Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **456,** 3827–3839 (2016) doi:10.1093/mnras/stv2919

The environments of high-redshift radio galaxies and quasars: probes of protoclusters

Álvaro A. Orsi,[1,2]★ Nikos Fanidakis,[3] Cedric G. Lacey[4] and Carlton M. Baugh[4]

[1]*Centro de Estudios de Física del Cosmos de Aragón, Plaza de San Juan 1, Teruel E-44001, Spain*
[2]*Instituto de Astrofísica, Pontificia Universidad Católica, Av. Vicuña Mackenna 4860, Santiago, Chile*
[3]*Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany*
[4]*Institute for Computational Cosmology, Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK*

Accepted 2015 December 9. Received 2015 November 23; in original form 2015 October 03

ABSTRACT

We use the GALFORM semi-analytical model to study high-density regions traced by radio galaxies and quasars at high redshifts. We explore the impact that baryonic physics has upon the properties of galaxies in these environments. Star-forming emission-line galaxies (Ly α and H α emitters) are used to probe the environments at high redshifts. Radio galaxies are predicted to be hosted by more massive haloes than quasars, and this is imprinted on the amplitude of galaxy overdensities and cross-correlation functions. We find that Ly α radiative transfer and active galactic nucleus feedback indirectly affect the clustering on small scales and also the stellar masses, star formation rates and gas metallicities of galaxies in dense environments. We also investigate the relation between protoclusters associated with radio galaxies and quasars, and their present-day cluster descendants. The progenitors of massive clusters associated with radio galaxies and quasars allow us to determine an average protocluster size in a simple way. Overdensities within the protoclusters are found to correlate with the halo descendant masses. We present scaling relations that can be applied to observational data. By computing projection effects due to the wavelength resolution of modern spectrographs and narrow-band filters, we show that the former have enough spectral resolution to map the structure of protoclusters, whereas the latter can be used to measure the clustering around radio galaxies and quasars over larger scales to determine the mass of dark matter haloes hosting them.

Key words: methods: numerical – galaxies: evolution – galaxies: high-redshift – large-scale structure of Universe.

1 INTRODUCTION

A fundamental ingredient of our understanding of galaxy formation and evolution is the cosmological growth of structure in the Universe. The hierarchical growth arises from the non-linear evolution of the dark matter (DM) density field under the action of gravity (for a review, see Springel, Frenk & White 2006). In this scenario, galaxies form in gravitational potential wells (i.e. DM haloes) that allow gas to cool and form stars, populating the cosmic web (White & Rees 1978). As a result, galaxies form and evolve in a diverse range of environments, from voids (or regions with densities well below the average), to field galaxies residing in average environments and to highly overdense regions.

Observational studies of galaxy properties in different environments have shown unequivocally that their properties are somehow connected with the environment in which they reside (e.g. Oemler 1974; Dressler 1980; Hashimoto et al. 1998; Kauffmann et al. 2004). Galaxy clusters, the most massive virialized structures in the Universe, have galaxy number densities of up to a few hundred times higher than the average. In such high-density environments, galaxies follow a well-known morphology–density relation (e.g. Dressler 1980; Balogh et al. 1997; Goto et al. 2003). This indicates that red, early-type galaxies are more abundant than blue, late-type galaxies in these environments.

Different baryonic processes are thought to be responsible for the transformation of galaxies in these environments, such as ram-pressure stripping and tidal interactions (Moore et al. 1996; Kawata & Mulchaey 2008; van den Bosch et al. 2008; Tecce et al. 2010). More recently, the quenching of star formation due to active galactic nucleus (AGN) feedback has been proposed as a key mechanism regulating the star formation of galaxies hosted by massive haloes (e.g. Bower et al. 2006; Croton et al. 2006; Lagos, Cora & Padilla 2008).

From a cosmological perspective, these large structures we see in the local Universe are expected to grow from the aggregation of

★E-mail: aaorsi@cefca.es

smaller haloes formed at earlier epochs (Press & Schechter 1974; Peebles & Shaviv 1982; Davis et al. 1985; Lacey & Cole 1993). The hierarchical structure formation scenario implies that massive DM haloes are likely to be embedded in overdense regions. This suggests that any galaxy property related to halo mass will also be correlated with environment on scales beyond its host halo, i.e. in superhalo scales (e.g. Mo et al. 2004). This is particularly important from the perspective of the formation of galaxy clusters, i.e. in the protoclusters regime (Overzier et al. 2006). These massive structures at high redshifts will eventually form a bound cluster at late times, and thus galaxies are expected to have properties that are connected to their host halo and its surroundings.

The search for protoclusters at high redshifts can thus be a key to unveiling the role of hierarchical merging and baryonic processes in galaxy formation and evolution. Observationally, detecting these structures has proven challenging. Common techniques used to detect nearby clusters, such as looking for the thermal X-ray emission from the intracluster medium are not sensitive enough at $z \gtrsim 2$, where clusters are expected to be in the early stages of formation. Searches for galaxy overdensities in wide-field surveys are more effective, but unless the galaxy redshifts are confirmed spectroscopically, the observed overdensities are likely to be affected by redshift uncertainties that are larger than the physical size of the protocluster itself (Chiang, Overzier & Gebhardt 2013). In addition, projection effects can smear out the real overdensity signal, or even enhance it depending on the viewing angle (Shattow et al. 2013). This makes the identification of overdense structures at high redshifts controversial.

Instead of looking for overdensities in a blind survey, it is also possible to map the environments around objects that are good candidates to lie in a density peak. Generally speaking, the most luminous galaxies are found to be highly clustered, meaning that these are hosted by massive haloes (e.g. Norberg et al. 2001). High-redshift quasars are among the brightest objects in the Universe, so it is commonly assumed that these trace the most massive structures as well (Steidel et al. 2005; Kashikawa et al. 2007; Overzier et al. 2009; Utsumi et al. 2010; Bañados et al. 2013; Husband et al. 2013). Other luminous objects used to find overdensities are the so-called Ly α blobs, which can typically reach luminosities above 10^{43} [erg s^{-1} h^{-2}] over extended regions of hundreds of kpc (Erb, Bogosavljević & Steidel 2011; Matsuda et al. 2011; Uchimoto et al. 2012). A third candidate for tracers of massive structures are radio galaxies, which can also have typical stellar masses of the order of 10^{12} [M$_\odot$ h^{-1}] (Kauffmann, Heckman & Best 2008; Miley & De Breuck 2008; Donoso et al. 2010; Falder et al. 2010; Ramos Almeida et al. 2013; Hatch et al. 2014; Karouzos, Jarvis & Bonfield 2014). Fanidakis et al. (2013) showed that, in the context of a hierarchical galaxy formation model, high-redshift quasars are neither hosted by the most massive haloes, nor are they the progenitors of the most massive clusters that we observe today. Instead, the model predicts that luminous radio galaxies reside in very massive haloes, and are thus better tracers of the most massive protoclusters. In this paper, we make use of the same galaxy formation model described in Fanidakis et al. (2013) to characterize the environments around these two types of active galaxies.

In order to map the environment around a protocluster candidate the redshifts of the objects must be known with sufficient accuracy. Typically, this is achieved with spectroscopic follow-up of a sample of photometric candidates, or by utilizing a narrow-band filter chosen to look for a specific emission line at the redshift of the protocluster. Using Ly α emitters to map the environment in a protocluster, Venemans et al. (2007) found that radio galaxies seem to pinpoint structures with masses $\sim 10^{14–15}$ M$_\odot h^{-1}$ over the redshift range $2 < z < 5$. Likewise, Saito et al. (2015) determine an overdensity of Ly α emitters around a radio galaxy at $z \sim 4$ that is only marginally reproduced in the galaxy formation model of Orsi et al. (2008).

Here, we explore the properties and evolution of overdense regions by focusing specifically on those traced by radio galaxies and quasars in a theoretical framework. Radio galaxies are expected to trace haloes that are subject to star formation quenching due to AGN feedback in massive haloes. Quasars, on the other hand, are characterized by a rapid accretion of cold gas triggered by mergers or disc instabilities, which places them in a broader range of halo masses. The most luminous quasars also experience star formation quenching due to AGN feedback (Fanidakis et al. 2013).

In this paper, we tackle two main problems related to these two types of active galaxies and their environments. First, we characterize the impact of baryonic processes on samples of emission-line galaxies (ELGs) populating overdense regions at high redshifts traced by radio galaxies and quasars. Secondly, by identifying the haloes that host radio galaxies and quasars we derive a simple way to define the size of protoclusters, and link their overdensities to the mass of the descendant haloes at $z = 0$.

The backbone of this study is the GALFORM semi-analytical model (Cole et al. 2000; Baugh et al. 2005; Bower et al. 2006; Lagos et al. 2011). This model incorporates state-of-the-art prescriptions for the evolution of galaxies and their central supermassive black holes (BH; Fanidakis et al. 2011, 2013), and at the same time uses a Monte Carlo radiative transfer code for Ly α photons to derive physically robust Ly α luminosities for high-redshift galaxies (Orsi, Lacey & Baugh 2012). GALFORM predicts observational properties of galaxies, attempting to include all relevant physical mechanisms in the galaxy formation and evolution process. Hence, our choice of specific tracers of high-density regions at high redshifts, and the galaxy populations used to measure the environment around them, are all predictions that our model provides within a robust framework, making it suitable to be confronted against observational measurements.

The structure of this paper is as follows. Section 2 describes the galaxy formation model used; Section 3 describes our results analysing both overdense regions and defining protoclusters traced by radio galaxies and quasars. Finally, Section 4 summarizes our main findings and discusses their implications.

2 THE THEORETICAL GALAXY FORMATION MODEL

This section describes the semi-analytical model of galaxy formation used, its main features, and the modelling of active galaxies and emission lines in star-forming galaxies.

We make use of the GALFORM semi-analytical model of galaxy formation to predict the properties of galaxies as a function of redshift. This model is described in detail elsewhere (Cole et al. 2000; Benson et al. 2003; Baugh et al. 2005; Bower et al. 2006; Lagos et al. 2011). The variant of GALFORM used here is presented in Lacey et al. (2015).

In short, GALFORM computes the formation and evolution of galaxies in the context of the hierarchical growth of DM structures. The properties and merging histories of DM haloes are extracted from the Millennium-*Wilkinson Microwave Anisotropy Probe (WMAP7)* DM only *N*-body simulation. The halo mass resolution is $\sim 10^{10}$ [M$_\odot$ h^{-1}] and its box-side length is 500 [Mpc h^{-1}].

