al. (2013, hereafter V13) found that, at least above $\log M_*/M_\odot \sim 10.25$ at $z = 0$, and above $\log M_*/M_\odot \sim 10.5$ at $z = 0.6$, the shape
of the mass distribution is very similar in clusters, groups, and the field. Similar conclusions have also been extended to \( z \approx 1 \) (van der Burg et al. 2013).

V13 also showed that the evolution of the mass function, in the interval \( z = 0.6–0.06 \), is similar in clusters and in the general field.

On the theoretical side, semi-analytic galaxy formation models (e.g., White & Frenk 1991; Kauffmann et al. 1993, 1999; Cole et al. 1994; Somerville & Primack 1999; Cole et al. 2000; Springel et al. 2001; Hatton et al. 2003; Baugh et al. 2005; De Lucia et al. 2007; Font et al. 2008) provide a powerful tool to interpret observational results in a cosmological context. However, it is important to keep in mind that models are usually normalized to a subset of low-\( z \) observations, and the field mass function has often been used by the most recent models as the primary constraint to tune the various model parameters.

Several studies have compared the observed field mass function to the one predicted by simulations. Semi-analytic models that include strong stellar feedback reproduce the \( z = 0 \) mass function well (e.g., Guo et al. 2011, hereafter G11; Bower et al. 2012) but they struggle in reproducing the mass function of low-mass galaxies at higher redshift (Fontanella et al. 2006; Marchesini et al. 2009; Drory et al. 2009; Fontanella et al. 2009, hereafter F09; Lo Faro et al. 2009; Cirasuolo et al. 2010, G11). In particular, they overproduce low-mass galaxies at \( z > 0.5 \), predicting almost no evolution in the number density of galaxies of mass \( \sim 10^{10} \, M_\odot \), since \( z \approx 2 \), in contrast with observational measurements for galaxies of the same mass whose number density is found to evolve by a factor \( \sim 6 \) over the same redshift range (F09). In the models, low-mass galaxies are predicted to form too early and have too little ongoing star formation at later times (e.g., F09; Firmani & Avila-Reese 2010; G11; Weinmann et al. 2012), so their present-day stellar populations are too old.

For low-mass galaxies, it is generally believed that the discrepancies with observational data are due to an incorrect treatment of the star formation and stellar feedback process. In addition, as pointed out by Weinmann et al. (2012), the problem is not limited to semi-analytic models but is also present in hydrodynamical simulations of galaxy formation.

A number of problems still affect model predictions for the most massive galaxies (log \( M_\star/ M_\odot > 11 \)). Their evolution from \( z \sim 1 \), which is driven by mergers, is marginally inconsistent with the observational results, being slightly too fast (F09). Models also underestimate both the number density and the SFR of massive galaxies at \( z > 2 \) (Marchesini et al. 2009; F09). In addition, models do not reproduce the observed chemical abundances: at the massive end, the predicted mass–metallicity relation turns over and is offset low with respect to the data (see, e.g., De Lucia & Borgani 2012).

The inaccuracy of the CDM paradigm seems to be related to the fact that model galaxies closely follow the evolution of DM halos, while it is necessary to find a way to decouple the halo accretion rate and the SFR of galaxies (Weinmann et al. 2012).

Recently, some studies have investigated the effect of assuming a warm dark matter (WDM) power spectrum. Menci et al. (2012) argued that the WDM scenario may solve the excess of low-mass galaxies, since it produces a smaller number of collapsed low-mass halos. However, Kang et al. (2013) argued that the claimed success might simply reflect a non-optimal parameterization of the physics of galaxy formation implemented in the model. This shows that a single observable (e.g., the stellar mass function) cannot constrain the effects of the warm component on galaxy formation, even though accurate measure-ments of the mass function and the link between galaxies and DM halos down to the very low-mass end can give very tight constraints on the nature of DM candidates.

All studies mentioned above focused on comparing model predictions with observational data for field galaxies. Very few studies have considered trends as a function of the environment (see, e.g., Liu et al. 2010 for an analysis of the local conditional stellar mass function). On the theoretical side, it is well established that the subhalo mass function depends only weakly on the mass of the parent halo (e.g., De Lucia et al. 2004b; Lee 2004; Giocoli et al. 2008). It has to be considered, however, that galaxies and subhalos are not simply related (Gao et al. 2004; Sawala et al. 2013) so that a similar subhalo mass function does not necessarily imply a similar galaxy mass function.

The aim of this paper is to analyze what semi-analytic models predict for the mass function at two different epochs (\( z = 0.06 \) and \( z = 0.62 \)), in the field and in halos of different masses. Our aim is to test whether simulations are able to reproduce the observational results and eventually help improve the available models by analyzing in detail where they fail.

The plan of the paper is as follows. Sections 2 and 3 introduce the observational and theoretical samples, respectively. Section 4 describes the method used for our analysis, while Section 5 shows the basic comparison between the observed and predicted mass functions. In Section 6, we discuss results from simulations, and in Section 7, we summarize and discuss our results. Finally, Section 8 gives our conclusions.

We assume \( H_0 = 73 \, \text{km s}^{-1} \, \text{Mpc}^{-1}, \Omega_0 = 0.25, \) and \( \Omega_\Lambda = 0.75. \) The adopted initial mass function (IMF) is that of Kroupa (2001) in the mass range \( 0.1–100 \, M_\odot. \) Magnitudes are in the Vega system.

2. OBSERVATIONS

In this paper, we exploit four different observed samples to measure the galaxy stellar mass function of low- and intermediate-\( z \) field and cluster galaxies.

2.1. Low \( z \)

We use the Padova-Millennium Galaxy and Group Catalogue (PM2GC; Calvi et al. 2011), which is a galaxy catalog representative of the general field population in the local universe. It is a database built on the basis of the Millennium Galaxy Catalogue (Liske et al. 2003), a deep and wide B-imaging survey along an equatorial strip of \( \sim 38 \, \text{deg}^2. \) The catalog contains only galaxies brighter than \( M_B \leq -18.7. \) By applying a friends-of-friends (FoF) algorithm, a catalog of galaxy groups with at least three members in the redshift range \( 0.04 \leq z \leq 0.1 \) has been created (see Calvi et al. 2011). Galaxies that after several iterations of the algorithm are within \( 1.5 \, R_{200} \) from the group center and \( 3 \sigma \) (velocity dispersion) from the group redshift are considered group members. \( R_{200} \) is defined as the radius delimiting a sphere with interior mean density 200 times the critical density of the universe at that redshift, and is commonly used as an approximation of the group/cluster virial radius. The \( R_{200} \) values for our structures are computed from the velocity dispersions using the formula (Finn et al. 2005)

\[
R_{200} = 1.73 \frac{\sigma}{1000 \, \text{km s}^{-1}} \left( \frac{1}{\sqrt{\Omega_\Lambda + \Omega_0 (1 + z)^3}} \right) h^{-1} \text{Mpc}.
\]

(1)

Galaxies that do not satisfy the group linking criteria adopted have been placed either in the catalog of single field galaxies,
which comprises the isolated galaxies, or in the catalog of binary field galaxies, which comprises the systems with two galaxies within 1500 km s\(^{-1}\) and 0.5 h\(^{-1}\) Mpc.

Stellar masses have been estimated following Bell & de Jong (2001; Calvi et al. 2011). Briefly, they were derived using the relation between \(M/L_B\) and the dust-uncorrected rest-frame \((B - V)\) color

\[
\log_{10}(M/L_B) = -0.51 + 1.45(B - V)
\]

valid for a Bruzual & Charlot model with solar metallicity and a Salpeter (1955) IMF (0.1–125 \(M_\odot\)). Then, they were converted to a Kroupa (2001) IMF, adding –0.19 dex to the logarithmic value of the masses. The typical scatter of the mass uncertainties is \(\sim 0.2–0.3\) dex (see, e.g., Kannappan & Gawiser 2007).

In this work, we refer to the galaxy sample described in C13: we consider galaxies at 0.04 \(\leq z \leq 0.1\) in the field, in groups, binary, and single systems. The mass completeness limit is \(\log M_*/M_\odot = 10.25\) and includes 1045 field galaxies.

The PM2GC covers a much smaller sky area than the Sloan Digital Sky Survey (York et al. 2000) but is characterized by a better imaging quality and higher spectroscopic completeness. In Figure 1 of their paper, C13 present a comparison between the PM2GC mass function and literature results (Bell et al. 2003; Cole et al. 2001; Baldry et al. 2012; Li & White 2009) and show a good agreement among the different estimates.

For clusters, we take advantage of the Wide-field Nearby Galaxy-clusters Survey (WINGS; Fasano et al. 2006), a multiwavelength survey at 0.04 \(\leq z \leq 0.07\). WINGS is based on deep optical \((B, V)\) wide field images (\(\sim 35'\times 35'\)) of 76 clusters. The clusters span a wide range in velocity dispersion (550 \(\leq \sigma \leq 1400\) km s\(^{-1}\)) and X-ray luminosity (0.2 \(\times 10^{44} \leq L_X < 5 \times 10^{44}\) erg s\(^{-1}\)). Besides the optical imaging data, a number of follow-ups were carried out to obtain additional homogeneous information for galaxies, such as optical spectroscopy (Cava et al. 2009), near-infrared \((J, K)\) data (Valentinuzzi et al. 2009), and U-band imaging (Omissolo et al. 2013). An OmegaCam/VST \(u, B, V\) follow-up of about 50 clusters is underway.

