

Clustering of Low-Redshift ($z \leq 2.2$) Quasars from the Sloan Digital Sky Survey

Nicholas P. Ross,^{1*} Yue Shen,² Daniel E. Vanden Berk,¹ Andrew J. Connolly,³
Gordon T. Richards,⁴ Donald P. Schneider,¹ Michael A. Strauss,²
Patrick B. Hall,⁵ Neta A. Bahcall,² Robert J. Brunner,⁶ et al.⁷

¹*Department of Astronomy and Astrophysics, The Pennsylvania State University, 525 Davey Laboratory, University Park, PA 16802, U.S.A.*

²*Princeton University Observatory, Princeton, NJ 08544, U.S.A.*

³*Department of Astronomy, University of Washington, Box 351580, Seattle, WA 98195, U.S.A.*

⁴*Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, U.S.A.*

⁵*Department of Physics and Astronomy, York University, Toronto, ON M3J 1P3, Canada*

⁶*Department of Astronomy, MC-221, University of Illinois, 1002 West Green Street, Urbana, IL 61801, U.S.A.*

10 September 2008

ABSTRACT

We present results for the Quasar Two-Point Correlation Function, ξ_Q , over the redshift range $0.3 \leq z \leq 2.2$ using data from the Sloan Digital Sky Survey (SDSS). Using nearly 50,000 quasars with spectroscopic redshifts from the Data Release 5 Quasar Catalogue, our study represents the largest sample used for this type of investigation to date. With our redshift range and an areal coverage of $\approx 5,700 \text{ deg}^2$, we sample nearly $40 h^{-3} \text{ Gpc}^3$ (comoving) of the Universe in volume, assuming the current ΛCDM cosmology.

Over our redshift range, we find that the redshift-space correlation function, $\xi(s)$, is well-fit by a single power-law, with $s_0 = 5.95 \pm 0.45 h^{-1} \text{ Mpc}$ and $\gamma_s = 1.16^{+0.11}_{-0.16}$ when fit over $1.0 \leq s \leq 25.0 h^{-1} \text{ Mpc}$. We find no evidence for deviation from $\xi(s) = 0$ at scales of $s > 100 h^{-1} \text{ Mpc}$ but do observe redshift-space distortions in the 2-D $\xi(r_p, \pi)$ measurement. Using the projected correlation function, $w_p(r_p)$, we calculate the real-space correlation length, $r_0 = 5.45^{+0.35}_{-0.45} h^{-1} \text{ Mpc}$ and $\gamma = 1.90^{+0.04}_{-0.03}$, over $1.0 \leq r_p \leq 130.0 h^{-1} \text{ Mpc}$ scales.

Splitting our sample up into redshift slices, we find evidence for very weak evolution of quasar clustering with the correlation length staying roughly constant at $s_0 \sim 6 - 7 h^{-1} \text{ Mpc}$ at $z \lesssim 2.5$, and only increasing at redshifts greater than this. Comparing our results to recent deep X-ray surveys, we find reasonable agreement with X-ray selected AGN at $z \sim 1$ and using SDSS-RASS matched catalogues, predict a clustering length of $r_0 > 30 h^{-1} \text{ Mpc}$ for RASS-SDSS quasars if clustering was dependent on X-ray luminosity.

We find the linear bias evolves from $b \sim 1.4$ at $z = 0.5$ to $b \sim 3$ at $z = 2.2$, with $b(z = 1.27) = 2.06 \pm 0.03$ for the full sample. Thus, we calculate a value of $\beta(z = 1.27) = 0.43 \pm 0.01$ for the dynamical infall parameter but caution this result does not take into account “Fingers of God” or redshift errors.

We compare our data to analytical models and find that quasars inhabit dark matter haloes of constant mass $M_{\text{halo}} = 5 \times 10^{11} - 4 \times 10^{12} h^{-1} M_\odot$ from redshifts $z \sim 2.5$ (the peak of quasar activity) to $z \sim 0$ and thus, the ratio of the halo mass for a typical quasar to the mean halo mass at the same epoch drops as you approach redshift $z = 0$. We also find that a model of “Inefficient feedback”, where quasars grow hierarchically from high- z until $z \sim 2$, with multiple epochs of activity at high redshifts, fits the current observational data out to $z = 6$, though deeper quasar surveys are needed to break model degeneracies. Finally, by comparing to recent high- z clustering measurements, of Lyman Break Galaxies, we suggest that these objects may be the progenitors of low, $z < 2$ quasars.

Key words: clustering – quasars: general – cosmology: observations – large-scale structure of Universe. observations – Sky Surveys

1 INTRODUCTION

Understanding how and when the structures we see in the local Universe formed from the initial conditions present in the early Universe is one of the fundamental goals of modern observational cosmology. By tracing the evolution of clustering with cosmic epoch, we have the potential to understand the growth of structure and its relation to the energy and matter content of the Universe, including the relationship between the dark matter and the luminous galaxies and quasars that we observe.

As such, one of the primary science goals of the Sloan Digital Sky Survey (SDSS; York et al. 2000) was to measure the large-scale distribution of galaxies and quasars, and in particular, to determine the spatial clustering of quasars as a function of redshift. Shen et al. (2007) report on the clustering of high, $z \geq 2.9$ quasars from the SDSS; in this paper, we investigate the spatial clustering from redshift $z = 2.2$ to the present day, i.e. the evolution of quasar clustering over nearly 80% of the age of the Universe (the gap in redshift being a consequence of the optical selection techniques used in the SDSS).

Due to their high intrinsic luminosity, quasars are seen to large cosmological distances, and are thus good probes of large-scale structure (LSS) and its evolution. However, until recently, quasar studies were plagued by low-number statistics, leading to shot noise, and samples covered only over small areas of sky, leading to sample variance. With the advent of large-area ($\gtrsim 1000 \text{ deg}^2$) surveys with efficient selection techniques, these limitations have been overcome and the number of known quasars has increased by more than an order magnitude in the last decade, thanks mainly to the 2dF QSO Redshift Survey (2QZ; Boyle et al. 2000; Croom et al. 2004) and the SDSS. At present, there are over 100 000 objects spectroscopically classified as quasars in the SDSS. The latest quasar catalogue (Schneider et al. 2007) contains nearly 80 000 objects. Using the data from these large surveys, we are now in a position to make high-precision measurements of quasar clustering properties.

The two-Point Correlation Function (2PCF), ξ , is a simple but powerful statistic commonly employed to quantify the clustering properties of a given class of object (Peebles 1980). The observed value of ξ for quasars can be related to the underlying (dark) matter density distribution via

$$\xi(r)_{\text{quasar}} = b_Q^2 \xi(r)_{\text{matter}} \quad (1)$$

where $\xi(r)_{\text{matter}}$ is the mass correlation function and b_Q is the linear bias parameter for quasars, which links the visible quasars to the underlying matter density fluctuations. Although equation 1 defines b_Q , and there are theoretical reasons to believe that b_Q is scale-independent on large-scale (e.g. Scherrer & Weinberg 1998), we do not know *a priori* if this is the case.

With certain assumptions, the measurement and interpretation of the bias can lead to determination of the dark matter halo properties of quasars and to quasar lifetimes (t_q , Martini & Weinberg 2001; Haiman & Hui 2001). In the standard scenario, quasar activity is triggered by accretion onto a central, supermassive black hole (SMBH, e.g. Salpeter 1964; Lynden-Bell 1969; Rees 1984). Given that the growth of the SMBH is theoretically related to that of the underlying dark matter halo (Baes et al. 2003; Wyithe & Loeb

2005; Wyithe & Padmanabhan 2006) and the halo properties are correlated with the local density contrast, clustering measurements provide an insight into quasar and black hole physics (da Ângela et al. 2008). This information includes constraining η , the fraction of the Eddington luminosity at which quasars shine (Wyithe & Loeb 2005).

Early measurements of the quasar 2PCF (e.g. Arp 1970; Hawkins & Reddish 1975; Osmer 1981; Shanks et al. 1983, 1987) measured statistically significant clustering on scales of a few $h^{-1} \text{ Mpc}$, for both the quasar auto-correlation function and cross-correlation with galaxies. This result has been confirmed with data from more recent surveys, (e.g. Croom et al. 2005; Porciani et al. 2004). The Quasar 2PCF is typically fit to a single power-law of the form,

$$\xi(r) = (r/r_0)^{-\gamma} \quad (2)$$

over the range $1 \text{ } h^{-1} \text{ Mpc} \leq r \leq 100 \text{ } h^{-1} \text{ Mpc}$. Here, r_0 is the correlation length quoted in comoving coordinates and γ is the power-law slope. Typical measured correlation lengths and slopes for quasars at redshift $z \sim 1.5$ are $r_0 = 5 - 6 \text{ } h^{-1} \text{ Mpc}$ and $\gamma \sim 1.5$, respectively.

The evolution of the quasar correlation function has been disputed for a long time, with some authors claiming r_0 either decreased or only weakly evolved with redshift (e.g. Iovino & Shaver 1988; Croom & Shanks 1996), while others reported an increase with redshift (e.g. Kundic 1997; La Franca et al. 1998). However, with the advent of the 2QZ Survey, r_0 has been shown to evolve with high significance, in the sense quasar clustering increases with redshift (Croom et al. 2001; Porciani et al. 2004; Croom et al. 2005). In particular, Croom et al. (2005) used over 20 000 objects from the final 2QZ dataset to measure the redshift-space two-point correlation function, $\xi(s)$, over the redshift range $0.3 < z < 2.2$ and found a significant increase in the clustering amplitude at high redshift. The quasar bias was found to be a strong function of redshift, with an empirical dependence of

$$b_Q(z) = (0.53 \pm 0.19) + (0.289 \pm 0.035)(1+z)^2. \quad (3)$$

These values were used to derive the mean dark matter halo (DMH) mass occupied by quasars, giving a redshift-independent value of $M_{\text{DMH}} = (3.0 \pm 1.6) \times 10^{12} h^{-1} M_\odot$. Independent analysis of the 2QZ data by Porciani et al. (2004) confirmed these findings.

Using the SDSS, Shen et al. (2007) found that redshift $2.9 \leq z \leq 5.4$ quasars are significantly more clustered than their $z \sim 1.5$ counterparts, having a real-space correlation length and power-law slope of $r_0 = 15.2 \pm 2.7 \text{ } h^{-1} \text{ Mpc}$ and $\gamma = 2.0 \pm 0.3$, respectively, over the scales $4 \text{ } h^{-1} \text{ Mpc} \leq r_p \leq 150 \text{ } h^{-1} \text{ Mpc}$ (where r_p is the separation from the projected correlation function, $w_p(r_p)$). Shen et al. (2007) also find that bias increases with redshift, with $b_Q \sim 8$ at $z = 3.0$ and $b_Q \sim 16$ at $z = 4.5$.

Myers et al. (2006, 2007), also using the SDSS, examined the clustering of photometrically selected quasar candidates over $\sim 50 \text{ } h^{-1} \text{ kpc}$ to $\sim 20 \text{ } h^{-1} \text{ Mpc}$ scales. In this sample, quasar redshifts were assigned from photometric rather than spectral information (Richards et al. 2001). They found that the linear bias, b_Q increases with redshift, from $b_Q = 1.93$ at redshifts $0.4 \leq z < 1.0$ to $b_Q = 2.84$ at $2.1 \leq z < 2.8$.

Padmanabhan et al. (2008) measured the clustering

of photometrically selected luminous red galaxies (LRGs) around a low, $0.2 < z < 0.6$, redshift sample of quasars, with both LRG and quasar samples coming from the SDSS. They determined a large-scale quasar bias $b_Q = 1.09 \pm 0.15$ at a median redshift of $z = 0.43$. After taking into account measurement and interpretation subtleties, the results from Padmanabhan et al. (2008), are in qualitative agreement with those from Serber et al. (2006), who find that $M_i \leq -22$, $z \leq 0.4$ quasars are located in higher local overdensities than typical L^* galaxies. Serber et al. (2006) suggest that quasars typically reside in L^* galaxies, but have a local excess of neighbors within $\sim 0.15 - 0.7 \ h^{-1}$ Mpc, which contributes to the triggering of quasar activity through mergers and other interactions. Hennawi et al. (2006), Myers et al. (2007) and Myers et al. (2008) reach similar conclusions by looking at pairs of quasars on $< 1 \ h^{-1}$ Mpc scales. The quasar correlation function shows an excess over a power law, and Hennawi et al. (2006) suggested that the small-scale excess can be attributed to dissipative interaction events that trigger quasar activity in rich environments.

Recently, da Ângela et al. (2008) combined data from the 2QZ and the 2SLAQ Survey (2dF-SDSS LRG And QSO Survey; Croom et al. 2008), to investigate quasar clustering and break the “ $L - z$ degeneracy”. Due to the evolution of the quasar luminosity function and the flux-limited nature of most quasar samples, there is a strong correlation between redshift and luminosity in these samples, i.e. a $L - z$ degeneracy. The extra dynamic range of the 2SLAQ QSO survey adds to the 2QZ by targetting fainter objects over the same redshift range. da Ângela et al. (2008) estimate the mass of the dark matter haloes which quasars inhabit to be $\sim 3 \times 10^{12} h^{-1} M_\odot$, in agreement with Croom et al. (2005), a value that does not evolve strongly with redshift or depend on QSO luminosity. Their results also suggest that quasars of different luminosities may contain black holes of similar mass.

There have also been recent advances in theoretical predictions of the quasar correlation function and its evolution with redshift (Lidz et al. 2006; Hopkins et al. 2007; Shankar et al. 2007; Hopkins et al. 2008; Basilakos et al. 2008). Hopkins et al. (2007) describe the aforementioned observational results well, with some important caveats, and we return to these models in detail in Section 4.

In this paper, we shall measure the quasar 2PCF for redshifts $z \leq 2.2$, using the largest sample of spectroscopically identified quasars to date. We will investigate the dependence of quasar clustering strength with redshift and luminosity, allowing us to test current quasar formation and evolution models.

This paper is organised as follows. In Section 2 we present our data sample, mentioning several effects that could give rise to systematics in the measurements. In Section 3 we briefly describe the techniques involved in measuring the two-point correlation function and in Section 4 we present our results. In Section 5 we compare and contrast our evolutionary results with recent observational results in the literature, and we conclude in Section 6. Appendix A gives technical details for the SDSS, Appendix B describes our error analysis and Appendix C carries out a series of systematic checks.

In our companion paper, Shen et al. (2008, in preparation), we shall expand our investigations on the cluster-

ing of SDSS quasars at redshifts $z < 2.5$. Using the same data as we examine here, Shen et al. studies the dependence of quasar clustering on luminosity, virial black hole mass, quasar colour and radio loudness.

We assume the currently preferred flat, “Lambda Cold Dark Matter” (Λ CDM) cosmology where $\Omega_b = 0.042$, $\Omega_m = 0.237$, $\Omega_\Lambda = 0.763$ (Sánchez et al. 2006; Spergel et al. 2007) and quote distances in h^{-1} Mpc to aid in ease of comparisons with previous results in the literature. Since we are measuring objects with redshifts resulting from the Hubble flow, all distances herein are given in comoving coordinates. Where a value of Hubble’s Constant is assumed e.g. for absolute magnitudes, this will be quoted explicitly. Our magnitudes are based on the AB zero-point system (Oke & Gunn 1983).

2 DATA

Much care must be taken when constructing a dataset that is valid for a statistical analysis. In this section, and Appendix A we describe the various samples we use to investigate potential systematic effects in our clustering measurements. Appendix A provides some of the technical details for the SDSS, discussing the Catalogue Archive Server (CAS) and the SDSS Survey geometry.

2.1 The Sloan Digital Sky Survey

The SDSS uses a dedicated 2.5m wide-field telescope (Gunn et al. 2006) to collect light for 30 2k×2k CCDs (Gunn et al. 1998) over five broad bands - *ugriz* (Fukugita et al. 1996) - in order to image $\sim \pi$ steradians of the sky. The imaging data are taken on dark photometric nights of good seeing (Hogg et al. 2001) and are calibrated photometrically (Smith et al. 2002; Ivezić et al. 2004; Tucker et al. 2006; Padmanabhan et al. 2008), and astrometrically (Pier et al. 2003), and object parameters are measured (Lupton et al. 2001; Stoughton et al. 2002).

Using the imaging data, quasar target candidates are selected for spectroscopic follow-up based on their colours, magnitudes and detection in the FIRST radio survey (Becker et al. 1995), as described by Richards et al. (2002). Unless stated otherwise, all quoted SDSS photometry has been corrected for Galactic extinction following Schlegel et al. (1998). Here we are concerned with only those quasars selected as primary targets, i.e. the main quasar selection (Richards et al. 2002). Low-redshift, $z \lesssim 3$, quasar targets are selected based on their location in *ugri*-colour space and the high-redshift, $z \gtrsim 3$, objects in *griz*-colour space. The quasar candidate sample is flux-limited to $i = 19.1$ but since high-redshift quasars are rare, objects lying in regions of colour-space corresponding to quasars at $z > 3$ are targetted to $i = 20.2$. Furthermore, if an unresolved, $i \leq 19.1$ SDSS object is matched to within $2''$ of a source in the FIRST catalogue, (Faint Images of the Radio Sky at Twenty centimetres; Becker et al. 1995), it is included in the quasar selection.

A tiling algorithm then assigns these candidates to specific spectroscopic plates, in order to maximise target completeness (Blanton et al. 2003). Each 3° spectroscopic plate

holds 640 fibres and quasar candidates are allocated approximately 18 fibres deg^{-2} . Targetting priority is required since no two fibres can be placed closer than 55" (corresponding to $\sim 0.7 h^{-1}$ Mpc at $\langle z \rangle = 1.27$, the mean redshift of our sample). The primary quasar candidates were given targetting priority over the MAIN galaxy and LRG survey targets (Strauss et al. 2002; Eisenstein et al. 2001, respectively). However, excluding subtle effects due to gravitational lensing (Scranton et al. 2005; Mountrichas & Shanks 2007), the LSS ‘footprint’ of these foreground galaxies should not affect our LSS quasar measurements. Some targets including brown dwarf and hot subdwarf standard star candidates, were given priority ahead of the primary quasar candidates. However, since the surface density of these Galactic objects is very low ($\ll 1 \text{ deg}^{-2}$), this should not have any significant impact on our results.

2.2 Quasar Samples

For our analysis, we use the Data Release Five (DR5; Adelman-McCarthy et al. 2007) and select quasars from the latest version of the quasar catalogue (DR5Q; Schneider et al. 2007). This catalog consists of spectroscopically identified quasars that have luminosities larger than $M_i = -22.0$ (measured in the rest frame) and at least one emission line with FWHM larger than 1000 km s^{-1} . Every object in the DR5Q had its spectrum manually inspected. There are 77 429 confirmed quasars over the $5\,740 \text{ deg}^2$ spectroscopic DR5 footprint, and the 65 660 DR5Q quasars with redshifts $z \leq 2.2$ will be the parent sample we use in this investigation.

The $z \leq 2.2$ limit is set due to the fact that at this redshift, the “ultra-violet excess” (UVX) method of selecting quasars begins to fail due to the Ly α -forest suppressing flux as it moves through the SDSS u -band and quasars have colours similar to those of F-stars (Fan 1999). Thus, for $2.2 < z \leq 2.9$, the completeness of the survey is dramatically lowered as is discussed in depth by Richards et al. (2006). Therefore, although we will present results in the redshift range $2.2 < z \leq 2.9$, we do so under caution, and will not include these measurements in any statistical analysis. The number of quasars used in this study represents a factor of > 2 increase over the previous largest quasar survey, the 2QZ (Boyle et al. 2000; Croom et al. 2005). This increase in data allows us to improve constraints on theoretical models, as well as to divide our sample in luminosity and redshift bins while still retaining statistical power.