Hence, this simulation is similar to the well-known Millennium run (Springel et al. 2005), except that it was run with a set of updated cosmological parameters taken from Komatsu et al. (2011) obtained using the *WMAP*7 data set, i.e. $\Omega_{\rm b} = 0.0455$, $\Omega_{\rm M} = 0.272$, $\Omega_\Lambda = 0.728$, $n_{\rm s} = 0.967$, $\sigma_8 = 0.810$, and $h = 0.704$.

The main baryonic processes that enter in the GALFORM calculation are (i) the shock heating and radiative cooling of gas inside haloes leading to the formation of a disc, (ii) quiescent star formation in the disc, and starbursts in a galactic bulge following galaxy mergers and disc instabilities, (iii) feedback due to supernovae, AGN and photoionization which regulate the star formation process, and (iv) the chemical enrichment of the gas and stellar component. Galaxy luminosities are computed using a population synthesis model (Cole et al. 2000; Gonzalez-Perez et al. 2014). Dust extinction for the stellar continuum is calculated self-consistently based on the radiative-transfer model described in Lacey et al. (2011).

The variant of GALFORM used here incorporates features from different versions of the model that were used to study specific problems into a single model to provide a powerful galaxy formation tool. Most notably, the model invokes a different initial mass function (IMF) for quiescent and starburst events. A top-heavy IMF is used in the latter case to explain the abundance of high-redshift galaxies detected in the sub-mm (Baugh et al. 2005). The model also includes a treatment of star formation in discs following the atomic and molecular hydrogen content of the gas (Lagos et al. 2011), and stellar luminosities using the Maraston (2005) stellar population synthesis model that incorporates the contribution from thermally pulsating asymptotic giant branch stars (see also Gonzalez-Perez et al. 2014).

2.1 Modelling ELGs

In order to obtain the line fluxes of several hydrogen recombination lines, GALFORM computes the total production rate of hydrogen ionizing photons (Lyman continuum photons) by integrating the composite spectral energy distribution of each galaxy over the extreme-ultraviolet continuum down to the Lyman break at $\lambda = 912$ Å. Then, by assuming that all of these ionizing photons are absorbed within the interstellar medium (ISM) of the galaxy (i.e. the escape fraction of ionizing photons is set to zero), case B recombination is used to convert a fraction of the Lyman continuum photons into different line fluxes (Osterbrock 1989; Dijkstra 2014). This intrinsic luminosity is later adjusted for the effect of dust attenuation by computing the continuum extinction at the wavelength of the line.

We apply the procedure outlined above to obtain the H$\,\alpha$ luminosities of galaxies. In the case of Ly$\,\alpha$, the intrinsic luminosity of Ly$\,\alpha$ photons is expected to be reduced by the scattering of Ly$\,\alpha$ photons by neutral hydrogen atoms in the ISM and their absorption by dust grains. The high-scattering cross-section of photons at the Ly$\,\alpha$ line centre makes these photons undergo numerous scattering events with hydrogen atoms, resulting in large path-lengths, and thus they are likely to be absorbed by dust grains present in the ISM. This results in a complex radiative transfer problem that cannot be accurately accounted for using analytical expressions, except for a few idealized configurations (Harrington 1973; Neufeld 1990; Dijkstra, Haiman & Spaans 2006).

Instead, we compute the escape of Ly$\,\alpha$ photons using a Monte Carlo radiative transfer code for Ly$\,\alpha$ photons that allows us to obtain a value for the Ly$\,\alpha$ escape fraction, $f_{\rm esc}$, which leads to the observed Ly$\,\alpha$ luminosity. A full description of the radiative transfer code and its implementation in GALFORM can be found in Orsi et al. (2012). In short, the code follows the scattering, absorption and escape histories of a large number of photons (in this case, of the order of 10^3 to 10^4 for each galaxy). Each photon can change its direction and frequency after interacting with hydrogen atoms and dust grains. If the former interaction occurs, then the photon is scattered, changing its direction, which leads to a change of frequency. If the latter takes place, then the photon can be scattered or absorbed, depending on the dust albedo. If a photon is absorbed by a dust grain, then it is discarded. If a photon escapes from the medium, its final frequency is recorded. The process is repeated until an accurate value for the escape fraction of Ly$\,\alpha$ photons is attained. Following Orsi et al. (2012), we restrict the total number of photons for a given run to be at least 10^3, and up to 5×10^4 when no photon was absorbed. Hence, the minimum escape fraction our model can compute is $f_{\rm esc} = 2 \times 10^{-5}$.

There is much observational evidence for the presence of outflows in Ly$\,\alpha$ emitters at high redshifts (e.g. Giavalisco, Koratkar & Calzetti 1996; Thuan & Izotov 1997; Kunth et al. 1998; Mas-Hesse et al. 2003; Shapley et al. 2003; Kashikawa et al. 2006; Hu et al. 2010; Kornei et al. 2010). Hence, Orsi et al. (2012) assumed simple outflow geometries for the Ly$\,\alpha$ photons to escape. For simplicity, two isothermal, spherically symmetric models of galactic-scale outflows were adopted: an expanding thin shell and an expanding wind. Both are similar and their properties, such as expansion velocity, size and metallicity are directly proportional to the galaxy's predicted cold gas mass, circular velocity, half-mass radius and cold gas metallicity, respectively. In addition, the wind geometry displays a gas density profile of the form $\rho(r) \propto r^{-2}$, where the normalization depends on the mass-ejection rate predicted by the supernova feedback implemented in GALFORM. For simplicity, we will hereafter use the thin shell geometry. We have checked that our predictions are not sensitive to the choice of the outflow geometry. Since the Orsi et al. (2012) model was developed using an earlier version of GALFORM (described in Baugh et al. 2005), we re-calibrated the free parameters controlling the relation between the half-mass radii of galaxies and the inner radius of the outflows for each galaxy. The new parameter values are found by matching the luminosity function of Ly$\,\alpha$ emitters in the redshift range $0.2 < z < 6.6$ (for details of the fitting procedure, see Orsi et al. 2012). For starbursts, this relation is described by

$$R_{\rm inner} = k(1+z)^{\gamma} \langle R_{1/2} \rangle, \quad (1)$$

where $k = 0.2$, $\gamma = 1$, and $\langle R_{1/2} \rangle$ is the average of the half-mass radii of the disc and bulge components of the galaxies, weighted by their intrinsic Ly$\,\alpha$ luminosity. For quiescent galaxies, $R_{\rm inner} = 2\langle R_{1/2} \rangle$.

2.2 Modelling of radio galaxies and quasars

To model radio galaxies, we use the AGN prescriptions described in Fanidakis et al. (2011). The Fanidakis et al. (2011) model follows the mass accretion rate on to the BHs and the evolution of the BH mass, $M_{\rm BH}$, and spin, a, allowing the calculation of a variety of predictions related to the nature of AGN. In this model, the evolution of BHs and their host galaxies is fully coupled: BHs grow during the different stages of the evolution of the host by accreting cold gas (merger/disc-instability driven accretion: starburst mode) and hot gas (diffuse halo cooling driven accretion: hot-halo mode) and by merging with other BHs. This builds up the mass and spin of the BH, and the resulting accretion power regulates the gas cooling and subsequent star formation in the galaxy. The resulting mass of the BH correlates with the mass of the galaxy bulge in agreement with the observations (see Fanidakis et al. 2011, 2012).

The BH spin distribution depends strongly on how the gas in a given accretion episode accretes on to the BH. Fanidakis et al. (2011) assume that the accretion flow fragments due to self-gravity into multiple accretion episodes (chaotic accretion; King 2005; King, Pringle & Hofmann 2008). In this case, star formation in the vicinity of the BH can randomize the angular momentum of the gas, resulting in a succession of randomly aligned accretion discs around the BH. The end effect of this process is typically a BH with a low spin. High-spin values occur only for the most massive BHs ($M_{\rm BH} > 10^8\,{\rm M}_\odot$), because the growth of these BHs is dominated by gas-poor BH–BH mergers which always result in fairly rapid spins of $a \sim 0.7$–0.9. Thus, in the chaotic accretion scenario there is a clear correlation of spin with BH mass and hence with host galaxy bulge mass. Massive BHs form in the most massive DM haloes, they are hosted by massive elliptical galaxies and have rapid spins, while lower mass BHs form in spiral galaxies and have much lower spins.

The gas accreted during a starburst episode is converted into an accretion rate $\dot{M}$. The bolometric luminosity of the accretion flow associated, $L_{\rm bol}$, will depend on the accretion rate in Eddington units $\dot{m} = \dot{M}/\dot{M}_{\rm Edd}$. If $\dot{m} \geq 0.01$, then the thin disc solution of Shakura & Sunyaev (1973) is used

$$L_{\rm bol} = \epsilon \dot{M} c^2, \tag{2}$$

where c is the speed of light, and ϵ is an adjustable parameter. Otherwise, the advection-dominated accretion flow (ADAF) thick disc solution is adopted (Narayan & Yi 1994),

$$L_{\rm bol,ADAF} = 0.44 \left(\frac{\dot{m}}{0.01}\right) \epsilon \dot{M} c^2. \tag{3}$$

Finally, if the accretion is super-Eddington ($L_{\rm bol} \geq \eta L_{\rm Edd}$), then the bolometric luminosity is obtained as (Shakura & Sunyaev 1973)

$$L_{\rm bol}(\geq \eta L_{\rm Edd}) = \eta[1 + \ln(\dot{m}/\eta)] L_{\rm Edd}, \tag{4}$$

where η is an adjustable parameter of the model.