We use the galaxy sample described in Vulcani et al. (2011). Briefly, we consider 21 clusters with spectroscopic completeness larger than 50\%. The brightest cluster galaxies (BCGs) have been excluded, while only spectroscopically confirmed members within 0.6\(R_{200}\) (the largest radius generally covered in WINGS clusters) have been included in the sample. Galaxies have been weighted for spectroscopic incompleteness using the ratio of the number of galaxies with a spectroscopic redshift to the number of galaxies in the parent photometric catalog, as a function of galaxy magnitude (Cava et al. 2009). Stellar masses have been estimated following Bell & de Jong (2001) (Vulcani et al. 2011). The mass complete sample includes all galaxies more massive than \(\log M_*/M_\odot = 9.8\), for a total of 1229 galaxies.

We use the WINGS mass function because, thus far, it is the only attempt to describe the galaxy stellar mass distribution in clusters at \(z \sim 0\). We note that Mercurio et al. (2012) investigated the Shapley supercluster but without explicitly isolating clusters. Furthermore, the cluster mass function presented in Baldry et al. (2008) is based on masses determined from luminosities using a fixed M/L ratio for all galaxies.\(^7\)

2.2. Intermediate \(z\)

For the field at 0.6 \(< z < 0.8\), we use the mass function presented in Drory et al. (2009), which is, among those published, the one with the lowest mass completeness limit. The authors exploited the COSMOS catalog with photometric redshifts derived from 30 broad and medium bands described in Capak et al. (2007) and Ilbert et al. (2009). They used galaxies with \(i'_{AB} < 25\), where the detection completeness is \(>90\%\) (Capak et al. 2007). They detected 36,885 galaxies at 0.6 \(< z < 0.8\) in an area of 1.73 deg\(^2\).

Drory et al. (2009) derived stellar masses comparing multiband photometry to a grid of stellar population models of varying star formation histories (SFHs), ages, and dust content (for further details, refer to Drory et al. 2009). They adopted a Chabrier (2003) IMF (0.1–100 \(M_\odot\)). To reduce the uncertainties due to the different method adopted by Drory et al. (2009), we apply a mean correction to the Drory et al. (2009) mass function, as described in Appendix A.1. Converting their stellar masses into our IMF and cosmology, their mass completeness limit is \(\log M_*/M_\odot = 10\).

We also use the ESO Distant Cluster Survey (EDisCS; White et al. 2005), a multiwavelength photometric and spectroscopic survey of galaxies in 20 fields containing galaxy clusters (400 \(< \sigma \leq 1100\) km s\(^{-1}\)) and groups (150 \(< \sigma \leq 400\) km s\(^{-1}\)) at 0.4 \(< z < 1\). Structures were selected as surface brightness peaks in smoothed images taken with a very wide optical filter (\(\sim 4500–7500\) Å) and have high-quality multiband optical and near-infrared photometry (White et al. 2005) and spectroscopy (Halliday et al. 2004; Milvang-Jensen et al. 2008; Vulcani et al. 2012). Photometric redshifts were computed using two independent codes, a modified version of the publicly available HyPerz code (Bolzonella et al. 2000) and the code presented in Rudnick et al. (2001, 2003, 2009). Photo-z membership was established using a modified version of the technique first developed in Brunner & Lubin (2000; De Lucia et al. 2004b, 2007; Pelló et al. 2009).

Stellar mass estimates are presented in Vulcani et al. (2010, 2011); they have been determined following Bell & de Jong (2001). The photometric magnitude limit \((I = 24)\) corresponds to a mass limit of \(\log M_*/M_\odot = 10.2\).

We use the EDisCS sample described in Vulcani et al. (2011), which includes all of the photo-z members of 14 clusters in the redshift range 0.4 \(< z < 0.8\), above the mass limit and within 0.6\(R_{200}\). BCGs have been excluded. The final sample consists of 489 galaxies. As discussed in the Appendix of Vulcani et al. (2011) and in V13, contamination from photo-z interlopers might alter the mass functions. However, Rudnick et al. (2009) found that at the magnitudes used in this paper, the photo-z counts are fully consistent with the statistically background subtracted counts. Anyway, it should be noted that the stellar mass function measured using photo-z membership becomes increasingly uncertain with decreasing

---

\(^{7}\) The lack of studies of cluster stellar mass functions might be quite surprising. An explanation for this might be that analyses instead focused mainly on conditional stellar mass function (see, e.g., Yang et al. 2008), which describes the average number of galaxies as a function of galaxy stellar mass in a dark matter halo of a given mass. This approach allows to disentangle the role of the galaxy formation model (which affects the relation between the mass of the halo and the galaxy stellar mass) from that of the adopted cosmology (which affects the number density) but strongly depends on the theoretical assumptions.
magnitude/mass, where our data lack a good spectroscopic coverage. We exploit the spectroscopic sample to characterize the maximum mass reached by galaxies in clusters, since photometric interlopers could bias this estimate.

As for the local universe, no other cluster galaxy mass functions are available at these redshifts.

3. THEORETICAL PREDICTIONS

We take advantage of publicly available galaxy catalogs from semi-analytic models run on the Millennium Simulation (Springel et al. 2005). This uses 10^{10} particles of mass 8.6 × 10^6 h^{-1} M_\odot to trace the evolution of the matter distribution in a cubic region of the universe of 500 h^{-1} Mpc, on a side from z = 127 until z = 0, and has a spatial resolution of 5 h^{-1} kpc.

We use two different semi-analytic models to investigate how different assumptions about the physical processes acting on the baryonic component impact the evolution of the galaxy stellar mass function.

The semi-analytic model discussed in De Lucia & Blaizot (2007, hereafter DLB07) builds on the methodology and prescriptions introduced in Springel et al. (2001), De Lucia et al. (2004b), and Croton et al. (2006) and has been the first variant of the “Munich” models family that has been made publicly available. The latest update is provided by the G11 model, that we use in the second part of this paper. Both models include prescriptions for supernova-driven winds, follow the growth of supermassive black holes, and include a phenomenological description of AGN feedback. The G11 model is based on the DLB07 one but it significantly differs in some areas. Several of the extensions involve a different treatment of some processes, such as the strangulation of satellite galaxies (that is instantaneous after accretion in DLB07), the supernova feedback (that is more efficient in G11), and the introduction of some processes which were not previously included, such as the disruption of galaxies to produce intracluster light. In particular, when a central galaxy first becomes satellite, that is, its halo is first linked to a more massive halo, G11 continue to treat it as a central galaxy, that is in the same manner as a galaxy at the center of a main subhalo, until it falls within the virial radius of the center of its new halo. At this point, they switch on tidal and ram-pressure stripping processes, which can remove gas from the galaxy or even disrupt it completely. This change leads to a reduction in the number of satellite galaxies with respect to predictions from DLB07. We refer to the original papers for more details.

Note that the DLB07 model was mainly constrained by the observed local K-band luminosity function. The G11 model, on the other hand, was tuned to reproduce the observed stellar mass function in the local universe.

In the following, we refer to an output value of the simulations either as a “sim-projected” or a “simulated” quantity. Sim-projected quantities (number of galaxies, velocity dispersion, etc.) are computed from the simulation with the same methods that would be used observationally and are projected on the x y plane, while simulated quantities are the three-dimensional (3D) estimates provided by the simulation.

As also discussed in Poggianti et al. (2010), the former quantities can be compared to observational measurements.

3.1. The Sim-projected Sample

To compare simulations to observations, we use only the DLB07 model and we assemble samples as similar as possible to the observed ones. The G11 model does not provide the magnitudes needed to compute the stellar masses as done with the data, and will be used only in the second part of this study.

To reproduce the observed clusters, using the available catalogs, we selected 150 halos with 10^{13} M_\odot ≤ M^{MS}_{200} ≤ 10^{15} M_\odot at z = 0.06 and 142 halos at z = 0.62, uniformly distributed in mass. M^{MS}_{200} masses are computed from the N-body simulation as the mass enclosed within R^{MS}_{200}, the radius of a sphere, which is centered on the most bound particle of the group and has an overdensity of 200 with respect to the critical density of the universe (the virial radius) at the redshift of interest. We then selected all galaxies within a box of 10 physical Mpc on a side, centered on each halo considered. 3D velocity dispersions have been computed using all galaxies within R^{MS}_{200} and more massive than M_\star = 10^9 M_\odot.

