The clustering of narrow-line AGN in the local Universe

Cheng Li,[1,2,3]⋆ Guinevere Kauffmann,[1] Lan Wang,[4,1] Simon D. M. White,[1] Timothy M. Heckman[5] and Y. P. Jing[3]

[1]*Max-Planck Institut für Astrophysik, D-85748 Garching, Germany*
[2]*Centre for Astrophysics, University of Science and Technology of China, Hefei, Anhui 230026, China*
[3]*The Partner Group of MPI für Astrophysik, Shanghai Astronomical Observatory, Nandan Road 80, Shanghai 200030, China*
[4]*Department of Astronomy, Peking University, Beijing 100871, China*
[5]*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

ABSTRACT

We have analysed the clustering of ∼90 000 narrow-line active galactic nuclei (AGN) drawn from the Data Release 4 (DR4) of the Sloan Digital Sky Survey. Our analysis addresses the following questions. (i) How do the locations of galaxies within the large-scale distribution of dark matter influence ongoing accretion on to their central black holes? (ii) Is AGN activity triggered by interactions or mergers between galaxies? We compute the cross-correlation between AGN and a reference sample of galaxies drawn from the DR4. We compare this to results for control samples of inactive galaxies matched simultaneously in redshift, stellar mass, concentration, velocity dispersion and mean stellar age, as measured by the 4000-Å break strength. We also compare near-neighbour counts around AGN and around the control galaxies. On scales larger than a few Mpc, AGN have almost the same clustering amplitude as the control sample. This demonstrates that AGN host galaxies and inactive control galaxies populate dark matter haloes of similar mass. On scales between 100 kpc and 1 Mpc, AGN are clustered more weakly than the control galaxies. We use mock catalogues constructed from high-resolution *N*-body simulations to interpret this antibias, showing that the observed effect is easily understood if AGN are preferentially located at the centres of their dark matter haloes. On scales less than 70 kpc, AGN cluster marginally more strongly than the control sample, but the effect is weak. When compared to the control sample, we find that only one in 100 AGN has an extra neighbour within a radius of 70 kpc. This excess increases as a function of the accretion rate on to the black hole, but it does not rise above the few per cent level. Although interactions between galaxies may be responsible for triggering nuclear activity in a minority of nearby AGN, some other mechanism is required to explain the activity seen in the majority of the objects in our sample.

Key words: galaxies: active – galaxies: distances and redshifts – cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

A major goal in the study of active galactic nuclei (AGN) has been to understand the physical mechanism(s) responsible for triggering accretion on to the central supermassive black hole and enhanced activity in the nucleus of the galaxy. From a theoretical standpoint, *N*-body simulations that treat the hydrodynamics of the gas have shown that interactions between galaxies can bring gas from the disc to the central regions of the galaxy, leading to enhanced star forma-

tion in the bulge (Barnes & Hernquist 1992; Mihos & Hernquist 1996). It is then natural to speculate that some of this gas will be accreted on to the central supermassive black hole and that this will trigger activity in the nucleus of the galaxy. However, there has been little clear observational evidence in support of this hypothesis.

Many observational studies have examined the correlations between AGN activity in galaxies and their local environment. These analyses have produced contradictory results. Early studies (see e.g. Petrosian 1982; Dahari 1984; Keel et al. 1985) noted that powerful Seyfert galaxies appear to show an excess of close companions relative to their non-active counterparts. More recent analyses of larger and more complete samples of Seyfert galaxies have reached

⋆E-mail: leech@ustc.edu.cn

the opposite conclusion (e.g. Schmitt 2001; Miller et al. 2003). Studies of X-ray selected AGN at intermediate redshifts also find no evidence for excess near-neighbour counts or enhanced levels of galaxy asymmetry (Grogin et al. 2005; Waskett et al. 2005). On the other hand, a recent study of the local environment of a sample of ~2000 quasars at $z < 0.4$ drawn from the Sloan Digital Sky Survey (SDSS) (Serber et al. 2006) concluded that quasars do have a significant local excess of neighbours when compared to L_* galaxies, but only on small scales (<0.2 Mpc). These authors found the excess to be significantly larger for the most luminous systems. This study suggests that the disagreement between different studies may reflect the fact that they targeted AGN with different intrinsic luminosities.

There have also been many studies of the large-scale clustering of AGN. This is usually quantified using the two-point correlation function (2PCF). In the standard model for structure formation, the amplitude of the 2PCF on scales larger than a few Mpc provides a direct measure of the mass of the dark matter haloes that host the AGN. The large redshift surveys assembled in recent years, in particular the 2dF QSO Redshift Survey (2QZ; Croom et al. 2001) and the SDSS (York et al. 2000), have allowed the clustering of quasars to be studied with unprecedented accuracy. The cross-correlation between QSOs in the 2QZ and galaxies in the two-degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001), measured by Croom et al. (2005), is found to be identical to the autocorrelation of L_* galaxies (a mean bias $b_{QG} = 0.97 \pm 0.05$). Measurements of the 2PCF of narrow-line AGN in the SDSS have been carried out by Wake et al. (2004). The results are similar to those for quasars – the amplitude of the AGN autocorrelation function is consistent with the autocorrelation function of luminous galaxies on scales from 0.2 to 100 h^{-1} Mpc. Similar results are found for X-ray selected AGN (Gilli et al. 2005; Mullis et al. 2005). On the other hand, radio-loud AGN appear to be significantly more clustered on large scales (Magliocchetti et al. 1999; Overzier et al. 2003; Magliocchetti et al. 2004), demonstrating that they reside in massive dark matter haloes. This is consistent with the fact that in the local Universe, radio-loud AGN are located in significantly more massive host galaxies than optical AGN (e.g. Best et al. 2005). Constantin & Vogeley (2006) analysed AGN in the SDSS Data Release 2 sample and find that low-ionization nuclear emission-line regions (LINERs) are more strongly clustered than Seyfert galaxies. Once again, this is consistent with the fact that LINERs are found in more massive host galaxies than Seyferts (Kauffmann et al. 2003; Kewley et al. 2006).

In this paper, we analyse the clustering properties of 89 211 narrow-line AGN selected from the Data Release 4 (DR4) of the SDSS using the procedure described in Kauffmann et al. (2003). Our methodology for computing correlation functions in the SDSS has been described in detail in Li et al. (2006), where the dependence of clustering on physical properties such as stellar mass, age of the stellar population, concentration and stellar surface mass density was studied. In this paper, we extend this analysis to AGN. Our approach differs from previous studies in the following ways.

(i) We compute AGN–galaxy cross-correlations from scales of a few tens of kpc to scales of ~10 Mpc. This allows us to study the detailed scale dependence of the AGN clustering amplitude.

(ii) We study how the clustering depends on both the black hole mass (estimated using the central velocity dispersion of the galaxy) and the accretion rate relative to the Eddington rate (estimated using L[O III], where L[O III] is the [O III]λ5007 line luminosity, and M_{BH} is the black hole mass estimated from the central stellar velocity dispersion of the host).

(iii) We wish to isolate the effect of the accreting black hole, so the clustering is always compared to the results obtained for control samples of inactive galaxies that are very closely matched to our AGN sample. We match simultaneously in redshift, mass and structural properties, and we test the effect of additionally matching the age of the stellar population as characterized by the 4000-Å break index $D_n(4000)$.

(iv) We have constructed mock catalogues using the high-resolution Millennium Run simulation (Springel et al. 2005). The mock catalogues match the geometry and selection function of galaxies in the DR4 large-scale structure sample, and they also reproduce the luminosity and stellar mass function of SDSS galaxies, as well as the shape and amplitude of the correlation functions in different bins of luminosity and stellar mass. We use these catalogues to explore in detail how AGN trace the underlying galaxy and halo populations.

Our paper is organized as follows. In Section 2, we describe the AGN and control samples that were used in the analysis. In Section 3, we describe how we construct mock catalogues from the Millennium Simulation and use these to correct for effects such as fibre collisions. Section 4 describes our clustering estimator; Section 5 describes the observational results and in Section 6, we describe how we use our mock catalogues to extract physical information from the data. Finally, in Section 7 we summarize our results and present our conclusions.

2 DESCRIPTION OF SAMPLES

2.1 The SDSS spectroscopic sample

The data analysed in this study are drawn from the SDSS. The survey goals are to obtain photometry of a quarter of the sky and spectra of nearly one million objects. Imaging is obtained in the *u, g, r, i, z* bands (Fukugita et al. 1996; Smith et al. 2002; Ivezić et al. 2004) with a special purpose drift scan camera (Gunn et al. 1998) mounted on the SDSS 2.5-m telescope (Gunn et al. 2006) at Apache Point Observatory. The imaging data are photometrically (Hogg et al. 2001; Tucker et al. 2005) and astrometrically (Pier et al. 2003) calibrated, and used to select stars, galaxies and quasars for follow-up fibre spectroscopy. Spectroscopic fibres are assigned to objects on the sky using an efficient tiling algorithm designed to optimize completeness (Blanton et al. 2003). The details of the survey strategy can be found in York et al. (2000), and an overview of the data pipelines and products is provided in the Early Data Release paper (Stoughton et al. 2002). More details on the photometric pipeline can be found in Lupton et al. (2001).