The mass, spin and mass accretion rate evolution are then coupled to the classic Blandford–Znajek jet model (Blandford & Znajek 1977). The jet power couples strongly to the accretion mode, as it most likely depends on the vertical (poloidal) magnetic field component $B_{\rm p}$ close to the BH horizon, $P_{\rm jet} \sim B_{\rm p}^2 M_{\rm BH}^2 a^2$. The accretion flow is assumed to form a geometrically thin disc for relatively high accretion rates (Shakura & Sunyaev 1973), switching at lower accretion rates to a geometrically thick disc in an ADAF (Narayan & Yi 1994). The expression for the mechanical jet energy in each regime is then (Meier 2002)

$$P_{\rm jet,ADAF} = 2 \times 10^{45} M_9 \left(\frac{\dot{m}}{0.01}\right) a^2\ {\rm erg\ s^{-1}}, \tag{5}$$

$$P_{\rm jet,TD} = 2.5 \times 10^{43} M_9^{1.1} \left(\frac{\dot{m}}{0.01}\right)^{1.2} a^2\ {\rm erg\ s^{-1}}, \tag{6}$$

where M_9 is the mass of the BH in units of $10^9\,{\rm M}_\odot$, a is the spin of the BH. The collapse by two orders of magnitude in scaleheight of the flow during the transition from an ADAF to a thin disc results in a similar drop in radio power. This already gives a dichotomy in radio properties which may explain some of the distinction between radio loud and radio quiet objects. Using these prescriptions for the accretion disc and jet, we calculate the optical and radio output from accreting BHs. The model fits the luminosity function of radio-loud AGN remarkably well when low-mass objects have lower jet powers than high-mass objects. This is achieved because the jet couples strongly to the BH spin and the lower mass BHs have lower spin than the most massive BHs. Overall, the model predictions for the AGN population can reproduce the diversity of nuclear activity seen in the local and high-z Universe (Fanidakis et al. 2011, 2012). The predictions for the radio-optical luminosities of Fanaroff Riley type I, broad emission line radio galaxy, Seyfert, and low-ionization nuclear emission-line region are in good agreement with the observed galaxy populations.

Based on the disc and jet luminosity the model calculates for every accreting BH, we define as a quasar any galaxy whose central BH produces a disc luminosity higher than $10^{46}\,{\rm erg\ s^{-1}}$. In contrast, a galaxy is defined as radio galaxy when its jet luminosity exceeds $P_{1.4{\rm GHz}} = 10^{23}\,{\rm W\,Hz^{-1}}$.

3 RESULTS

We now explore the predictions of GALFORM in the redshift range $2 < z < 6$. This is the redshift interval where the bulk of the observational work about the environments around radio galaxies and quasars using Ly α emitters has been carried out. Compiling statistical samples of Ly α emitters is challenging at low redshifts ($z \lesssim 2$), and has only been possible with the *GALEX* satellite (e.g. Deharveng et al. 2008; Cowie, Barger & Hu 2010; Wold, Barger & Cowie 2014). At $z > 6$, our model predicts that radio galaxies are very rare, such that it is not possible to robustly characterize the population.

Throughout this paper, we measure overdensities around radio galaxies and quasars in redshift space. This means that the peculiar velocity of galaxies in the radial direction contributes to the observed redshift of galaxies, producing a distortion of their derived comoving distances. To introduce this effect into our model, we take the comoving z-coordinate, r_z, to represent the line-of-sight direction and replace it by its value in redshift space s:

$$s = r_z + \frac{v_z}{aH(z)}, \tag{7}$$

where v_z is the peculiar velocity of the galaxy along the line of sight, a is the expansion factor and $H(z)$ the Hubble parameter evaluated at redshift z.

Most of our predictions focus on four redshifts $z = 2.2$, 3.0, 4.5, and 5.7. These are the redshifts at which Ly α has been detected from the ground with negligible atmospheric contamination. In addition, $z = 2.2$ is particularly important since it is also the redshift at which a ground-based near-infrared instrument can typically search for H α emitters (e.g. Geach et al. 2008; Koyama et al. 2013), although H α searches have extended up to $z \sim 2.5$ (Cooke et al. 2014).

Fig. 1 shows the environment around selected radio galaxies and quasars within the redshift range discussed above, as predicted by GALFORM. Each image displays a cuboid of $20 \times 20 \times 10\,[{\rm Mpc}^3 h^{-3}]$, which illustrates the complicated filamentary structure of the DM, shown in grey. Ly α emitters with luminosities $\log(L_{{\rm Ly}\,\alpha}[{\rm erg\ s^{-1}}\ h^{-2}]) > 41$ are shown in blue.

Radio galaxies and quasars trace different environments as a consequence of their different triggering mechanisms. Quasar activity is triggered in gas-rich galaxies, typically associated with haloes of mass $10^{12}[\,{\rm M}_\odot\,h^{-1}]$ (Fanidakis et al. 2013). For radio galaxies, we find that the brightest objects live in the most massive haloes because this is where the spin and mass of BHs are higher, but also the accretion rate is low enough to form an ADAF which gives powerful jets, as discussed in the previous section.

Fig. 2 shows the mass function of haloes that host radio galaxies and quasars at different redshifts. Both types of AGN can be found in massive haloes. However, they represent a small subset of

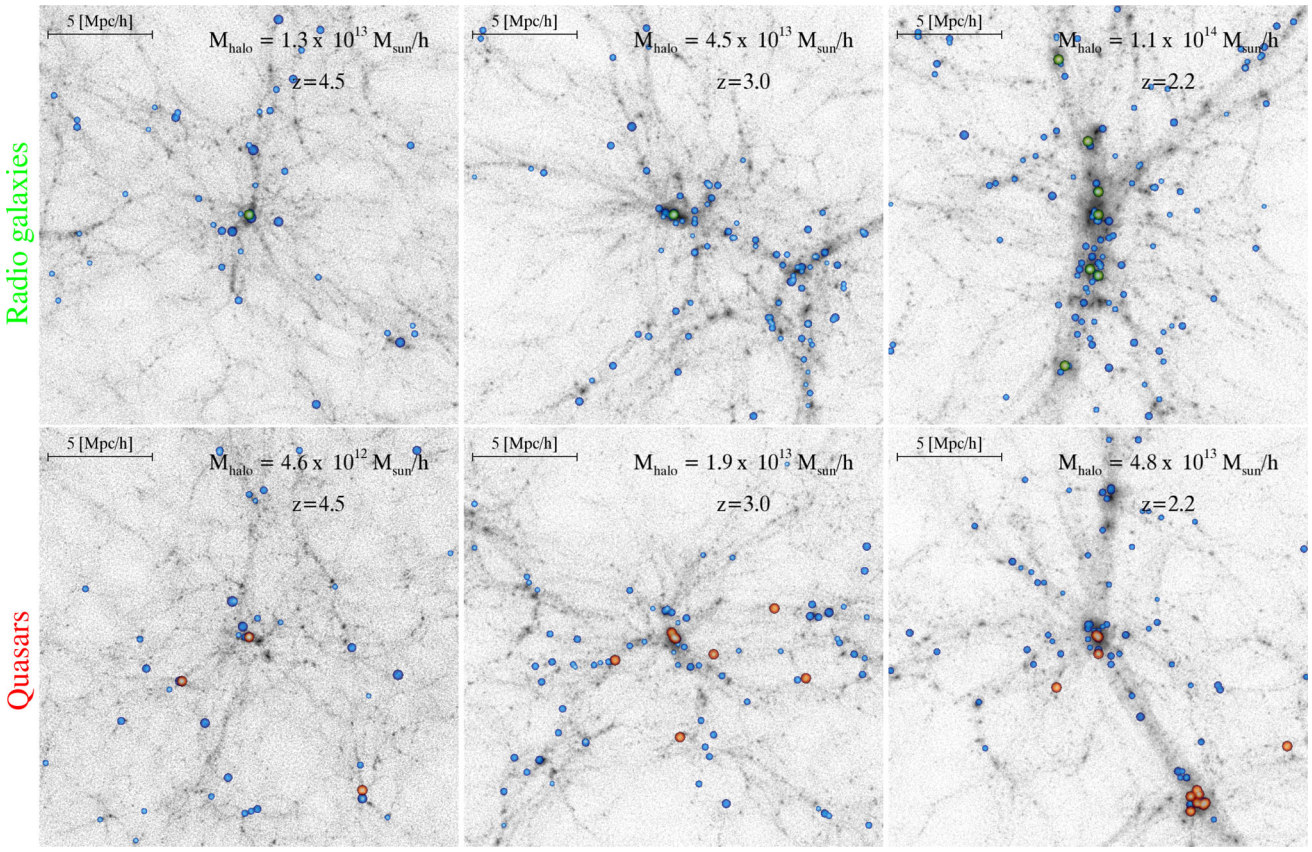

Figure 1. The predicted spatial distribution of Ly α emitters around typical radio galaxies (top) and quasars (bottom) at redshifts 4.5, 3.0, and 2.2 (from left to right). Each box shows a slice of 20 [Mpc h^{-1}] on a side and 10 [Mpc h^{-1}] depth. The DM density distribution is displayed in grey-scale. Darker pixels indicate a higher density of DM. Green circles show the position of the radio galaxies. Orange circles show the position of the quasars. Blue circles show the spatial distribution of Ly α emitters with $L_{\mathrm{Ly}\alpha} > 10^{41}$[erg s^{-1} h^{-2}].

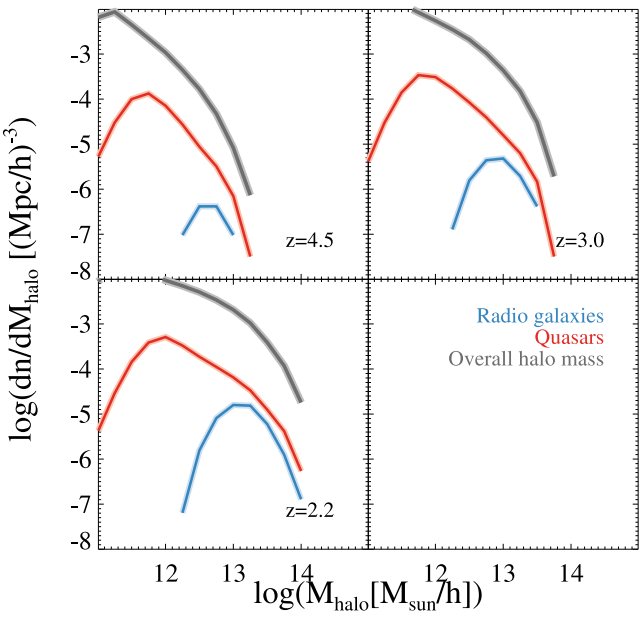

Figure 2. The halo mass function of radio galaxies (blue) and quasars (red) in our model. The grey line corresponds to the halo mass function of the overall galaxy population. Panels correspond to redshifts $z = 4.5$, 3.0 and $z = 2.2$ as labelled.

the population of massive haloes at any redshift. Furthermore, the fraction of haloes hosting a radio galaxy or quasar at a given halo mass is predicted to increase with redshift: at $z = 2.2$, radio galaxies and quasars account for only 4.5 per cent of the haloes with mass above $M_{\mathrm{halo}} > 10^{13}[\mathrm{M}_{\odot}\, h^{-1}]$. At $z = 3.0$ this fraction increases to 5.2 per cent, and at $z = 4.5$ to 12 per cent.