To reproduce the observed field, we selected portions of “simulated sky” corresponding to square boxes of ~38 deg^2 (30 × 30 Mpc) at z = 0.06, and to ~1 deg^2 (17 × 17 Mpc) at z = 0.62. At low z, to match the observed field sample, the boxes extracted are 323 physical Mpc deep, while at higher redshift, the maximum depth of the simulated box (500 h^{-1} Mpc) is used, which corresponds to 422 physical Mpc. To account for cosmic variance, 10 simulated field samples were selected at each redshift. No preselection on halo mass was applied to these boxes.

The model provides stellar masses and rest-frame Vega magnitudes, which include the effect of the dust in the Buser (1978) system, calculated using the models of Bruzual & Charlot (2003). Instead of using the stellar masses provided by the models, we recomputed them using the Bell & de Jong (2001) formulation (Equation (2)), as done for the observations. To this aim, we first converted the model magnitudes to the Johnson–Cousins system, as defined in Bessell (1990). A comparison between these stellar masses and those provided by the model is provided in Appendix A.2.

In the following, we consider a conservative mass limit of log M_\star/M_\odot = 9.4.

The sim-projected quantities are computed using the σ along the z axis. Considering a different projection of the velocity dispersion or the average of the three does not considerably affect the results. Then, assuming the virial theorem is valid, the sim-projected R_{200} can be estimated from the velocity dispersion, as given in Equation (1). We note that Poggianti et al. (2010) showed that both sim-projected velocity dispersions and sim-projected R_{200} similarly deviate from the velocity dispersion obtained from the halo mass and from the theoretical R^{MS}_{200} radius, underestimating them at low values and overestimating them at high values (see also Biviano et al. 2006).

Our sim-projected samples are defined as follows.

Field: we consider galaxies from the snapshot corresponding to z = 0.06 and z = 0.62. To take into account the cosmic variance (or, more accurately, sample variance), we compute a mean of the ten representations of the fields and consider the scatter among them in the errors. Above the PM2GC (COSMOS) mass limit, we are left with a mean value of 1439 (5093) galaxies.

Clusters: at z = 0.06, we consider the 47 halos with 550 < σ < 1400 km s^{-1} and the 77 halos at z = 0.62 halos with...
400 < \sigma < 1100 \text{ km s}^{-1}.\)\(^{10}\) To measure the mass functions from the sim-projected samples, we randomly extract 21 clusters at low \(z\) and 14 clusters at intermediate redshift. We repeat the sampling ten times, to take into account the cluster variance. Galaxies whose velocity \(v_z\) is within \(3\sigma\) are considered members of the halos. We note that this is much larger than the size of the boxes we extracted: at low \(z\), \(3\sigma\) spans from 20 to 50 Mpc, while at higher \(z\), it spans from 10 to 25 Mpc. For the mass functions, we exclude the central galaxy of each halo and take into account only galaxies within the projected \(0.6R_{200}\), where \(R_{200}\) has been computed from the \(\sigma\) as in Equation (1). Above the WINGS (EDisCS) mass limit, we are left with a mean value of 2877 (532) galaxies.

### 3.2. The Simulated Sample

When analyzing only simulations, we take advantage of all of the information coming from the Millennium Simulation and use 3D values. We exploit both the DLB07 and G11 models.

We extract all galaxies from the snapshot corresponding to \(z = 0.06\) and \(z = 0.62\) to characterize the field. At both low and intermediate \(z\), we also extract all halos in three different halo mass bins: 76853 and 62333 halos with 13.25 < \(\log M_{\text{halo}}\) < 13.55 (least massive halos), 2981 and 1910 halos with 13.9 < \(\log M_{\text{halo}}\) < 14.25 (intermediate massive halos), and 41 and 6 halos with 14.9 < \(\log M_{\text{halo}}\) < 15.25 (most massive halos), where \(\log M_{\text{halo}} = \log(M_{\text{halo}}/M_\odot)\), respectively. As done for the sim-projected sample, for each halo, we extracted a box of 10 physical Mpc on a side, centered on the most bound particle. 3D velocity dispersions have been computed using all galaxies within \(R_{200}\) and more massive than \(M_\star = 10^9 M_\odot\).

To compute the galaxy stellar mass function, we have considered all galaxies that are members and excluded central galaxies. Finally, we separately consider galaxies at different halo-centric distances: within \(0.6R_{200}\), within \(R_{200}\), and within 1–3\(R_{200}\).

The number of galaxies in the different samples are given in Table 1. The G11 model always has a lower number of galaxies than the DLB07 one, at any redshift and in any environment. This is just due to the stronger stellar feedback assumed in the G11 model, which reduces the number of low-mass galaxies.

### 4. METHODS

To quantify the similarities between the mass functions in different environments, we analyze several aspects of the mass distributions.

1. We inspect the shape of the data distributions. We consider bins of 0.2 dex in stellar mass and express the mass functions in units of comoving volume (in units of number per \(h^{-3} \text{ Mpc}^3\) dex\(^{-1}\)). To compute the volume for clusters, we consider the sphere (or part of the sphere for observations and sim-projections), defined in terms of virial radius, which includes galaxies in the considered sample. Details on how clusters mass functions have been normalized can be found in Appendix B. In all of the figures, we do not plot errors along the \(x\)-direction but errors on stellar masses are typically of 0.2–0.3 dex. Errors along the \(y\)-direction are computed adding in quadrature the Poissonian errors (Gehrels 1986) and the uncertainties due to cosmic variance or cluster-to-cluster variations, computed from the sim-projections. Following Marchesini et al. (2009), we measured \(\phi\) separately for each of the \(n\) fields or clusters, and then we estimated the contribution to the error budget of \(\phi\) from cosmic/cluster variance using \(\sigma_{cv} = \text{rms}(\Phi)\sqrt{n}\). The values obtained are then also used for observations. When considering only simulations, neither cosmic nor cluster variances are considered, since we are using all of the available galaxies.

2. We use the Kolmogorov–Smirnov (K-S) test, which quantifies the probability that two data sets are drawn from the same parent distribution.\(^{11}\) A “low probability” (\(P_{K-S} < 5\%) means that two samples are different; on the contrary a “high probability” means that the test is unable to find differences. In the following figures, only significant K-S results will be given. For the WINGS sample, we consider completeness weights when running the test. When we investigate only simulations, we do not use the test, since, given the very large numbers of galaxies, it turns out to be meaningless.

3. We consider analytical fits to the mass functions using a Monte Carlo Markov chain method. Assuming that the number density \(\phi(M)\) of galaxies can be described by a Schechter (1976) function, the mass function can be written as

\[
\phi(M) = (\ln 10) \Phi^* 10^{(M - M^*)/(1 + \alpha)} \exp(-10^{(M - M^*)}),
\]

where \(M = \log(M_\star/M_\odot)\), \(\alpha\) is the low-mass end slope, \(\Phi^*\) is the normalization, and \(M^* = \log(M_\star^*/M_\odot)\) is the characteristic mass. Schechter function fits are computed only above the completeness limits.

Our Schechter fits are used to characterize mass functions, but the procedure does not take into account uncertainties on the estimates of the stellar mass. However, errors and uncertainties associated with masses are not negligible. For example, Marchesini et al. (2009) showed that both systematic and random errors can arise from the unknown true SFHs, metallicities, and dust corrections, and also from photometric redshift errors, differences in stellar population models, the unknown stellar IMF and its evolution.

---

\(^{10}\) The velocity dispersion distribution of simulated clusters is consistent with that of the observed ones, so that trends cannot be driven by the size of the clusters.

\(^{11}\) We note that the K-S test does not take into account the errors.
and cosmic variance. The error convolution can have a significant impact upon the shape of the high-mass end of the mass distribution (e.g., F09 and references therein).

In addition, the high-mass end is affected by the Eddington (1913) bias: in the mass range where the number counts are steep, more objects of intrinsic low mass will be scattered to higher values of stellar mass than systems of intrinsically high mass being scattered to lower masses. This results in an overestimate of the true galaxy counts at high stellar masses.

As a consequence, when we use stellar masses determined using Bell & de Jong (2001; Section 5), we should take into account the impact of errors on the mass estimates. We assume that the error has a Gaussian distribution (independent of mass and redshift). Following an approach similar to that described in Behroozi et al. (2010), we perform additional fits to our mass functions, with the aim of estimating the Schechter parameters that describe the “intrinsic” mass function ($\phi_{\text{true}}$). The observed mass function, convolved with the errors, is

$$\phi(M) = (\ln 10) \int \phi_{\text{true}}(m') \frac{1}{\sqrt{2\pi} \sigma} \exp \left( -\frac{(\ln 10 M - m')^2}{2\sigma^2} \right) dm',$$

where $m = \ln(M/M_\odot)$, and $\sigma$ is the uncertainty on the stellar mass, assumed to be $\sim 0.2$ dex. The effect of the convolution is

$$\phi(M) = (\ln 10) \frac{1}{\sqrt{2\pi} \sigma} \int \Phi^* \exp \left[ \frac{(\alpha' + 1)(m' - M^*)}{2\sigma^2} \right] \exp \left[ -\exp(m' - M^*) \right] \exp \left[ -\frac{(m' - M^*)^2}{2\sigma^2} \right] dm',$$

where $\Phi^*$, $\alpha'$ and $M^* = \ln(M_\star/M_\odot)$ are the best-fit parameters we can insert in the Schechter function to get $\phi_{\text{true}}$.