We construct two subsamples from DR5Q. The first is designated as the “PRIMARY” Sample, which will include those objects in the DR5Q which were targetted as primary quasar candidates (Richards et al. 2002), having satisfied one, or more, of the `TARGET_QSO`, `TARGET_HIZ` or `TARGET_FIRST` selections (see Stoughton et al. (2002), Section 4.8, for more details on these flags). The SDSS quasar survey was designed to be complete in the primary sample, and no attempt was made at completeness for the quasars selected by other means. In total there are 55 577 quasars in the DR5Q that had their target flags set to one (or more) of these primary flags, with 46 272 quasars satisfying our redshift cut-off (Table 1).

The SDSS quasar selection algorithm was in flux in the early part of the survey, and was only finalised after DR1.

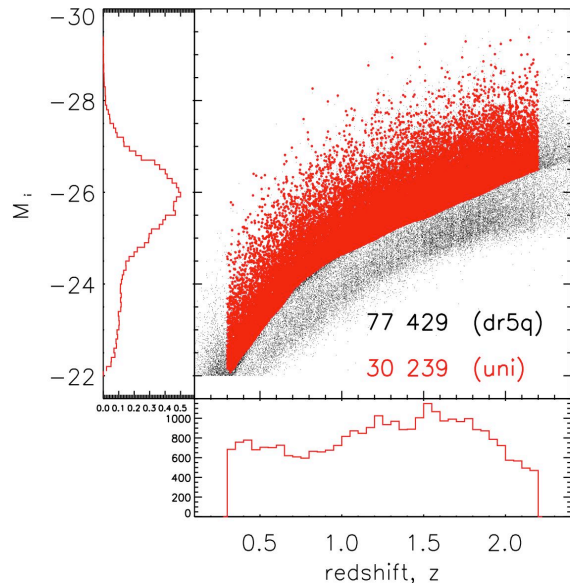


Figure 1. The SDSS DR5 Quasar $L-z$ plane for the PRIMARY sample. The affect of the apparent $i = 19.1$ magnitude limit can clearly be seen, and the objects that are fainter than this were those selected by the $z > 3$ algorithm.

We define the “UNIFORM” sample to be those primary objects selected with this final version. The UNIFORM sample contains 38 208 objects, dropping to 31 290 when a redshift cut of $z \leq 2.2$ is applied. We show the distribution of objects in the redshift-luminosity plane for the full DR5Q and $0.30 \leq z \leq 2.2$ UNIFORM sample in Figure 1. We will use both the PRIMARY and the UNIFORM samples in what follows and find some inconsistent results at scales $\gtrsim 60 h^{-1}$ Mpc. This is investigated further in Appendix C.

The quasar correlation function is sensitive to a number of potential systematic effects, including bad photometry and and improperly corrected dust reddening. Since quasars are selected by their optical colors, we shall perform checks on both our PRIMARY and UNIFORM samples to see what effect regions with poor photometry (as defined by Richards et al. 2006; Shen et al. 2007) has on our clustering measurements.

While all selection for the quasar sample is undertaken using dereddened colors (Richards et al 2001), if there remain systematic errors in the reddening model they can induce excess power into the clustering in a number of different ways. Appendix C describes how these effects affect our $\xi(s)$ measurements and the interpretations based thereon.

3 TECHNIQUES

In this section we describe the techniques we shall use to calculate the Quasar $z \leq 2.2$ 2PCF. The 2PCF has been described in detail by e.g. Peebles (1980, 1993); Peacock (1999); Coles & Lucchin (2002); Martínez & Saar (2002) and the interested reader is referred to those comprehensive texts for full details on the 2PCF.

Sample Description	Area /deg ²	Number in sample	z_{\min}	z_{\max}	z_{med}
DR5Q	≈ 5740	77 429	0.078	5.414	1.538
" $z \leq 2.9$		71 375	0.078	2.900	1.372
" $0.3 \leq z \leq 2.9$		69 692	0.300	2.900	1.400
" $z \leq 2.2$		65 660	0.078	2.200	1.278
" $0.3 \leq z \leq 2.2$		63 977	0.300	2.200	1.306
PRIMARY	5713	55 577	0.080	5.414	1.543
" $z \leq 2.9$		50 062	0.080	2.900	1.326
" $0.3 \leq z \leq 2.9$		48 526	0.300	2.900	1.360
" $z \leq 2.2$		46 272	0.080	2.200	1.234
" $0.3 \leq z \leq 2.2$		44 736	0.300	2.200	1.268
UNIFORM	4013	38 208	0.084	5.338	1.575
" $z \leq 2.9$		33 699	0.084	2.900	1.319
" $0.3 \leq z \leq 2.9$		32 648	0.300	2.900	1.234
" $z \leq 2.2$		31 290	0.084	2.200	1.354
" $0.3 \leq z \leq 2.2$		30 239	0.300	2.200	1.269

Table 1. The SDSS Spectroscopic Quasar Samples used in our analysis, with minimum, maximum and median redshifts. The DR5Q is the catalogue presented in Schneider et al. (2007), while the PRIMARY and UNIFORM samples are described in Section 2. The results for the bold samples are given in Section 4.

3.1 Estimating the 2-Point Quasar Correlation Function

In practice, ξ is measured by comparing the actual quasar distribution to a catalogue of “random” points, which have the same selection function, angular mask and radial distribution as the data, but are spatially distributed in a “random” manner - i.e. are not clustered. The construction of this random sample shall be described in Section 3.3.

We use the estimator of Landy & Szalay (1993) to calculate ξ , as this has been found to be the most reliable estimator for 2PCF studies (Kerscher et al. 2000). We do, however, compare our results using the ‘standard’ estimator of Davis & Peebles (1983) and the estimator from Hamilton (1992). The three estimators are given by,

$$\xi_{LS}(s) = 1 + \left(\frac{N_{rd}}{N} \right)^2 \frac{DD(s)}{RR(s)} - 2 \left(\frac{N_{rd}}{N} \right) \frac{DR(s)}{RR(s)} \quad (4)$$

$$\equiv \frac{\langle DD \rangle - \langle 2DR \rangle + \langle RR \rangle}{\langle RR \rangle} \quad (5)$$

$$\xi_{Std}(s) = \left(\frac{N_{rd}}{N} \frac{DD(s)}{DR(s)} \right) - 1, \quad (6)$$

and

$$\xi_{Ham}(s) = \frac{DD(s) \cdot RR(s)}{DR(s)^2} - 1 \quad (7)$$

for the Landy-Szalay (LS), ‘Standard’ and Hamilton estimators respectively, and these three estimators differ mostly in their sensitivity to errors in the mean density. Here N and N_{rd} are the number of data and random points in the sample, $DD(s)$ is the number of data-data pairs with separation between s and $s + \delta s$ in the given catalogue, $DR(s)$ is the number of data-random pairs and $RR(s)$ the number of random-random pairs. The angled brackets denote the suitably normalised pair counts, since we employ at least twenty times more random points than data in order to reduce Pois-

son noise. We choose our bins to be logarithmically spaced, with widths of $\Delta \log(s / h^{-1} \text{ Mpc}) = 0.1$.

The measurement of a quasar redshift will not only have a (large) component due to the Hubble expansion, but also components due to the intrinsic peculiar velocities and redshift errors associated with the individual quasar. The peculiar velocities can be seen in the redshift-space correlation function, both at small and large scales (see Section 4). However, as noted in Schneider et al. (2007) and discussed in detail in Shen et al. (2007, Appendix A), quasar redshift determination can have uncertainties of $\sigma_v = 500 - 1450 \text{ km s}^{-1}$ and hence $\sigma_z = 0.003 - 0.01$, and these redshift-errors will dominate any precise determination of the peculiar velocity signal.

3.2 The 2-D 2PCF, $\xi(r_p, \pi)$, and the Projected Correlation Function, $w_p(r_p)$

The real-space correlation function, $\xi(r)$, is what would be measured in the absence of any redshift-space distortions and we can measure $\xi(r)$ by projecting out the effects of peculiar velocities and redshift errors along the line of sight.

One can resolve the redshift-space separation, s , between two quasars into two components, r_p and π , where r_p is the separation between two objects *perpendicular* to the line-of-sight and π is the separation *parallel* to the line-of-sight. Thus,

$$s^2 = r_p^2 + \pi^2 \quad (8)$$

(where $r_p \equiv \sigma$ is also found in the literature). The ‘2-D’ redshift-space correlation function, $\xi(r_p, \pi)$, can be calculated as before,

$$\xi_{LS}(r_p, \pi) = \frac{\langle DD(r_p, \pi) \rangle - \langle 2DR(r_p, \pi) \rangle + \langle RR(r_p, \pi) \rangle}{\langle RR(r_p, \pi) \rangle} \quad (9)$$

where the bin sizes are now chosen to be $\Delta \log(r_p / h^{-1} \text{ Mpc}) = \Delta \log(\pi / h^{-1} \text{ Mpc}) = 0.2$.

Redshift-space distortions affect only the radial component of $\xi(r_p, \pi)$; thus by integrating along the line-of-sight direction, π , we obtain the projected correlation function,

$$w_p(r_p) = 2 \int_0^\infty \xi(r_p, \pi) d\pi. \quad (10)$$

In practice we set the upper limit on the integral to be $\pi_{\max} = 10^{1.8} = 63.1 \ h^{-1} \text{ Mpc}$ and show that although varying this limit does cause some difference to the deduced $w_p(r_p)$, it does not cause significant changes to the 2PCF over the scales of interest for our studies (Appendix C8).

The integral in equation 10 can be rewritten in terms of $\xi(r)$ (Davis & Peebles 1983),

$$w_p(r_p) = 2 \int_0^{\pi_{\max}} \frac{r \xi(r)}{\sqrt{(r^2 - r_p^2)}} dr. \quad (11)$$

If we assume that $\xi(r)$ is a power law of the form, $\xi(r) = (r/r_0)^{-\gamma}$ (which, as we shall find later, is a fair assumption), then equation 11 can be integrated analytically, such that with $\pi_{\max} = \infty$,

$$w_p(r_p) = r_0^\gamma r_p^{1-\gamma} \left[\frac{\Gamma(\frac{1}{2}) \Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})} \right] \equiv r_0^\gamma r_p^{1-\gamma} A(\gamma), \quad (12)$$

where $\Gamma(x)$ is the Gamma function.

In linear theory and in the absence of small-scale velocities and redshift errors, the redshift-space and real-space correlation function can be related via

$$\xi(s) = \xi(r) \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \right), \quad (13)$$

where

$$\beta(z) = \frac{\Omega_m(z)^{0.55}}{b(z)} \quad (14)$$

parametrizes the ‘flattening’ at large scales of the correlation function due to the infall of matter from underdense to overdense regions. The value of $\beta(z)$ has traditionally been measured via fits to observed data (e.g. Kaiser 1987; Fisher et al. 1994; Peacock et al. 2001; Hawkins et al. 2003; Ross et al. 2007; Guzzo et al. 2008).

3.3 Construction of the Random Catalogue

As mentioned above, to calculate ξ in practice, one needs to construct a random catalogue of points that mimics the data in every way, bar its clustering signal. Here we describe the construction of such a “random” catalogue for our samples. The angular mask and completeness for the PRIMARY and UNIFORM sample is described in detail in Appendix A.

The radial distribution of the sample is measured from the data themselves. Figure 2 shows the $N(z)$ distribution of the DR5Q quasars from our samples. We fit a tenth-order polynomial to both the PRIMARY and UNIFORM samples, which we use to generate the random sample redshift distribution. This method has proved reliable in previous quasar clustering studies (e.g. Croom et al. 2005; da Ângela et al. 2008).

Due to the design of the SDSS fibres and plates, no two spectroscopic fibres can be separated by less than $55''$. Thus, if two spectroscopic objects are closer than this separation, they will not both be fibred in one plate observation. We use

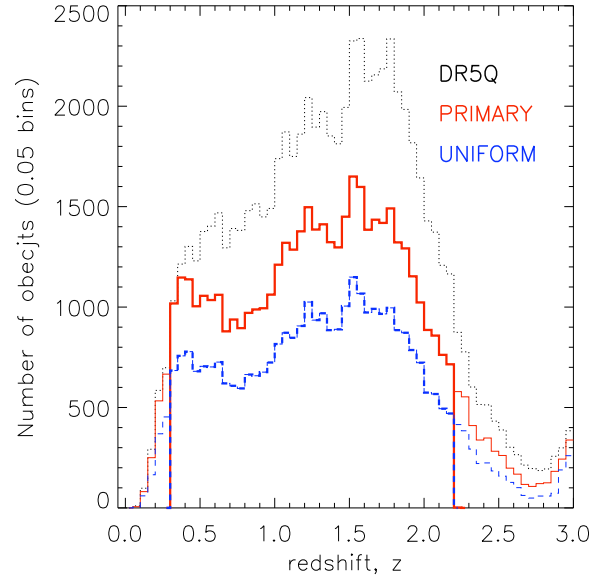


Figure 2. The SDSS DR5 Quasar $N(z)$. The solid (red) histogram shows the quasar redshift distribution for the PRIMARY sample, while the dashed (blue) histogram shows the redshift distribution for the UNIFORM sample. The thin lines for both PRIMARY and UNIFORM do not include the $0.3 \leq z \leq 2.2$ cuts. As a comparison, the full DR5Q sample is given by the dotted (black) histogram.

the PhotoObjAll catalogue (POA, Appendix A) to investigate which potential quasars would not have been observed due to fibre collisions and report our findings in Appendix C.

3.4 Errors and Covariances

Recent studies (e.g. Scranton et al. 2002; Zehavi et al. 2002; Myers et al. 2006; Ross et al. 2007) have employed three main methods, *Poisson*, *Field-to-Field* and *Jackknife* to estimate errors in correlation function measurements. The ‘simplest’ of these is the Poisson error described by Peebles (1973), and is the Poisson noise due to the number of pairs in the sample,

$$\sigma_{\text{Poi}} = \frac{1 + \xi(s)}{\sqrt{\text{DD}(s)}}. \quad (15)$$

This should be valid at smaller scales where the number of pairs is small and most pairs are independent (i.e. few quasars are involved in more than one pair Shanks & Boyle 1994; Croom & Shanks 1996). However, as reported in Myers et al. (e.g. 2005) and Ross et al. (2007), the Poisson error under-estimates measurement error when compared to e.g. the Field-to-Field or Jackknife errors at larger scales, where quasar pairs are not independent. For this work, we will not report any Field-to-field errors, but instead concentrate on a jackknife resampling procedure in order to calculate the full covariance matrix. Full details of the jackknife procedure, including the geometry of the subsamples used, are given in Appendix B.

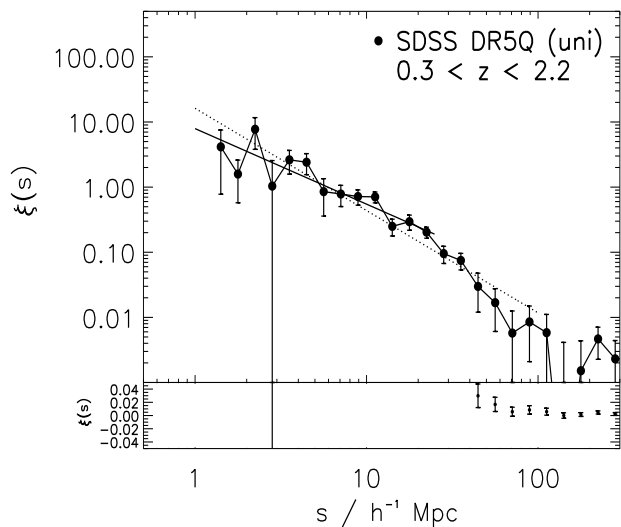


Figure 3. The SDSS Quasar redshift-space 2PCF, $\xi(s)$, from the UNIFORM sample (filled circles). The solid line shows the best fit single power-law model over $1 \leq s \leq 25.0 \ h^{-1} \text{ Mpc}$, while the dotted line shows the best fit single power-law model over $1 \leq s \leq 100.0 \ h^{-1} \text{ Mpc}$. The lower panel shows the $\xi(s)$ behaviour near zero on a linear scale. The quoted errorbars are jackknife errors from the diagonalised Covariance matrix.

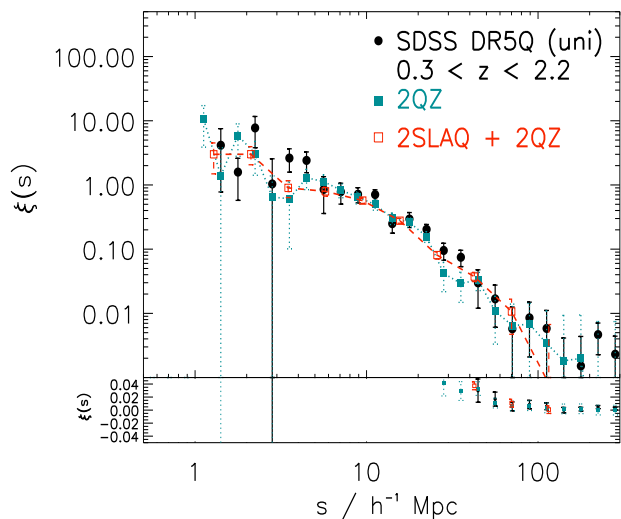


Figure 4. The Quasar redshift-space 2PCF, $\xi(s)$, from the UNIFORM sample as in Fig. 3. Also shown are the redshift-space correlation functions from the 2QZ (cerulean) filled squares, dotted line, (Croom et al. 2005) and the 2SLAQ QSO (red) open squares, dashed line (da Ângela et al. 2008) surveys.

4 RESULTS

4.1 SDSS Quasar Redshift-Space Two-Point Correlation Function, $\xi(s)$ ($0.30 \leq z \leq 2.2$)

The two-point redshift-space correlation function for the SDSS DR5Q UNIFORM sample over the redshift interval $0.3 < z < 2.2$ is given in Figure 3. The errorbars are jackknife errors from the diagonal elements of the Covariance matrix, i.e. $\sigma_i^2 = C_{ii}$.

We start by fitting a simple, single power-law model

of the form in Equation 2. We find that a single power law with a redshift-space correlation length of $s_0 = 5.95 \pm 0.45 \ h^{-1} \text{ Mpc}$ and power-law slope of $\gamma_s = 1.16^{+0.11}_{-0.08}$ provides a good description of the data over the scales $1.0 \leq s \leq 25.0 \ h^{-1} \text{ Mpc}$ (solid line, Fig. 3). Here a value of $\chi^2 = 11.5$ is obtained with 11 degrees of freedoms (dof). A less suitable fit is found at larger scales. Over the range $1.0 \leq s \leq 100.0 \ h^{-1} \text{ Mpc}$, the best fit model has a similar correlation length, $s_0 = 5.90 \pm 0.30 \ h^{-1} \text{ Mpc}$ but a steeper power-law slope, $\gamma_s = 1.57^{+0.04}_{-0.05}$ (dotted line, Fig. 3). The χ^2 for this model is 32.8 with 15 dof. Note for the power-law model fit over $1.0 \leq s \leq 100.0 \ h^{-1} \text{ Mpc}$, the data systematically deviate from the model, possibly due to the effects of redshift-distortions, with a “flattening” of the data compared to the model at both small, $s \lesssim 5 \ h^{-1} \text{ Mpc}$, and large, $s \gtrsim 40 \ h^{-1} \text{ Mpc}$, scales.

In Figure 4, we show the redshift-space correlation function $\xi(s)$ from two other recent surveys, the 2QZ (Croom et al. 2005) and the 2SLAQ QSO (da Ângela et al. 2008) surveys. The analysis by da Ângela et al. (2008) uses data from both the 2QZ and 2SLAQ QSO surveys and thus the samples are not completely independent.