Quasars are significantly more abundant than radio galaxies at all redshifts and halo masses, and they also span a larger range of halo mass, peaking at $M_{\mathrm{halo}} \sim 10^{11.5}-10^{12}[\mathrm{M}_{\odot}\, h^{-1}]$. Radio galaxies, on the other hand, populate a small subset of the most massive haloes, peaking above $M_{\mathrm{halo}} \sim 10^{13}[\mathrm{M}_{\odot}\, h^{-1}]$.

The environments of these two types of AGN are mostly dominated by this fundamental difference in their halo mass distribution. Furthermore, baryonic processes that are important over these halo mass ranges can also have an impact on the properties of the galaxy population used to trace the environment and overdensities.

Observations and models have shown that star-forming galaxies tend to *avoid* the centres of massive structures (e.g. Orsi et al. 2010; Contreras et al. 2013). This is due to the star formation quenching mechanisms that are expected to act in overdense regions. The typical star formation time-scale that is traced by nebular emission is of the order ∼10 Myr. Hence, any baryonic process that results in an abrupt quenching of star formation, such as AGN feedback, is expected to affect the line luminosities of galaxies, given its short time-scale. Hence, the properties of ELGs around a massive structure will be related not only to the environment itself, but also to the depth (i.e. the limiting flux) of the galaxy sample used to trace such

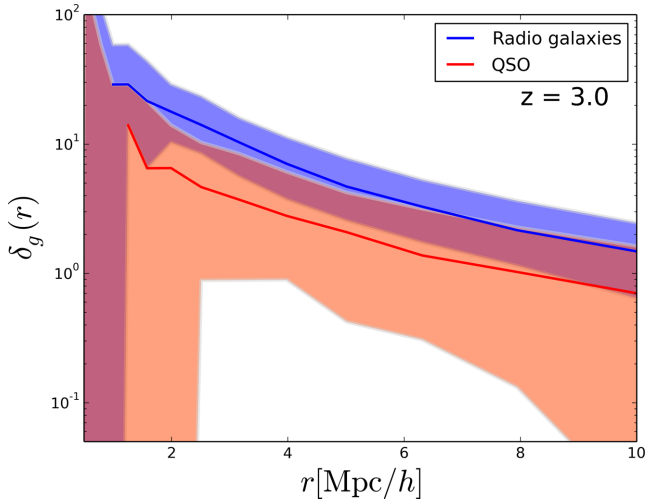

Figure 3. The overdensity of Ly α emitters, δ_g, around radio galaxies and quasars at redshift $z = 3$, as a function of the distance from the central object. The solid curve and the shaded regions show the median and 10–90 percentile of the overdensity distribution for a sample of Ly α emitters with $L_{\mathrm{Ly}\alpha} > 10^{41}$[erg s^{-1} h^{-2}].

environments. Hereafter, we study the dependence of a number of properties on environment by splitting our ELG sample into 'faint', i.e. those galaxies with line luminosity $L > 10^{41}$[erg s^{-1} h^{-2}] which represent a deep survey, and 'bright', with $L > 10^{42}$[erg s^{-1} h^{-2}], which represents a shallow survey. From the observational perspective, these represent complementary strategies. A deep and small survey can characterize the properties of galaxies within their host halo in a small volume, whereas a shallower and wider survey could be used to measure statistical properties, such as the galaxy clustering around these massive structures.

3.1 The environment of overdense regions traced by radio galaxies and quasars

We start by characterizing the properties of overdense regions traced by radio galaxies and quasars. The detection of these overdense regions at high redshifts typically requires long exposures and dedicated observations via narrow-band imaging or spectroscopic follow-up. It is thus common for observational studies to focus on one or only a handful of central objects (e.g. Steidel et al. 2005; Venemans et al. 2005, 2007; Overzier et al. 2006; Kuiper et al. 2011; Bañados et al. 2013; Husband et al. 2013; Adams et al. 2015; Chiang et al. 2015; Saito et al. 2015). Their interpretation, from a galaxy formation perspective, is, thus, limited by cosmic variance and the lack of statistical samples. In addition, projection effects can smear out the overdensity significantly, even in narrow-band surveys (Chiang et al. 2013).

In Fig. 3, we illustrate the predicted overdensities of faint Ly α emitters around radio galaxies and quasars at $z = 3$ as a function of distance in redshift space. We define the galaxy overdensity as $\delta_g(r) = (n(r) - \bar{n})/\bar{n}$, where $n(r)$ is the number density of galaxies within a sphere of radius r around a central object, and $\bar{n}$ is the galaxy number density averaged over the simulation volume.

These predictions represent the ideal case in which no projection effects affect the measured overdensities. Overall, the median overdensity around radio galaxies is predicted to be higher than that around quasars. This is consistent with the correlation between halo mass local overdensities. However, the scatter of δ_g around the median (shown in the plot as the 10 and 90 percentile range of the distribution of overdensities around each radio galaxy and quasar in the model) is significant. For example, at distances below $\sim 1\ h^{-1}$Mpc, the distribution of overdensities around radio galaxies and quasars can span several orders of magnitude due to the small number of galaxies at these distances. At larger distances from the central object, overdensities of quasars display a larger scatter than around radio galaxies. This is likely to be caused by the larger range of halo masses that are hosts of quasars, as opposed to radio galaxies that are found in a much narrower halo mass range (see Fig. 2).

By studying a sample of overdensities around radio galaxies above $z > 2$, Venemans et al. (2007) reported that two out of eight of their radio galaxies have environments that are consistent with being equivalent to the field. Our model predicts that the fraction of radio galaxies hosted in average or underdense environments increases towards lower redshifts and lower host halo masses. Also, this fraction is significantly lower when measuring the overdensity of galaxies with the faint samples of galaxies instead of the bright ones. At $z = 2.2$, for instance, 20 per cent of radio galaxies hosted by haloes with mass $M_{\mathrm{halo}} \sim 10^{12}$ are in environments with $\delta_{\mathrm{gal}} \leq 0$ of bright Ly α emitters. This fraction is reduced to $\sim$9 per cent when measuring δ_{gal} with faint galaxies. In quasars, these fraction are generally higher, reaching up to 35 per cent for the same halo masses. This population of underdense radio galaxies and quasars arise partly due to cosmic variance and, as shown in Section 4, due to the fact that most of the radio galaxies and quasars with $\delta_{\mathrm{gal}} \approx 0$ do not evolve to become massive clusters, but instead become average haloes at $z = 0$.

A better way to quantify the clustering is to compute the cross-correlation function between central quasars or radio galaxies and ELGs, ξ_{cc}. This is estimated as

$$\xi_{\mathrm{cc}}(r) = \frac{DD(r)}{N_c n_{\mathrm{gal}} \Delta V(r)} - 1, \quad (8)$$

where $DD(r)$ is the total number of galaxies around central objects at a distance $r \pm \Delta r/2$, N_c is the total number of central objects in the simulation box, n_{gal} is the mean number density of galaxies, and $\Delta V(r)$ is the volume of a spherical shell of radius r and width Δr. This width corresponds to the bin size used to compute ξ_{cc}. Equation (8) is suitable for computing the cross-correlation function because our simulation box is periodic, and thus, the pair counts are not affected by edge effects. As a result, there is no need to make use of estimators that rely on random sets of objects.

Fig. 4 shows the predicted cross-correlation function between radio galaxies and Ly α emitters, and quasars and Ly α emitters at three redshifts spanning $2.2 \leq z \leq 4.5$. The amplitude of ξ_{cc} is higher when using radio galaxies as the central objects than it is for quasars for any redshift, as expected. An interesting feature arises at small scales (i.e. $r \lesssim 2\ h^{-1}$Mpc), where ξ_{cc} has a higher amplitude when computed using the sample of faint Ly α emitters than when using the bright sample. This might seem counterintuitive at first, since we would expect brighter objects to be more clustered than faint ones.

To investigate the origin of this significant difference in the amplitude of clustering at small scales, we look at the properties of the galaxies surrounding radio galaxies and quasars. The mean halo mass of Ly α emitters as a function of distance is shown in the middle panels of Fig. 4. The mean halo mass of Ly α emitters around radio galaxies is about 0.5 dex higher than that of Ly α emitters around quasars. This is a consequence of the hierarchical growth of structures, in which massive haloes are surrounded by other massive haloes. In both samples surrounding quasars and radio galaxies,

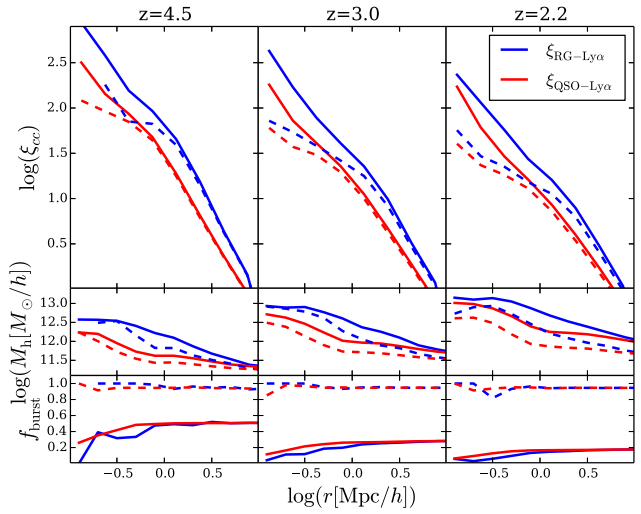

Figure 4. The clustering and properties of Ly α emitters around radio galaxies (blue) and quasars (red) as a function of distance from the central object for redshifts $z = 4.5$, 3.0, and $z = 2.2$. Solid lines correspond to predictions for the faint sample of Ly α emitters, whereas dashed lines correspond to the bright sample. Top: the cross-correlation function between radio galaxies, quasars, and Ly α emitters. Middle: the mean halo mass of Ly α emitters as a function of distance to the central object. Bottom: the mean fraction of starburst galaxies as a function of distance to the central object.