Our parent sample for this study is composed of 397 344 objects which have been spectroscopically confirmed as galaxies and have data publicly available in the SDSS DR4 (Adelman-McCarthy et al. 2006). These galaxies are part of the SDSS 'main' galaxy sample used for large-scale structure studies (Strauss et al. 2002) and have Petrosian *r* magnitudes in the range $14.5 < r < 17.77$ after correction for foreground galactic extinction using the reddening maps of Schlegel, Finkbeiner & Davis (1998). Their redshift distribution extends from ~0.005 to 0.30, with a median *z* of 0.10.

The SDSS spectra are obtained with two 320-fibre spectrographs mounted on the SDSS 2.5-m telescope. Fibres 3 arcsec in diameter are manually plugged into custom-drilled aluminium plates mounted at the focal plane of the telescope. The spectra are exposed for 45 min or until a fiducial signal-to-noise ratio (S/N) is reached. The median S/N per pixel for galaxies in the main sample

is $\sim$14. The spectra are processed by an automated pipeline, which flux and wavelength calibrates the data from 3800 to 9200 Å. The instrumental resolution is $R \equiv \lambda/\delta\lambda = 1850$–$2200$ (FWHM $\sim$ 2.4 Å at 5000 Å).

2.2 The AGN and control samples

We have performed a careful subtraction of the stellar absorption-line spectrum before measuring the nebular emission lines. This is accomplished by fitting the emission-line-free regions of the spectrum with a model galaxy spectrum computed using the new population synthesis code of Bruzual & Charlot (2003), which incorporates a high-resolution (3-Å FWHM) stellar library. A set of 39 model template spectra was used spanning a wide range in age and metallicity. After convolving the template spectra to the measured stellar velocity dispersion of an individual SDSS galaxy, the best fit to the galaxy spectrum is constructed from a non-negative linear combination of the template spectra. Further details are given in Tremonti et al. (2004). Physical parameters, such as stellar masses, metallicities and star formation rates, have been estimated using the spectra, and these are publicly available at http://www.mpa-garching.mpg.de/SDSS/. The reader is referred to Tremonti et al. (2004) and Brinchmann et al. (2004) for more details.

AGN are selected from the subset of galaxies with S/N > 3 in the four emission lines [O III]λ5007, Hβ, [N II]λ6583 and Hα. Following Kauffmann et al. (2003), a galaxy is defined to be an AGN if

$$\log([\text{O\,{\sc iii}}]/\text{H}\beta) > 0.61/[\log([\text{N\,{\sc ii}}]/H\alpha) - 0.05] + 1.3. \tag{1}$$

We divide all the AGN into three subsamples according to logarithmic stellar velocity dispersion $\log_{10} \sigma_*$. It is also interesting to study how clustering depends on the strength of nuclear activity in the galaxy. In order to address this issue, we follow Heckman et al. (2004) and use the [O III] emission-line luminosity as an indicator of the rate at which matter is accreting on to the central supermassive black hole, and we use the relation given in Tremaine et al. (2002) to estimate black hole masses from the stellar velocity dispersion measured within the fibre aperture. We then use the ratio $L[\text{O\,{\sc iii}}]/M_{\text{BH}}$ as a measure of the accretion rate relative to the Eddington rate, to define subsamples of 'powerful' and 'weak' AGN. (Note that in the current analysis, $L[\text{O\,{\sc iii}}]$ is corrected for dust extinction.) The AGN in each $\log_{10} \sigma_*$ subsample are ordered by decreasing $L[\text{O\,{\sc iii}}]/M_{\text{BH}}$. The top 25 per cent are defined as 'powerful', and the bottom 25 per cent as 'weak'. For each AGN sample, we construct 20 control samples of non-AGN by simultaneously matching four physical parameters: redshift, stellar mass, concentration and stellar velocity dispersion. We have also constructed control samples where the 4000-Å break strength is matched in addition to these parameters. The matching tolerances are $\Delta cz < 500$ km s^{-1}, $\Delta \log M_* < 0.1$, $\Delta \sigma_* < 20$ km s^{-1}, $\Delta C < 0.1$ and $\Delta D_n(4000) < 0.05$.

We correct the [O III] luminosities of the AGN in our sample for dust using the difference between the observed Hα/Hβ emission-line flux ratios and the case-B recombination value (2.86). We assumed an attenuation law of the form $\tau_\lambda \propto \lambda^{-0.7}$ (Charlot & Fall 2000). This procedure has clear physical meaning in the 'pure' Seyfert 2's and LINERs. In the case of the transition objects, the lines will arise both in the NLR and in the surrounding H II regions, with a greater relative AGN contribution to [O III] than to the Balmer lines. Thus, a dust correction to [O III] based on the ratio Hα/Hβ should be regarded as at best approximate.

2.3 Reference galaxy sample

We use the New York University Value Added Galaxy Catalogue (NYU-VAGC)[1] to construct a reference sample of galaxies, which are cross-correlated with the AGN sample. The original NYU-VAGC is a catalogue of local galaxies (mostly below $z \approx 0.3$) constructed by Blanton et al. (2005) based on the SDSS DR2. Here, we use a new version of the NYU-VAGC (SAMPLE DR4), which is based on SDSS DR4. The NYU-VAGC is described in detail in Blanton et al. (2005).

We have constructed two reference samples: (i) a *spectroscopic* reference sample, which is used to compute the projected AGN–galaxy cross-correlation function $w(r_p)$, and (ii) a *photometric* reference sample, which is used to calculate counts of close neighbours around AGN.

The spectroscopic reference sample is constructed by selecting from SAMPLE DR4 all galaxies with $14.5 < r < 17.6$ that are identified as galaxies from the main sample. (Note that r-band magnitude has been corrected for foreground extinction.) The galaxies are also restricted to the redshift range $0.01 \leqslant z \leqslant 0.3$, and the absolute magnitude range $-23 < M_{0.1_r} < -17$. The spectroscopic reference sample contains 292 782 galaxies. We do not consider galaxies fainter than $M_{0.1_r} = -17$ because the volume covered by such faint samples is very small and the results are subject to large errors as a result of cosmic variance (see e.g. fig. 6 of Li et al. 2006). The faint apparent magnitude limit of 17.6 is chosen to yield uniform galaxy sample that is complete over the entire area of the survey.

The photometric reference sample is also constructed from SAMPLE DR4 by selecting all galaxies with $14.5 < r < 19$. The resulting sample includes 1065 183 galaxies. Throughout this work, we adopt standard λ cold dark matter (λCDM) cosmological parameters: $\Omega = 0.3, \Omega_\Lambda = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}.

3 MOCK CATALOGUES

In this section, we describe how we construct a large set of mock galaxy samples with the same geometry and selection function as the spectroscopic samples described in the previous section. We will use these mock catalogues to test our method for correcting for the effect of fibre collisions on the measurement of the AGN–galaxy cross-correlation function on small scales. We will also use these mock catalogues to construct models of AGN clustering for comparison with the observations.

3.1 Galaxy properties in the Millennium Simulation

Our mock catalogues are constructed using the Millennium Simulation (Springel et al. 2005), a very large simulation of the concordance ΛCDM cosmogony with 10^{10} particles. The chosen simulation volume is a periodic box of size $L_{\text{box}} = 500 \, h^{-1}$ Mpc on a side, which implies a particle mass of $8.6 \times 10^8 \, 4 \, h^{-1} \, \text{M}_\odot$. Haloes and subhaloes at all output snapshots are identified using the SUBFIND algorithm described in Springel et al. (2001), and merger trees are then constructed that describe how haloes grow as the Universe evolves. Croton et al. (2006) implemented a model of the baryonic physics in these simulations in order to simulate the formation and evolution of galaxies and their central supermassive black holes [see Croton et al. (2006) for more details]. This model produced a catalogue of 9 million galaxies at $z = 0$ down to a limiting absolute magnitude
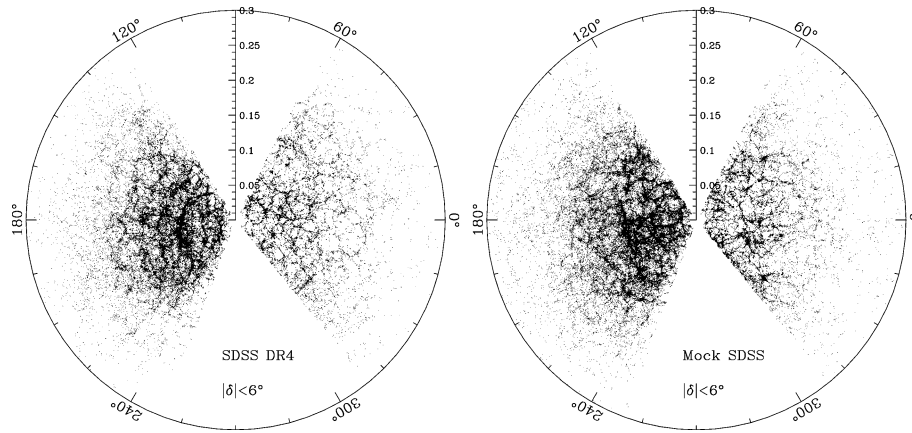
---