In our figures, we report fits of $\phi_{\text{true}}$ to give an indication of how much the mass functions change when comprehensive error estimates are taken into account. However, in the comparison between observations and sim-projections, since we treat simulations in the same way as observations, we simply use $\phi(M)$ because this is the approach commonly used in observational studies (e.g., Bundy et al. 2006; Pozzetti et al. 2010). Our choice is also justified by the fact that errors inherent to the transformation from magnitudes to masses obtained with spectrophotometric models (which dominate the total error) are common to both observations and sim-projections in our approach.

4. We also analyze the highest mass reached by galaxies in each environment, hereafter “maximum mass,” both including and excluding BCGs in clusters. This is another important aspect that can help us to characterize the environment–mass relation.

5. RESULTS: COMPARISON BETWEEN SIM-PROJECTED AND OBSERVED MASS FUNCTIONS

In this section, we use the samples described in Sections 2.1, 2.2, and 3.1.

We begin with investigating central cluster galaxies, which are excluded from our mass function analysis. Both in simulations and observations, “total” magnitudes are used to compute stellar masses, and hence they may include some intra cluster light. Figure 1 shows that the mass of the central galaxy correlates with the velocity dispersion of the hosting halo at both redshifts considered. Observations and sim-projections are in agreement, showing a similar slope of the relations. This is a sign that the environment has a strong effect on the mass of the central galaxy. This result resembles that presented for EDisCS’ BCGs in Whiley et al. (2008): clusters with large velocity dispersions tend to have BCGs with larger stellar masses. The strong link between central galaxies and the hosting halo is known and has been discussed by several authors (e.g., Shankar et al. 2006; Wang et al. 2006; Moster et al. 2010; Leauthaud et al. 2012).

The detected trend of more massive galaxies in higher velocity dispersion halos may be attributable to the environment. However, there is a statistical effect that can cause such a trend. If stellar masses are randomly drawn from the mass function of galaxies, then massive halos, that host a larger number of galaxies, have a higher probability to host more massive galaxies (Tremaine & Richstone 1977; Bhavsar & Barrow 1985; Lin et al. 2010; Dobos & Csabai 2011; Paranjape & Sheth 2012; More 2012). To disentangle the statistical effect from the environmental effect, we adopt the following approach. For each cluster, we randomly sample $N$ galaxies from the observed field stellar mass function at similar redshift, where $N$ is the number of galaxies (including the BCG) in that cluster. We perform the random sampling 1000 times in order to sample the probability distribution of the maximum mass. The results are plotted in Figure 1. At both redshifts, the random sampling agrees quite well with the data for low-mass structures, while there is a deviation for the most massive halos. In this regime, at any fixed velocity dispersion, the observed and sim-projected masses are systematically higher than those coming from the random sampling. The effect is more noticeable at low $z$ than at higher $z$. This test suggests that the relation can indeed be ascribed to the environment, in particular, for very massive structures (in agreement with the results presented in Rodríguez-Puebla et al. 2013).

We now focus only on cluster satellites and field galaxies and investigate whether simulations can reproduce the observed mass function in different environments and at different redshifts. We stress that comparisons make sense only above the mass limit of each observed sample. The fits presented in the figures are computed above the observed mass limit, hence comparisons between samples with different completeness limits are not meaningful.

As shown in Figure 2 and in Table 2,14 the agreement between sim-projections and observations is only partial.

At low $z$, in the field (upper left panel), sim-projected and observed mass functions are very similar ($M^*$ and $\alpha$ are in agreement), as also pointed out by the K-S test. Sim-projections reproduce the observed number density, at any mass above the limit. On the other hand, they do not predict the right maximum mass reached by galaxies in the observed field: in the

12 Using the halo mass given by simulations instead of the velocity dispersion gives similar results.
13 For sim-projections, we use the central galaxy of all halos with $550 < \sigma < 1400$ km s$^{-1}$ at low $z$ and with $400 < \sigma < 1100$ km s$^{-1}$ at high $z$.
14 The values listed are not the same as those given in Vulcani et al. (2011), C13, and V13 for the same samples due to the different cosmology, normalization adopted, and fitting routine. However, the reported values are in agreement within the errors with those already published. Drory et al. (2009) use a double Schechter fit data, so their best-fit values are not directly comparable to ours.
At log $M/\text{M}_\odot$ and the Schechter fit confirm these visual findings: in the sim-projected sample, there are galaxies with log $M/\text{M}_\odot$ smaller (parameters not compatible at 3 $\sigma$) from photo-$z$ interlopers does not alter the shape of the mass function. We note that EDisCS is the only non-spectroscopic sample, and, although as discussed in Section 2.2, the contamination from photo-$z$ interlopers does not alter the shape of the mass function, the number density is influenced by it. To account for this, we correct the EDisCS number density using the ratio of the counts weighted for incompleteness in the spectroscopic sample, in the mass range across which they overlap. This factor is independent of mass and it is equal to 2.8.16 At log $M/\text{M}_\odot < 11$, all of the models overpredict the observed mass function, with the discrepancy increasing with increasing redshift.

At similar redshifts in clusters (upper right panel), sim-projected and observed mass functions are different: the WINGS mass distribution always lies below the sim-projected one, while at higher masses, the sim-projected function extends to higher masses than the observed mass function. Both the K-S test and the Schechter fit confirm these visual findings: in the sim-projected sample, $M*$ is much larger than in WINGS, while $\alpha$ is smaller (parameters not compatible at 3$\sigma$ level).

At higher $z$, simulations do not reproduce well the observed field mass function (bottom left panel). At log $M/\text{M}_\odot < 11$, the sim-projected mass function rises more steeply than the observed one, while at higher masses, there are hints of an excess of very massive galaxies in the observed mass function compared to the sim-projected function. However, we remind that in this regime, uncertainties are large, and the correction we applied to the Drory et al. (2009) mass function may not be very accurate at these masses. The results of the Schechter fit confirm that the two samples are not drawn from comparable distributions.15 Our results are in agreement with those discussed, e.g., in FO9: comparing three different semi-analytic models to observations, they found that for, log $M/\text{M}_\odot < 11$, all of the models overpredict the observed mass function, with the discrepancy increasing with increasing redshift.

At similar redshifts, sim-projected and observed clusters (bottom right panel) show different shapes of the mass functions. From Drory et al. (2009) at our disposal.

15 We cannot perform the K-S test since we do not have the unbinned counts from Drory et al. (2009) at our disposal.

16 This renormalization does not influence the shape of the mass functions.
massive clusters considerably deviate from the trends found for the entire galaxy population, the second brightest galaxies are simply the statistical extreme of such a population (Lin et al. 2010; Shen et al. 2013), and it is thus interesting to consider whether their mass correlates with halo mass. For EDisCS, we use the spectroscopic sample to have a more reliable indication on the cluster members. The agreement between observations and sim-projections is very good. At both low $z$ (left panel) and intermediate $z$ (right panel), and for both sim-projected and observed cluster samples, there is a strong dependence of the second most massive galaxy on the mass of the halo: the more massive the halo, the larger the stellar mass of the second most massive galaxy. Comparing the trends at the two redshifts, we find that the relation is shifted toward slightly higher masses at low redshift, suggesting that, in the local universe, structures host more massive galaxies.

As for the central galaxies, the detected trend might be attributable to the environment. We perform the same random sampling described above to disentangle the statistical effect from the environmental effect. In this case, we exclude the central galaxies from the samples. The results are plotted in Figure 3. At low $z$, both for the observed and sim-projected clusters, the random sampling predicts a flatter relation than that obtained using the real samples, suggesting that the environment is playing a role in the most massive halos. In contrast, at higher redshift, smaller differences between the randomly sampled galaxies and real samples are detected, so that at this redshift, the relation found can be explained by a purely statistical effect. A larger number of halos with high velocity dispersion would be useful to place stronger constraints on this finding.

5.1. A More Careful Analysis of the Field Mass Function

In the previous section, we have shown how well simulations reproduce the general field (meant as a wide portion of the sky, including all environments) mass function in the local

Figure 2. Observed (blue symbols) and sim-projected (red symbols) mass functions for galaxies in the field and in clusters at low and intermediate $z$, respectively. Mass functions are normalized to the comoving volume probed by the samples. Error bars on the $y$ axis are computed combining the Poissonian errors (Gehrels 1986) and the uncertainties due to cosmic variance or cluster-to-cluster variations. For the Drory et al. (2009) sample, errors also include the uncertainty related to photometric redshifts. Points and crosses represent the measured mass functions, solid lines and shaded areas represent Schechter fits with 1$\sigma$ errors, and dashed lines represent the mass function ($\phi_{\text{true}}$) obtained from the deconvolution with the uncertainties on masses (see text for details). The K-S probabilities are also shown as percentages. The bottom left insets in each panel show the $M^\ast$ and $\alpha$ Schechter fit parameters with 1$\sigma$, 2$\sigma$, 3$\sigma$ error contours. The dotted vertical line shows the mass limit of each sample.