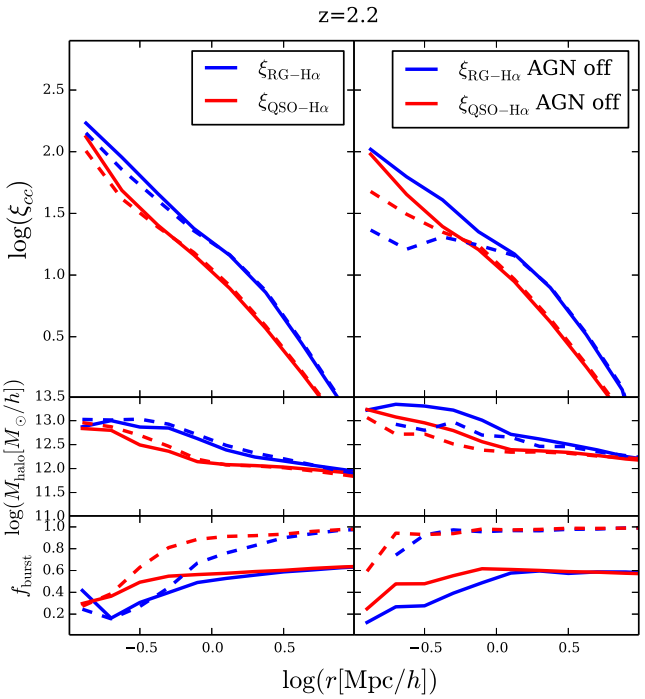

Figure 5. Same as in Fig. 4 but using H α emitters at $z = 2.2$. The left-hand panel shows results for the fiducial model, whereas the right-hand panels corresponds to the case in which AGN feedback has been turned off.

faint Ly α emitters are hosted by haloes that are slightly more massive than brighter ones. At separations of $r \gtrsim 10\,h^{-1}$Mpc, the mean halo mass of both faint and bright galaxies tend to converge. This is translated into their cross-correlation functions having virtually the same amplitude.

To understand why faint Ly α emitters are hosted by more massive haloes than bright Ly α emitters are, we look at the relation between line luminosity and halo mass. Overall, both quantities are correlated. However, as shown in Orsi et al. (2012), the correlation differs depending on whether galaxies are forming stars quiescently or in starbursts. In this variant of GALFORM, quiescent galaxies form stars following a Kennicutt (1983) IMF. Starbursts, on the other hand, form stars with a top-heavy-like IMF (see Baugh et al. 2005). This IMF produces about four times more ionizing photons for each star formation episode than a 'normal' IMF (Le Delliou et al. 2005). This implies that lower mass haloes that experience a starburst can be brighter in Ly α luminosity than a more massive halo forming stars quiescently. The bottom panels of Fig. 4 show precisely this: the bright sample of Ly α emitters consists almost entirely of starbursts, whereas the faint sample contains at most 40 per cent of starbursts.

The difference in the fraction of starbursts between faint and bright samples of Ly α emitters is also enhanced by the Ly α radiative transfer model, which assigns low escape fractions to quiescent galaxies, and higher ones to starbursts, as shown in Orsi et al. (2012). Hence, the resulting cross-correlation functions shown in Fig. 4 are the result of a combination of AGN modelling (which results in radio galaxies populating more massive haloes than quasars), the hierarchical clustering of haloes (which leads to more massive haloes surrounding radio galaxies than quasars), the choice of the IMF for quiescent galaxies and starbursts (which is responsible for most of the bright Ly α emitters in lower mass haloes than quiescent Ly α emitters), and finally the Ly α radiative transfer model favouring the escape of Ly α photons in starbursts.

Unlike Ly α emitters, H α emitters are not subject to complex radiative transfer due to resonant scattering, and their attenuation by dust is smaller, making these galaxies excellent tracers of the instantaneous star formation rate (SFR; Kennicutt 1998; Calzetti 2013). Apart from the radiative transfer effects, however, H α emitters are essentially equivalent to Ly α emitters in nature. Hence, by comparing the properties of the same environments traced by Ly α and H α emitters we can obtain a better picture of how environmental processes affect the galaxy properties.

We show the cross-correlation functions between radio galaxies, quasars and samples of faint and bright H α emitters at $z = 2.2$ in Fig. 5. As with Ly α emitters, the cross-correlation function involving radio galaxies has a higher amplitude than that with quasars as central objects. However, unlike with Ly α emitters, there is little difference between the correlation functions when using the bright and faint samples of H α emitters, even at small scales. The middle panels of Fig. 5 show that the mean halo mass for the two samples of galaxies is also very similar. Also, the fraction of starbursts, shown by the bottom panels of Fig. 5 shows that there is a similar fraction of starbursts at small scales, although the fraction of starbursts increases towards larger scales in the bright sample of H α emitters. At large scales, there is only a weak dependence between H α luminosity and clustering amplitude, measured by the bias factor b (Orsi et al. 2008).

In order to understand what is causing galaxies in the faint and bright samples of H α emitters to have similar halo masses and starburst fractions, we explore the effect of an environmental process that plays a role only on small scales. Since central objects are active galaxies, we run a variant of the GALFORM model in which AGN feedback is switched off, leaving everything else the same. Interestingly, this variant of the model mimics the suppression of the clustering amplitude on small scales that is also evident in the

Ly α emitter samples, although in this case the difference between the amplitude of the cross-correlations involving faint and bright samples is much stronger when the central objects are radio galaxies. This is consistent with the expectation that AGN feedback plays a stronger role in the environments of radio galaxies than in quasars. In the absence of AGN feedback, there are two noticeable changes in the properties of galaxies. On one hand, galaxies in massive haloes that would be quenched by AGN feedback are now part of both the faint and the bright samples of Hα emitters, thus increasing the average halo mass of both populations. On the other hand, the bulk of the quiescent galaxies do not reach luminosities above 10^{42}[erg s^{-1} h^{-2}], so the bright sample of galaxies consists almost entirely of starbursts. Those galaxies, as discussed earlier, span a larger range of halo mass because of their higher production of ionizing photons. Therefore, the bright sample has, on average, smaller halo masses than the average fainter ones. As a result, at small scales, the mean halo mass of faint H α emitters is higher than that of bright ones, resulting in a significant difference in their clustering amplitude.

The analysis above shows that although the properties of galaxies in protoclusters are fundamentally determined by the properties of their host haloes, the role of baryonic effects on the properties of galaxies in overdense regions is also predicted to be important.

To gain deeper insight into the effect of environment on galaxy properties, we compute how other galaxy properties change as a function of distance from their central objects. In particular, we focus on the stellar mass $M_{\rm stellar}$, SFR and the cold gas metallicity $Z_{\rm cold}$. Fig. 6 shows the mean values of these properties as a function of distance from their central objects for both Ly α and H α emitters at $z = 2.2$. We also compute these quantities using two variants of our fiducial model, one in which there is no Ly α radiative transfer, and another in which AGN feedback is turned off.

The mean stellar mass of Ly α emitters in the field is predicted to be about $10^{9.6}$ [M$_\odot$ h^{-1}], and increases about 0.2 dex in the inner regions of the overdense regions. The stellar mass of galaxies in radio galaxy or quasar environments is predicted to be indistinguishable, even when comparing faint and bright samples. When Ly α radiative transfer is switched off, the stellar mass is predicted to be about 0.6 dex higher than in the previous case. For H α emitters, bright galaxies are more massive than faint ones, and the effect of the environment is more important. When AGN feedback is off, galaxies can grow in stellar mass reaching $\sim 10^{10.8}$ [M$_\odot$ h^{-1}] for small distances in bright H α emitters.

The SFR of faint Ly α emitters tends to decrease by up to about 0.4 dex in overdense environments. Bright Ly α emitters, as expected, have higher SFRs, but they display no environmental dependence. When radiative transfer is disabled, the environmental effect on the faint sample is erased, and there is a slight increase of the SFR in bright Ly α emitters. H α emitters, on the other hand, also present a very small decrease of their SFR in overdense environments. Interestingly, the variants with AGN feedback turned off also present a small decrease in the SFR. This occurs because in GALFORM when a galaxy becomes a satellite its hot halo is stripped. Such environmental effect reduces the total reservoir of gas to form stars, thus producing a quenching of star formation. Recently, Peng, Maiolino & Cochrane (2015) found evidence that this effect is a primary mechanism for quenching star formation in a sample of local galaxies.

Finally, gas metallicities tend to increase as a function of environment density for both Ly α and H α emitters. Interestingly, our model predicts that faint Ly α emitters should have a higher gas metallicity than bright ones, but the opposite is true for H α emitter samples. This is due to Ly α radiative transfer effects. When AGN feedback is off, galaxies are predicted to have much higher gas metallicities. This is a natural consequence of them having also larger stellar masses and SFRs.

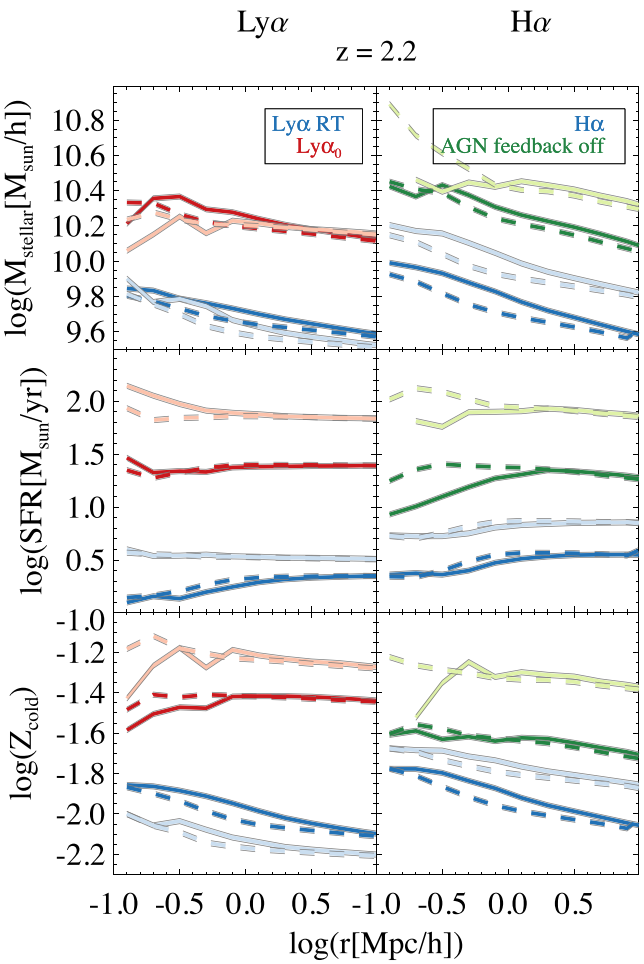

Figure 6. The mean stellar mass (top), SFR (middle), and gas metallicity (bottom) as a function of distance from radio galaxies and quasars at $z = 2.2$. Solid curves are computed using environments around radio galaxies, whereas dashed curves represent the same around quasars. The left-hand column shows the properties of Ly α emitters. Blue and red corresponds to the fiducial model and a variant without radiative transfer effects, respectively. The right-hand column shows the properties of H α emitters. Here, blue and green are the fiducial and a variant with no AGN feedback, respectively. In both columns, faint colours correspond to the sample of bright galaxies, whereas darker colours correspond to the faint samples.