[1] http://wassup.physics.nyu.edu/vagc/

Figure 1. Equatorial distribution of right ascension and redshift for galaxies within 6° of the equator in the SDSS (left-hand side) and in one of our mock catalogues (right-hand side).

limit of $M_r - 5 \log h = -16.6$. This catalogue is well-matched to many properties of the present-day galaxy population (luminosity–colour distributions, clustering, etc.). It is publicly available at http://www.mpa-garching.mpg.de/galform/agnpaper. In our work, we adopt the positions and velocities of the galaxies given in the Croton et al. catalogue. The *r*-band luminosities and stellar masses are assigned to each model galaxy, using the parametrized functions described by Wang et al. (2006). These functions relate the physical properties of galaxies to the quantity M_{infall}, defined as the mass of the halo at the epoch when the galaxy was the last central dominant object in its own halo. They were chosen so as to give close fits to the results of the physical galaxy formation model of Croton et al. (2006), but their coefficients were adjusted to improve the fit to the SDSS data, in particular the galaxy mass function at the low-mass end. Extensive tests have shown that the adopted parametrized relations allow us to accurately match the luminosity and stellar mass functions of galaxies in the SDSS, as well as the shape and amplitude of the 2PCF of galaxies in different luminosity and stellar mass ranges (Li et al. 2006; Wang et al. 2006).

3.2 Constructing the catalogues

Our aim is to construct mock galaxy redshift surveys that have the same geometry and selection function as the SDSS DR4. A detailed account of the observational selection effects accompanies the NYU-VAGC release. The survey geometry is expressed as a set of disjoint convex spherical polygons, defined by a set of 'caps'. This methodology was developed by Andrew Hamilton to deal accurately and efficiently with the complex angular masks of galaxy surveys (Hamilton & Tegmark 2002).[2] The advantage of using this method is that it is easy to determine whether a point is inside or outside a given polygon (Tegmark et al. 2002). The redshift sampling completeness is then defined as the number of galaxies with redshifts divided by the total number of spectroscopic targets in the polygon. The completeness is thus a dimensionless number between 0 and 1, and it is constant within each of the polygons. The limiting magnitude in each polygon is also provided. (It changes slightly across the survey region.)

We construct our mock catalogues using the methods described in Yang et al. (2004), except that we position the virtual observer

randomly inside the simulation and not at the centre of the box. Because the survey extends out to $z \sim 0.3$, this implies that we need to cover a volume that extends to a depth of $900 \, h^{-1}$ Mpc, that is, twice that of the Millennium catalogue. We thus create $5 \times 5 \times 5$ periodic replications of the simulation box and place the observer randomly within the central box, so that the required depth can be achieved in all directions for the observer.

We produce 20 mock catalogues by following the procedure described below.

(i) We randomly place a virtual observer in the stack of boxes described above. We define a (α, δ)-coordinate frame and remove all galaxies that lie outside the survey region.

(ii) For each galaxy, we compute the redshift as 'seen' by the virtual observer. The redshift is determined by the comoving distance and the peculiar velocity of the galaxy.

(iii) We compute the *r*-band apparent magnitude of each galaxy from its absolute magnitude M_r and its redshift, applying a (negative) *K*-correction but neglecting any evolutionary correction. We then select galaxies according to the position-dependent magnitude limit (provided in the SAMPLE DR4) and apply a (positive) *K*-correction to compute $M_{0.1_r}$, the *r*-band absolute magnitude of the galaxy at $z = 0.1$.

(iv) To mimic the position-dependent completeness, we randomly eliminate galaxies using the completeness masks provided in SAMPLE DR4.

(v) Finally, we mimic the actual selection criteria of our own reference sample by restricting galaxies in the mock catalogue to $0.01 < z < 0.3$, $14.5 < r < 17.6$ and $-23 < M_{0.1_r} < -17$.

Fig. 1 shows the equatorial distribution of galaxies in one of our mock catalogues, compared to that in the observational sample. The average number of galaxies in our mock catalogues is $\sim$320 000, with a rms dispersion of $\sim$9000, in good agreement with the observed number.

3.3 Fibre collisions

The procedure described above does not account for the fact that the spectroscopic target selection becomes increasingly incomplete in regions of the sky where the galaxy density is high, because two fibres cannot be positioned closer than 55 arcsec from each other. In order to mimic these fibre 'collisions', we modify Step (iv) above. We no longer randomly sample galaxies using the completeness

---

[2] http://casa.colorado.edu/~ajsh/mangle/

masks. Instead, we assign fibres to our mock galaxies using a procedure that is designed to mimic the tiling code that assigns spectroscopic fibres to SDSS target galaxies (Blanton et al. 2003). We run a friends-of-friends grouping algorithm on the mock galaxies with a 55-arcsec linking length. Isolated galaxies, that is those in single-member groups, are always assigned fibres. If two targets collide, we simply pick one at random to be the spectroscopically targeted galaxy. For groups with more than two members, we need to determine which members will be assigned fibres and which will not. The resulting groups are almost always of sufficiently low multiplicity so that in principle, one could simply check all possibilities to find the best possible combination of galaxies that would eliminate fibre collisions. (This is the procedure adopted in the SDSS target selection.) In our work, we use a somewhat more efficient method. For each group, we calculate the geometric centre. The group members located closer to the centre are preferentially eliminated. For example, in a triple collision this algorithm will keep the outer two members rather than the middle one.

This procedure is complicated by the fact that some fraction of the sky will be covered with overlaps of different tiles. (Each spectroscopic fibre plug plate is referred as a 'tile', which has a circular field of view with a radius of $1°.49$.) About 30 per cent of the sky is covered in such overlaps. This means that if, for example, a binary group is covered by two or more tiles, both of the group members can be assigned fibres. We take this into account by iteratively repeating the procedure described above as follows.

(i) First, we determine the number of tiles that cover a given galaxy and set the quantity n_{chances} equal to this number. For example, $n_{\mathrm{chances}} = 2$ for a galaxy covered by two tiles; this galaxy has two chances to be assigned a fibre.

(ii) We assign fibres by applying the algorithm described above to those galaxies with $n_{\mathrm{chances}} > 0$. If a galaxy obtains a fibre in this procedure, then the quantity n_{chances} is set to zero. If not, we set $n_{\mathrm{chances}} = n_{\mathrm{chances}} - 1$.

(iii) Step (ii) is repeated until n_{chances} reaches zero for all galaxies. All galaxies that are not assigned fibres are then removed from our mock catalogue.

(iv) Finally, we remove a number of galaxies at random so that the resulting mock sample has the same overall position-dependent completeness as the real SDSS sample.

4 CLUSTERING MEASURES

In order to compute the two-point cross-correlation function $\xi(r_{\mathrm{p}}, \pi)$ between the AGN host (or matched control) sample and the reference galaxy sample, we have constructed random samples that are designed to include all observational selection effects. This is described in detail in Li et al. (2006). $\xi(r_{\mathrm{p}}, \pi)$ is then calculated using the estimator

$$\xi(r_{\mathrm{p}}, \pi) = \frac{N_R}{N_D} \frac{QD(r_{\mathrm{p}}, \pi)}{QR(r_{\mathrm{p}}, \pi)} - 1, \tag{2}$$

where r_{p} and π are the separations perpendicular and parallel to the line of sight; N_D and N_R are the number of galaxies in the reference sample and in the random sample; $QD(r_{\mathrm{p}}, \pi)$ and $QR(r_{\mathrm{p}}, \pi)$ are the cross pair counts between AGN (or control) and the reference sample, and between AGN (or control) and the random sample, respectively.

In what follows, we focus on the projection of $\xi(r_{\mathrm{p}}, \pi)$ along the line of sight:

$$w_{\mathrm{p}}(r_{\mathrm{p}}) = \int_{-\infty}^{+\infty} \xi(r_{\mathrm{p}}, \pi)\, \mathrm{d}\pi = \sum_i \xi(r_{\mathrm{p}}, \pi_i)\Delta\pi_i. \tag{3}$$

Here, the summation for computing $w_{\mathrm{p}}(r_{\mathrm{p}})$ runs from $\pi_1 = -39.5\, h^{-1}$ Mpc to $\pi_{80} = 39.5\, h^{-1}$ Mpc, with $\Delta\pi_i = 1\, h^{-1}$ Mpc.