(A color version of this figure is available in the online journal.)
et al. 2011) to the sim-projected field samples and obtained

The choice of adopting the Schechter fit of the PM2GC data and not the

same FoF we used for the PM2GC (see Section 2 and Calvi

this finer division of environments. To do this, we applied the

simulations are able to reproduce the galaxy mass function of

universe. Since the general field sample is the sum of group,

binary systems, and single galaxies, we aim to understand if

simulations are able to reproduce the galaxy mass function of

this finer division of environments. To do this, we applied the

same FoF we used for the PM2GC (see Section 2 and Calvi et al. 2011) to the sim-projected field samples and obtained samples that describe the single galaxies, binary systems, and galaxies in groups. From both sim-projections and observations, we excluded galaxies in groups with a velocity dispersion $\sigma > 500 \text{ km s}^{-1}$, to eliminate possible contamination from clusters.

Figure 4 shows the comparison between the sim-projected and observed mass function in the three different environments. The agreement is striking; in any given environment, mass distributions are indistinguishable. The result is also supported by the analysis of the Schechter fits (see Table 2).

Hence, the semi-analytic model correctly includes all of the processes that are responsible for the assembly of galaxies in these finer environments, groups included. This finding is important above all in the light of the fact that the cluster mass function at the same redshift is not reproduced, suggesting an inaccuracy in the treatment of cluster-specific processes, as will be discussed later, in Section 7.4.

5.2. Dependence on the Environment and Evolution

The two main results of C13 and V13 were that (1) both at low and intermediate $z$, the shape of the mass distribution of galaxies in clusters and in the field is similar, and (2) the evolution of the shape of the mass function from $z \sim 0.6$ to $z \sim 0.06$ is the same in clusters and in the field. We now wish to test whether samples drawn from simulations are able to reproduce these observational findings.

In Figure 5, we plot the ratio of the observed and sim-projected mass functions to the best-fit Schechter function of our observed field mass function in the local universe (PM2GC) ($\Phi/\Phi_{\text{PM2GC}}-\text{bf}$), used for reference. This is another way of detecting differences between distributions. If mass functions

have similar shapes, trends have to be similar. Flat trends mean agreement with the shape of the low-$z$ observed field mass function.

Discrepancies between observations and sim-projections are evident, as already discussed in the previous section.

In the observed samples, at a given redshift, cluster and field trends are similar. This indicates that the environment does not strongly affect the shape of the mass distributions, at both redshifts considered, at least above $\log M_*/M_\odot = 10.25$, our most conservative mass limit (see also C13 and V13). At masses larger than $\log M_*/M_\odot = 11.4$, there are differences among the different samples but error bars are too large to draw statistically robust conclusions. The same main result is found when considering the sim-projected samples. Hence, sim-projections, though not reproducing all of the environments separately, can reproduce the observed invariance of the shape of the mass function with the environment above $\log M_*/M_\odot = 10.25$. We note that, in the low-$z$ clusters, the exclusion of galaxies less massive than $\log M_*/M_\odot = 10.25$ reduces the discrepancies between the observed and sim-projected mass functions, making this result possible. This highlights the importance of reaching a broad mass dynamical range to robustly determine the shape of the galaxy mass function. We do not have observational data at lower masses for galaxies in all the environments; however, sim-projections predict similar mass functions for galaxies in different environments, at a given redshift.

G11, exploiting the increased resolution provided by the Millennium-II Simulation (Boylan-Kolchin et al. 2009), showed that in their simulations, the stellar mass function of galaxies in rich clusters at low $z$ is predicted to be very similar in shape to that in the general field, even down to $\log M_*/M_\odot \sim 7$. As far as the evolution is concerned, in observations, the number density of cluster galaxies increases at increasing redshift, although not systematically as a function of stellar mass (the blue solid line is steeper than the red line). In the field, the opposite is found: the number density of galaxies tends to decrease with increasing redshift, with the decrease being more significant at lower stellar masses. However, at both redshifts, the relative slopes at low masses of clusters and in the field are

17 The choice of adopting the Schechter fit of the PM2GC data and not the data itself entails that the PM2GC line does not trace $y = 1$ for all masses.
similar, highlighting a similar change of the shape of the mass function with redshift.\footnote{This agrees with the results presented in Vulcani et al. (2011), V13, where the authors focused on the shape of the mass function and did not discuss the evolution of the number density.}

For the sim-projected samples, both environments exhibit an increase in the number density of galaxies with redshift. The evolution does not depend strongly on the stellar mass (i.e., the lines are shifted vertically by about the same factor as a function of stellar mass); hence, the shape of the mass functions stays almost constant over cosmic time, and the evolution seen in observations is not reproduced.

6. RESULTS: THE SIMULATED MASS FUNCTIONS

In the previous section, we considered observations and sim-projections, finding that, in the local universe, the semi-analytic model well reproduces the general field and its finer environments, while it fails in reproducing the cluster mass function. So there appears to be a break between groups and clusters. Observationally, groups and clusters have similar mass functions (C13), while in sim-projections, they do not (plot not shown). In this section, we try to more carefully analyze the transition between groups and clusters, exploiting only simulations. In this way, we can control the mass of the halos and inspect whether and at which halo mass differences emerge.

We consider three different halo mass bins: halos of $\log M_{\text{halo}} \sim 13.4$ (least massive halos) roughly correspond to groups, halos of $\log M_{\text{halo}} \sim 14.1$ (intermediate massive halos) correspond to intermediate-mass clusters, and halos of $\log M_{\text{halo}} \sim 15.1$ correspond (most massive halos) to massive clusters. Using only simulations, we can also extend to larger cluster-centric distance and inspect whether galaxies at different distances are regulated by different mass distributions.\footnote{An analysis of this kind is not possible with our observations, since we do not yet have information on the external regions of these structures.} For completeness, we also present a similar analysis at higher redshift. We take advantage of two different semi-analytic models to inspect whether their different recipes produce different results. We consider the DLB07 and G11 models and the samples presented in Section 3.2 and consider all galaxies more massive than $\log M_*/M_\odot = 9.4$.

In Figure 6, we show how the mass function changes across different environments, from the field to massive halos, separately for the two models. We remind the reader that the field includes galaxies in all halos above the resolution of the Millennium at that redshift.

For log $M_*/M_\odot < 11$, in both models and at both redshifts, mass functions in different environments are almost indistinguishable. The DLB07 and G11 models give different predictions for the shape of the mass function: the former predicts...
a steeper mass function than the latter, which flattens out at log $M_\star/M_\odot \approx 10.6$ in all the environments (see also Table 3). In addition, with the adopted normalization, the G11 field shows a systematically lower number density than the other environments considered, in particular at high $z$. The same is true for the field at high $z$ in the DLB07 model, although the effect is less significant.

The different shape of the mass function in the two models is mainly due to the stronger stellar feedback adopted in the G11 model, that, together with the introduction of a model for stellar stripping, contributes to reduce the number of intermediate to low-mass galaxies with respect to the DLB07 model.

Mass distributions differentiate with environment at high stellar masses: the maximum mass reached by galaxies in the most massive halos is higher than that reached by galaxies in lower-mass halos. This is more noticeable at low $z$, where, in both models, the maximum mass varies from log $M_\star/M_\odot = 11.5$ in the least massive systems to log $M_\star/M_\odot = 12.1$ in the most massive ones. At higher redshift, the difference is only 0.2 dex. The field mass function reaches a mass of log $M_\star/M_\odot = 12.5$ at low redshift in both models, and 12.1 (DLB07) and 12.3 (G11) at higher redshift.

The number density of the most massive galaxies increases with the mass of the halos. This is particularly evident in the G11 model and at low redshift where the mass function of halos with log $M_{\text{halo}} \sim 15.1$ exhibits a strong excess of massive galaxies and deviates significantly from the exponential shape provided by the best-fit Schechter function. In this case the analytical fits (Table 3) do not fully describe the measured mass distributions.

In general, Schechter fits show a variation of the parameters from the least massive environment, characterized by larger $\alpha$ and smaller $M^\star$, to the most massive one, characterized by smaller $\alpha$ and larger $M^\star$. The best-fit parameters describing the mass functions are always statistically different at the 3σ level, except for the field, intermediate massive, and most massive halos in the G11 model, where differences are significant only at the 1σ level (see Table 3).

In both models and at all redshifts, the field mass function is not the steepest one: it resembles the function of intermediate
massive halos. This may be the result of the fact that these particular fields are dominated by halos of this mass.\textsuperscript{20}

Hence, in the simulated samples the environment is able to influence both the shape of the mass function and the maximum mass.

### 6.1. Halo-centric Distance

We now focus only on galaxies in halos and analyze if, at any halo mass, simulations predict a dependence of the mass distribution on cluster-centric distance. Some change is expected because going from the center toward the outskirts, galaxies suffer different processes. The external parts of halos can be seen as intermediate and transitional regions between the cores and the field. They contain both galaxies just fallen in the structure and galaxies that, lying on very eccentric orbits, might have experienced more than one passage through the cluster core region.