Despite the difference in halo masses hosting radio galaxies and quasars, and AGN feedback being less common in quasars than radio galaxies, there is very little difference in the environmental dependence of the properties studied here between the two tracers. AGN feedback acts to shut down the SFR, thereby preventing the formation of more massive galaxies and also the chemical enrichment of their gas component. Ly α radiative transfer favours the escape of Ly α photons in galaxies with lower gas metallicities, stellar masses and SFRs than is the case for H α emitters.

3.2 Protoclusters and their descendants

We now study how the overdense regions discussed before are related to progenitors of massive structures at the present day.

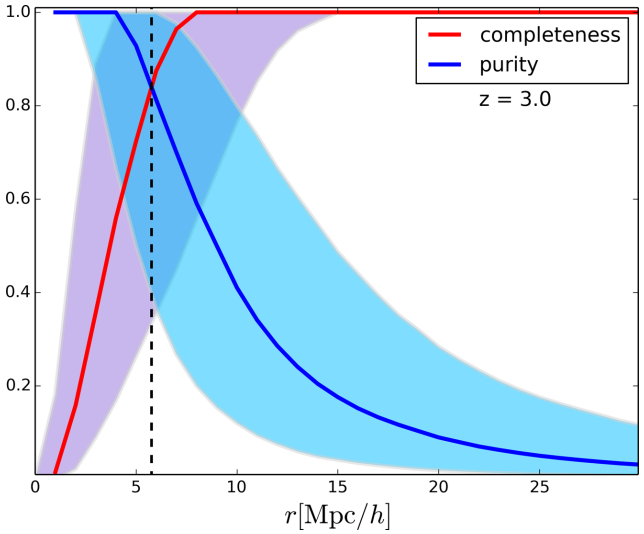

Figure 7. The protocluster completeness (red) and purity (blue) of the sample of faint Ly α emitters around radio galaxies at $z = 3.0$. The solid lines show the median completeness and purity. The shaded regions show the 10–90 percentile of the distribution of completeness and purity of Ly α emitters at a given distance from a central radio galaxy. The vertical dashed line corresponds to the distance at which the completeness and purity are equal.

GALFORM can track the progenitors of any present-day halo by retrieving its merger history. This is useful for determining which galaxies surrounding a central object at high redshift actually form part of a massive cluster at late times. We determine the progenitors associated to a protocluster-like structure as those galaxies that surround a central object (radio galaxy or quasar) that at $z = 0$ are within the same (friends-of-friends) massive DM halo. Furthermore, we identify protoclusters as those massive structures that become part of a DM halo with mass $M_{\rm halo} \geq 10^{14}\,[{\rm M}_{\odot}\,h^{-1}]$. As we discuss below, a significant fraction of radio galaxy and quasar environments do not form a massive cluster by $z = 0$. Nevertheless, in the following we apply the same analysis to all environments around these central objects.

A practical problem in observations is to determine the size of a protocluster. Since we can identify the progenitors of a massive halo, here we devise a method to determine the average size of overdensities around both types of central objects. We take the simple approach of considering protocluster members as those galaxies that lie within a sphere of a given radius around the central object. If that radius coincides with the most distant progenitor, then the sphere will contain all true progenitors, but also a significant fraction of galaxies that are not progenitors of the same halo. Hence, to find an appropriate radius, we measure the completeness of a protocluster at a radius r, defined as the fraction of progenitors enclosed within that radius r. Likewise, we measure the purity of the sample at a radius r, defined as the ratio of the number of galaxies that are progenitors to the total number of galaxies enclosed within r. In the following, for simplicity, we will restrict the analysis of overdensities and progenitors to those determined using the sample of faint Ly α emitters.

Fig. 7 illustrates this by showing the distribution of completeness and purity as a function of distance from radio galaxies at $z = 3$. As expected, at small distances from the central objects the purity is at its highest, and the completeness is low. At a given radius r^*, both quantities intersect, typically at high values. We call r^* the typical protocluster radius at that redshift, and C the corresponding completeness and purity at that radius.

Table 1 shows r^* and the completeness and purity at that radius for radio galaxies and quasars at different redshifts. Overall, r^* decreases towards lower redshifts. For radio galaxies, for instance, r^* goes from 9.37 to 5.12 h^{-1}Mpc for redshifts $z = 5.7$ to $z = 2.2$. Radio galaxies at lower redshifts are, on average, hosted by more massive haloes than those at higher redshifts, and this is effectively translated into larger structures (in physical coordinates) at low redshifts than at high redshifts. However, this also implies a decrease of the radius of protoclusters towards lower redshifts when sizes are expressed in comoving units.

Also, the size of the protostructures traced by radio galaxies have completeness and purities around ∼0.8 at all redshifts. Quasars, on the other hand, have lower values of completeness and purity, ranging from 0.5 at $z = 2.2$ to 0.7 at $z = 5.7$. This suggests that the r^* values shown in Table 1 are reasonable for radio galaxies, but do a poorer job of representing the size of protostructures traced by quasars. This is likely to be related to the large range of halo masses that host quasars, which is translated into a large variety of sizes and environments traced by quasars, unlike radio galaxies that are significantly more restricted in host halo mass (see Fig. 2). This also results in our model predicting radio galaxies to trace more massive descendant halo masses compared to quasar descendants. The fraction of protoclusters (i.e. those structures with $z = 0$ halo mass $M_{\rm halo} \geq 10^{14}\,[{\rm M}_{\odot}\,h^{-1}]$) traced by radio galaxies and quasars differs significantly. At $z = 2.2$, half of radio galaxies are predicted to trace protoclusters. On the other hand, only 19 per cent of quasars pinpoint the location of protoclusters at this redshift. These fractions grow considerably towards higher redshifts, since both radio galaxies and quasars trace increasingly higher and rarer peaks of the matter density distribution of the Universe towards high redshifts.

Since the overdensity of galaxies is a biased tracer of the matter content, we explore how the overdensity within protoclusters correlate with their descendant $z = 0$ halo mass. To do this, we compare the overdensity of galaxies measured within r^* with the $z = 0$ descendant mass for protoclusters traced by radio galaxies and by quasars. This relation is shown in Fig. 8. Overall, protoclusters traced by radio galaxies become haloes at $z = 0$ with masses ranging over $10^{13.5} \lesssim M_{\rm halo}^{z=0}\,[{\rm M}_{\odot}\,h^{-1}] \lesssim 10^{15}$. Quasar descendants also reach those high halo masses, but a significant fraction have masses

Table 1. Protocluster radius r^* [Mpc h^{-1}], median completeness/purity C, mean descendant halo mass $M(z = 0)[{\rm M}_{\odot}\,h^{-1}]$ and fraction of cluster descendants $f(M > 10^{14}[{\rm M}_{\odot}\,h^{-1}])$ of Ly α samples around radio galaxies and quasars for different redshifts.

Redshift	$r^*_{\rm radio}$	$C_{\rm radio}$	$\log M_{\rm radio}(z = 0)$	$f_{\rm radio}(M > 10^{14})$	$r^*_{\rm qso}$	$C_{\rm qso}$	$\log M_{\rm qso}(z = 0)$	$f_{\rm qso}(M > 10^{14})$
5.7	9.37	0.88	14.59	0.76	6.88	0.73	14.07	0.34
4.5	6.97	0.85	14.48	0.69	6.04	0.69	14.01	0.29
3.0	5.77	0.84	14.34	0.58	5.08	0.61	13.91	0.23
2.2	5.12	0.78	14.26	0.50	5.00	0.50	13.83	0.19

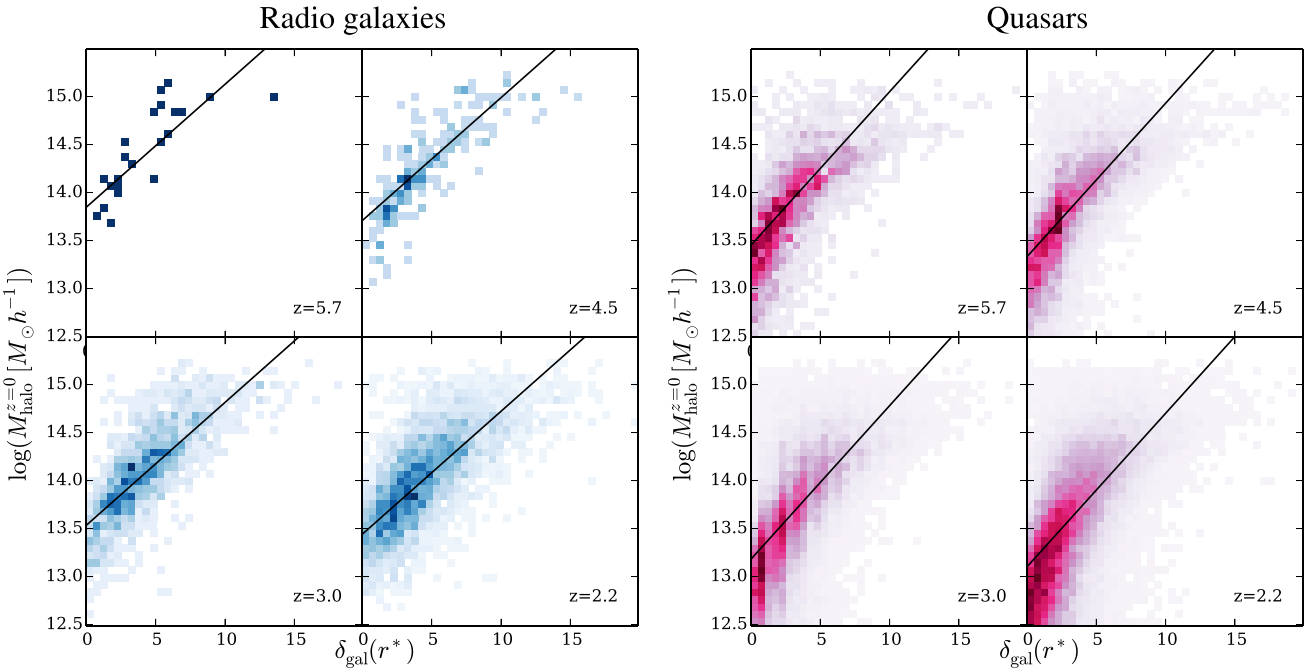

Figure 8. The relation between protocluster galaxy overdensity, δ_g, measured with the faint sample of Ly α emitters within r^* and their $z = 0$ descendant halo mass $M^{z=0}_{\rm halo}$. The left-hand panels show this relation for protoclusters traced by radio galaxies at different redshifts, and the right-hand panels with quasars. Each panel correspond to a different protocluster redshift, ranging from $z = 4.5$ down to $z = 2.2$ as labelled.

below 10^{13} [$\rm M_\odot\,h^{-1}$] (for a discussion on quasar descendants see Fanidakis et al. 2013). The overdensity of galaxies within r^* correlates well with the descendant halo masses for both radio galaxies and quasar descendants. This shows that our choice of radius r^* to determine the average extent of protoclusters at a given redshift is a reasonable one.