The 2QZ and 2SLAQ QSO surveys both cover very similar redshift ranges to our $z < 2.2$ SDSS DR5Q UNIFORM sample. The 2QZ covers a much smaller area, $\approx 750\text{deg}^2$, than the SDSS but has 2/3 as many quasars as our sample, since it reaches to a deeper limiting magnitude of $b_J = 20.85$ (corresponding to $g \approx 20.80$ and $i \approx 20.42$). The 2SLAQ QSO survey has a smaller area again, $\approx 180\text{deg}^2$ and goes a magnitude deeper than the 2QZ to $g = 21.85$ ($i \approx 21.45$) resulting in 8 500 quasars with $0.3 < z < 2.2$.

The agreement in the correlation function between surveys over $1 \ h^{-1} \text{ Mpc} \leq s \leq 100 \ h^{-1} \text{ Mpc}$ scales is impressive but not necessarily unexpected, since we are essentially sampling the same type of objects i.e. luminous AGN, powered by supermassive black holes accreting at or near their Eddington limits, quite possibly in similar mass environments (see Section 5). However, the fact that the samples have quite different luminosities suggests that variation in quasar luminosity is due to a variation in SMBH fueling, rather than a variation in SMBH mass. We shall return to this subject in Section 5.

In Figure 5 we show the very large scale $\xi(s)$ using the LS estimator (we present $\xi(s)$ using the Standard and Hamilton estimators at these very large scales in Appendix C). We see that apart from one data point at $s \approx 400 \ h^{-1} \text{ Mpc}$, the redshift-space correlation function is consistent with $\xi(s) = 0$ at scales greater than $\sim 300 \ h^{-1} \text{ Mpc}$, to within one sigma. A χ^2 test comparing the data to $\xi(s) = 0$ over the range of $100 \leq s < 1000 \ h^{-1} \text{ Mpc}$ and $100 \leq s < 3000 \ h^{-1} \text{ Mpc}$ gives $\chi^2 = 8.2$ (18 dof) and $\chi^2 = 25.3$ (54 dof) respectively. Our rms scatter is ± 0.001 , which compares well to the 2QZ value of ± 0.002 ; with a sample $\sim 50\%$ larger, we have roughly doubled the pair counts at these very large scales.

Even with our increased area coverage over the 2QZ, the dimensions of our sample does not allow us to probe separations beyond $3000 \ h^{-1} \text{ Mpc}$.

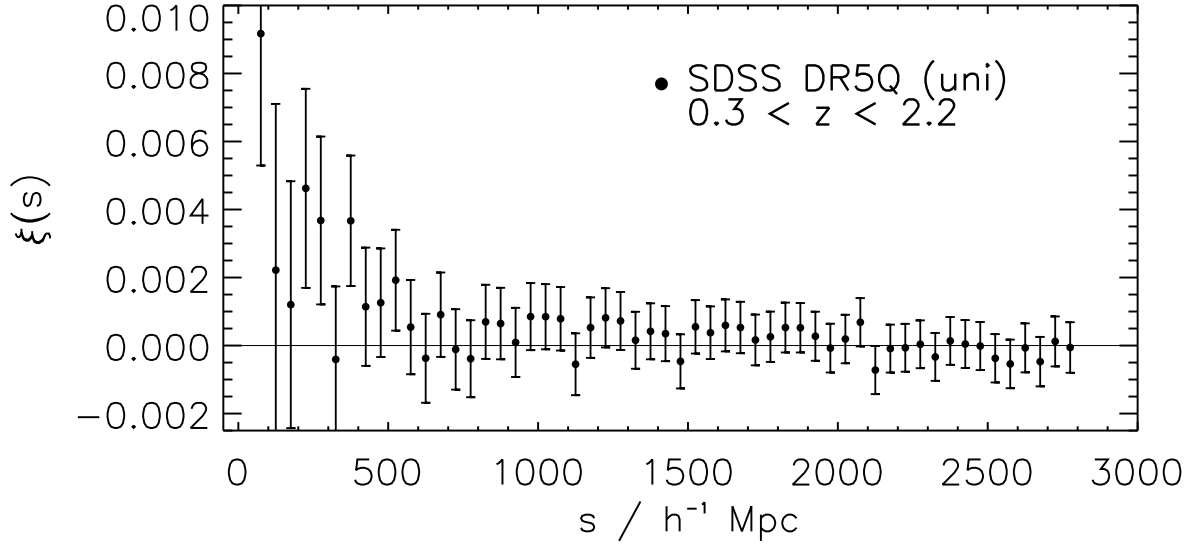


Figure 5. The SDSS Quasar redshift-space 2PCF, $\xi(s)$, for our UNIFORM sample over the redshift range $0.3 \leq z \leq 2.2$ at very large scales. Jackknife errors are plotted and the data are consistent with $\xi(s) = 0$ out to scales of $s \sim 3000 \ h^{-1} \text{ Mpc}$. Above these scales, there is a rapid drop-off in Quasar-Quasar ‘DD’ pairs.

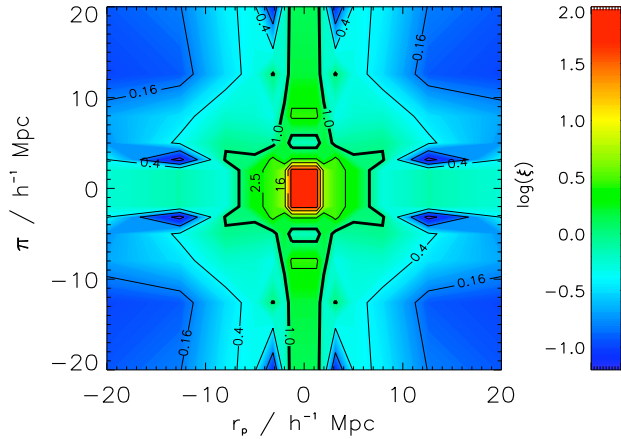


Figure 6. The SDSS DR5 Quasar $\xi(r_p, \pi)$. The contours give lines of constant ξ , and $\delta \log \xi = 0.4$ between contours, with $\log \xi = 1.6$ the highest value at the centre of the plot. The thick contour is $\xi = 1.0$. This is the first time $\xi(r_p, \pi)$ has been reported for the SDSS Quasar sample.

4.2 SDSS Quasar 2-D 2-Point Correlation Function, $\xi(r_p, \pi)$ ($0.30 \leq z \leq 2.2$)

Figure 6 shows the SDSS DR5 Quasar 2-D redshift-space correlation function $\xi(r_p, \pi)$ for the UNIFORM sample, over $0.3 \leq z \leq 2.2$. The redshift-space distortions in the clustering signal - seen as deviations from isotropy - are immediately apparent. At small r_p , the random peculiar motions and redshift errors of quasars cause an elongation of the clustering signal along the line-of-sight direction, π . This is the well-known ‘Fingers-of-God’ effect Jackson (1972). Cosmological information can be extracted from the Quasar 2D $\xi(r_p, \pi)$ measurement Hoyle et al. (e.g. 2002); da Ángela et al. (e.g. 2005, 2008). However, full treatment of the separation of the effects of large-scale ‘squashing’ in r_p (used

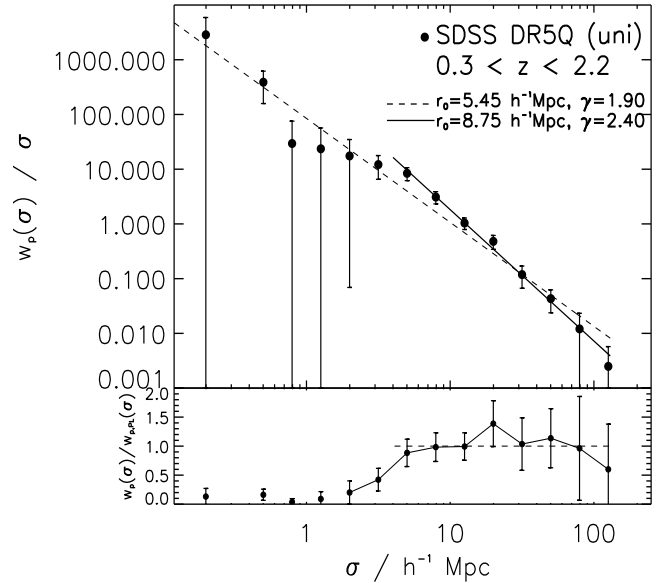


Figure 7. The SDSS Quasar redshift-space 2PCF, $w_p(r_p)$. The broken line shows the best fit single power-law to the data over our full range of scales, $0.1 < r_p < 130.0 \ h^{-1} \text{ Mpc}$. Here, the real-space correlation length is $r_0 = 5.45^{+0.35}_{-0.45}$ with a slope, $\gamma = 1.90^{+0.04}_{-0.03}$. Restricting the range to $4.0 < r_p < 130.0 \ h^{-1} \text{ Mpc}$, the best-fit values now become $r_0 = 8.75^{+0.35}_{-0.50}$ and $\gamma = 2.40^{+0.07}_{-0.10}$. In the lower panel, we show the ratio of the data divided by our power-law model over $4.0 < r_p < 130.0 \ h^{-1} \text{ Mpc}$.

to determine $\beta(z)$ in equation 14) and the substantial contribution from the Fingers-of-God at small-scales is left to a future paper.

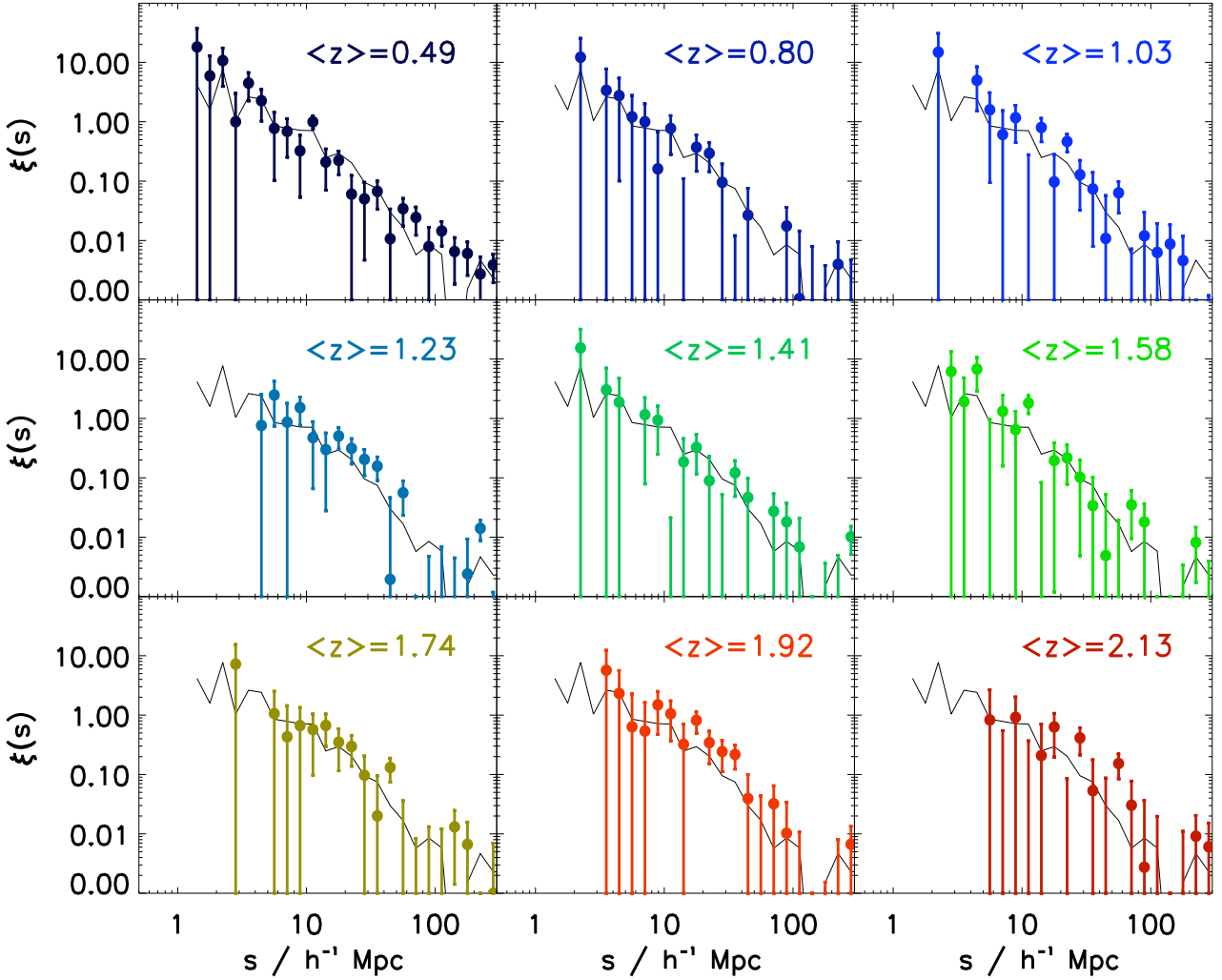


Figure 8. The SDSS DR5 Quasar redshift-space 2PCF, $\xi(s)$, and its evolution with redshift. The thin (black) line in each panel is $\xi(s)$ for the full DR5Q UNIFORM sample, over $0.30 < z < 2.20$. The quoted errorbars are Poisson (see text for justification).

4.3 SDSS Quasar Projected 2-Point Correlation Function

In Figure 7, we show the projected 2-point correlation function, $w_p(r_p)$, calculated using equation 11. The reported error bars are jackknife errors, using the same jackknife area subsamples as for the $\xi(s)$ calculation (Appendix B). Since we want to fit power-laws of the form, $\xi(r) = (r/r_0)^\gamma$ (equation 12), we plot $w_p(r_p)/r_p$ on the ordinate. We find the best fitting single power-law to the SDSS Quasar $w_p(r_p)/r_p$ data to be $r_0 = 5.45^{+0.35}_{-0.45} h^{-1} \text{ Mpc}$ and $\gamma = 1.90^{+0.04}_{-0.03}$ over our full range of scales, $0.1 < r_p < 130.0 h^{-1} \text{ Mpc}$. This provides a somewhat adequate fit giving a value of $\chi^2 = 22.02$ with 12 degrees of freedom. We remind the reader that due to fibre collisions, measurements at scales of $r_p \lesssim 1 h^{-1} \text{ Mpc}$ are biased low. Restricting the range to $4.0 < r_p < 130.0 h^{-1} \text{ Mpc}$, we find the best fit power-law has an increased real-space correlation length of $r_0 = 8.75^{+0.35}_{-0.50} h^{-1} \text{ Mpc}$ and a steeper slope of $\gamma = 2.40^{+0.07}_{-0.10}$. This power-law is a more acceptable fit, having $\chi^2 = 3.47$ with 6 dof. In the bottom panel we divide the data by this

steeper, higher amplitude power-law and see negligible deviation over the scales we fit. Comparisons of our $w_p(r_p)/r_p$ results to those of Shen et al. (2007) for the high, $z > 2.9$ redshift quasar measurements, shows the high redshift SDSS quasars having a much larger clustering amplitude than the lower redshift sample. The consequences of this are discussed in detail in Shen et al. (2007).

4.4 Evolution of the SDSS Quasar Correlation Function

Figures 8 and 9 show the evolution of the redshift-space, $\xi(s)$, and the projected, $w_p(r_p)$, 2PCF, using the SDSS DR5 UNIFORM Quasar sample.

We plot both $\xi(s)$ and $w_p(r_p)$ for sub-samples of the UNIFORM data, with the relevant redshift limits given in Table 2. Here we choose the redshift slices so that we match those of the 2QZ Survey given by Croom et al. (2005). Our survey generally has 50% more data in each redshift bin. However, since the 2QZ selects QSO candidates on the basis of their stellar appearance on photographic plates,

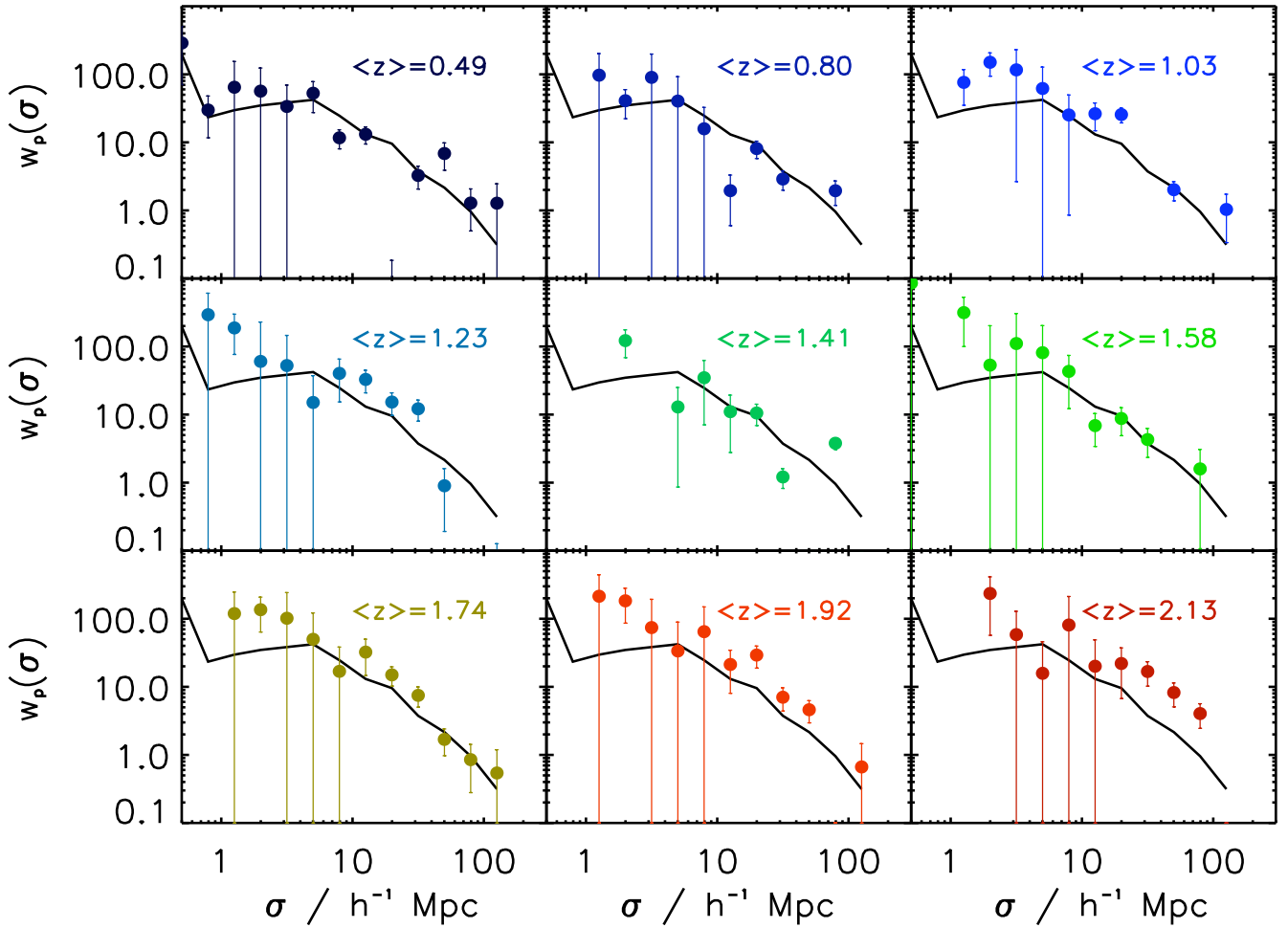


Figure 9. The SDSS DR5 Quasar projected 2PCF, $w_p(r_p)$, and its evolution with redshift. The thin (black) line in each panel is $w_p(r_p)$ for the full DR5Q UNIFORM sample, over $0.30 < z < 2.2$. The quoted errorbars are scaled jackknives.

low-redshift quasars with detectable host galaxies on the plate are preferentially deselected from the final 2QZ catalogue, and the SDSS Quasar UNIFORM sample has a larger proportion of low, $z \lesssim 0.5$, redshift quasars. We fit power-law models of the form given by equation 2, over the ranges $1.0 \leq s \leq 25.0 \ h^{-1} \text{ Mpc}$ (except for our $2.02 \leq z < 2.20$ bin, where to get finite constraints, we fit to $s_{\text{max}} = 100 \ h^{-1} \text{ Mpc}$). The best fit parameters and corresponding 1σ errors are given in Table 2.