The errors on the clustering measurements are estimated using the bootstrap resampling technique (Barrow, Bhavsar & Sonoda 1984). We generate 100 bootstrap samples from the observations and compute the correlation functions for each sample using the weighting scheme (but not the approximate formula) given by Mo, Jing & Boerner (1992). The errors are then given by the scatter of the measurements among these bootstrap samples. More details about our procedures and tests of our methodology can be found in Li et al. (2006). We also obtain robust estimates of our uncertainties from the scatter between the results obtained from disjoint areas of the sky.

We are particularly interested in the amplitude of the AGN–reference galaxy cross-correlation on small scales (<100 kpc), because this allows us to evaluate whether mergers and interactions play a role in triggering AGN activity. A careful correction for the effect of fibre collisions when measuring the clustering is thus very important. As described in Li et al. (2006), we correct for fibre collisions by comparing the angular 2PCF of the spectroscopic sample with that of the parent photometric sample. Here, we use our mock SDSS catalogues to test the correction method. We calculate the angular correlation functions $w_z(\theta)$ and $w_{\mathrm{p}}(\theta)$ using mock catalogues with and without fibre collisions. The function

$$F(\theta) = \frac{1 + w_z(\theta)}{1 + w_{\mathrm{p}}(\theta)} \tag{4}$$

is then used to correct for collisions by weighting each pair by $1/F(\theta)$. Fig. 2 (the top panel) shows measurements of $w_{\mathrm{p}}(r_{\mathrm{p}})$ for galaxies in one of our mock catalogues. The solid line is the 'true' correlation function calculated for the mock catalogues that do not include fibre collisions. Circles show the results when fibre collisions are included. Triangles show the results that are obtained
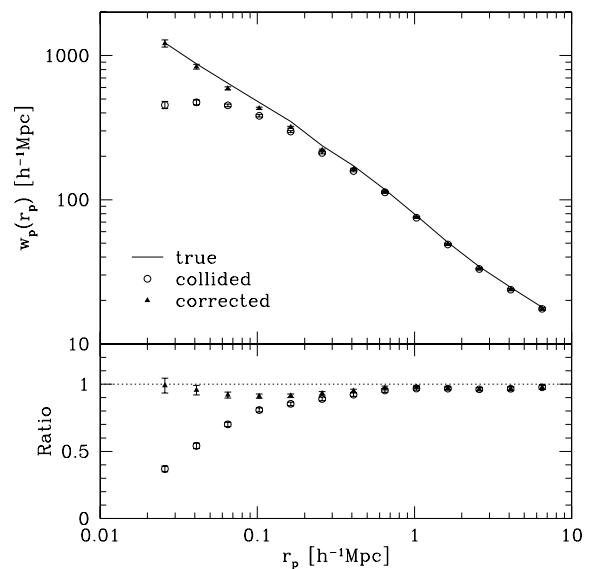

Figure 2. Top: projected 2PCF $w_{\mathrm{p}}(r_{\mathrm{p}})$ measured for the mock catalogue without including fibre collisions (solid line) and for the mock catalogue with fibre collisions (circles). The filled triangles show the measured $w_{\mathrm{p}}(r_{\mathrm{p}})$ for the mock after correcting the effect of fibre collisions using the method described in the text. Bottom: the ratios of the uncorrected and the corrected $w_{\mathrm{p}}(r_{\mathrm{p}})$ relative to the 'true' $w_{\mathrm{p}}(r_{\mathrm{p}})$.

when the 2PCF is corrected for the effect of fibre collisions using the method described above. In the bottom panel, we plot the ratios of the uncorrected and the corrected $w_p(r_p)$ relative to the 'true' correlation function. As can be seen, our correction procedure works well. The turnover in the amplitude on small physical scales resulting from the lower sampling of galaxies in dense regions disappears and the corrected $w_p(r_p)$ is very close to the real one. It is notable that there is still a very small deficit in the corrected $w_p(r_p)$ on scales between 0.05 and 1 Mpc. This should not be a significant contribution to the bias between AGN and normal galaxies (see below), because fibre collisions are expected to affect the AGN and the reference galaxies in the same way.

5 OBSERVATIONAL RESULTS

5.1 AGN bias

We first compute $w_p^{\mathrm{AGN-ref}}(r_p)$, the cross-correlation of the AGN sample with respect to the reference sample. As described in Section 2, we have constructed two sets of 20 control samples. The first set is constructed by simultaneously matching redshift, stellar mass, concentration and stellar velocity dispersion, and the second set by additionally matching the 4000-Å break strength. We then compute $\bar{w}_p^{\mathrm{contr-ref}}(r_p)$, the average cross-correlation of the control samples with respect to the reference sample. The quantity $w_p^{\mathrm{AGN-ref}}(r_p)/\bar{w}_p^{\mathrm{contr-ref}}(r_p)$ then measures the *bias* of the AGN sample with respect to the control sample of non-AGN as a function of projected radius r_p.

The results are shown in Fig. 3. In the top panel, we plot $w_p^{\mathrm{AGN-ref}}(r_p)$ as circles. $\bar{w}_p^{\mathrm{contr-ref}}(r_p)$ is evaluated for the two sets of control
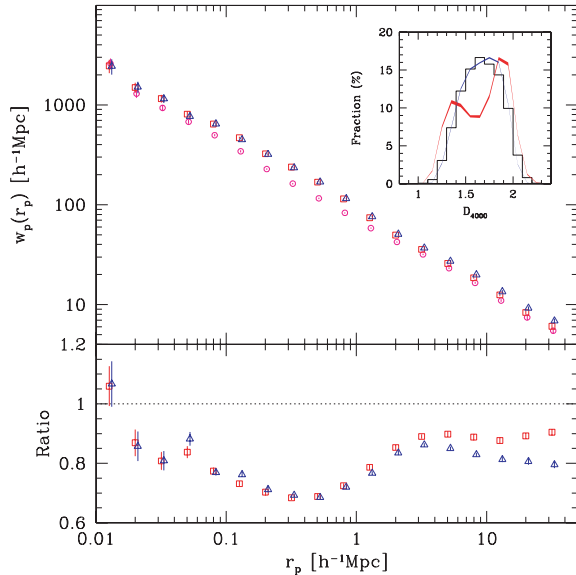

Figure 3. Top: projected cross-correlation function $w_p^{\mathrm{AGN-ref}}(r_p)$ between AGN and reference galaxies is plotted as circles. The average cross-correlation between two sets of 20 control samples of non-AGN and the same reference galaxies is plotted as squares and triangles (squares: redshift, stellar mass, stellar velocity dispersion and concentration are matched; triangles: the control samples are constructed by additionally matching the 4000-Å break strength). The inset compares the D_{4000} distribution for the two sets of control samples, with red for the former set and blue for the latter. The histogram shows the $D_n(4000)$ distribution for the AGN. Bottom: ratio of the $w_p(r_p)$ measurement of AGN to that of non-AGN. Symbols are the same as in the top panel.
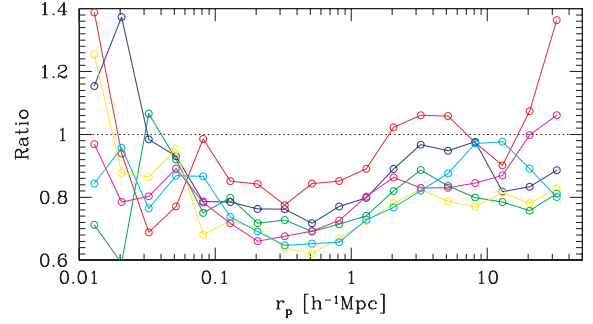

Figure 4. Ratio of the $w_p(r_p)$ measurement of AGN to that of non-AGN for six disjoint regions of the survey.

samples, and the results are plotted as squares for the first set and triangles for the second set. The measurement errors are estimated using the bootstrap resampling technique described in the previous section. In the bottom panel, we plot the ratio $w_p^{\mathrm{AGN-ref}}(r_p)/\bar{w}_p^{\mathrm{contr-ref}}(r_p)$ for the two control samples. The errors are estimated in the same manner as in the top panel. For clarity, squares and triangles in both panels have been slightly shifted along the r_p-axis.

Fig. 3 shows that there exists a *scale-dependent bias* in the distribution of AGN relative to that of normal galaxies. In particular, the ratio between the two cross-correlations appears to exhibit a pronounced 'dip' at scales between 100 kpc and 1 Mpc. We note that error bars estimated using the bootstrap resampling technique do not take into account effects due to cosmic variance. The coherence length of the large-scale structure is large, and even in a survey as big as the SDSS, this can induce significant fluctuations in the amplitude of the correlation function from one part of the sky to another. The difference in the clustering amplitude of AGN and non-AGN shown in Fig. 3 is only a 10–30 per cent effect, so it is important to test whether the dip seen in Fig. 3 is truly robust.