In Figure 7, we show the mass function of galaxies within $0.6R_{200}$, within $0.6-1R_{200}$, and within $1-3R_{200}$, at low (left panels) and intermediate $z$ (right panels), for halos of different mass and for both models.

In the least massive systems (upper panels), and at both redshifts considered, both the DBL07 and G11 mass functions seem to slightly depend on the halo-centric distance; the innermost regions ($r < 0.6R_{200}$) are characterized by flatter mass functions at intermediate masses ($10 < \log M_*/M_\odot < 11$) with respect to the other regions. In addition, galaxies at $1 < r/R_{200} < 3$ always show a lower number density but this might be due to the adopted normalization (see Appendix B). The parameters of the Schechter fit (see Table 3) support these discrepancies. The mass range spanned by galaxies at the different halo centric distances is the same. Comparing the two models, at high and low masses, the G11 model predicts a smaller number of galaxies than the DBL07 one, and instead at intermediate masses, it predicts a slightly higher number density, in all environments. The consequence is that the DBL07 mass functions are steeper.

In halos of intermediate mass (central panels), at both redshifts and for both models separately, the shapes of the mass function seem very slightly dependent on cluster-centric distance, as also supported by the Schechter fit. Intermediate regions show an excess of low-mass galaxies, making the mass

---

\textsuperscript{20} Given the resolution limit of the simulations, halos of $\log M_{\text{halo}} \sim 12$ are barely resolved in the simulations. If one considers a higher resolution simulation, the result might change.
Figure 7. Simulated mass function of galaxies in structures of the same halo mass but at different cluster-centric distances at low (left panels) and intermediate $z$ (right panels) for the DLB07 (solid symbols and darker shaded areas) and the G11 (empty symbols and lighter shaded areas) models. The mass of the halos is indicated in the upper part of each panel. Red symbols represent galaxies within $0.6R_{200}$, blue symbols represent galaxies within $0.6-1R_{200}$, and green symbols represent galaxies within $1-3R_{200}$. Normalization, error bars on the $y$ axis, lines, shaded areas, and labels are as in Figure 2. Given that the uncertainties on the best-fit parameters are very small, we do not show error contours.

(A color version of this figure is available in the online journal.)
function steeper. Only the external regions are systematically characterized by a smaller number density. This result is qualitatively in agreement with the observational results: V13, exploiting EDisCS data ($M_{\text{halo}} \sim 14$), showed how galaxies within and outside the virialized regions of clusters have a similar mass function. Both models predict that very massive galaxies ($\log M_*/M_\odot \sim 11.9$ at low $z$, $\log M_*/M_\odot \sim 11.7$ at higher) might be found only in the outer regions. This result might be surprising, however, similar findings have been found by De Lucia et al. (2004a), who analyzed the radial distribution of substructures. The radial profile of substructure number density is “anti-biased” relative to the DM profile in the inner regions of halos. The most massive substructures reside preferentially in the outer regions of halos. They suggest that this might be due to the fact that substructures undergo substantial tidal stripping in the dense inner regions of halos.

In the most massive halos (bottom panels), mass function shapes are similar up to $\log M_*/M_\odot \sim 11$, then they differentiate: there is an excess of very massive galaxies in the halo cores. Both the excess and the maximum masses are similar in the two models: $\log M_*/M_\odot \sim 12.1$ at low $z$ and $\log M_*/M_\odot \sim 11.7$ at high $z$. Best-fit parameters are in agreement at the $1\sigma$–$2\sigma$ level, probing on an analytical ground the similarities of these mass functions.

To summarize, the DLB07 and G11 models qualitatively lead to similar results. However, the mass functions obtained from the two models are different. In the low-mass regime, (1) the G11 mass function is flatter than the DLB07 one, (2) the G11 model predicts similar mass functions for galaxies within $0.6R_{200}$ and within $0.6$–$1R_{200}$, while the DLB07 model shows an excess of low-mass galaxies in the outer regions, in the least massive and intermediate massive systems. Finally, the shape of the mass function does not strongly depend on the halo-centric distance, at fixed redshift and halo mass. However, some differences are found for galaxies in the cores of the least massive and most massive systems considered.

### 6.2. Evolution in Different Environments

The last step is to test whether simulations predict an evolution for galaxies in the different environments. For galaxies in halos, we consider only galaxies within the virial radius.

In the field (upper left panel of Figure 8), both models predict an evolution of the high-mass end of the mass functions. In the local universe, galaxies with $\log M_*/M_\odot > 12.2$ are more numerous than in the distant one. Also, the maximum mass evolves, increasing of $\sim 0.4$ dex in the DLB07 model to $\sim 0.2$ dex in the G11 one. In contrast, at low masses, no significant evolution is detected. An analogous evolution in the high-mass end of the mass function, and consequently in $M^*$, is detected in halos in all mass bins considered, though this is less significant probably due to the lower number statistics.

In the lowest halo mass bin considered, no evolution of the maximum mass is detected. In both models and at both redshifts, there are no galaxies more massive than $\log M_*/M_\odot \sim 11.5$. The evolution of the massive end is somewhat more important in the G11 model.

In the intermediate mass halos, the evolution of the shape of the mass function is very similar in the two models but a significant evolution of the maximum mass (from $\log M_*/M_\odot \sim 11.5$ to $\sim 11.9$) is found only in the DLB07 model.

In the most massive halos, differences between mass functions emerge only for $\log M_*/M_\odot > 11.5$. In these systems, the maximum mass also depends on redshift: in the distant universe, there are no galaxies more massive than $\log M_*/M_\odot = 11.7$, while in the local universe, galaxies with $\log M_*/M_\odot \sim 12.1$ are found.

Where the fits describe the mass functions well, they support the results, showing a quite strong evolution of $M^*$ but very little evolution of $\alpha$. As already mentioned, this is not the case for the most massive halos.

To conclude, the shape of the massive end of the mass function evolve, in all environments considered. In contrast, no evolution is found for low-mass galaxies.

### 7. DISCUSSION

#### 7.1. Observations and Sim-projections

In the first part of the paper, we have tested whether simulations are able to reproduce the observational results for the galaxy stellar mass function in different environments, at $z \sim 0$ and $z \sim 0.6$.

Being able to match the field (and its finer environments) mass function in the local universe down to the completeness limit of the observational sample ($\log M_*/M_\odot > 10.25$), the DLB07 model fails in reproducing the observed mass function of field galaxies at higher $z$ and that of clusters at both epochs. In all three cases, the sim-projected mass function is steeper than the observed one in the low-mass regime (the mass at which it occurs is different in the different environments and at the different redshifts), indicating an excess of low-mass galaxies. As far as the high-mass end is concerned, sim-projections predict a slightly higher number density in clusters at low $z$ and a slightly lower number density in both environments at higher $z$ than those observed. We stress that in this mass regime, uncertainties on estimates are large, hence comparisons have to be interpreted with caution.

Our results for the field are in line with several studies. As already mentioned in Section 1, G11 and Bower et al. (2012) have shown that the semi-analytic models can be tuned to reproduce well the $z = 0$ mass function. At higher redshift, e.g., Fontana et al. (2006), Marchesini et al. (2009), Drory et al. (2009), F09, Pozzetti et al. (2010), and G11 have already largely discussed the existing tension between simulations and observations and showed that semi-analytic models reproduce the number density of low and intermediate mass functions. Indeed, they predict a different shape at any mass and redshift explored.

In the studies above, observed and simulated stellar masses have been computed in different ways. Hence, even though F09 pointed out that the disagreement can be only partially resolved by the bias introduced by uncertainties in the mass determination, the comparisons might have been influenced by the different methods followed. In a recent paper, Croton (2013) discussed the riskiness of comparing quantities obtained using different cosmologies and how delicate the conversion is between different $h$ values (see the paper for details). In our approach, we compute stellar masses in the same way for sim-projections and observations, so any discrepancies due to the method are better under control.

The cluster environment has not yet been carefully investigated using semi-analytic models. Weinmann et al. (2011), using the G11 model, compared some properties of galaxies in nearby clusters, without considering stellar masses. They found that abundances, velocity dispersions, and number density profiles are reproduced well by the model. However, simulated clusters could reproduce the red fraction of galaxies only for Coma and
Figure 8. Evolution of the simulated mass function in each environment for the DLB07 (solid symbols and darker shaded areas) and G11 (empty symbols and lighter shaded areas) models. The environment is indicated in the upper part of each panel. Red symbols represent galaxies at $z = 0.06$, and blue symbols represent galaxies at $z = 0.62$. Normalization, error bars on the y axis, lines, shaded areas, and labels are as in Figure 2. Given that the uncertainties on the best-fit parameters are very small, we do not show error contours.

Previous studies did not compare any cluster property at higher redshift. Our results show how the overestimate of low-mass galaxies is even more evident than in the local universe.