In Fig. 8, we show measurements for $\xi(s)$ for 9 of 11 redshift slices. We exclude the lowest and highest redshift bins due to issues with low number statistics and completeness. The measurement of $\xi(s)$ for the full redshift range measurement is given by the thin line in each panel. We quote Poisson errors as these are approximately equal to jackknife errors on scales where the number of DD pairs is less than the number of quasars in the (sub-)sample. This scale is $s \sim 50 - 100 \ h^{-1} \text{ Mpc}$ for the sub-samples given here. The $\xi(s)$ data show a trend to ‘lose’ quasar-quasar DD pairs at the smallest separation, as the redshift increases. Keep in mind that the length scale suppressed due to the $55''$ fi-

bre collision limitation increases from $s \sim 0.2 \ h^{-1} \text{ Mpc}$ at $z = 0.5$ to $s \sim 1 \ h^{-1} \text{ Mpc}$ at $z = 2$, giving rise to the apparent depression in the correlation function on small-scales.

Fig. 9 has the same form as Fig. 8. However, here we quote jackknife errors (scaled using the $w_p(r_p)/r_p$ measurements from the full sample) as we find Poisson errors vastly underestimate the errorbars for projected correlation function measurements.

Figure 10 shows the evolution of the redshift-space correlation length, s_0 , with both redshift and the age of the Universe (adopting the cosmology given at the end of Section 1). Since there is a covariance between the best-fit s_0 and γ_s , here we fix γ_s to the best-fit value of the whole sample ($\gamma_s = 1.16$) and then measure the best-fit s_0 . We find the clustering strength remains reasonably constant with redshift out to $z \sim 3$, which is equivalent to approximately 80% of the history of the Universe. This trend was also seen in Fig. 8. The correlation length is measured to be $s_0 = 5 - 7 \ h^{-1} \text{ Mpc}$ for bright, optically identified quasars in the SDSS, up to $z \sim 3$.

z-interval	\bar{z}	N_{gals}	Bolo. Lum ($10^{46} \text{ ergs}^{-1}$)	$s_0 / h^{-1} \text{ Mpc}$	γ_s	χ^2	ν	$s_0 / h^{-1} \text{ Mpc}$ ($\gamma_s = 1.16$)	$r_0 / h^{-1} \text{ Mpc}$ ($\gamma = 2.0$)
0.30,2.20	1.269	30 239	3.73	5.95 ± 0.45 $^a 5.90 \pm 0.30$	$1.16^{+0.11}_{-0.16}$ $1.57^{+0.04}_{-0.05}$	11.5 32.8	11 17	5.95 ± 0.45	$5.45^{+0.35}_{-0.45}, \gamma = 1.90^{+0.04}_{-0.03}$ $8.75^{+0.35}_{-0.50}, \gamma = 2.40^{+0.07}_{-0.10}$
0.08,0.30	0.235	1 051	0.17	$6.90^{+1.35}_{-1.50}$	$1.37^{+0.41}_{-0.31}$	2.6	9	6.20 ± 1.55	8.80 ± 0.85
0.30,0.68	0.488	5 404	0.49	$6.05^{+0.45}_{-0.65}$	$1.67^{+0.23}_{-0.24}$	12.2	11	$4.60^{+0.80}_{-0.75}$	6.44 ± 0.57
0.68,0.92	0.801	3 001	1.45	$7.05^{+1.15}_{-1.45}$	$1.90^{+0.60}_{-0.60}$	6.2	8	$5.40^{+1.60}_{-1.70}$	6.40 ± 0.64
0.92,1.13	1.029	3 365	2.18	$2.68^{+1.42}_{-1.28}$	$0.57^{+0.14}_{-0.15}$	9.4	7	$6.30^{+1.60}_{-1.65}$	8.20 ± 0.84
1.13,1.32	1.228	3 623	2.94	$7.10^{+1.45}_{-1.65}$	$1.00^{+0.30}_{-0.25}$	2.5	6	$7.75^{+1.50}_{-1.60}$	7.31 ± 0.82
1.32,1.50	1.412	3 332	3.81	$6.05^{+1.35}_{-1.85}$	$2.13^{+0.87}_{-0.78}$	6.1	7	$3.65^{+1.70}_{-1.80}$	6.78 ± 0.93
1.50,1.66	1.577	3 405	4.69	$6.10^{+1.40}_{-1.60}$	$1.67^{+0.81}_{-0.50}$	13.5	8	$4.65^{+1.55}_{-1.70}$	7.71 ± 0.82
1.66,1.83	1.744	3 240	5.81	$7.70^{+1.70}_{-1.90}$	$1.11^{+0.39}_{-0.31}$	1.0	6	$7.90^{+1.80}_{-1.85}$	7.34 ± 0.71
1.83,2.02	1.917	2 970	7.37	$7.43^{+2.37}_{-2.43}$	$0.84^{+0.41}_{-0.30}$	2.5	7	$8.70^{+2.05}_{-2.15}$	9.38 ± 0.79
2.02,2.20	2.104	1 899	9.67	$^a 3.65^{+1.60}_{-1.85}$	$1.10^{+0.29}_{-0.15}$	8.7	10	$4.10^{+1.75}_{-1.90}$	10.51 ± 0.96
2.20,2.90	2.462	2 409	13.49	$10.75^{+2.15}_{-3.42}$	$2.60^{+0.6}_{-1.1}$	0.2	4	$7.15^{+5.50}_{-6.45}$	13.51 ± 1.80

Table 2. Evolution of the redshift-space, s_0 and real-space, r_0 correlation lengths. For s_0 , both the correlation length and power-law slope were allowed to vary. All samples were fitted over the range $1.0 \leq s \leq 25.0 \text{ } h^{-1} \text{ Mpc}$, unless otherwise noted with a , where the range was $1.0 \leq s \leq 100.0 \text{ } h^{-1} \text{ Mpc}$. For the redshift slice subsamples, the calculation of r_0 was made by fitting a single power-law to $w_p(r_p)$, over the scales $1.0 \leq r_p \leq 130.0 \text{ } h^{-1} \text{ Mpc}$, while keeping the power-law index fixed at $\gamma = 2.0$.

5 EVOLUTION OF GALAXY, AGN AND QUASAR CLUSTERING

5.1 The Redshift-Space Evolution

The study of quasar clustering most comparable to our own is that presented by Croom et al. (2005) for the 2QZ survey. Thus, we start by comparing our results with theirs (Fig. 10). We immediately notice the results from our study using the SDSS DR5Q UNIFORM quasar sample and the 2QZ are in very good agreement over the full redshift range, given the associated uncertainties, with minimal evolution upto $z \sim 3$ being seen by both surveys. The similarity of these results again suggest that quasar clustering is luminosity-independent, a topic that we discuss further in Shen et al. (2008, in preparation).

The filled (red) star in Fig. 10 is from the study by Wake et al. (2004) who use a sample of 13 605 narrow-line AGNs in the redshift range $0.055 < z < 0.2$ from the first Data Release of the SDSS. They find that the AGN autocorrelation function is consistent with the observed galaxy autocorrelation function over $s = 0.2 - 100 \text{ } h^{-1} \text{ Mpc}$ scales. Furthermore, they show that the AGN 2PCF is dependent on the luminosity of the narrow [O III] emission line ($L_{[\text{OIII}]}$), with low $L_{[\text{OIII}]}$ AGNs having a higher clustering amplitude than high $L_{[\text{OIII}]}$ AGNs. This measurement suggests that lower activity AGNs reside in more massive DM haloes than higher activity AGNs do, as $L_{[\text{OIII}]}$ provides a good indicator of AGN fueling rate (e.g. Miller et al. 2003; Kauffmann et al. 2003). As such, it is very interesting to note that our lowest redshift quasar clustering data point is, within the uncertainties, consistent with the measurement from Wake et al. (2004). We use the term ‘quasar’ here loosely, as for our lowest redshift bin, the mean bolometric luminosity is $1.7 \times 10^{45} \text{ ergs s}^{-1}$, a factor of 20 lower than our full sample (Table 2).

Although not plotted in Fig. 10, Constantin & Vogele (2006) study the clustering of specific classes of AGN, namely Seyfert galaxies and LINERs (low-ionization nuclear

emission-line regions) with the classes being separated on the basis of emission-line diagnostic diagrams (e.g. Baldwin et al. 1981; Kewley et al. 2001). They find that LINERs, which exhibit high $s_0 = 7.82 \pm 0.64$ clustering show the lowest luminosities and obscuration levels, suggesting that these objects harbor relatively massive black holes that are weakly active or inefficient in their accretion, potentially due to the insufficiency of their fuel supply. Seyfert galaxies however, have lower clustering, $s_0 = 5.67 \pm 0.62$, are very luminous and show large emitting gas densities, suggesting that their black holes are less massive but accrete quickly and efficiently enough to dominate the ionization. Therefore, based on our lowest redshift clustering results, the stronger link for our low-luminosity ‘quasars’ is to Seyfert galaxies rather than LINERs. We shall leave further exploration of this connection as future work.

5.2 The Real-Space Evolution

In Figure 11, we compare our measurements of the real-space correlation length, r_0 , to those recently published in the literature. We calculate our values for r_0 by fitting our $w_p(r_p)$ measurements using equation 12. Motivated by the fits in Fig. 7, we hold the power-law index fixed at $\gamma = 2.0$.

Myers et al. (2006) were the first to make a measurement of the clustering of quasars using SDSS data. These authors use $\sim 80\,000$ quasars photometrically classified from the catalogue of Richards et al. (2004). The r_0 measurements from Myers et al. (2006) are given by the open (red) squares in Fig. 11, and are in very good agreement with our own data (we plot the data from their Table 1, from the ‘Deprojected r_0 ’ section and the $0.75 \leq r < 89 \text{ } h^{-1} \text{ Mpc}$ row).

Coil et al. (2007) calculate the cross-correlation between $\sim 30\,000$ redshift $0.7 < z < 1.4$ galaxies observed as part of the DEEP2 galaxy redshift survey (Davis et al. 2001, 2003), and quasars over the same redshift range. In total there are 36 SDSS quasars and 16 quasars identified from the DEEP2

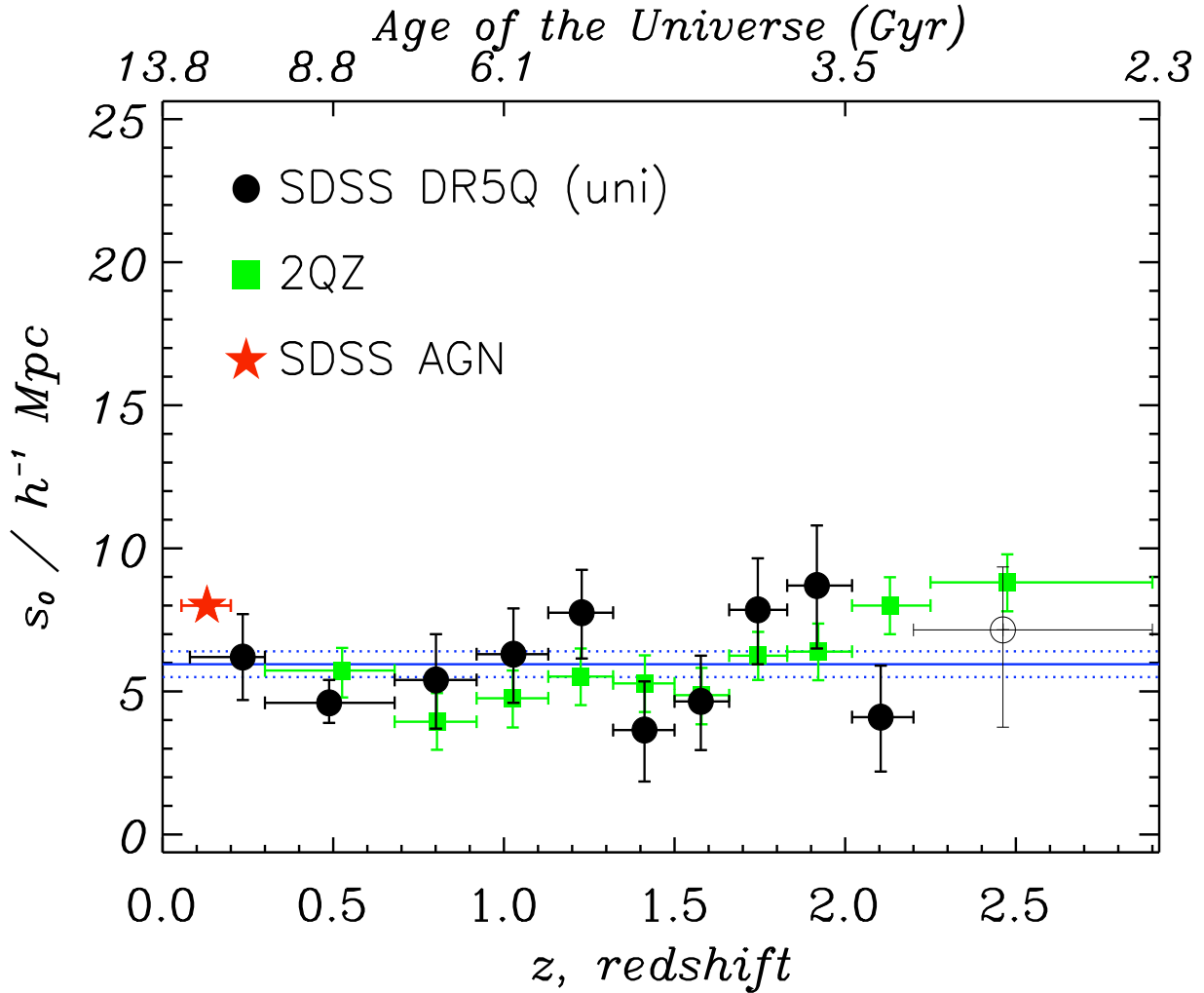


Figure 10. Evolution of the redshift-space correlation length, s_0 , up to redshift $z = 2.9$. The (black) filled circles are from the DR5Q UNIFORM sample and the (blue) line gives the best-fit value for the whole sample with associated 1σ errors. The (green) filled squares are from the 2QZ (Croom et al. 2005), while the (red) filled star is from a measurement of AGN clustering at low, $z < 0.2$, redshift by Wake et al. (2004).

survey itself over the 3 deg^2 covered by the DEEP2. Coil et al. (2007) find that $r_0 \sim 3.4 \pm 0.7 \text{ h}^{-1} \text{ Mpc}$ for the quasar-galaxy cross-correlation (ξ_{QG}). These authors then find $r_0 \sim 3.1 \pm 0.6 \text{ h}^{-1} \text{ Mpc}$ for the inferred quasar clustering scale length, assuming γ does not change between galaxy and the quasar samples and also a linear bias, such that $\xi_{QG} = \sqrt{(\xi_{QQ} \cdot \xi_{GG})}$. We show this as an open (purple) diamond in Fig. 11. Although still consistent with the low-redshift measurement of Myers et al. (2006), it is at odds with our measurements. One possible explanation is that the model for ξ_{QG} to be the geometric mean of ξ_{GG} and ξ_{QQ} assumes that δ_G and δ_Q are perfectly correlated spatially, i.e. the correlation coefficient is $r = \pm 1$ (e.g. Blanton et al. 1999; Swanson et al. 2008)¹. Thus, as is quite plausi-

ble, if different luminous tracers (in this case $z \sim 1$ quasars and galaxies) sample the underlying mass density field differently, then one can reconcile the difference in correlation lengths by invoking a correlation coefficient that is modestly different from unity. This is a key point that needs to be kept in mind for all cross-correlation studies.

Further cross-correlation comparisons are from Adelberger & Steidel (2005a,b), who study the clustering of Lyman Break galaxies (LBGs) around $2 \lesssim z \lesssim 3$ AGN. The dynamic range in luminosity for this sample is nearly 10 magnitudes ($16 \lesssim G_{AB} \lesssim 26$, Adelberger & Steidel 2005b) and is thus much greater than for our SDSS DR5 UNIFORM sample. These authors report a value of $r_0 = 5.27^{+1.59}_{-1.36} \text{ h}^{-1} \text{ Mpc}$ for a sample of 38 AGN with central SMBH masses of $10^{5.8} < M_{BH}/M_\odot < 10^8$ and $r_0 = 5.20^{+1.85}_{-1.16} \text{ h}^{-1} \text{ Mpc}$ for a sample of 41 AGN with $10^8 < M_{BH}/M_\odot < 10^{10.5}$. If we do assume the correlation coefficient is $r = 1$ and the power-law slopes are constant between samples, we find (with $r_{0,LBG-LBG} = 4.0 \pm 0.6$ at $z = 2.9$, Adelberger et al.

¹ The simplest and frequently assumed relationship between δ_1 and δ_2 is “deterministic linear bias”, $\delta_1 = b_{lin}\delta_2$ where b_{lin} is a constant parameter, $\delta_1 = \rho_1(\mathbf{x})/\bar{\rho}_1 - 1$ and $\delta_2 = \rho_2(\mathbf{x})/\bar{\rho}_2 - 1$, e.g. Peebles (1980); Swanson et al. (2008).