We have thus divided the survey into six different areas on the sky. Each subsample includes ∼12 000 AGN. We recompute the AGN bias for each of these subsamples, and the results are shown in Fig. 4. Note that in this plot, we only use a single control sample to compute the bias, not the average of 20 control samples as in Fig. 3. The scatter between the different curves in Fig. 4 thus provides an upper limit to the true error in the measurement of the bias. As can be seen, on small scales (<50 kpc), the different subsamples scatter in bias above and below unity. On scales between 0.1 and 1 Mpc, however, all six subsamples lie systematically below this line. On scales larger than 1–2 Mpc, five out of six subsamples show bias values below unity, but the effect is clearly less significant than on scales between 0.1 and 1 Mpc. In Fig. 5, we examine the dispersion in the bias measurement for the AGN sample as a whole caused by differences between the control samples. As can be seen, the scatter in the measurement of the bias between different control samples is considerably smaller than the scatter between different survey regions, showing that the cosmic variance is, in fact, the dominant source of uncertainty in our results. Once again, there is clear indication that AGN are antibiased relative to the control galaxies on scales larger than 100 kpc.

We conclude that on scales between 0.1 and 1 Mpc, AGN are significantly antibiased relative to non-AGN of the same stellar mass, concentration and stellar velocity dispersion. Fig. 3 shows that this antibias persists even when the mean age of the stellar population is matched in addition to stellar mass and structural parameters. We note that these scales are comparable to the diameters of the dark
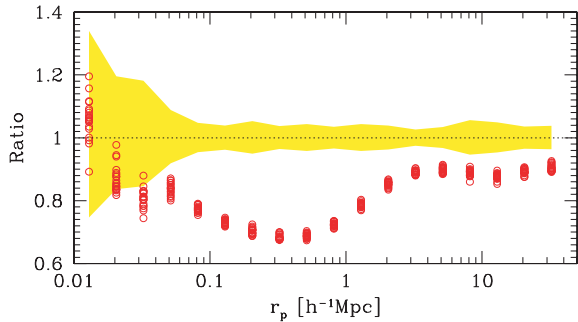
Figure 5. Ratio of the $w_p(r_p)$ measurement of AGN to that of non-AGN for 20 different non-AGN control samples. The yellow-shaded region shows the ratio of the $w_p(r_p)$ measurements of the different non-AGN control samples relative to each other.

matter haloes that are expected to host galaxies with stellar masses comparable to the objects in our sample. In Section 6, we construct halo occupation (HOD) models using mock catalogues constructed from the Millennium Simulation that can explain the antibias on these scales. As we will show, the same models naturally predict a smaller, but significant, antibias on scales larger than 1 Mpc.

5.2 Dependence on black hole mass and AGN power

It is interesting to study how AGN clustering depends on black hole mass and the strength of nuclear activity in the galaxy. As described above, we use the stellar velocity dispersion as an indicator of the black hole mass and divide all AGN into three subsamples according to $\log_{10}\sigma_*$. We then use the ratio $L[O\,\textsc{iii}]/M_{BH}$ as a measure of the accretion rate relative to the Eddington rate. We rank order all the AGN in a given interval of stellar velocity dispersion according to $L[O\,\textsc{iii}]/M_{BH}$ and define subsamples of 'powerful' and 'weak' AGN as those contained within the upper and lower 25th percentiles of the distribution of this quantity.

The results are shown in Fig. 6. Panels from left- to right-hand side correspond to different intervals of $\log_{10}\sigma_*$, as indicated at the top of the figure. The first two rows show the $w_p(r_p)$ measurements for the AGN and the corresponding control samples. The third row shows the ratio between the two. Red (blue) lines correspond to the powerful (weak) subsamples. Black lines show results for the sample as a whole.

As mentioned previously, there are two sets of control samples: Sample 1 is constructed by simultaneously matching redshift, stellar mass, concentration and stellar velocity dispersion; Sam-

ple 2 is constructed by additionally matching the 4000-Å break strength. For clarity, the main panels in Fig. 6 show the results only for Sample 1. The insets in the bottom row compare the ratio $w_p^{AGN-ref}(r_p)/\bar{w}_p^{contr-ref}(r_p)$ for the two control samples (black corresponds to set 1 and red to set 2). The two different control samples give very similar results on scales less than a few Mpc, but the large-scale antibias is more pronounced for Sample 2. (Note that this is also seen in Fig. 2.)

As can be seen, the 'dip' in clustering on scales between 0.1 and 1 Mpc is most pronounced for AGN with the largest central stellar velocity dispersions and the highest accretion rates. On scales smaller than 0.1 Mpc, more powerful AGN appear to be somewhat more strongly clustered than weaker AGN and more strongly clustered than galaxies in the control sample. The error bars on the measurements are large, however, and effect is not of high significance. In Fig. 7, we plot results for AGN of all velocity dispersions, but now in four different intervals of $L[O\,\textsc{iii}]/M_{BH}$, as indicated at the top of the figure. Once again, we see a marginal tendency for AGN with higher values of $L[O\,\textsc{iii}]/M_{BH}$ to be more strongly clustered on small scales.

5.3 Close neighbour counts

As already discussed, one of the problems with computing the galaxy–AGN cross-correlation function on very small physical scales in the SDSS is that corrections for the effect of fibre collisions are required. These corrections are statistical in nature and even if they are correct on average, they may still introduce systematic effects in our analysis. An alternative approach is to count the number of galaxies in the vicinity of AGN in the photometric sample, which is not affected by incompleteness. The disadvantage of using the photometric sample is that many of the close neighbours will not be truly nearby systems, but rather chance projections of foreground and background galaxies that lie along the line-of-sight. We can make a statistical correction for this by evaluating the counts around randomly placed 'galaxies' with the same assumed joint distribution of apparent magnitude and redshift as the AGN (control) samples.

In the upper panel of Fig. 8, we plot the average correlated neighbour count (i.e. after statistical correction for uncorrelated projected neighbours) within a given value of the projected radius R_p for the AGN sample (red) and the control samples (blue). The lower panel and its insets show the difference between the average correlated counts for the AGN and control samples as a function of R_p. The variance in the counts around the control galaxies estimated from
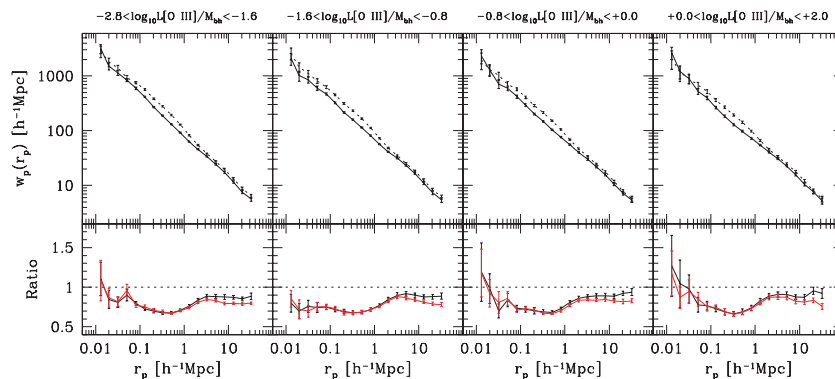