The correlation between the galaxy overdensity and the halo descendant mass shown in Fig. 8 can be described with a simple linear form:

$$\log\left(M^{z=0}_{\rm rad}[\,{\rm M}_\odot\,h^{-1}]\right) = 13.07 + 0.11(1+z) + 0.12\delta_{\rm gal}(r^*), \quad (9)$$

$$\log\left(M^{z=0}_{\rm QSO}[\,{\rm M}_\odot\,h^{-1}]\right) = 12.78 + 0.10(1+z) + 0.16\delta_{\rm gal}(r^*), \quad (10)$$

where $M^{z=0}$ is the descendant halo mass in units of [$\rm M_\odot\,h^{-1}$]. Although the constants in equations (9) and (10) have similar values, their difference reflects the smaller descendant masses of quasars with respect to radio galaxies. Also, the scatter of the descendant mass as a function of $\delta_{\rm gal}$ is larger in quasars than in radio galaxies, reflecting the broader diversity of environments that are traced by quasars. Despite the apparent broad scatter around the scaling relations, we have checked that, for radio galaxies, 68 per cent of the distribution is within 0.2 dex of the relation given by equation (9). For quasars, this remains the same except at $z = 2.2$, where 68 per cent of the distribution is within 0.6 dex of the scaling relation of equation (10).

Observational samples could make use of equations (9) and (10) using r^* predicted by GALFORM to obtain an approximate value of the $z = 0$ descendant that a protocluster is expected to evolve into. The scaling relations obtained when using the bright sample of Ly α emitters is consistent with the ones obtained for the faint sample within the $1-\sigma$ region, making equations (9) and (10) suitable for samples of galaxies limited by these two luminosity ranges.

The above analysis results in simple scaling relations that could be easily applied to observational data. However, observations generally give projected measurements of the overdensity of galaxies, and these are expected to differ significantly with respect to the spherically averaged overdensities (e.g. Muldrew et al. 2012; Chiang et al. 2013; Shattow et al. 2013). We can use our model to determine what overdensity values typical instruments would measure at different redshifts. To do this, we project positions along the redshift space s coordinate (see equation 7), and convert the spectral resolution of an instrument into a comoving distance probed along the line of sight. We then compute the overdensity around a given protocluster by using a cylinder of radius r^* and depth given by the spectral resolution of the instrument. In the case of very high resolution instruments, r^* can be larger than the comoving distance range probed by the spectral resolution. Such cylinders would then probe only a fraction of the galaxies in a protocluster overdensity. To include all galaxies within the protocluster radius, we stack several cylinders of equal depth and decreasing radius around both sides of a central cylinder of radius r^* to probe a volume that approximates that of a sphere of the same radius.

Fig. 9 shows a direct comparison between the overdensity computed within a sphere, $\delta^{\rm 3D}$ and that within cylinders by projecting distances along the line of sight, $\delta^{\rm 2D}$, both measured at r^* at their respective redshifts according to Table 1. To illustrate the effect of different spectral resolutions, we choose four instruments, one for each redshift spanning $5.7 \leq z \leq 2.2$.

Multi Object Optical and Near-infrared Spectrograph (MOONS; Cirasuolo et al. 2014) is a future multi-object spectrograph proposed for the very large telescope (VLT) that, given its spectral coverage, can be used to find Ly α emitters at $z = 5.7$ and beyond. Its wavelength resolution, $\Delta\lambda = 0.9$ Å at the Ly α observed wavelength, translates into a line-of-sight comoving resolution at $z = 5.7$ of $\Delta r_z = 0.24\ h^{-1}$Mpc. Fig. 9 shows that our model

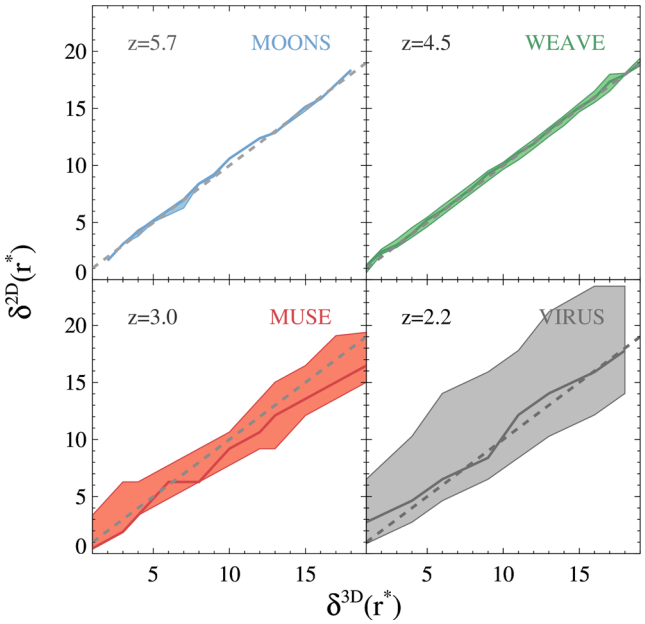

Figure 9. The overdensity of protoclusters traced by radio galaxies and faint Ly α emitters at different redshifts. The panels compare the overdensity measured within a sphere, δ^{3D} and that measured with projected distances along the line of sight within a cylinder, δ^{2D}. The distance probed along the line of sight varies depending on the spectral resolution of the instrument simulated, shown in the legend of each panel.

predicts a remarkably accurate projected overdensity compared to the real, spherical one. Similarly, WEAVE (Dalton et al. 2012), another future multi-object spectrograph at the 4.2 metre William Herschel Telescope on La Palma can be used to search for Ly α emitters at $z = 4.5$. Its spectral resolution at the Ly α wavelength translates into a comoving resolution of $\Delta r_z = 0.12\,h^{-1}$Mpc. This allows us to obtain very accurate projected overdensities, similarly to MOONS at $z = 5.7$.

At lower redshifts, a given spectral resolution is translated into larger comoving distance ranges, meaning larger projection effects. At $z = 3.0$, for instance, we predict the projected overdensities that could be obtained with Multi Unit Spectroscopic Explorer (MUSE), a panoramic integral field unit (IFU) mounted on VLT (Bacon et al. 2010). Its wavelength resolution at $z = 3.0$ translates into $\Delta r_z = 1.33\,h^{-1}$Mpc. The median value of δ^{2D} falls very close to the exact, spherical overdensity δ^{3D}, although the scatter around is larger than in the previous cases due to the larger comoving distance resolution. Finally, for $z = 2.2$, we illustrate the results of using Visible Integral-Field Replicable Unit Spectrographs (VIRUS), an array of IFUs mounted on the Hobby–Eberly Telescope to carry out the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX) cosmological survey (Hill et al. 2008). The spectral resolution for Ly α emitters at $z = 2.2$ translates into a comoving resolution of $\Delta r_z = 5.08\ h^{-1}$Mpc, which is comparable to the typical size of protoclusters at this redshift. The scatter is, as expected, larger than in the previous cases. It is worth noting that both VIRUS and MUSE can obtain projected overdensities with much better accuracies for higher redshifts than shown here.

On the other hand, narrow-band filters with a typical width of $\sim$100 Å can deliver projected overdensities with resolutions along the line of sight of about 50−100 h^{-1}Mpc over the redshift range shown here (e.g. Venemans et al. 2007; Kuiper et al. 2011; Saito et al. 2015). Therefore, such projected overdensities are not suitable for characterizing the scales of protoclusters as shown here. However, given the wider areas that are accessible with a photometric survey, such studies could characterize the clustering around radio galaxies and quasars over scales beyond the protocluster typical radii to put constraints on the DM halo masses hosting these two central objects.

4 CONCLUSIONS

In this paper, we have presented predictions for the properties of overdense regions around radio galaxies and quasars at high redshifts using a model that includes a physical treatment for quasars, radio galaxies, and star-forming ELGs.

Most of our results here follow from a key prediction of our model, namely that, at a given redshift, radio galaxies are hosted on average by DM haloes that are significantly more massive than quasars. Furthermore, most quasars are hosted by DM haloes of masses $\sim 10^{12}$ [$M_\odot\,h^{-1}$] almost irrespective of redshift, whereas radio galaxies populate the most massive haloes present at any redshift. This crucial distinction is translated into their clustering properties being different, and the impact of AGN feedback being stronger in radio galaxy environments.

This paper discusses two problems. First, we study the properties of the overdense regions traced by radio galaxies and quasars at $z > 2$, measuring the environment with samples of ELGs and exploring how the physics of baryons impacts the galaxies in these environments at high redshift. Secondly, we identify the progenitors of the descendant halo that is traced by radio galaxies and quasars to define protoclusters and link their properties to their descendant haloes.

The analysis of individual overdense regions at high redshifts is subject to significant cosmic variance that limits their interpretation in terms of general galaxy formation physics. This explains, for instance, why some authors have found peculiar environments that are average or underdense around both radio galaxies and quasars (Venemans et al. 2007; Bañados et al. 2013; Husband et al. 2013).