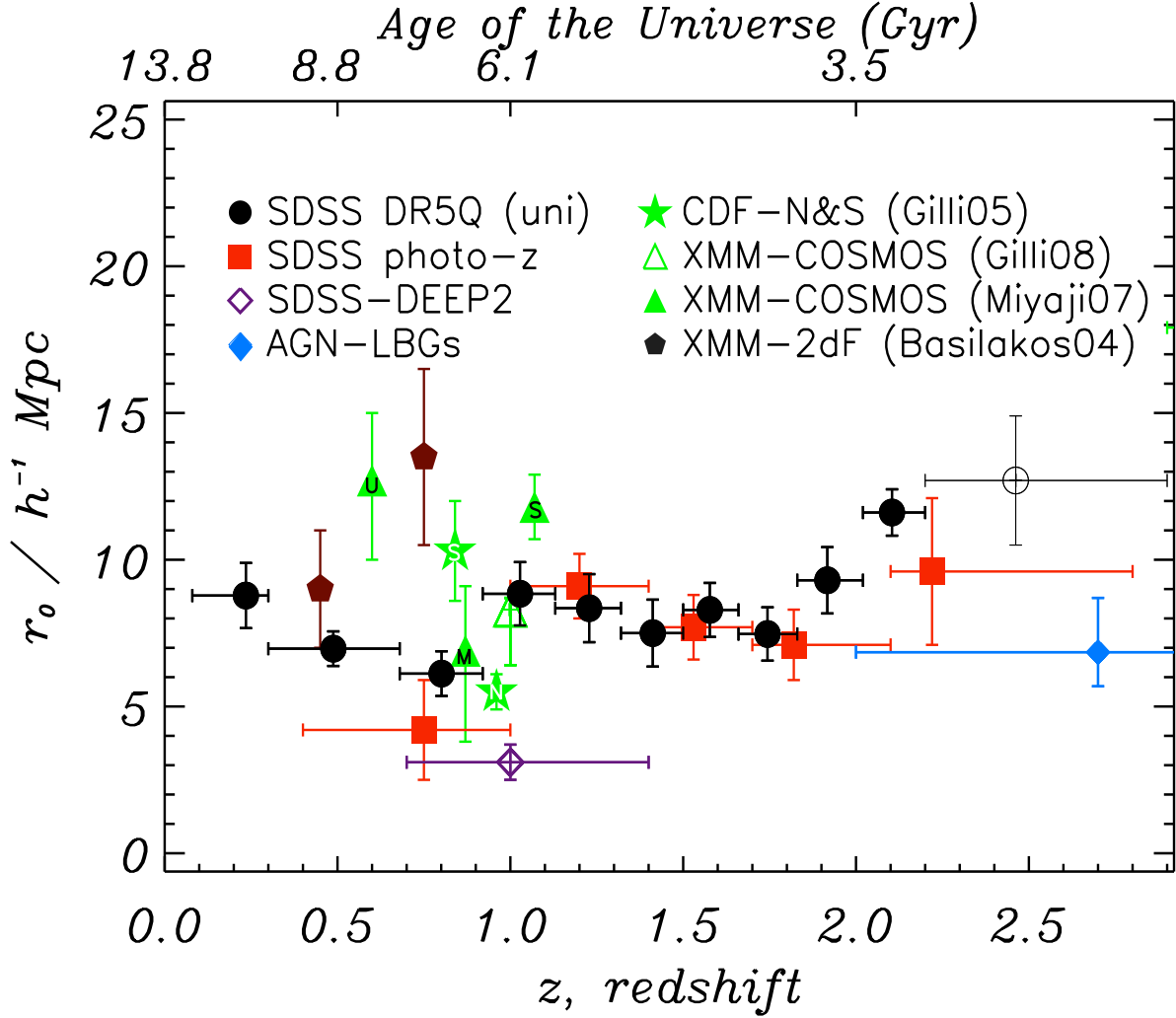


Figure 11. Evolution of the real-space correlation length, r_0 , up to redshift $z = 2.9$. The filled (black) circles are from the DR5Q UNIFORM sample (this work); the open (red) squares are from the photometric sample of SDSS quasars from Myers et al. (2006); The open (purple) diamond is the quasar-galaxy cross-correlation from the DEEP2 Survey (Coil et al. 2007) and the solid (blue) diamond is the AGN-galaxy cross-correlation using Lyman Break galaxies (LBGs, Adelberger & Steidel 2005b). X-ray data from the *Chandra* Deep Fields (Gilli et al. 2005) are shown by filled (green) 5-pointed stars (with the North and South fields denoted ‘N’ and ‘S’ respectively); the *XMM-Newton*-2dF survey (Basilakos et al. 2004) are indicated by filled (dark red) pentagons; the *XMM-Newton* COSMOS survey (Miyaji et al. 2007), is shown by filled (green) triangles, with the ‘SFT’ (0.5-2 keV), ‘MED’ (2-4.5 keV) and ‘UHD’ (4.5-10 keV) band measurements denoted as S, M and U respectively. Very recent measurements from Gilli et al. (2008, in prep.), also using the *XMM-Newton* COSMOS survey are given by the open (green) triangle. Our measurement of the X-ray bright SDSS quasars using the *Rosat* All Sky Survey (Anderson et al. 2007) is given by the filled (light blue) diamond.

2005) that $r_{0,\text{AGN-AGN}} \approx 6.9 \ h^{-1} \text{ Mpc}$. This result is very broadly consistent with Myers et al. (2006) but in tension with our SDSS DR5 UNIFORM results. Realising that Adelberger & Steidel (2005b) samples very different luminosity ranges and find the clustering does not vary significantly, immediately rules out a luminosity dependence explaining the different clustering amplitudes. Again, the assumption of perfect correlation is called into question, with a non-unity correlation coefficient r reconciling potentially both these and the Coil et al. (2007) DEEP2 results. Further investigations into these discrepancies and comparisons to other recent high redshift galaxy survey clustering measurements are left to a future paper (Ross et al. 2008, in prep.).

A very interesting comparison can also be made to clustering measurements of recent deep X-ray surveys, which are particularly well suited to finding intrinsically less luminous, potentially obscured objects at high redshift (Brandt & Hasinger 2005). An immediate caveat we place in the following comparison is that the SDSS DR5Q surveys $\sim 4000 \text{ deg}^2$, while the largest solid angle of the current deep X-ray surveys is of order 1 deg^2 .

Basilakos et al. (2004) estimate r_0 using the angular autocorrelation function, $w(\theta)$, of hard, (2-8 keV) X-ray selected sources detected in a $\approx 2 \text{ deg}^2$ field using a shallow ($f_X[2-8 \text{ keV}] \approx 10^{-14} \text{ ergs cm}^{-2} \text{ s}^{-1}$) and contiguous *XMM-Newton* survey. The area surveyed consisted of 13 us-

able pointings, overlapping that of the 2QZ survey and resulted in the detection of 171 sources. Various models for the redshift distribution are given in Basilakos et al. (2004, see their Table 1) and for our comparison, we take the r_0 values calculated using $(\Omega_m, \Omega_\Lambda) = (0.3, 0.7)$, and which either assume “pure luminosity evolution” (PLE, Boyle et al. 1998) or “luminosity-dependent density evolution” (LDDE, Ueda et al. 2003). As such, the PLE and LDDE models give different mean redshifts of $\bar{z} = 0.45$ and $\bar{z} = 0.75$ respectively for the AGN sample and Basilakos et al. (2004) find $r_0 = 9.0 \pm 2.0 \ h^{-1}$ Mpc for the PLE model and $r_0 = 13.5 \pm 3 \ h^{-1}$ Mpc for the LDDE model, fixing the power-law slope at $\gamma = 2.2$. These observations are given by filled (dark red) pentagons in Fig. 11.

Gilli et al. (2005) obtained a sample of nearly 260 AGN in the *Chandra* Deep Field North (CDF-N, Alexander et al. 2003; Barger et al. 2003) and South (CDF-S, Rosati et al. 2002) with spectroscopic redshifts. They report that in both fields the AGN have $\bar{z} \sim 0.9$ and a median 0.5–10 keV luminosity of $\bar{L}_X \sim 10^{43}$ erg s $^{-1}$, i.e. in the local Seyfert galaxy luminosity regime. Correlation lengths and slopes of $r_0 = 5.5 \pm 0.6 \ h^{-1}$ Mpc, $\gamma = 1.50 \pm 0.12$ and $r_0 = 10.3 \pm 1.7 \ h^{-1}$ Mpc, $\gamma = 5.5 \pm 0.6$ are found for the CDF-N and CDF-S respectively (Gilli et al. 2005, their Table 2), shown as filled (green) stars in Fig. 11.

Miyaji et al. (2007) measured the angular autocorrelation function of X-ray point sources detected by *XMM-Newton* in the $\sim 2\text{deg}^2$ COSMOS field (Scoville et al. 2007). The measurements for the 0.5–2 (SFT), 2–4.5 (MED) and 4.5–10 (UHD) keV bands are given by filled (green) triangles in Fig. 11. Very recently, Gilli et al. (2008, in prep.) also report on the spatial clustering of AGN in the COSMOS field using ~ 550 spectroscopically identified AGN at a median redshift of $z \sim 1$. They find a value of $r_0 = 8.3 \pm 0.4 \ h^{-1}$ Mpc (Fig. 11, open green triangle) and a power-law slope of $\gamma = 1.89 \pm 0.07$. However, this result is affected by a coherent structure of 40 AGN at $z \sim 0.36$. Removing this structure causes r_0 to drop to $\sim 6 \ h^{-1}$ Mpc, similar to that of the previous deep X-ray AGN measurements.

We find that our clustering measurements are in good agreement with the lower correlation lengths found by some of the deep X-ray surveys, e.g. Gilli et al. (2005) for the CDF-N, Miyaji et al. (2007) for their MED (2–4.5 keV) band and XMM-COSMOS (Gilli et al. 2008 in prep.). However, there is still much scatter in the deep X-ray data, potentially due to cosmic variance and the small samples used for these analyses.

Plionis et al. (2008) however, suggest that the scatter in the X-ray clustering results can be explained if one takes into account the observed flux limit, leading to a X-ray luminosity dependence for clustering of AGN. Motivated by this explanation, we use the matched *Rosat* All Sky Survey (RASS) and SDSS catalogues from Anderson et al. (2003, 2007) to examine the X-ray properties of our quasars. In particular we take the > 6 200 objects found in the Anderson et al. (2007) catalogue and match them to the 30 239 quasars in our $0.3 \leq z \leq 2.2$ UNIFORM sample, finding 2 717 matches. The mean bolometric luminosity of the X-ray matched quasars is $L_{\text{bol}} = 2.33 \times 10^{46}$ ergs s $^{-1}$ which is very similar to that of the overall UNIFORM quasar sample. This sample has a mean redshift of $\bar{z} = 0.72$ and we can see that optically identified SDSS quasars at this redshift have

redshift	$\bar{\xi}_Q(s, z)$	$\bar{\xi}_\rho(r, z)$	b
1.27	0.391 ± 0.011	0.069	2.06 ± 0.03
0.24	0.462 ± 0.104	0.176	1.41 ± 0.18
0.49	0.363 ± 0.028	0.138	1.38 ± 0.06
0.80	0.311 ± 0.133	0.104	1.45 ± 0.38
1.03	0.383 ± 0.118	0.085	1.83 ± 0.33
1.23	0.524 ± 0.095	0.072	2.37 ± 0.25
1.41	0.309 ± 0.134	0.062	1.92 ± 0.50
1.58	0.411 ± 0.119	0.054	2.42 ± 0.40
1.74	0.472 ± 0.141	0.049	2.79 ± 0.47
1.92	0.674 ± 0.166	0.043	3.62 ± 0.49
2.10	0.425 ± 0.442	0.039	2.99 ± 1.42

Table 3. The evolution of the linear bias for the SDSS Quasar UNIFORM sample.

a real-space correlation length of $r_0 = 6 - 7 \ h^{-1}$ Mpc. However, these SDSS DR5Q UNIFORM-RASS objects are generally very X-ray bright and have a calculated mean 0.1–2.4 keV flux of $f_X = 7.9 \times 10^{-13}$ ergs cm $^{-2}$ s $^{-1}$ and a mean X-ray luminosity of $L_X = 9.6 \times 10^{44}$ ergs s $^{-1}$. Thus, extrapolation of the Plionis et al. (2008, their Fig. 4) result would predict a correlation length for these objects of $r_0 > 30 \ h^{-1}$ Mpc, which seems unrealistically high. We leave the actual clustering measurement of the RASS-SDSS DR5Q matched sources for future investigation, since describing the complex RASS-SDSS selection function is outside the scope of this present paper.

5.3 Evolution of Bias

One key reason for measuring the correlation function as a function of redshift, $\xi(s, z)$, is to determine the linear bias, b , defined by the model of equation 1. We shall assume that b is independent of scale on the scales and redshift range under investigation here.² We follow the method in Croom et al. (2005) and da Ângela et al. (2008) to determine b using our redshift-space correlation function $\xi(s, z)$ measurements from Section 4.

The volume-averaged correlation function, $\bar{\xi}$ is defined as

$$\bar{\xi} = \frac{\int_{s_{\min}}^{s_{\max}} 4\pi s'^2 \xi(s') ds'}{\int_{s_{\min}}^{s_{\max}} 4\pi s'^2 ds'} \quad (16)$$

$$= \frac{3}{s_{\max}^3} \int_{s_{\min}}^{s_{\max}} \xi(s') s'^2 ds'. \quad (17)$$

where $s_{\min} = 1.0 \ h^{-1}$ Mpc is set in practice. Unless explicitly stated otherwise, s_{\max} is always chosen to be $20 \ h^{-1}$ Mpc, so that non-linear effects in the sample should be small due to the s^2 weighting and for ease of comparison with Croom et al. (2005) and da Ângela et al. (2008). In the

² The precise way in which galaxies/luminous AGN trace the underlying matter distribution is still poorly understood. Blanton et al. (2006), Schulz & White (2006), Smith et al. (2007) and Coles & Erdogdu (2007) all suggest that bias is potentially scale dependent. We do not take this into account in the current analysis.

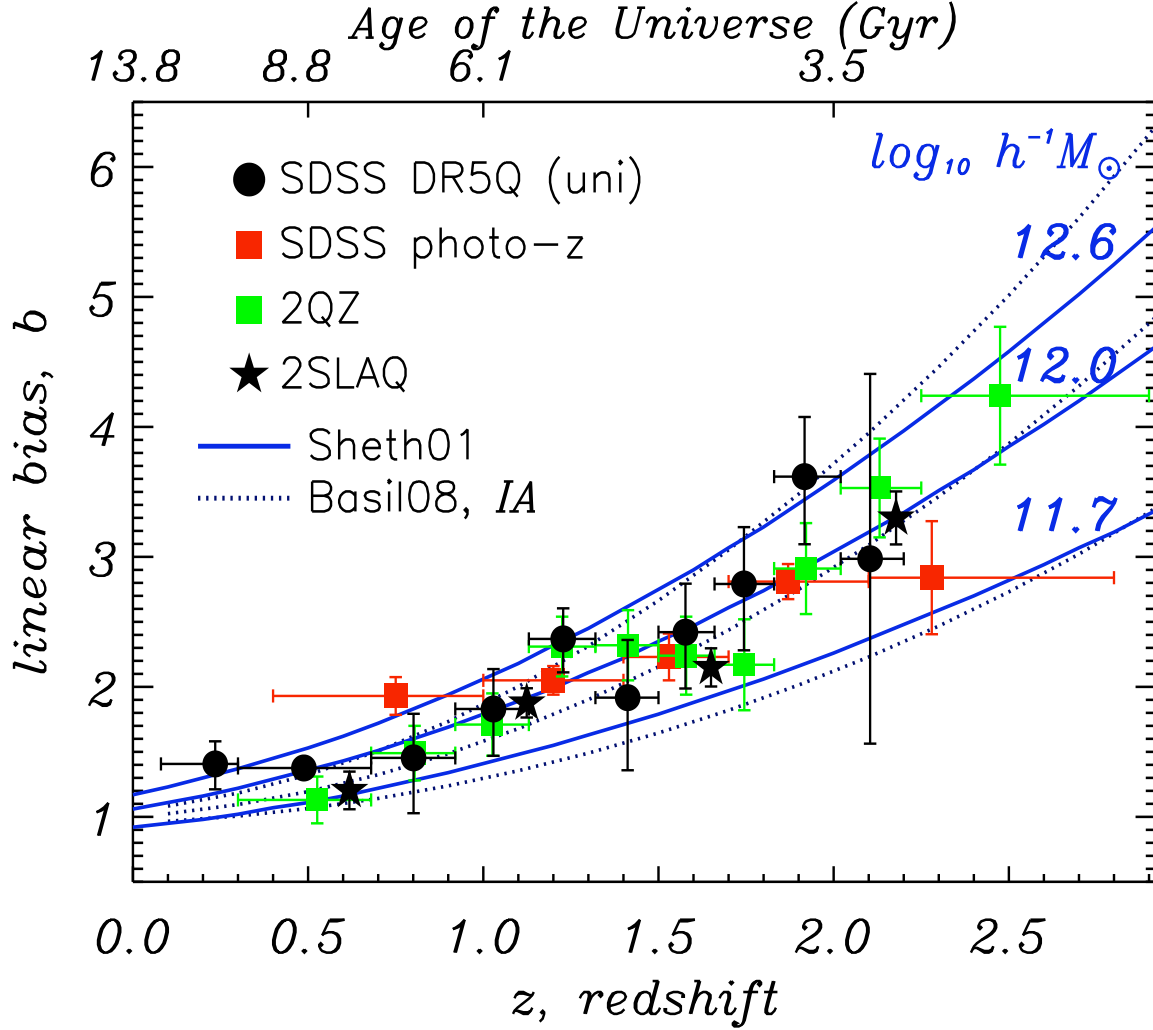


Figure 12. Evolution of the linear bias of quasars, b_Q , with redshift, upto $z = 3$. The filled (black) circles, are from the SDSS DR5Q UNIFORM sample (this work); the open (red) squares, from the photometric SDSS quasar measurements (Myers et al. 2006); the filled (green) squares from the 2QZ survey (Croom et al. 2005); the open (black) stars are from the 2SLAQ QSO survey (da Ângela et al. 2008); The lines give dark halo masses from the models of Basilakos et al. (2008).

linear regime, the z -space and real-space correlation functions can be given by equation 13. Thus, combining equations 1 and 13, and recognising that $\beta = \Omega_m^{0.55}/b$ leaves us with a quadratic equation in b . We are assuming a flat, cosmological-constant model and hence the exponent of Ω_m is 0.55 (Linder 2005; Guzzo et al. 2008), rather than the value of 0.6 (though we find this makes virtually no difference to our bias measurements). Solving the quadratic in b leads to

$$b(z) = \sqrt{\frac{\bar{\xi}_Q(s, z)}{\bar{\xi}_\rho(r, z)} - \frac{4\Omega_m^{1.1}(z)}{45} - \frac{\Omega_m^{0.55}(z)}{3}}. \quad (18)$$

We now use our measured $\bar{\xi}_Q(s, z)$ together with a theoretical estimate of $\bar{\xi}_\rho(r, z)$ and $\Omega_m(z)$ to determine the bias.

To estimate $\bar{\xi}_\rho(r, z)$, we follow Myers et al. (2007) and da Ângela et al. (2008), and use the non-linear estimate of $P(k)$ given by Smith et al. (2003). The models of Smith et al. (2003) predict the non-linear power spectrum of dark matter for a range of CDM cosmologies over a wide range

of scale and we thus Fourier transform these $P(k)$ models, and integrate over $s = 1 - 20 \ h^{-1}$ Mpc to compute $\bar{\xi}_\rho(r, z)$. The cosmological parameters used in our chosen model are $\Omega_m(z=0) = 0.3$, $\Omega_\Lambda(z=0) = 0.7$, $\Gamma = 0.17$ and $\sigma_8 = 0.84$. We find the simple form,

$$\bar{\xi}_\rho(r, z) = A \exp(Bz) + C \quad (19)$$

where $A = 0.20413$, $B = -1.0823$, and $C = 0.0178$ models the evolution of $\bar{\xi}_\rho(r, z)$ extremely well, for $1 \ h^{-1}$ Mpc $\leq s \leq 20 \ h^{-1}$ Mpc.

With $\Omega_m(z = 1.27) = 0.81$, our value for $b_Q(z)$ from the full SDSS DR5Q UNIFORM sample is $b_Q(z = 1.27) = 2.06 \pm 0.03$ and the values for our redshift sub-samples are shown in Fig. 12 with the filled (black) circles and given in Table ???. We estimate our errors by using the variations in $\bar{\xi}(s)$ from our 21 jackknife estimates, scaled using the number of DD pairs in each redshift slice subsample. Previous measurements from the 2QZ Survey (filled green circles, Croom et al. 2005), the 2SLAQ QSO Survey (open black

stars, da Ângela et al. 2008) and photometrically selected SDSS quasars (open red squares, Myers et al. 2007) are again in excellent agreement to our data. We compare these bias estimates with various models in Section 5.4.