Figure 6. $w_p(r_p)$ in different $L[L\,\textsc{iii}]/M_\bullet$ bins (indicated above each panel), for all AGN (solid) and for control samples of non-AGN (dashed). The small panels give the ratio between the above two (black). The red lines are results where the 4000-Å break strength is also matched when constructing control samples.
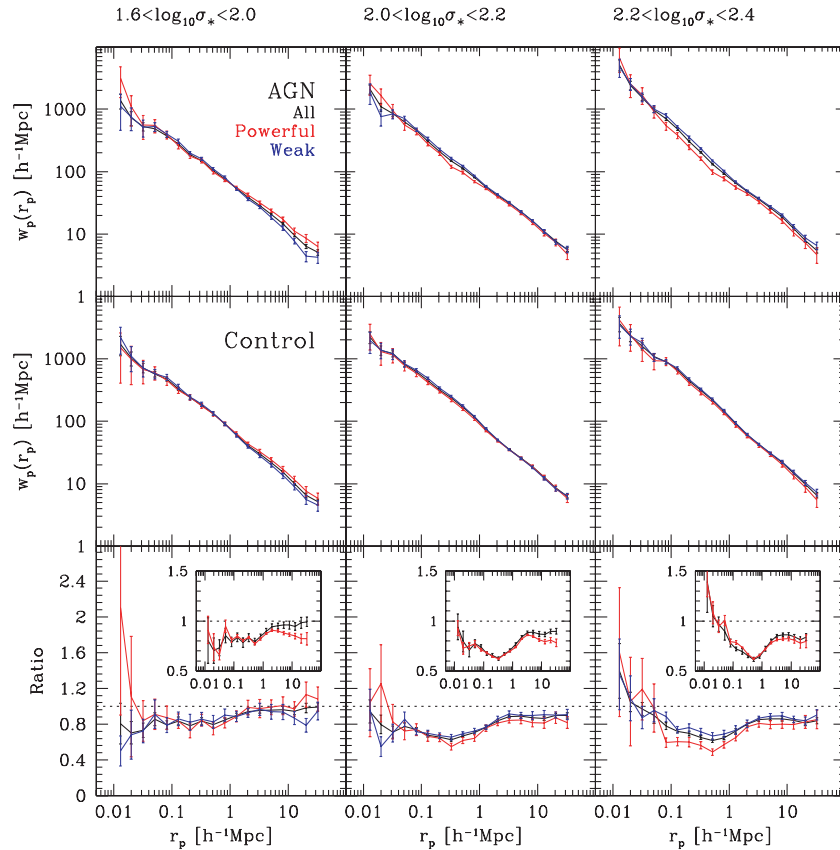
Figure 7. Top: projected cross-correlation $w_p(r_p)$ in different σ_* bins (indicated above each panel), for all AGN (black), powerful (red) and weak (blue) AGN. The powerful (weak) AGN are defined as the top (bottom) 25 per cent objects ordered by decreasing $L[\text{O\,III}]/M_\bullet$. The middle row is for control samples of non-AGN and the bottom row shows the ratio between the results for the AGN and the control samples. The insets in the bottom panels compare results for all AGN using different control samples. Black is for control samples constructed by matching redshift, stellar mass, stellar velocity dispersion and concentration, while red is for control samples where the 4000-Å break strength is also matched.

20 different control samples is shown in yellow. The AGN sample has an *r*-band limiting magnitude of 17.6, and the photomteric reference sample that we use is limited at $r = 19.0$. In order to ensure that we are counting similar neighbours at all redshifts, the counts only include those galaxies with $r < r_{\text{AGN}} + 1.4$ mag. In this analysis, the control sample is matched in *r*-band *apparent* magnitude as well as redshift, stellar mass, velocity dispersion and concentration. This ensures that we are counting galaxies to the same limiting magnitude around both the AGN and the control galaxies.

Fig. 8 shows that the counts around the AGN and the control galaxies match well on large scales. On small scales, there is a small but significant excess in the number of neighbours around AGN out to scales of $\sim$70 kpc. As may be seen from the bottom panel of Fig. 8, AGN are approximately twice as likely to have a near neighbour as galaxies in the control sample. This does not mean, however, that every AGN has a close companion. Fig. 8 also shows that only one in 100 AGN has an additional close ($R_p < 70$ kpc) neighbour as compared to the control galaxies. On scales larger than 100 kpc, the pair counts around the AGN dip below the counts around the control samples, leading to the 'antibias' discussed in the previous section. This may be compensated on scales larger than several Mpc, although such compensation is not required with our present statistics.

Fig. 9 shows the counts around AGN in four different ranges of $L[\text{O\,III}]/M_{\text{BH}}$. As can be seen, the excess on small scales increases

as a function of the accretion rate on to the black hole. However, the excess affects only a few per cent of the AGN, even for the objects in our highest $L[\text{O\,III}]/M_{\text{BH}}$ bin. We note that Serber et al. (2006) analysed galaxy counts around quasars compared to L_* galaxies at the same redshift and found a clear excess on scales less than 100 kpc, very similar to the $\sim$70 kpc scales where we see the upturn in the counts around our sample of narrow-line AGN. Serber et al. also found that the excess was largest for the most luminous quasars; the excess count reached values of $\sim$1 (i.e. significantly larger than the excess found for the most powerful narrow-line AGN in our sample) for quasars with *i*-band magnitudes brighter than -24. If we use the relation between [O\,III] line luminosity and quasar continuum luminosity of Zakamska et al. (2003) to compare the AGN in our sample to the quasars studied by Serber et al., we find that the luminosities where quasars begin to exhibit a significant excess count lie just beyond those of the AGN that populate our highest $L[\text{O\,III}]/M_{\text{BH}}$ bin.

Our conclusion, therefore, is that we do not find strong evidence that interactions and mergers are playing a significant role in triggering the activity in typical AGN in the local Universe. One caveat that should be mentioned is that if the activity is triggered *after* the merger has already taken place, our pair count statistics would not be a good diagnostic. In order to assess this, more work is needed to assess whether AGN exhibit any evidence for disturbed morphologies or structural peculiarities.
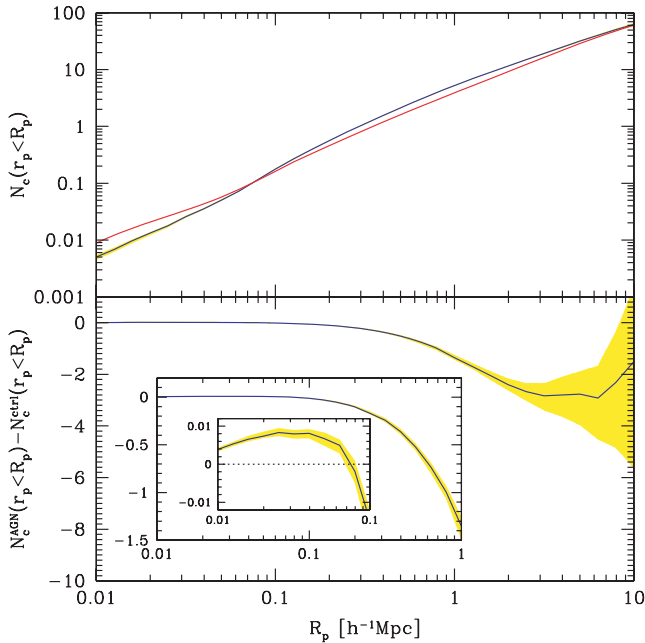
matter haloes where conditions are more favourable for continued fuelling of the central black hole. One obvious preferred location would be the halo centre where gas is expected to be able to reach high enough overdensities to cool via radiative processes. In the main body of the more massive dark matter haloes, most of the surrounding gas will have been shock heated to the virial temperature of the halo and will no longer be able to cool efficiently. In addition, the vast majority of galaxy–galaxy mergers within a halo will occur with the galaxy that is located at the halo centre (Springel et al. 2001).

In this section, we use our mock catalogues to test whether a model in which AGN are preferentially located at the centres of dark matter haloes can fit our observational results. As mentioned in Section 2, we have used the methodology introduced by Wang et al. (2006) to assign stellar masses to the galaxies in the catalogues. These authors adopted parametrized functions to relate galaxy properties such as stellar mass to the quantity $M_{\rm infall}$ defined as the mass of the halo at the epoch where the galaxy was the last central dominant object in its own halo. It was demonstrated that these parametrized relations were able to provide an excellent fit to the basic statistical properties of galaxies in the SDSS, including the stellar mass function and the shape and amplitude of the 2PCF function evaluated in different stellar mass ranges. We now introduce a simple model in which $p_{\rm AGN}$, the probability of a galaxy to be an AGN, depends only on whether it is the central galaxy of its own halo.

In order to create mock AGN and control catalogues that we can compare directly with the observational data, we follow the following procedure. For every AGN in our sample, we select galaxies from the mock catalogue that have the same stellar mass and redshift. We then choose an AGN from among these galaxies based on whether they are central or satellite systems. The control galaxies are selected at random from the same set. The AGN and control samples are then cross-correlated with a reference sample that is drawn from the mock catalogue in exactly the same way as our real SDSS reference sample. The top panel in Fig. 10 shows how the AGN–reference galaxy cross-correlation function changes as a function of the fraction of AGN that are central galaxies. Note that if the probability of being an AGN is *independent* of whether the galaxy is a central or a satellite system, 73 per cent of the AGN will be central galaxies. The bottom panel of Fig. 10 shows how the ratio between the AGN and control galaxy cross-correlation functions varies as this fraction changes.