The cross-correlation functions ξ_{cc} between overdensity tracers (radio galaxies and quasars) and ELGs at high redshifts offer different information on small and large scales. At large scales ($r \gtrsim 10\ h^{-1}$Mpc) the amplitude of ξ_{cc} is larger when the central objects are radio galaxies, because these are hosted by more massive haloes. More interestingly, if we split the sample of ELGs into faint and bright galaxies, we find that the clustering on small scales is very different for both. For Ly α emitters, radiative transfer effects and the higher production rate of ionizing photons in starbursts invoked in GALFORM makes faint Ly α emitters have a higher ξ_{cc} than bright ones. In the case of H α emitters, which are not affected by radiative transfer effects, the difference between faint and bright samples is insignificant. In this case, we find that AGN feedback prevents starbursts from dominating the galaxy abundance at small separations. This effect is stronger in radio galaxies.

The next generation of large redshift surveys will be able to characterize the environments of overdense regions with unprecedented detail. In particular, a number of these will rely on ELGs to map the matter distribution. Spectroscopic surveys such as HETDEX (Hill et al. 2008) and Dark Energy Spectroscopic Instrument (Levi et al. 2013) have sufficient spectral resolution to probe the scales on which our model predicts that baryonic effects are noticeable on small scales. Other multinarrow-band surveys such as Javalambre Physics of the Accelerating Universe Astrophysical Survey (Benitez et al. 2014) and Physics of the Accelerating Universe (Castander et al. 2012) will provide large samples at $z \gtrsim 2$, and could measure

the amplitude of ξ_{cc} on larger scales, thus allowing tight constraints to put on the halo masses hosting radio galaxies and quasars.

Ly α radiative transfer and AGN feedback also have an impact on the physical properties of ELGs in overdense regions. Overall, the stellar mass, SFR and gas metallicity of Ly α emitters have a small environmental dependence, which is stronger in the absence of Ly α radiative transfer. AGN feedback produces average values of these properties in overdense regions that are different in radio galaxies than quasars, but the difference is too small to be probed observationally.

The progenitors of a present-day cluster allow us to define in the model an average protocluster radius r^* in overdensities traced by radio galaxies and quasars, by determining the distance at which the average completeness and purity of the predicted protoclusters coincide. This simple definition allows us to correlate the galaxy overdensity within the protoclusters with the halo descendant mass at $z = 0$. We compute scaling relations between these two quantities that should provide with a good approximation the halo descendant mass of observed protoclusters at high redshifts.

Current and planned high-resolution multi-object spectrographs and IFUs are predicted to be able to measure projected overdensities that are very close to the ones computed in 3D, thus making it feasible to explore the predictions of this work with regard to observational data.

The current generation of instrumentation with resolutions above $R \sim 3000$ should be able to resolve the inner structure of protoclusters, where the strongest departures of the galaxy properties with respect to the field are predicted to occur. On the other hand, wide-area narrow-band surveys should be able to provide large statistical samples to probe the clustering properties of protoclusters traced by quasars and radio galaxies.

ACKNOWLEDGEMENTS

We would like to thank Nelson Padilla and Andrea Maccio for encouraging discussions about this project. AO acknowledges support from Fundación ARAID and FONDECYT project 3120181. CMB acknowledges support from STFC Consolidated grant ST/L00075X/1. Part of the calculations of this paper were carried out by the Geryon-2 supercluster at the Centro de Astro-Ingenieria UC. This work also made extensive use of the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, and STFC DiRAC Operations grant ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure.

REFERENCES

Adams S. M., Martini P., Croxall K. V., Overzier R. A., Silverman J. D., 2015, MNRAS, 448, 1335
Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1997, ApJ, 488, L75
Bañados E., Venemans B., Walter F., Kurk J., Overzier R., Ouchi M., 2013, ApJ, 773, 178
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Benitez N. et al., 2014, preprint (arXiv:1403.5237)
Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38
Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Calzetti D., 2013, in Falcn-Barroso J., Knapen J. H., eds, Secular Evolution of Galaxies. Cambridge Univ. Press, Cambridge, p. 419
Castander F. J. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84466D
Chiang Y.-K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127
Chiang Y.-K. et al., 2015, ApJ, 808, 37
Cirasuolo M. et al., 2014, in Ramsay S. K., McLean I. S., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V. SPIE, Bellingham, p. 91470N
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Contreras C. et al., 2013, MNRAS, 430, 924
Cooke E. A., Hatch N. A., Muldrew S. I., Rigby E. E., Kurk J. D., 2014, MNRAS, 440, 3262
Cowie L. L., Barger A. J., Hu E. M., 2010, ApJ, 711, 928
Croton D. J. et al., 2006, MNRAS, 365, 11
Dalton G. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460P
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Deharveng J. et al., 2008, ApJ, 680, 1072
Dijkstra M., 2014, PASA, 31, 40
Dijkstra M., Haiman Z., Spaans M., 2006, ApJ, 649, 14
Donoso E., Li C., Kauffmann G., Best P. N., Heckman T. M., 2010, MNRAS, 407, 1078
Dressler A., 1980, ApJ, 236, 351
Erb D. K., Bogosavljević M., Steidel C. C., 2011, ApJ, 740, L31
Falder J. T. et al., 2010, MNRAS, 405, 347
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Fanidakis N. et al., 2012, MNRAS, 419, 2797
Fanidakis N., Macciò A. V., Baugh C. M., Lacey C. G., Frenk C. S., 2013, MNRAS, 436, 315
Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, MNRAS, 388, 1473
Giavalisco M., Koratkar A., Calzetti D., 1996, ApJ, 466, 831
Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, MNRAS, 439, 264
Goto T., Yamauchi C., Fujita Y., Okamura S., Sekiguchi M., Smail I., Bernardi M., Gomez P. L., 2003, MNRAS, 346, 601
Harrington J. P., 1973, MNRAS, 162, 43
Hashimoto Y., Oemler A., Jr, Lin H., Tucker D. L., 1998, ApJ, 499, 589
Hatch N. A. et al., 2014, MNRAS, 445, 280
Hill G. J. et al., 2008, in Kodama T., Yamada T., Aoki K., eds, ASP Conf. Ser. Vol. 399, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Francisco, p. 115
Hu E. M., Cowie L. L., Barger A. J., Capak P., Kakazu Y., Trouille L., 2010, ApJ, 725, 394
Husband K., Bremer M. N., Stanway E. R., Davies L. J. M., Lehnert M. D., Douglas L. S., 2013, MNRAS, 432, 2869
Karouzos M., Jarvis M. J., Bonfield D., 2014, MNRAS, 439, 861
Kashikawa N. et al., 2006, ApJ, 648, 7
Kashikawa N., Kitayama T., Doi M., Misawa T., Komiyama Y., Ota K., 2007, ApJ, 663, 765
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Kauffmann G., Heckman T. M., Best P. N., 2008, MNRAS, 384, 953
Kawata D., Mulchaey J. S., 2008, ApJ, 672, L103
Kennicutt R. C., Jr, 1983, ApJ, 272, 54
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
King A., 2005, ApJ, 635, L121
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Komatsu E. et al., 2011, ApJS, 192, 18

Kornei K. A., Shapley A. E., Erb D. K., Steidel C. C., Reddy N. A., Pettini M., Bogosavljević M., 2010, ApJ, 711, 693
Koyama Y. et al., 2013, MNRAS, 434, 423
Kuiper E. et al., 2011, MNRAS, 417, 1088
Kunth D., Mas-Hesse J. M., Terlevich E., Terlevich R., Lequeux J., Fall S. M., 1998, A&A, 334, 11
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, MNRAS, 412, 1828
Lacey C. G. et al., 2015, preprint (arXiv:1509.08473)
Lagos C. D. P., Cora S. A., Padilla N. D., 2008, MNRAS, 388, 587
Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649
Le Delliou M., Lacey C., Baugh C. M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S. L., 2005, MNRAS, 357, L11
Levi M. et al., 2013, preprint (arXiv:1308.0847)
Maraston C., 2005, MNRAS, 362, 799
Mas-Hesse J. M., Kunth D., Tenorio-Tagle G., Leitherer C., Terlevich R. J., Terlevich E., 2003, ApJ, 598, 858
Matsuda Y. et al., 2011, MNRAS, 410, L13
Meier D. L., 2002, New Astron. Rev., 46, 247
Miley G., De Breuck C., 2008, A&AR, 15, 67
Mo H. J., Yang X., van den Bosch F. C., Jing Y. P., 2004, MNRAS, 349, 205
Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613
Muldrew S. I. et al., 2012, MNRAS, 419, 2670
Narayan R., Yi I., 1994, ApJ, 428, L13
Neufeld D. A., 1990, ApJ, 350, 216
Norberg P. et al., 2001, MNRAS, 328, 64
Oemler A., Jr, 1974, ApJ, 194, 1
Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, MNRAS, 391, 1589
Orsi A., Baugh C. M., Lacey C. G., Cimatti A., Wang Y., Zamorani G., 2010, MNRAS, 405, 1006
Orsi A., Lacey C. G., Baugh C. M., 2012, MNRAS, 425, 87
Osterbrock D. E., 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, CA
Overzier R. A. et al., 2006, ApJ, 637, 58
Overzier R. A., Guo Q., Kauffmann G., De Lucia G., Bouwens R., Lemson G., 2009, MNRAS, 394, 577
Peebles P. J. E., Shaviv G., 1982, Space Sci. Rev., 31, 119
Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192
Press W. H., Schechter P., 1974, ApJ, 187, 425
Ramos Almeida C., Bessiere P. S., Tadhunter C. N., Inskip K. J., Morganti R., Dicken D., González-Serrano J. I., Holt J., 2013, MNRAS, 436, 997
Saito T. et al., 2015, MNRAS, 447, 3069
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shattow G. M., Croton D. J., Skibba R. A., Muldrew S. I., Pearce F. R., Abbas U., 2013, MNRAS, 433, 3314
Springel V. et al., 2005, Nature, 435, 629
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Steidel C. C., Adelberger K. L., Shapley A. E., Erb D. K., Reddy N. A., Pettini M., 2005, ApJ, 626, 44
Tecce T. E., Cora S. A., Tissera P. B., Abadi M. G., Lagos C. D. P., 2010, MNRAS, 408, 2008
Thuan T. X., Izotov Y. I., 1997, ApJ, 489, 623
Uchimoto Y. K. et al., 2012, ApJ, 750, 116
Utsumi Y., Goto T., Kashikawa N., Miyazaki S., Komiyama Y., Furusawa H., Overzier R., 2010, ApJ, 721, 1680
van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79
Venemans B. P. et al., 2005, A&A, 431, 793
Venemans B. P. et al., 2007, A&A, 461, 823
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wold I. G. B., Barger A. J., Cowie L. L., 2014, ApJ, 783, 119

This paper has been typeset from a TEX/LATEX file prepared by the author.