From equation 14, we can estimate the parameter $\beta(z)$, which takes into account the large-scale effects of the linear redshift-space distortions. With $b(z = 1.27) = 2.06 \pm 0.03$ the linear bias for the DR5Q UNIFORM, thus $\beta(z = 1.27) = 0.43 \pm 0.01$. This is consistent with the value measured from the 2QZ survey, $\beta(z = 1.4) = 0.50^{+0.13}_{-0.15}$ (Outram et al. 2001; da Ângela et al. 2005). Due to the much smaller space density of quasars than that of galaxies, the errors on the (e.g. $\xi(r_p, \pi)$) clustering measurement are much larger than for galaxy surveys (cf. Hawkins et al. 2003; Zehavi et al. 2005; Ross et al. 2007; Guzzo et al. 2008). Furthermore, as discussed in Section 4.2, we have not included the effects from the “Fingers-of-God” in the present calculation of $\beta(z)$ and it is known that the peculiar velocities at small (transverse r_p) scales will very strongly affect the value of β (Fisher et al. 1994; da Ângela et al. 2005). Thus we caution that although this is the highest S/N quasar measurement of β to date, our formal errorbar is most likely very under-representative of the true error.

5.4 Models of bias and dark matter halo mass estimation

We compare our bias measurements with those of recent models.

Basilakos et al. (2008) use linear perturbation theory in the matter-dominated epoch and reasonably assume that the mass-tracer number density (i.e. quasars in our case) is not conserved with time. These authors introduce an interaction term, Ψ , where Ψ is proportional to $(1+z)^\nu$. There are three general cases: $\nu > 2$ implies merging of dark matter haloes; $\nu < 2$ implies the destruction of dark matter haloes of a particular mass and $\nu = 2$ is the no-interaction scenario (where merging and destruction cancel out). N -body simulations are used to calibrate ν and the best-fit value is found to be $\nu = 2.56$ over the halo mass range $M_{\text{DMH}} = 7.7 \times 10^{11} - 6.0 \times 10^{13} h^{-1} M_\odot$.

We plot these models in Fig. 12 as dotted lines for 3 assumed halo masses of (top to bottom) $M_{\text{DMH}} = 1.0 \times 10^{13} h^{-1} M_\odot$, $7.7 \times 10^{12} h^{-1} M_\odot$ and $1.0 \times 10^{12} h^{-1} M_\odot$. Comparing these models to our data, the mass at which a ‘typical SDSS’ quasar resides remains *constant* (given associated errors) with redshift, at a value of $M_{\text{DMH}} = 5 \times 10^{12} - 1 \times 10^{13} h^{-1} M_\odot$ from redshift $z \sim 2.5$ to the present day, i.e. over 80% the assumed age of the Universe. Therefore, as dark matter halo masses generally grow with time, the ratio of the halo mass for a typical quasar to the mean halo mass at the same epoch drops as you approach redshift $z = 0$.

Sheth et al. (2001) also provide fitting functions for the halo bias from and we show bias predictions from these authors in Fig. 12 (solid lines). Here the 3 assumed halo masses of (top to bottom) $M_{\text{DMH}} = 4.0 \times 10^{12} h^{-1} M_\odot$, $1.0 \times 10^{12} h^{-1} M_\odot$ and $5.0 \times 10^{11} h^{-1} M_\odot$. Comparing our results to the Sheth et al. (2001) models, we find that a ‘typical SDSS quasar’ inhabits a halo $M_{\text{DMH}} = 5 \times 10^{11} - 4 \times 10^{12} h^{-1} M_\odot$ with this mass not significantly changing from $z \sim 2.5$ to the present day. This dark matter halo mass is appreciably

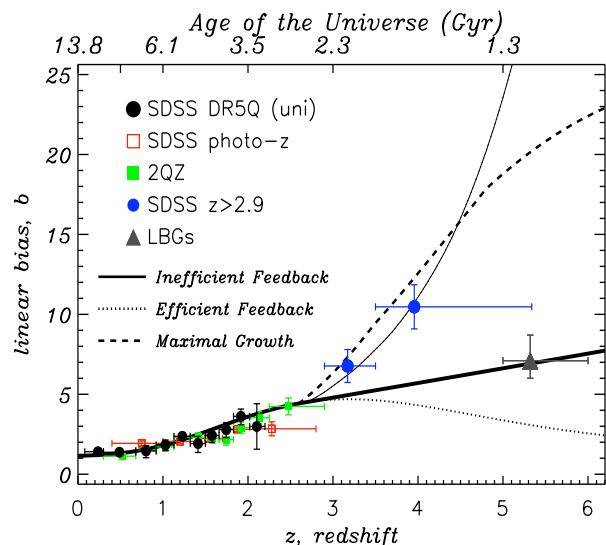


Figure 13. Evolution of the linear bias of quasars, b_Q , with redshift to $z = 6$. Filled (black) circles, this work; Filled (green) squares, Croom et al. (2005); Open (red) squares, Myers et al. (2006); Open (blue) circles, Shen et al. (2007); Solid lines show the “Uniform growth” models for a magnitude limited survey of $i = 20.2$ (thin solid line) and an effectively infinitely deep survey (thick solid line). Dashed lines are for the “Efficient feedback” survey (though note at $i = 20.2$ these are identical to the Uniform growth case. “Maximal growth” at $i = 20.2$ is given by the dashed line. The solid (grey) triangle gives the recent measurements of LBG bias from McLure et al. (2008).

smaller than that of the Basilakos et al. (2008) models. However, since the Sheth et al. (2001) prescription is more often quoted in the literature, to ease comparisons with previous work, we favour these lower values.

We next compare with the models of Hopkins et al. (2007, e.g. their Fig. 13). Here three models are described with all the models having the same $z < 2$ behaviour due to quasars (and their SMBH central engines) “shutting down” at $z \lesssim 2$.

The first of the Hopkins et al. (2007) models is the “Inefficient Feedback” model, given by the solid lines in Fig. 13. Here $z \sim 6$ quasars grow either continuously or episodically with their host systems until the epoch where “downsizing” begins (i.e. $z \sim 2$). Thus, at redshifts $z > 2$ any feedback from quasar activity is not enough to completely shutdown the quasar, hence the term “inefficient feedback”. In this model, the quasar luminosity function (QLF) at all redshifts $z > 2$ represents the *same* systems building up hierarchically.

The second of the Hopkins et al. (2007) models is the “Extreme Feedback” model, given by the dotted line in Fig. 13. Here each SMBH only goes through one episode of quasar activity, after which the quasar will completely shut down, even if this occurs at high ($z > 2$) redshifts. BH growth will cease after this one-off quasar phase. If objects cannot grow after their quasar epoch even at high redshifts, then the subsequent decline of the QLF L^* traces a decline in characteristic active masses, and the linear bias of active systems “turns over”.

The third of the Hopkins et al. (2007) models is the “Maximal Growth” model, given by the dashed line in

Fig. 13. In this model the BHs grow mass at their maximal rate. For example, a $\sim 10^8 M_{\text{dot}}$ BH at $z \sim 6$ will grow to $\sim 5 \times 10^9 M_{\text{dot}}$ at $z \sim 2$ at which point the growth ceases and the BH mass remains constant until $z = 0$.

Further details of the models can be found in Hopkins et al. (2007) but the limiting factor in our ability to discriminate between models is the luminosity and redshift ranges, hence the need for high redshift information. As such, we extend our redshift baseline up to $z = 6$ in Figure 13 and now also plot the bias estimates for the $z > 2.9$ SDSS quasar clustering measurements of Shen et al. (2007), given by the closed (blue) circles.

The three Hopkins et al. (2007) models (“Inefficient Feedback”, “Efficient Feedback” and “Maximal Growth”) are shown for two limiting magnitude ranges, $m_i < 20.2$ and $m_i < 30$, given by thin and thick lines in Fig. 13 respectively.

The brighter $m_i < 20.2$ magnitude limit is chosen to match the SDSS high-redshift quasar selection. However, at this limit the QLF break L^* is only marginally resolved at $z \sim 2 - 3$ (e.g. Richards et al. 2006) and so above this redshift surveys are systematically biased to more massive $L > L^*$ BHs with higher clustering and larger linear biases. Subsequently, the models have no discriminating power at $z > 2$ with the $m_i < 20.2$ limit and the predicted behaviour for the linear bias from the “Inefficient”, “Efficient” and “Maximal Growth” models is identical. As can be seen in Fig 13, all models match the observed clustering data well at all redshifts (note the $m_i < 20.2$ “Efficient” and “Maximal Growth” models are obscured by the “Inefficient” model tracks in Fig. 13).

In order to break this model degeneracy at $i = 20.2$, models with an effectively infinitely deep flux limit of $i = 30$ are plotted in Fig. 13. Although there are currently no quasar surveys with clustering measurements at $z > 2$ deeper than Shen et al. (2007) to discriminate the models, we plot recent results from a study of high redshift Lyman Break Galaxies (LBGs) from a deep survey by McLure et al. (2008), given by the solid (grey) triangle in Fig. 13. The “Inefficient” and “Efficient Feedback” models are shown as the thick solid and thick dotted lines respectively, and describe the bias evolution that could be observed from a survey deep enough to see L^* objects up to $z = 6$. Interestingly, the “Inefficient feedback” model is now heavily favoured, with this model going straight through the observational point. This result might suggest that LBGs could be the progenitors of $z < 2$ quasars, having grown hierarchically with multiple epochs of quasar activity at high redshifts. However, further investigations into the link between AGN/Quasar activity, the build-up of SMBH mass and the formation and evolution of galaxies using clustering measurements are left to a future paper.

6 CONCLUSIONS

We have used the SDSS Quasar Survey to calculate the two-point correlation function over $\sim 5000 \text{ deg}^2$ of the sky, covering a redshift range of $0 < z \leq 2.9$, thus representing a measurement over the largest volume of the Universe ever sampled. We find that:

- The 2-Point redshift-space correlation function is well described by a single power-law of the form $\xi = (s/s_0)^{-\gamma}$ where $s_0 = 5.95 \pm 0.45 \text{ h}^{-1} \text{ Mpc}$ and $\gamma_s = 1.16^{+0.11}_{-0.16}$ over $1 \leq s \leq 25 \text{ h}^{-1} \text{ Mpc}$.
- We have no evidence for deviation from $\xi(s) = 0$ at scales of $s > 100 \text{ h}^{-1} \text{ Mpc}$.
- There are strong redshift-space distortions present in the 2-D $\xi(r_p, \pi)$ measurement with “Fingers of God” being seen at small-scales.
- We find very weak or even no evolution of clustering amplitude of the SDSS quasar correlation function increasing with redshift to $z \sim 2.5$.
- Comparing our results to those of recent deep X-ray surveys, we find that our clustering measurements are in general agreement these surveys, e.g. Gilli et al. (2005), Miyaji et al. (2007) and XMM-COSMOS (Gilli et al. in prep.). However, there is still much scatter in the deep X-ray data, potentially due to cosmic variance and the small samples used for these analyses.
- Using SDSS-RASS matched catalogues for our quasar sample and extrapolating the X-ray luminosity:AGN clustering strength relation leads us to predict a correlation length of $r_0 > 30 \text{ h}^{-1} \text{ Mpc}$ for this SDSS X-ray sample, compared to $r_0 = 6 - 7 \text{ h}^{-1} \text{ Mpc}$ for the optical selection, at $z = 0.7$. We suggest this is an unfeasibly large correlation length but defer measurement of the RASS-SDSS sample for future work.
- The linear bias for SDSS quasars over the redshift range of $0.3 \leq z \leq 2.2$ is $b(z = 1.27) = 2.06 \pm 0.03$. Using this bias measurement but not taking into account velocity dispersion due to “Fingers of God” or redshift errors, we find $\beta(z = 1.27) = 0.43 \pm 0.01$. Both these values are consistent with measurements from previous surveys, i.e. the 2QZ.
- Using models which relate dark halo mass to clustering strength, we find that the dark halo mass at which a ‘typical SDSS quasar’ resides, remains roughly constant with redshift at $M = 5 \times 10^{11} - 4 \times 10^{13} M_{\odot}$. Therefore, as dark halo masses grow with time, the typical halo mass for a quasar becomes *relatively less massive*, as compared with other haloes at the same epoch, as you approach redshift $z = 0$.
- Using current observational data, we are unable to discriminate between the “Inefficient Feedback”, “Efficient Feedback” and “Maximal Growth” models proposed by Hopkins et al. (2007). However, using recent observations of high redshift Lyman Break Galaxies, we tentatively suggest that high, $z = 5 - 6$ LBGs are the progenitors of low, $z < 2$ redshift quasars and AGN.

In Shen et al. (2008, in prep.) we shall continue our investigations into the SDSS Quasar population and study the clustering properties of DR5 quasars as a function of luminosity, virial mass, colour and radio-loudness.

The SDSS is now complete and the final quasar catalogue, Data Release 7, is being prepared. This catalogue will contain over 100,000 quasars with spectroscopic observations and will appreciably ($\gtrsim \times 2$) the number of quasars used in our type of study. DR7 will not change the luminosity dynamic range of the SDSS quasar survey but with final analysis of data from e.g. the 2SLAQ QSO Survey (Croom et al., 2008 submitted), and extension of the deep X-ray surveys (e.g. Extended CDF-S, Lehmer et al. 2005) connec-

tions between the “luminous” and “average” AGN luminosity regimes should begin to converge.

Looking further afield, even with the dramatic increase in data that surveys such as the 2QZ and SDSS have provided, the desire to increase dynamic range continues. For instance, due to the steepness of the faint end of the quasar luminosity function, low luminosity quasars should be relatively plentiful, as long as one can identify these objects. This will be a strong challenge for the next generation of

quasar redshift surveys (e.g. the Baryon Oscillation Spectroscopic Survey (BOSS), Schlegel et al. 2007) but one that will lead to another significant increase in our understanding of quasars, supermassive black holes, galaxy formation and evolution and the properties of the Universe.

ACKNOWLEDGMENTS

This work was partially supported by National Science Foundation grants AST-0607634 (N.P.R. and D.P.S.) and AST-0707266 (Y.S. and M.A.S.). We warmly thank S.M. Croom for providing the 2QZ data points, S. Basilakos, A. Lidz and P. Hopkins for providing their model data shown in Section 5 and R. Gilli for allowing us to report the most recent XMM-Newton COSMOS results prior to publication. P. Allen, W.N. Brandt, A.D. Myers and R. Nemmen provided very useful discussion. The JavaScript Cosmology Calculator was used whilst preparing this paper (Wright 2006).

Funding for the creation and distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

APPENDIX A: SDSS TECHNICAL DETAILS

A1 The Catalog Archive Server

The SDSS database can be interrogated through the Catalog Archive Server³ (CAS) using standard Structured Query Language (SQL) queries. When querying the CAS, one has a choice to query either the **best** or **target** database for a given Data Release (in our case, DR5). The former database contains information on all the photometric and spectroscopic objects obtained using the latest versions (and i.e.

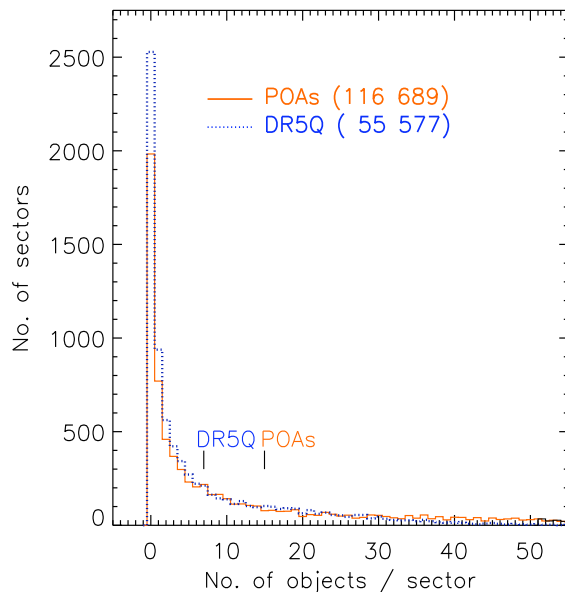


Figure A1. Distribution of the number of objects per sector. The 116 689 POA objects have a mean of 14.9 objects per sector, while 55 577 DR5Qs have a mean of 6.97 objects per sector (shown by the two horizontal lines respectively).

the “best”) of the data reduction and analysis pipelines (Section 3, Abazajian et al. 2004). The **target** database however, contains the information on objects available at the time that the targeting algorithm pipelines were run, which, in general, pre-dates the current **best** data by a significant amount. An object’s measured properties, such as magnitude or colour, can be subtly different between target allocation and the most recent data processing, and some objects change their target selection status between the two. More details regarding the CAS, **best** and **target** are given in the relevant SDSS Data Release papers (Stoughton et al. 2002; Abazajian et al. 2004; Adelman-McCarthy et al. 2007).

Thus, in order to create a statistical data sample, or indeed to mimic it for a comparative ‘random’ sample, we need to know the properties of our chosen objects *at the time of targetting*, i.e. which objects were selected as quasar candidates. Thus here, we only use information from **target**.

A2 SDSS Survey Geometry

The first SQL query we run simply asks the CAS to return all the objects in the Photometric database that were targetted as being “primary” candidate quasars. When run on DR5, this returns 203 185 unique objects, from the PhotoObjAll table.

We next calculate which primary ‘PhotoObjAll’ objects (POAs) fall within the spectroscopic survey plate boundaries. We use the parameter file, `maindr5spectro.par` found on the SDSS website⁴, which contains the plate number, Modified Julian Date (MJD), and plate centre (in J2000 Right Ascension and Declination). We do not use any of the

³ <http://cas.sdss.org>

⁴ <http://www.sdss.org/dr5/coverage/index.html>

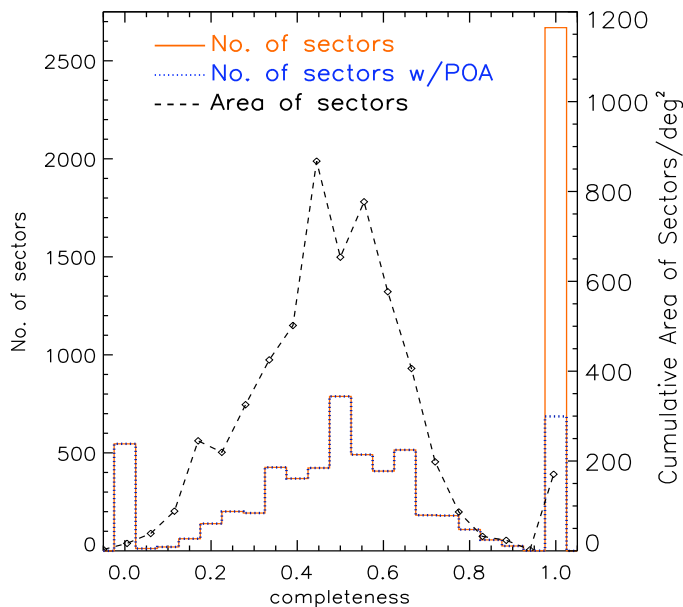


Figure A2. Histogram showing the completeness of the DR5 Quasar survey by sector. The solid (orange) line shows completeness distribution for all 7814 sectors, while the dotted (blue) line shows the completeness distribution for the 5831 sectors which have one or more POAs in them. The summed area of sectors with given completenesses is shown by the dashed line.

“Extra”, “Special”, or “ExtraSpecial” plates for our analysis as these plates were not targetted with the normal quasar algorithm, or are duplicates. There are 145 524 POA objects that fall within 1.49 degrees of a given DR5 plate centre, noting that since plates overlap due to the tiling scheme, an object can be in more than one plate.