As the fraction of centrally located AGN increases, the 'dip' on scales smaller than 1 Mpc becomes more and more pronounced. There is also a small decrease of the ratio $w^{\rm AGN-ref}(r_{\rm p})/w^{\rm control-ref}(r_{\rm p})$ on large scales. The latter effect arises because, as more and more AGN are required to be central galaxies in their own haloes, they also shift into lower mass haloes. High-mass haloes are less abundant
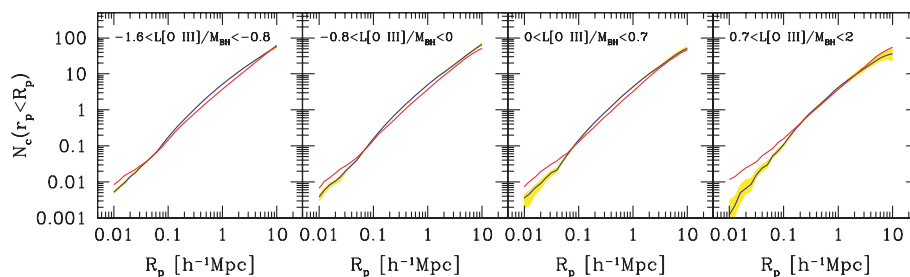

Figure 8. Top: average counts of galaxies in the photometric sample ($r_{\rm lim} < 19$) within a given projected radius $R_{\rm p}$ from the AGN (red) and from the control galaxies (blue). Bottom: the difference between the counts around the AGN and the control galaxies is plotted as a function of $R_{\rm p}$. The yellow bands indicate the variance in the results between the 20 different control samples.

6 INTERPRETATION OF AGN CLUSTERING USING HALO OCCUPATION MODELS

In the previous section, we showed that the main difference in the AGN–galaxy cross-correlation function with respect to that of a closely matched control sample of non-AGN is that AGN are more weakly clustered on scales between 100 kpc and 1 Mpc. On larger scales, there is a much smaller difference in the clustering signal of AGN and non-AGN. The clustering amplitude of AGN on large scales provides a measure of the mass of the dark matter haloes that host these objects. The fact that there is only a small difference between the AGN and the control sample tells us that AGN are found in roughly similar dark matter haloes to non-AGN with the same stellar masses and structural properties.

The physical scales of 0.1–1 Mpc, where we do see strong differences in the clustering of AGN and non-AGN, are similar to the virial diameters of the dark matter haloes that are expected to host galaxies with luminosities of $\sim L_*$ (Mandelbaum et al. 2006). The simplest interpretation of the AGN antibias on these scales, therefore, is that AGN occupy preferred positions within their dark

Figure 9. Same as the top panel of Fig. 8, but for four subsamples of AGN with different L[O III]/$M_\bullet$, as indicated in each panel.
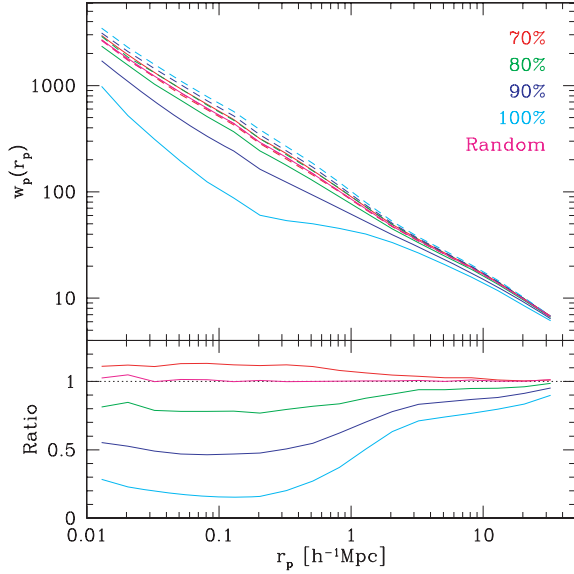
Figure 10. Top: the AGN–reference cross-correlation function calculated from the mock catalogues is plotted as solid curves, while the control/reference cross-correlations are plotted as dashed curves. The different colours indicate models in which a given percentage (as indicated on the plot) of the AGN are located at the centres of their own dark matter haloes. Bottom: the ratio between the AGN–reference galaxy cross-correlation functions and the control galaxy–reference galaxy correlation functions are plotted for the same set of models.

Figure 11. Top: the AGN–reference cross-correlation function calculated from the SDSS is plotted as solid circles. The open circles show the control galaxy–reference galaxy cross-correlation function from the SDSS. Results for AGN and control galaxies for our best-fitting model are plotted as solid and dashed red lines. Bottom: solid circles show the ratio $w^{\mathrm{AGN-ref}}(r_p)/w^{\mathrm{control-ref}}(r_p)$ for the SDSS sample. The solid red curve shows the result for our best-fitting models. The error bars indicate the uncertainty due to cosmic variance as estimated from 20 different mock catalogues (see text for more details). The dashed red curves indicate the variance between different control samples from different mock catalogues (see text).

than low-mass haloes and by definition, each halo can only contain one central galaxy. As the central galaxy criterion on AGN becomes more stringent, fewer AGN will reside in haloes of 10^{14}–10^{15} M$_\odot$ and more in haloes with 10^{12}–10^{13} M$_\odot$. The biggest effect, however, the dip on scales less than 1 Mpc, results from the fact that central galaxies in lower mass haloes have fewer neighbours than non-central galaxies within rich groups.

In Fig. 11, we compare the results of our simple model with the observational data. In the top panel, the solid and open circles show the cross-correlation functions of the AGN and the control samples, respectively. The solid and dashed lines show our best-fitting models, in which 84 per cent of all AGN are located at the centres of their own dark matter haloes. In the bottom panel, we compare the ratio of $w(r_p)$ for the observed AGN and control galaxies with that obtained for the model. The error bars plotted on the model curve provide an estimate of the uncertainty in the result due to *cosmic variance effects*. In order to estimate these errors, we have created 20 different mock catalogues by repositioning the virtual observer at random within the simulation volume. For each mock catalogue, we repeat our computations of the AGN and control galaxy cross-correlation functions. The errors are then calculated by looking at the variance in the ratio $w^{\mathrm{AGN-ref}}(r_p)/w^{\mathrm{control-ref}}(r_p)$ for AGN and control galaxies selected from different catalogues. The errors estimated in this way are similar in size to the errors estimated by calculating the variance in $w^{\mathrm{control-ref}}(r_p)$ from different mock catalogues; these errors are indicated by the dashed red lines in Fig. 11. As can be seen, the model provides a good fit to the observations from scales of ∼30 kpc out to scales beyond 10 Mpc. On scales smaller than 30 kpc, the AGN show a small, but significant, excess in clustering with respect to the model. This is nicely in line with the results presented in the previous section.
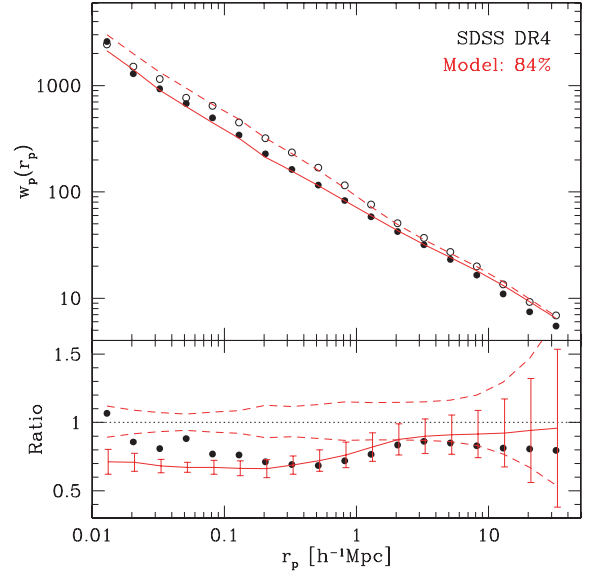
7 SUMMARY AND CONCLUSIONS

In this paper, we have analysed the clustering of type 2 narrow-line AGN in the local Universe using data from the SDSS. The two physical questions we wish to address are as follows. (i) How do the locations of galaxies within the large-scale distribution of dark matter influence ongoing accretion on to their central black holes? (ii) Is AGN activity triggered by interactions and mergers between galaxies? To answer these questions, we analyse the scale-dependence of the AGN–galaxy cross-correlation function relative to control samples of non-AGN that are closely matched in stellar mass, redshift, structural properties and mean stellar age as measured by the 4000-Å break strength. This close matching is important because previous work has established that the clustering of galaxies depends strongly on properties such as luminosity, stellar mass, colour, spectral type, mean stellar age, concentration and stellar surface mass density (Norberg et al. 2002; Zehavi et al. 2002, 2005; Li et al. 2006). Previous work has also established that AGN are not a random subsample of the underlying galaxy population. Rather, they are found in massive, bulge-dominated galaxies; powerful AGN tend to occur in galaxies with smaller black holes and younger-than-average stellar populations for their mass (Kauffmann et al. 2003; Heckman et al. 2004). If we wish to understand whether there is a real physical connection between the location of a galaxy and the accretion state of its central black hole, it is important that we normalize out these zeroth-order trends with galaxy mass, structure and mean stellar age.

When we compare the clustering of AGN relative to carefully matched control samples, and take the errors due to cosmic variance into account, we obtain the following results.

(i) On scales larger than a few Mpc, the clustering amplitude of AGN hosts does not differ significantly from that of similar but inactive galaxies.