7.2. Satellites and Orphans

We note that in clusters, we consider all non-central galaxies as satellites, even though, as explained in Springel et al. (2001) and De Lucia et al. (2004a), models make a distinction between satellites and orphans. The former are galaxies attached to DM substructures, and they were previously the central galaxy of a halo that merged to form the larger system in which they currently reside. The latter are galaxies no longer associated with distinct DM substructures, and their stellar mass is assumed not to be affected by the tidal stripping that reduces the mass of their parent halos. They may later merge into the central galaxy of their halo.

Figure 3 showed that the model also reproduces the observed $M_\ast - \sigma$ relation for satellite+orphan galaxies. A more accurate analysis showed that half of the most massive galaxies are
orphans, and half are satellites. On the other hand, Figure 2 showed that the model does not reproduce the mass function for the same galaxies. Indeed, inspecting the mass distribution of orphans and satellites separately (Figure 9), we find that they are characterized by different mass distributions. Orphan galaxies are characterized by a very steep mass function and they dominate at low masses, while satellite galaxies show a mass function very similar to the observed one (see also Saro et al. 2010). It has been shown that the presence of orphan galaxies is fundamental to reproduce several properties well, i.e., the clustering of the structures and the differences between the galaxy and subhalo profiles in the inner regions of clusters (Gao et al. 2004; Wang et al. 2006). However, the fact that excluding orphans reduces the discrepancies between the observed and sim-projected mass functions is a tantalizing result that could suggest the treatment of orphan galaxies might be improved to help solve the problem.

7.3. The Reasons for the Discrepancies

The excess of low-mass galaxies at high z in the models might be due to the fact that these galaxies are predicted to form too early and have too little ongoing star formation at later times (e.g., F09; Firmani & Avila-Reese 2010; G11; Weinmann et al. 2012). The same discrepancy has also been reported by Weinmann et al. (2012) in two state-of-the-art cosmological hydrodynamical simulations, highlighting that the problem is fundamental. Indeed, they showed that a key problem is the presence of a positive instead of a negative correlation between a specific SFR and stellar mass. A similar correlation also characterizes the specific DM halo accretion rate and the halo mass, indicating that in the models, the growth of galaxies follows the growth of their host halos too closely. Therefore, it is necessary to find a mechanism that decouples the growth of low-mass galaxies, which occurs primarily at late times, from the growth of their host halos, which occurs primarily at early times. Weinmann et al. (2012) then argued that the current form of star formation driven feedback implemented in most galaxy formation models is unlikely to achieve this goal, owing to its fundamental dependence on host halo mass and time.

It has also been suggested that AGN feedback could provide a solution to the “downsizing problem” (Bower et al. 2006; Croton et al. 2006). Indeed, the suppression of late gas condensation in massive halos causes shorter formation time-scales for more massive galaxies (De Lucia et al. 2006), in qualitative agreement with the observed trends. However, as also discussed also by Somerville et al. (2008) showed that the predicted trends may not be as strong as the observed ones, even in the presence of AGN feedback.

F09 showed that the main contributors at all redshifts to the low-mass end excess are low-mass (10^{11} < M_\text{h}/M_\odot < 10^{12}) DM halos. Models predict a roughly constant number density of low-mass central galaxies, while the low-mass satellite population shows a gradual increase which F09 argue is due to the infall of galaxies from the surrounding areas into clusters. This implies that small objects are overproduced while they are central galaxies, and the excess is not primarily due to inaccuracies in the modeling of satellites, even though it manifests in their mass distribution. Therefore, mechanisms that only impact satellite galaxies (such as ram pressure stripping) cannot solve this problem. On the contrary, it would be important to find a physical process that can suppress the star formation in central galaxies hosted by intermediate to low-mass halos (M_\text{h}/M_\odot < 10^{12}).

Hopkins et al. (2013) have recently improved the treatment of the ISM and stellar feedback in a series of high-resolution cosmological simulations. They included cold molecular through atomic, ionized, and hot diffuse ISM and took the stellar feedback inputs (energy, momentum, mass, and metal fluxes) from stellar population models, without free/adjustable parameters. These simulations lead to a reasonable concordance with the stellar mass function for log M_*/M_\odot < 11, suggesting that the effects of stellar feedback on galaxy SFHs and stellar masses might actually be correctly modeled.

7.4. Clusters and Groups

We have shown that at low z, the semi-analytic model correctly reproduces the observed field finer divisions, groups included, at least above the observed completeness limit. In contrast, it cannot reproduce the cluster mass function at similar redshift. This suggests an inaccuracy in the treatment of cluster-specific processes.

Clusters and groups are environments where the decline of gas accretion in galaxies influences galaxy properties but the specific processes that operate in them might be different. However, in the semi-analytic models, they are treated in a similar way. Motivated by the starvation scenario suggested by Larson et al. (1980), all hot gas around the satellites is immediately removed upon infall. In the DLB07 model, the stripped gas is then made available for cooling to the central galaxy of the corresponding structure. This simple prescription leads to satellite galaxies that are too red (e.g., Weinmann et al. 2006b; Wang et al. 2007) compared to observations, suggesting that part of the hot gas should remain with the satellite galaxy. This is not a minor issue, since satellite galaxies constitute a significant fraction of the total galaxy population and since changing the prescription for gas stripping in satellites may have a considerable impact on the central galaxy population. First of all, satellites eventually merge with the central galaxies in their halo. If they can grow to higher stellar masses, central galaxies will become more massive as well. However, in simulations, the merging history of central galaxies seems to be modeled correctly, as shown in Figure 1, where the M_\text{s}−σ relation for central galaxies is well reproduced.
by the model. Then, if part of the hot gas stays attached to satellites, the amount of gas available for cooling to the central galaxy is reduced. Finally, if satellite galaxies merging with the central galaxy still contain cold gas, the ensuing star burst will make the central galaxy bluer for a certain period of time and will result in a higher final stellar mass.

Some attempts to treat environmental effects more realistically have been made, e.g., by Kang & van den Bosch (2008) and Font et al. (2008). The former showed that the decrease of the efficiency of gas stripping in satellites leads to a fraction of blue central galaxies, which is higher than observed but that can be counterbalanced by the inclusion of an additional prescription for the disruption of satellites. The latter implemented a model based on the hydrodynamical simulations of ram-pressure stripping by McCarthy et al. (2008), to obtain better agreement with observed environmental effects than previous semi-analytic models. Okamoto & Nagashima (2003) and Lanzoni et al. (2005) investigated the effect of ram-pressure stripping of the cold disk gas in satellites, concluding that it has a negligible impact on the results, because the complete stripping of the hot gas halo already makes the satellites passive. Weinmann et al. (2010) used the DLB07 model letting the diffuse gas halo around satellites be stripped at the same rate at which the DM subhalo loses mass due to tidal effects. They found that observations at \( z = 0 \) can be reproduced by a simple recipe in which 10–20% of the initial gas reservoir in the halo is stripped per Gyr. They did not focus on higher redshift galaxies, where the DM stripping seems to be more efficient. The same authors also implemented a model in which environmental effects arise solely due to ram-pressure stripping. They showed that this assumption is not exhaustive. Indeed, stripping effects in clusters for low-mass galaxies are too strong to explain the observations, while they are not strong enough to reproduce the observations for high-mass galaxies. These discrepancies could indicate that the DLB07 model, like other models, overestimates the hot gas mass in groups with masses between \( 10^{12} \) and \( 10^{14} h^{-1} M_\odot \), especially in the central, most dense regions of these systems (Bower et al. 2008). This could lead to too strong pre-processing of cluster galaxies in groups.

7.5. Simulations

In the second part of this study, we made use of the DLB07 and G11 semi-analytic models to investigate the dependence of the mass function on environment. We considered field galaxies and galaxies in halos with log \( M_{\text{halo}} \sim 13.4 \) (least massive halos), log \( M_{\text{halo}} \sim 14.1 \) (intermediate massive halos), and log \( M_{\text{halo}} \sim 15.1 \) (most massive halos).

The two models always reach similar conclusions.

In the mass range \( \log M_*/M_\odot = 9.4–10.5 \), galaxies in all of the environments share very similar mass distributions. At higher stellar masses, differences emerge, suggesting an effect of the environment only for \( \log M_*/M_\odot > 10.5 \). Very massive galaxies are hosted only in the most massive halos. As a consequence, the mass distribution characterizing the most massive halos is flatter than those describing galaxies in the other environments. Comparing mass functions at different redshifts, simulations find an evolution only in the massive end, in all environments. Both the shape and the extension of the mass functions change: in the local universe, galaxies can reach higher stellar masses and the number density of massive galaxies is higher than in the distant one.

The dependence of the high-mass end of the mass function on environment and its evolution in the high-mass end are at odds with the observational results (e.g., C13; V13; van der Burg et al. 2013; Ilbert et al. 2010; Pozzetti et al. 2010; Baldry et al. 2012). If they are both real, the discrepancies with the observational results might suggest that, in observations, uncertainties at high masses are still too large to firmly detect the simulated results.