Of these 145,524 objects, we would next like to know how many were (a) designated as spectroscopic (“tilable”) targets by the process of ‘Tiling’ and (b) allocated fibres. A tile is a 1.49 degree radius circle on the sky which contains the locations of up to 592 tilable targets and other science targets (the other 48 fibres are assigned to calibration targets and blank sky). For each tile, one, or more, physical aluminum plates will be created. The plates will have holes drilled in them for fibres to be plugged, in order to observe the tiled targets. Thus, one tile can be associated with several plates, with common centres. The goal of the tiling procedure, described in detail by Blanton et al. (2003)⁵, is to maximise the total number of targets being fibred. Due to the large-scale structure in the quasar/galaxy distribution the optimal tiling procedure overlaps individual tiles with one other.

As described in Blanton et al. (2003); Tegmark et al. (2004); Blanton et al. (2005); Percival et al. (2007), a “sector” is defined as a set of tile overlap regions (spherical polygons) observed by a unique combination of tiles and survey “chunks”. A ‘chunk’ is a unit of SDSS imaging data and is a part of an SDSS ‘stripe’, which is a 2.5° wide cylindrical segment aligned along a great circle between the survey poles. These sectors are the natural areas on which to define the completeness of our sample, with there being 7 814 sectors

for DR5, 5 831 of which have one or more POA objects in them.

Using the RegionID field in the **target** table (which gives the sector identification number if set, zero otherwise) we match the positions (R.A.’s and Decs) of objects in **target** to those that are in PhotoObjAll and the DR5Q. There are 116 689 POAs with a non-null RegionID and 55 577 DR5Qs (ie. the “PRIMARY” DR5Q Sample). The distribution of the number of (POA and DR5Q) objects per sector are given in Fig. A1. The mean number of POAs per sector is 14.9 and the mean number of DR5 (primary) quasars per sector is 6.97. The POA distribution has a long tail out to just under 300 objects / target, though only 65 sectors have > 100 POAs.

With the efficiency of the targetting algorithm being ~ 95% (Vanden Berk et al. 2005), for our purposes we can define two functions for the primary sample, which have dependence on angular position in the sky only, in order to calculate the completeness of the survey:

- Coverage Completeness, f_Q . The coverage completeness is the ratio of the number of quasar targets that are assigned a spectroscopic fibre to the total number of quasar candidates in a given sector. Fibre collisions will be one contributing factor in the coverage completeness, however, we do not include the fact that targets less than 100 arcsec from the tile center are not tiled.

- Spectroscopic Completeness, $f_s(\theta)$. This is the ratio of the number of high-quality spectra obtained in a sector to the number of spectroscopically observed. Due to the nature of the SDSS quasar survey, this ratio tends to be very high.

The ‘overall completeness’, f_O , is defined as $f_O = f_Q \times f_s$ and the distribution of this overall sector completeness is shown in Fig. A2. Although there are peaks in the number of sectors with 0, 0.5 and 1.0 completeness, 50% of the area of the survey is contributed by sectors with overall completenesses of 0.40-0.60.

APPENDIX B: JACKKNIFE ERRORS

Here we follow Scranton et al. (2002, §§3.4.5, 11.3 and their Eq. 10), Zehavi et al. (2002, §3.4 and equation 7) and Myers et al. (2007, Appendix A) to calculate the jackknife error estimates on our quasar clustering data. We also note the recent work of Loh (2008) on this issue who performed a detailed investigation into estimating the two-point correlation function and its errors.

Myers et al. (2007) estimate errors using an “inverse variance” weighted jackknife technique. This method divides the data into N sub-samples and then recalculates the given statistic (e.g. $\xi(s)$) using the Landy-Szalay estimator 5, *leaving out* one sub-sample area at one time. Following their convention we denote subsamples by the subscript L and recalculate $\xi(s)_L$ in each jackknife realization via equation 5. The inverse-variance-weighted covariance matrix, $C(s_i, s_j) = C_{ij}$, is

$$C_{ij} = \sum_{L=1}^N \sqrt{\frac{RR_L(s_i)}{RR(s_i)}} [\xi_L(s_i) - \xi(s_i)] \sqrt{\frac{RR_L(s_j)}{RR(s_j)}} [\xi_L(s_j) - \xi(s_j)] \quad (B1)$$

⁵ see also <http://www.sdss.org/dr6/algorithms/tiling.html>

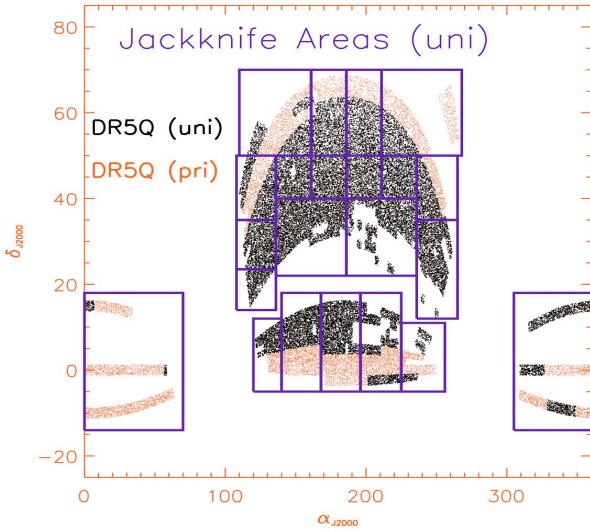


Figure B1. Geometry of the SDSS DR5Q Jackknife areas, showing the location of the DR5Q PRIMARY (orange/grey dots) and the UNIFORM (black dots) samples. The jackknife areas were chosen to follow the overall geometry of the SDSS Quasar survey. The number of quasars in each area is approximately equal. Note the sparse coverage of the UNIFORM sample in the Southern Stripes.

where ξ denotes the correlation function for all data and ξ_L denotes the correlation function for subsample L . Jackknife errors σ_i are obtained from the diagonal elements ($\sigma_i^2 = C_{ii}$), and the normalized covariance matrix, also known as the regression matrix, is

$$|C| = \frac{C_{ij}}{\sigma_i \sigma_j} \quad (\text{B2})$$

We divide the sample into 21 sub-samples. The number of subdivisions is chosen such that each represents a cosmologically significant volume, while retaining sufficient numbers of objects that shot noise will not dominate any subsequent analysis. The detailed boundaries of the sub-samples are given in Table B1 and B1.

We find, as in previous quasar clustering work e.g. Shanks & Boyle (1994); Croom & Shanks (1996), that Poisson errors are a good description on scales where $N_q \lesssim DD_q$, where N_q is the number of quasars in a given sample and DD_q is the number of quasar pairs in a given bin. On larger scales, the Poisson error tends to underestimate the Jackknife error, see Fig. B4. The scale where $N_q \approx DD_q$ is $\sim 70 \ h^{-1} \text{ Mpc}$ for the SDSS UNIFORM Quasar sample.

Using

$$\chi^2 = \sum_{ij} [\xi(s_i) - \xi_{\text{mod}}(s_i)] C_{ij}^{-1} [\xi(s_j) - \xi_{\text{mod}}(s_j)] \quad (\text{B3})$$

where C_{ij}^{-1} is the inverse covariance matrix, and $\xi_{\text{mod}}(s) = (s/s_0)^{-\gamma_s}$ is our model, where we vary s_0 over the range $s_0 = 0.0 - 15.0 \ h^{-1} \text{ Mpc}$ in steps of $0.05 \ h^{-1} \text{ Mpc}$ and γ_s over the range $\gamma_s = 0.00 - 3.00$ in steps of 0.01 , we now estimate χ^2 and determine errors on fits from $\Delta\chi^2$ over the $1 \ h^{-1} \text{ Mpc} < s < 25.0 \ h^{-1} \text{ Mpc}$ scales.

Here we find that using the full Covariance Matrix does change our estimates of the redshift-space correlation length and power-law slope to $s_0 = 6.35^{+0.40}_{-0.35} \ h^{-1} \text{ Mpc}$ and $\gamma_s =$

Region	RA min	RA max	Dec min	Dec max	No. of Quasars	No. of Randoms
N01	120.	140.	-5.	12.	29 445	870 558
N02	140.	168.	-5.	15.	28 445	841 193
N03	168.	196.	-5.	18.	27 904	825 442
N04	196.	225.	-5.	18.	28 717	846 926
N05	225.	256.	-5.	11.	29 891	879 837
N06	108.	136.	14.	23.5	29 614	873 778
N07	108.	136.	23.5	35.	28 646	845 871
N08	136.	186.	22.	40.	26 942	798 307
N09	186.	236.	22.	40.	27 957	820 491
N10	236.	265.	12.	35.	28 253	831 920
N11	108.	136.	35.	50.	29 003	856 576
N12	136.	161.	40.	50.	28 875	855 021
N13	161.	186.	40.	50.	28 857	853 908
N14	186.	211.	40.	50.	28 917	854 055
N15	211.	236.	40.	50.	28 924	854 070
N16	236.	265.	35.	50.	29 246	863 221
N17	110.	161.	50.	70.	29 253	863 420
N18	161.	186.	50.	70.	28 899	853 792
N19	186.	211.	50.	70.	28 911	853 175
N20	211.	268.	50.	70.	29 404	868 561
S	0✓305	70✓360	-14	18	28 675	842 497

Table B1. Details of the regions used for the Jackknife subsamples.

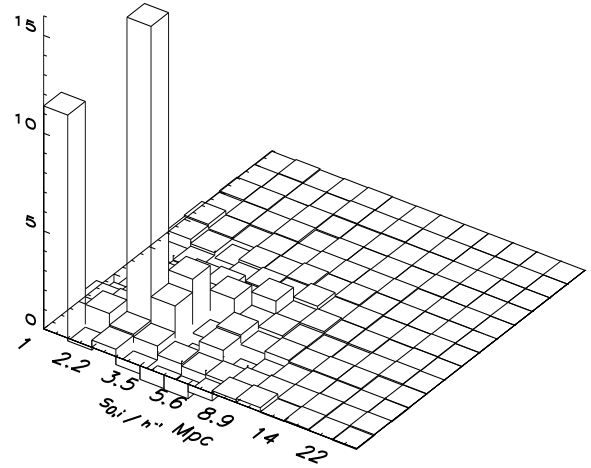


Figure B2. The Covariance matrix, \mathbf{C} , from jackknife error analysis on 21 sub-samples of the UNIFORM DR5Q, for scales $1 \ h^{-1} \text{ Mpc} < s < 25 \ h^{-1} \text{ Mpc}$.

$1.16^{+0.11}_{-0.08}$ respectively (cf. $s_0 = 5.95 \pm 0.45 \ h^{-1} \text{ Mpc}$ and $\gamma_s = 1.16^{+0.11}_{-0.08}$ found in Section 4.1). By studying the best-fit contours in the (s_0, γ_s) plane we find that our results are consistent for $1.0 \ h^{-1} \text{ Mpc} < s < 25.0 \ h^{-1} \text{ Mpc}$. However, fitting over $1.0 \ h^{-1} \text{ Mpc} < s < 100.0 \ h^{-1} \text{ Mpc}$ scales, we find there is some tension between the best-fit values given in Section 4.1 of $s_0 = 5.90 \pm 0.30 \ h^{-1} \text{ Mpc}$ and $\gamma_s = 1.57^{+0.04}_{-0.05}$ and the best-fit values using the covariance matrix of $s_0 = 6.95^{+0.45}_{-0.55} \ h^{-1} \text{ Mpc}$ and $\gamma_s = 1.53 \pm 0.09$. We

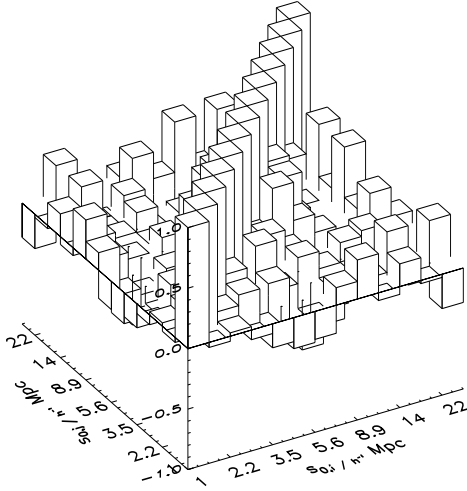


Figure B3. The normalised Covariance Matrix, (the Regression Matrix) for $\xi(s)$ from jackknife error analysis on 21 sub-samples of the UNIFORM DR5Q, for scales $1 \ h^{-1} \text{ Mpc} < s < 25 \ h^{-1} \text{ Mpc}$.

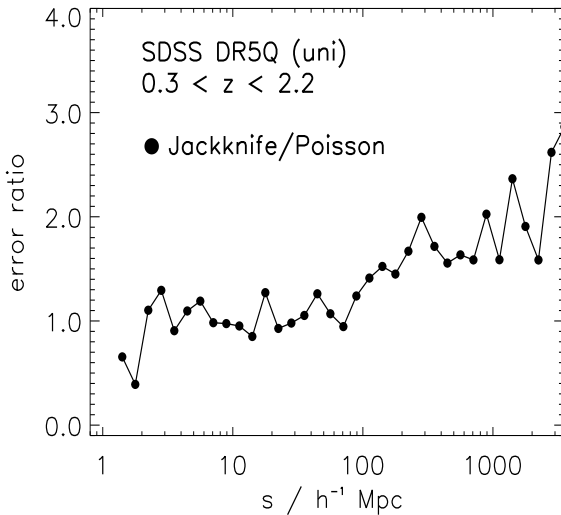


Figure B4. Comparison of Poisson and Jackknife errors for the UNIFORM DR5 Quasar sample. The ratio between the Poisson and Jackknife errors is very close to one at $s \lesssim 70 \ h^{-1} \text{ Mpc}$, while at $s \gtrsim 70 \ h^{-1} \text{ Mpc}$, the Poisson errors are \sim double that of the Jackknifes.

believe this is due to the noisy matrix inversion, where small values at large scales in the covariance matrix will dominate the signal in the inverse matrix. However, we are confident that using the diagonal elements of the Covariance matrix only for our model fits does not change the interpretation of our results.

APPENDIX C: SYSTEMATICS IN THE SDSS QUASAR 2PCF.

Here we produce evidence showing how various changes to our data and random methodology described in Sections 2 and 3 affect our main results. We shall check on the

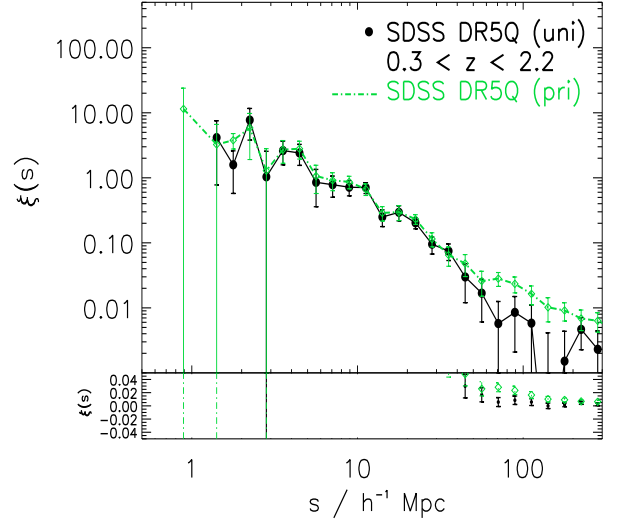


Figure C1. The SDSS DR5 Quasar $\xi(s)$ for the PRIMARY and UNIFORM samples. The lower panel shows the behaviour of $\xi(s)$ near zero on a linear scale.

effects of different samples (Sec. C1), different estimators (Sec. C2), changing the high-redshift cut (Sec. C3), Reddening (Sec. C6) and Fibre Collisions (Sec. C7). We shall report on $\xi(s)$ and $w_p(r_p)/r_p$ and find that when using the UNIFORM sample, our overall results (and subsequent interpretations) do not change given the assumptions in Section 3.

C1 Effects of Different Samples on $\xi(s)$

Figure C1 shows the difference in the redshift-space correlation function, $\xi(s)$, for the PRIMARY sample, versus that of the UNIFORM sample using the LS estimator. We see excellent agreement of the two samples at small scales $s \leq 20 \ h^{-1} \text{ Mpc}$, but the PRIMARY sample exhibits a higher clustering strength at large scales, $s \geq 40 \ h^{-1} \text{ Mpc}$. One possible explanation for this discrepancy is due to the differing radial distributions in PRIMARY and UNIFORM resulting from the different target selection used earlier on in the SDSS. This result provides our main motivation for using the UNIFORM sample exclusively in sections 3 and 4.

C2 The Use of Different Estimators for the 2PCF $\xi(s)$

Figures C2, C3 and C4 show the redshift-space 2PCF, $\xi(s)$ for the UNIFORM sample, using the different estimators of Davis & Peebles (1983), Hamilton (1992) and Landy & Szalay (1993). While the two estimators that use random-random (RR) pair counts are in extremely good agreement, the ‘Standard’ estimator, has too much power on large, $s \geq 40 \ h^{-1} \text{ Mpc}$ scales as this estimator is less robust to errors in the estimation of mean density.

C3 High redshift cut-offs

Figure C5 shows the redshift-space 2-point correlation function $\xi(s)$ for the UNIFORM sample with the high-redshift

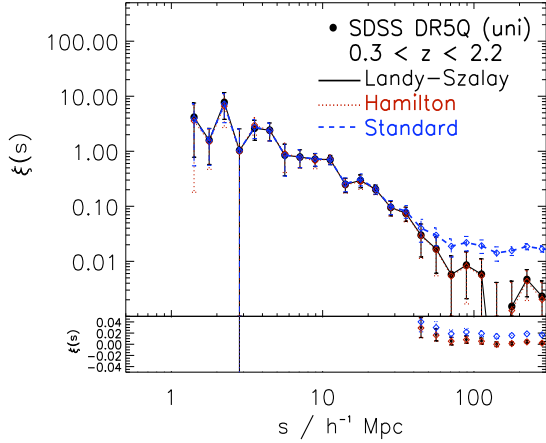


Figure C2. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample using the ‘Standard’, Hamilton and Landy-Szalay estimators. The lower panel shows the behaviour of $\xi(s)$ near zero on a linear scale.

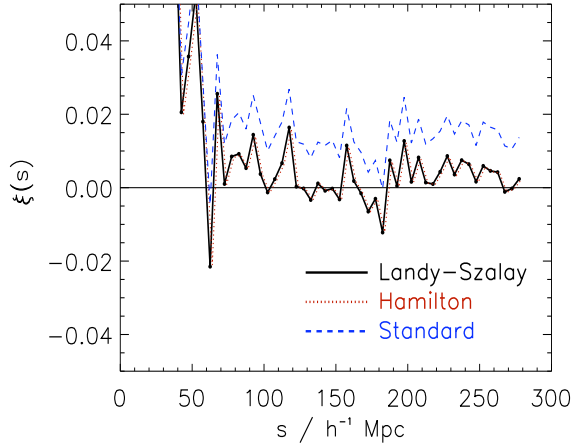


Figure C3. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample using the ‘Standard’, Hamilton and Landy-Szalay estimators, now on a linear scale

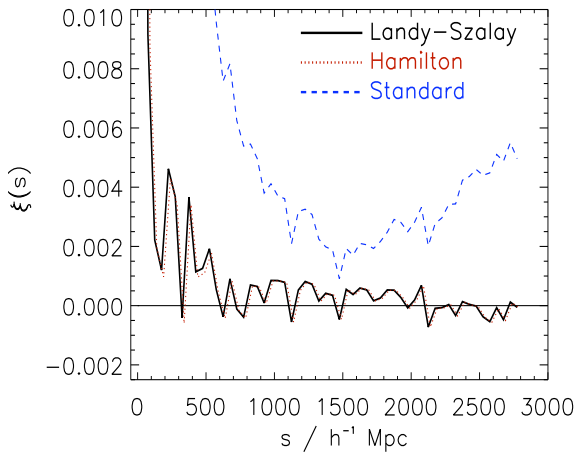


Figure C4. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample using the ‘Standard’, Hamilton and Landy-Szalay estimators, now on a much larger scale.