(ii) On scales between 100 kpc and 1 Mpc, AGN hosts are clustered more weakly than control samples of similar but inactive galaxies.

(iii) On scales less than 70 kpc, AGN cluster more strongly than inactive galaxies, but the effect is weak. The excess number of close companions is only one per 100 AGN.

Our clustering results on large scales demonstrate that the host galaxies of AGN are found in similar dark matter haloes to inactive galaxies with the same structural properties and stellar masses. We have used mock catalogues constructed from high-resolution *N*-body simulations to show that the AGN antibias on scales between 0.1 and 1 Mpc can be explained by AGN residing preferentially at the centres of their dark matter haloes. Our result on small scales indicates that although interactions may be responsible for triggering AGN activity in a minority of galaxies, an alternative mechanism is required to explain the nuclear activity in the majority of these systems.

As we have already mentioned, it is easy to understand why dark matter halo centres may be preferential places for ongoing growth of black holes. These are the regions where gas would be expected to cool and settle through radiative processes. In addition, dynamical friction will erode the orbits of satellite galaxies within a dark matter halo until they sink to the middle and merge with the central object. Both these processes may bring fresh gas to the central galaxy and fuel episodes of nuclear activity and black hole growth. As we have seen, however, the evidence for an excess number of close neighbours around AGN is rather weak, perhaps because in most cases the offending satellite has already been swallowed. We also note that even the most powerful AGN in our sample are less luminous than the quasars with $M(i) < -24$ for which Serber et al. (2006) detected an excess number of companions on small scales. What about the evidence for cooling?

Direct observational evidence for cooling from hot X-ray emitting gas at the centres of dark matter haloes has also been elusive. Benson et al. (2000) used *ROSAT* Position-Sensitive Proportional Counter (PSPC) data to search for extended X-ray emission from the haloes of three nearby, massive spiral galaxies. Their 95 per cent upper limits on the bolometric X-ray luminosities of the haloes show that the present-day accretion from any hot virialized gas surrounding the galaxies is very small. Recently, Pedersen et al. (2006) detected a gaseous halo around the quiescent spiral NGC 5746 using *Chandra* observations, but this remains the only spiral galaxy with evidence for ongoing accretion from an extended reservoir of hot gas. In clusters, X-ray spectroscopy has shown that most of the gas does not manage to cool below 10^7 K (e.g. David et al. 2001; Peterson et al. 2001).

Gas accretion in the form of cold H I-emitting clouds is, however, much less well-constrained. In recent work, Kauffmann et al. (2006) studied a volume-limited sample of bulge-dominated galaxies with data from both the SDSS and the Galaxy Evolution Explorer (GALEX) satellite. Almost all galaxies with bluer-than-average NUV-*r* colours were found to be AGN. By analysing GALEX images, these authors demonstrated that the excess ultraviolet (UV) light is nearly always associated with an extended disc. They then went on to study the relation between the UV-bright outer disc and the nuclear activity in these galaxies. The data indicate that the presence of the UV-bright disc is a necessary but not sufficient condition for strong AGN activity in a galaxy. They suggest that the disc pro-

vides a reservoir of fuel for the black hole. From time to time, some event transports gas to the nucleus, thereby triggering the observed AGN activity.

The GALEX results indicate that the extended discs of galaxies play an important role in the fuelling of AGN. The clustering results from the SDSS indicate that AGN are preferentially located at the centres of dark matter haloes. In theoretical models, rotationally supported discs are *expected* to form at the centres of dark matter haloes (Mo, Mao & White 1998). After the galaxy is accreted by more massive haloes and becomes a satellite system, the discs may lose their gas via processes such as ram-pressure stripping (e.g. Cayatte et al. 1994). Discs located at halo centres are likely to survive for longer periods. Dynamical perturbations driven by the dark matter near the centres of the haloes may result in gas inflows and fuelling of the central black hole (see Gao & White 2006, for a recent discussion). Further progress in understanding the AGN phenomenon in the local Universe will require detailed modelling of the observable components of galaxies within evolving dark matter haloes, as well as further investigation of the connection between AGN activity and phenomena such as bars, warps, lopsided images and asymmetric rotation curves.

REFERENCES

Adelman-McCarthy J. K. et al., 2006, ApJS, 162, 38
Barnes J. E., Hernquist L., 1992, ARA&A, 30, 705
Barrow J. D., Bhavsar S. P., Sonoda D. H., 1984, MNRAS, 210, 19P
Benson A. J., Bower R. G., Frenk C. S., White S. D. M., 2000, MNRAS, 314, 557

Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25

Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, AJ, 125, 2276

Blanton M. R. et al., 2005, AJ, 129, 2562

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Cayatte V., Kotanyi C., Balkowski C., van Gorkom J. H., 1994, AJ, 107, 1003

Charlot S., Fall S. M., 2000, ApJ, 539, 718

Colless M. et al., 2001, MNRAS, 328, 1039

Constantin A., Vogeley M. S., 2006, preprint (astro-ph/0601717)

Croom S. M., Shanks T., Boyle B. J., Smith R. J., Miller L., Loaring N. S., Hoyle F., 2001, MNRAS, 325, 483

Croom S. M. et al., 2005, MNRAS, 356, 415

Croton D. J. et al., 2006, MNRAS, 365, 11

Dahari O., 1984, AJ, 89, 966

David L. P., Nulsen P. E. J., McNamara B. R., Forman W., Jones C., Ponman T., Robertson B., Wise M., 2001, ApJ, 557, 546

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Gao L., White S. D. M., 2006, preprint (astro-ph/0605687)

Gilli R. et al., 2005, A&A, 430, 811

Grogin N. A. et al., 2005, ApJ, 627, L97

Gunn J. E. et al., 1998, AJ, 116, 3040

Gunn J. E. et al., 2006, AJ, 131, 2332

Hamilton A. J. S., Tegmark M., 2002, MNRAS, 330, 506

Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109

Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2001, AJ, 122, 2129

Ivezić Ž. et al., 2004, AN, 325, 583

Kauffmann G. et al., 2003, MNRAS, 346, 1055

Kauffmann G. et al. 2006, MNRAS, preprint (astro-ph/0609436)

Keel W. C., Kennicutt R. C., Jr, Hummel E., van der Hulst J. M., 1985, AJ, 90, 708

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

Lupton R., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., Yasuda N., 2001, ASPC, 238, 269

Magliocchetti M., Maddox S. J., Lahav O., Wall J. V., 1999, MNRAS, 306, 943

Magliocchetti M. et al., 2004, MNRAS, 350, 1485

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

Mihos J. C., Hernquist L., 1996, ApJ, 464, 641

Miller C. J., Nichol R. C., Gómez P. L., Hopkins A. M., Bernardi M., 2003, ApJ, 597, 142

Mo H. J., Jing Y. P., Boerner G., 1992, ApJ, 392, 452

Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

Mullis C. R., Rosati P., Lamer G., Böhringer H., Schwope A., Schuecker P., Fassbender R., 2005, ApJ, 623, L85

Norberg P. et al., 2002, MNRAS, 332, 827

Overzier R. A., Röttgering H. J. A., Rengelink R. B., Wilman R. J., 2003, A&A, 405, 53

Pedersen K., Rasmussen J., Sommer-Larsen J., Toft S., Benson A. J., Bower R. G., 2006, New Astron., 11, 465

Peterson J. R. et al., 2001, A&A, 365, L104

Petrosian A. R., 1982, Afz, 18, 548

Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, AJ, 125, 1559

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schmitt H. R., 2001, AJ, 122, 2243

Serber W., Bahcall N., Ménard B., Richards G., 2006, ApJ, 643, 68

Smith J. A. et al., 2002, AJ, 123, 2121

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

Springel V. et al., 2005, Nat, 435, 629

Stoughton C. et al., 2002, AJ, 123, 485

Strauss M. A. et al., 2002, AJ, 124, 1810

Tegmark M. et al., 2002, ApJ, 571, 191

Tremaine S. et al., 2002, ApJ, 574, 740

Tremonti C. A. et al., 2004, ApJ, 613, 898

Tucker D. et al. 2005, AJ, submitted

Wake D. A. et al., 2004, ApJ, 610, L85

Wang L., Li C., Kauffmann G., De Lucia G., 2006, MNRAS, 371, 537

Waskett T. J., Eales S. A., Gear W. K., McCracken H. J., Lilly S., Brodwin M., 2005, MNRAS, 363, 801

Yang X., Mo H. J., Jing Y. P., van den Bosch F. C., Chu Y., 2004, MNRAS, 350, 1153

York D. G. et al., 2000, AJ, 120, 1579

Zakamska N. et al., 2003, AJ, 126, 2125

Zehavi I. et al., 2002, ApJ, 571, 172

Zehavi I. et al., 2005, ApJ, 630, 1

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