However, mass functions obtained from the DLB07 and G11 models are always different. In general, at \( \log M_*/M_\odot < 10 \), the G11 mass functions are flatter than the DLB07 ones, indicating that in the G11, prescription low-mass galaxies are better treated and their excess is reduced. We remind that the G11 model is an update of the DLB07 model and differs from it mainly for a different treatment of satellite evolution and a more efficient stellar feedback, as presented in Section 3 (see also G11). These implementations reduce the tensions with observations but there are still residuals, indicating that the truncation mechanisms that influence the evolution of central and satellite galaxies still have to be better implemented.

8. SUMMARY AND CONCLUSIONS

In this paper, we exploited two semi-analytic models (DLB07 and G11) to carefully investigate mass functions at different redshifts and in different environments, defined by the halo masses. We have compared the theoretical findings of the DLB07 model with the observational results published in Drory et al. (2009), Vulcani et al. (2011), C13, and V13.

The main results can be summarized as follows.

1. Being able to match the field mass function (and its finer environments) at \( z = 0 \), the DLB07 model fails to reproduce the observed mass function of clusters at low \( z \) and overpredicts the number of low-mass galaxies in both clusters and field at \( z \sim 0.6 \).

2. In sim-projections, the observed invariance of the mass function with the environment is reproduced for galaxies more massive than \( \log M_*/M_\odot = 10.25 \), the mass limit probed by the observations. On the other hand, the observed evolution of the shape of the mass function is not reproduced, neither in the field nor in clusters.

3. Sim-projections reproduce well the observed stellar mass–velocity dispersion relation, both for central galaxies and for the second most massive galaxies. The relation can be ascribed to the environment for the central galaxies at both redshifts and for the second most massive galaxies at low \( z \). On the other hand, it might simply be a statistical effect for the second most massive galaxies at higher redshift.

4. Inspecting only simulations, our results show that in both models, the mass function depends on the mass of the halo, both at \( z = 0.06 \) and \( z = 0.62 \). In very massive halos, there are proportionally more massive galaxies, with the result that the mass function is flatter and \( M^* \) shifted toward higher values compared to lower mass halos. Simulations also detect a mass segregation with the environment: low-mass halos do not host massive galaxies.

5. In both models, the overall shape of the mass function does not strongly depend on the halo-centric distance, once redshift and halo mass are controlled. However, subtle differences might be found when carefully inspecting the least massive systems and the high-mass end of the most massive halos.

6. In both models, the shape of the mass function for \( \log M_*/M_\odot < 11.2 \) does not evolve, in any environment. In contrast, there is an evolution in the number of the most
massive galaxies, which are more numerous in the local universe.

The fact that at low $z$, simulations fail to reproduce the cluster mass function while they are successful in groups, binary systems, and isolated galaxies suggests an incorrect inclusion of environmental processes in the models, such as tidal disruption, ram-pressure, and strangulation. However, the disagreement between observations and simulations in all environments at higher redshifts reveals that cluster physical processes are not the only problem in the models and the redshift-dependent evolution still needs to be better modeled.

This work was supported by World Premier International Research Center Initiative (WPI), MEXT, Japan. The Millennium Simulation databases used in this paper and the web application providing online access to them were constructed as part of the activities of the German Astrophysical Virtual Observatory (GAVO). G.D.L. acknowledges financial support from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement No. 202781. We are grateful to Ulisse Munari for providing us the conversions between different photometric systems. We thank Niv Drory for providing COSMOS stellar masses and the field mass functions, and Fabio Fontanot for useful discussions.

APPENDIX A

CONSISTENCY CHECKS ON STELLAR MASSES

A.1. Drory et al. (2009) Stellar Masses

Drory et al. (2009) derived stellar masses comparing multiband photometry to a grid of stellar population models of varying SFHs, ages, and dust content (for further details, refer to Drory et al. 2009). They adopted a Chabrier (2003) IMF (0.1–100 $M_\odot$).

To reduce the uncertainties due to the different method adopted by Drory et al. (2009), we computed stellar masses for the COSMOS sample using Bell & de Jong (2001) and then compared them to the stellar masses (N. Drory 2013, private communication) used by Drory et al. (2009). The comparison of the different mass estimates is presented in Figure 10. In addition to a large scatter, Drory et al. (2009) mass estimates are systematically smaller. Discrepancies are around $\sim 0.2$ dex, which corresponds to the typical error on the mass estimates. In our analysis, we apply a mean correction to the Drory et al. (2009) mass function as a function of stellar mass, taking into account the median difference between the two methods.

A.2. DLB07 Stellar Masses

In our analysis, when semi-analytic model data are compared to observations, we adopt stellar masses computed using the Bell & de Jong (2001) formulation (Equation (2)), as done for the observations.

Figure 11 shows the comparison between the two mass estimates, both obtained assuming a Kroupa (2001) IMF. Above $\log M_*/M_\odot = 9.8$ (our lowest observed mass limit), the estimates are in agreement within the typical scatter of the mass uncertainties ($\sim 0.2–0.3$ dex, see, e.g., Kannappan & Gawiser 2007). The median difference of the masses is $\sim -0.08$ at $z = 0.06$ and $\sim +0.05$ at $z = 0.62$.

APPENDIX B

THE NORMALIZATION OF THE MASS FUNCTIONS

To compare mass functions in different environments, it is important to adopt normalizations that will allow a meaningful comparison. This is not always straightforward, since the number densities in clusters and in the field are expected to be very different. In numerical simulations of collisionless matter, cluster halos are defined as spherical regions enclosing a density that is 200 times the critical density $\rho_c(z)$ of the universe. Therefore, the density of matter measured within cluster-sized halos will be much larger than the average matter density. If galaxies are assumed to follow the DM distribution, then clusters should also have the same overdensity of galaxies within their boundaries with respect to the density of galaxies in the field. Our aim here is to remove this principal normalization difference to compare galaxy stellar mass functions in different environments. There could be additional corrections if galaxies do not exactly follow the DM distribution (e.g., if the number density profile around a cluster does not follow a Navarro–Frenk–White profile or shows luminosity segregation), which we ignore for simplicity.

The volume normalization adopted for the field allows us to express the counts per unit comoving volume. To obtain the normalization for the cluster counts, note that $\rho_c(z)$ is related to the matter density $\bar{\rho}(z)$ as

$$\rho_c(z) = \bar{\rho}(z) \times [\Omega_0(1+z)^3 + \Omega_\Lambda]/\Omega_0(1+z)^3 ,$$  

where $\Omega_0$ is the matter density parameter and $\Omega_\Lambda$ is the dark energy density parameter, and we have assumed a flat universe. The number density of galaxies within $R_{200}$, should therefore be diluted by a factor

$$200 \times [\Omega_0(1+z)^3 + \Omega_\Lambda]/\Omega_0(1+z)^3 ,$$  

while comparing to the number density in the field.
When we consider smaller regions in clusters (e.g., $r < 0.6R_{200}$), we have to also take into account the density variation as a function of the radius, as described by a Navarro et al. (1997) profile. The mass enclosed within a radius $X = r/r_\text{c}$ is

$$M(< X) = M_{200} \times \frac{\mu(X)}{\mu(c_{200})} = \frac{4\pi}{3} R_{200}^3 \rho_\text{c}(z) \times \frac{\mu(X)}{\mu(c_{200})},$$

where $c_{200} = R_{200}/r_\text{c}$ is the concentration parameter, $r_\text{c}$ is a scale factor, and $\mu(X) = \ln(1 + X) - X/(1 + X)$. For relaxed halos, $c_{200}$ mildly depends on the halo mass (see, e.g., Macciò et al. 2008). However, for our purpose, we can assume $c_{200} = 5$.

The density of matter within a radius $X$ is

$$\rho(< X) = \frac{3}{4\pi} \frac{M(< X)}{X^3} = \frac{3}{4\pi} \frac{M_{200} \times \mu(X)}{M_{200} \times \mu(c_{200})} = \frac{\rho_\text{c}(z) \times \mu(X)}{\mu(c_{200})}.$$ 

(B4)

Therefore, using Equation (B1), the number counts within radius $X$ should be scaled by the factor

$$200 \times \frac{3}{4\pi} \frac{\mu(X)}{\mu(c_{200})} \times \frac{\rho_\text{c}(z) \times \mu(X)}{\mu(c_{200})},$$

(B5)

to compare with the number density in the field.

Similarly, when we consider a shell delimited by $X_2$, $X_1$ (e.g., $1R_{200} < r < 3R_{200}$), the appropriate normalization factor to use is

$$200 \times \frac{\rho_\text{c}(z) \times \mu(X_2) - \rho_\text{c}(z) \times \mu(X_1)}{\mu(c_{200})} \times \Omega_{\Lambda}/\Omega_0 (1+z)^3.$$

(B6)

**REFERENCES**

Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469

Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, MNRAS, 421, 621

Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, MNRAS, 388, 945

21 In our halo mass range $c_{200}$ varies from 4 to 5.5, and this does not change our results.