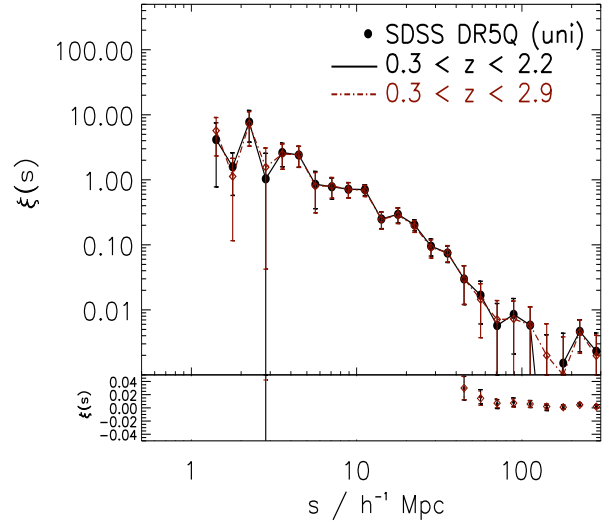


Figure C5. The SDSS DR5 Quasar UNIFORM $\xi(s)$ with upper redshift cut-offs of $z \leq 2.2$ and $z \leq 2.9$. The lower panel shows the behaviour of $\xi(s)$ near zero on a linear scale.

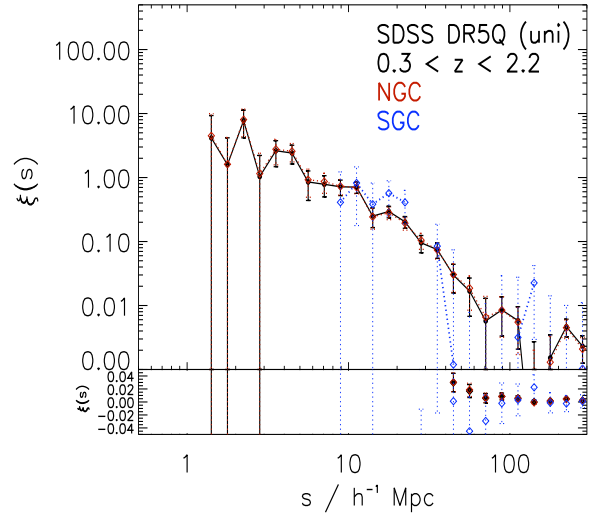


Figure C6. The SDSS DR5 Quasar $\xi(s)$ the UNIFORM sample with the sample split into the NGC and SGC.

cut-off being changed from $z \leq 2.2$ to $z \leq 2.9$. Although it is reassuring that the change between $\xi(s)$ is minimal, this is somewhat unexpected since our the optical selection for the quasar sample is known to be very low in completeness between $z = 2.2$ and $z = 2.9$ (Richards et al. 2006).

C4 The NGC vs. the SGC

Figure C6 shows the redshift-space 2-point correlation function $\xi(s)$ for the UNIFORM sample, split into quasars from the North Galactic Cap (NGC) and the South Galactic Cap (SGC). Note the data is heavily dominated by the NGC in the UNIFORM sample, with there being no signal in the SGC $\xi(s)$ below $s \approx 10 h^{-1} \text{ Mpc}$.

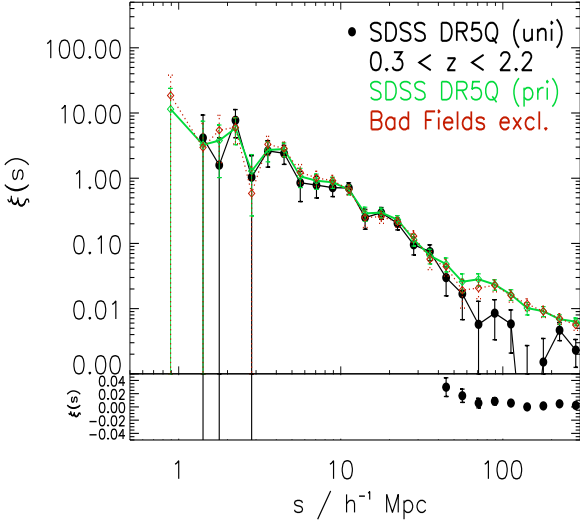


Figure C7. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample is given by the solid (black) circles. $\xi(s)$ for the PRIMARY sample is given with including, green (solid) line and excluding, dashed (red) line bad imaging fields.

C5 Bad Fields

In the SDSS, a “field” is an image in all five bands, with approximate dimensions of $13' \times 10'$. Since the quasar target selection algorithm searches for outliers from the stellar locus in colour space it is very sensitive to data with large photometric errors due to problems in photometric calibration or in point-spread function (PSF) determination (Richards et al. 2006). Thus, using the definitions of “bad fields” given by Richards et al. (2006) and Shen et al. (2007), we calculate the correlation function both including and excluding data from these areas.

Figure C7 shows the redshift-space 2-point correlation function $\xi(s)$ for the UNIFORM sample (solid black circles). Also shown is $\xi(s)$ for the PRIMARY sample including, (solid green) and excluding, (dashed red) lines, the “Bad Fields” as defined by Shen et al. (2007). Here we can see that there is minimal difference (for the PRIMARY DR5Q sample) between the $\xi(s)$ estimates both when including and excluding the bad fields.

C6 Reddening

While all selection for the quasar sample is undertaken using dereddened colors Richards et al. (2002), if there remain systematic errors in the reddening model they can induce excess power into the clustering in a number of different ways. The most obvious possibility comes from a modulation in the angular density of quasars as a function of position on the sky. A more subtle effect may arise in the redshift distribution of the quasars whereby the color dependence of the reddening correction may preferentially exclude quasars at specific redshifts. As we currently assume a common $N(z)$ for all quasars in the UNIFORM sample, an $N(z)$ that is reddening dependent can also induce excess clustering. For this analysis we will assume that any artificial signal that might be induced by the reddening correction will scale with the magnitude of the reddening correction itself. We therefore

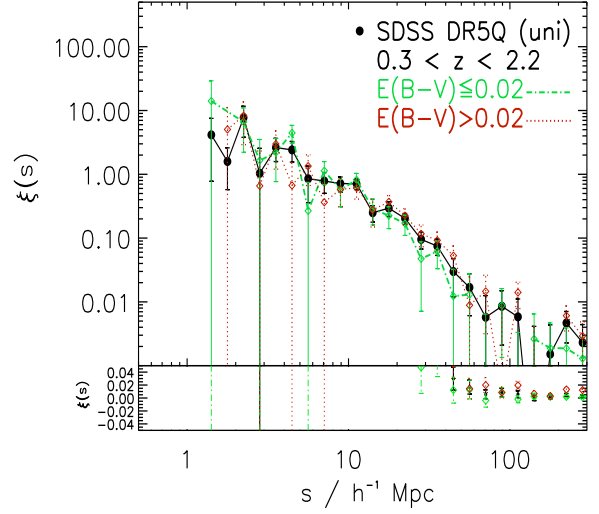


Figure C8. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample with the sample split by $E(B - V)$.

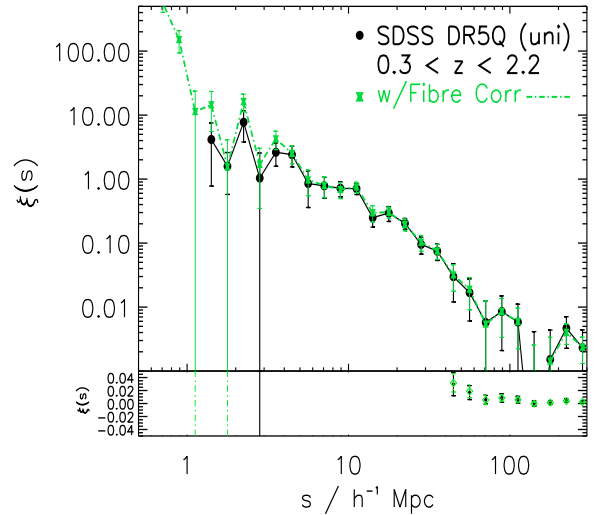


Figure C9. The SDSS DR5 Quasar $\xi(s)$ for the UNIFORM sample taking into account fibre collisions given by the (green) hour-glass points and dot-dashed line.

subdivide the UNIFORM quasar sample into two subsets, of approximately equal number, a low reddening sample, with $0.0028 < E(B - V) \leq 0.0217$, and a high reddening sample $0.0217 < E(B - V) \leq 0.2603$. The reddening estimates are derived from the reddening maps of Schlegel et al. (1998).

Figure C8 shows the redshift-space 2-point correlation function $\xi(s)$ for the full UNIFORM sample, with the reddening split sub-samples. The low reddening component, dot-dashed (green) line and the high reddening sample, dotted (red) line are consistent within the error for all scales out to $\sim 250 h^{-1}$ Mpc and there is no evidence for a systematic difference in the clustering signal on large scales that might be induced by any modulation due to variations in the reddening.

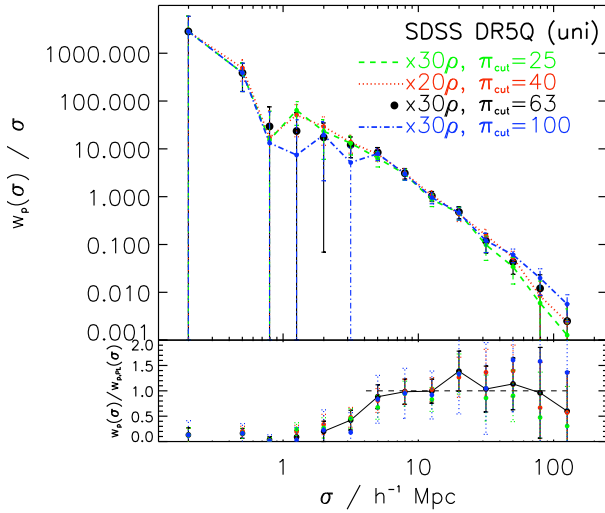


Figure C10. The projected correlation function $w_p(r_p)/r_p$ for the SDSS DR5Q UNIFORM sample with $0.30 < z < 2.2$, varying π_{\max} from equation 10.

C7 Fibre Collisions

As was noted in Section 3.3, due to the design of the SDSS fibres and plates, no two spectroscopic fibre can be separated by less than $55''$. To investigate this effect on our correlation function estimates we use the 145 524 POA objects and find which of these objects were not observed due to fibre collisions resulting from observations of the UNIFORM sample. We find there are 431 objects that were within $55''$ of a UNIFORM quasar that were not observed. Using the new version of the SDSS Quasar photometric catalogue, (Richards et al., 2008), we assign the redshift of the nearest photometric quasar to the “collided” objects. We assume that all the collided objects are in fact quasars, though in reality this is not the case. We then recalculate the 2PCF with these additional objects.

As we can see from Fig. C9, the inclusion of these collided objects makes very little difference to our measurement of $\xi(s)$ at scales $\gtrsim 5 h^{-1}$ Mpc. However, we do measure increased values of $\xi(s)$ at $s = 1 - 5 h^{-1}$ Mpc. Therefore, we again fit a single power-law to the data which has been corrected for fibre collisions over the scales $1 < s < 25 h^{-1}$ Mpc and find $s_0 = 6.70^{+0.45}_{-0.30} h^{-1}$ Mpc and $\gamma_s = 1.29^{+0.12}_{-0.10}$ (cf. $s_0 = 5.95 \pm 0.45 h^{-1}$ Mpc and $\gamma_s = 1.16^{+0.11}_{-0.08}$ found in Section 4.1). Unsurprisingly with the inclusion of more data at small separations, the fibre-corrected $\xi(s)$ has a higher s_0 value and steeper slope, but we find these results are consistent with our measurement of $\xi(s)$ without the fibre collision corrections, given the errors. We are thus satisfied that fibre collisions do not impact the results presented herein and refer the reader to Hennawi et al. (2006) and Myers et al. (2008) for more detailed investigations of quasar clustering and quasar binaries on these very small scales.

C8 Varying π_{\max} limits for $w_p(r_p)$

Figure C10 shows the projected correlation function $w_p(r_p)/r_p$ for the SDSS DR5Q UNIFORM sample with $0.30 < z < 2.2$, varying π_{\max} from equation 10. We vary

π_{\max} in intervals of $10^{0.2}$ from $\pi_{\max} = 10^{1.4-2.0} = 25.1 - 100.0 h^{-1}$ Mpc. We see that although changing the π_{\max} cut does produce a noticeable effect in estimates of w_p/r_p , we are confident that the integration limit of $\pi_{\max} = 63.1 h^{-1}$ Mpc provides a good balance between larger π values which would add noise to our w_p/r_p estimate and lower π values which might not recover the full signal at the largest separations.

REFERENCES

- Abazajian K., et al., 2004, *AJ*, 128, 502
 Adelberger K. L., Steidel C. C., 2005a, *ApJ*, 630, 50
 Adelberger K. L., Steidel C. C., 2005b, *ApJ Lett.*, 627, L1
 Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, *ApJ*, 619, 697
 Adelman-McCarthy J. K., et al., 2007, *ApJS*, 172, 634
 Alexander D. M., et al., 2003, *AJ*, 126, 539
 Anderson S. F., et al., 2003, *AJ*, 126, 2209
 Anderson S. F., et al., 2007, *AJ*, 133, 313
 Arp H., 1970, *AJ*, 75, 1
 Baes M., Buyle P., Hau G. K. T., Dejonghe H., 2003, *MNRAS*, 341, L44
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
 Barger A. J., et al., 2003, *AJ*, 126, 632
 Basilakos S., Georgakakis A., Plionis M., Georgantopoulos I., 2004, *ApJ Lett.*, 607, L79
 Basilakos S., Plionis M., Ragone-Figueroa C., 2008, *ApJ*, 678, 627
 Becker R. H., White R. L., Helfand D. J., 1995, *ApJ*, 450, 559
 Blanton M., Cen R., Ostriker J. P., Strauss M. A., 1999, *ApJ*, 522, 590
 Blanton M. R., Eisenstein D., Hogg D. W., Zehavi I., 2006, *ApJ*, 645, 977
 Blanton M. R., et al., 2005, *AJ*, 129, 2562
 Blanton M. R., Lin H., Lupton R. H., Maley F. M., Young N., Zehavi I., Loveday J., 2003, *AJ*, 125, 2276
 Boyle B. J., Georgantopoulos I., Blair A. J., Stewart G. C., Griffiths R. E., Shanks T., Gunn K. F., Almaini O., 1998, *MNRAS*, 296, 1
 Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N., Heymans C., 2000, *MNRAS*, 317, 1014
 Brandt W. N., Hasinger G., 2005, *ARA&A*, 43, 827
 Coil A. L., Hennawi J. F., Newman J. A., Cooper M. C., Davis M., 2007, *ApJ*, 654, 115
 Coles P., Erdogdu P., 2007, *Journal of Cosmology and Astro-Particle Physics*, 10, 7
 Coles P., Lucchin F., 2002, *Cosmology: The Origin and Evolution of Cosmic Structure*, Second Edition. ISBN 0-471-48909-3, Wiley-VCH.
 Constantin A., Vogeley M. S., 2006, *ApJ*, 650, 727
 Croom S. M., et al., 2005, *MNRAS*, 356, 415
 Croom S. M., et al., 2008, *MNRAS*, in prep.
 Croom S. M., Shanks T., 1996, *MNRAS*, 281, 893
 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., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, *MNRAS*, 349, 1397
 da Ângela J., et al., 2008, *MNRAS*, 383, 565

- da Ângela J., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Loaring N. S., Miller L., Smith R. J., 2005, *MNRAS*, 360, 1040
- Davis M., et al., 2003, in Guhathakurta P., ed., *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II*. Vol. 4834 of *Proceedings of the SPIE*, p161,
- Davis M., Newman J. A., Faber S. M., Phillips A. C., 2001, in Cristiani S., Renzini A., Williams R. E., eds, *Deep Fields: Proceedings of the ESOorkshop Held, Garching, Germany, 9-12 October 2000*, Springer-Verlag, p. 241, *The DEEP2 Redshift Survey*,
- Davis M., Peebles P. J. E., 1983, *ApJ*, 267, 465
- Eisenstein D. J., et al., 2001, *AJ*, 122, 2267
- Fan X., 1999, *AJ*, 117, 2528
- Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1994, *MNRAS*, 267, 927
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, 111, 1748
- Gilli R., et al., 2005, *Astron. & Astrophys.*, 430, 811
- Gunn J. E., et al., 1998, *AJ*, 116, 3040
- Gunn J. E., et al., 2006, *AJ*, 131, 2332
- Guzzo L., et al., 2008, *Nature*, 451, 541
- Haiman Z., Hui L., 2001, *ApJ*, 547, 27
- Hamilton A. J. S., 1992, *ApJ Lett.*, 385, L5
- Hawkins E., et al., 2003, *MNRAS*, 346, 78
- Hawkins M. R. S., Reddish V. C., 1975, *Nature*, 257, 772
- Hennawi J. F., et al., 2006, *AJ*, 131, 1
- Hogg D. W., Finkbeiner D. P., Schlegel D. J., Gunn J. E., 2001, *AJ*, 122, 2129
- Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, *ApJS*, 175, 356
- Hopkins P. F., Lidz A., Hernquist L., Coil A. L., Myers A. D., Cox T. J., Spergel D. N., 2007, *ApJ*, 662, 110
- Hoyle F., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Smith R. J., 2002, *MNRAS*, 332, 311
- Iovino A., Shaver P. A., 1988, *ApJ Lett.*, 330, L13
- Ivezić Ž., et al., 2004, *Astronomische Nachrichten*, 325, 583
- Jackson J. C., 1972, *MNRAS*, 156, 1P
- Kaiser N., 1987, *MNRAS*, 227, 1
- Kauffmann G., et al., 2003, *MNRAS*, 346, 1055
- Kerscher M., Szapudi I., Szalay A. S., 2000, *ApJ Lett.*, 535, L13
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121
- Kundic T., 1997, *ApJ*, 482, 631
- La Franca F., Andreani P., Cristiani S., 1998, *ApJ*, 497, 529
- Landy S. D., Szalay A. S., 1993, *ApJ*, 412, 64
- Lehmer B. D., et al., 2005, *ApJS*, 161, 21
- Lidz A., Hopkins P. F., Cox T. J., Hernquist L., Robertson B., 2006, *ApJ*, 641, 41
- Linder E. V., 2005, *Phys. Rev. D*, 72, 043529
- Loh J. M., 2008, *ApJ*, 681, 726
- Lupton R., Gunn J. E., Ivezić Z., Knapp G. R., Kent S., 2001, in Harnden Jr. F. R., Primini F. A., Payne H. E., eds, *Astronomical Data Analysis Software and Systems X* *Astronomical Society of the Pacific Conference Series*, 238, 269. p. 269
- Lynden-Bell D., 1969, *Nature*, 223, 690
- Martínez V. J., Saar E., 2002, *Statistics of the Galaxy Distribution*. Chapman & Hall/CRC
- Martini P., Weinberg D. H., 2001, *ApJ*, 547, 12
- McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2008, *ArXiv:0805.1335*
- Miller C. J., Nichol R. C., Gómez P. L., Hopkins A. M., Bernardi M., 2003, *ApJ*, 597, 142
- Miyaji T., et al., 2007, *ApJS*, 172, 396
- Mountrichas G., Shanks T., 2007, *MNRAS*, 380, 113
- Myers A. D., Brunner R. J., Nichol R. C., Richards G. T., Schneider D. P., Bahcall N. A., 2007, *ApJ*, 658, 85
- Myers A. D., Brunner R. J., Richards G. T., Nichol R. C., Schneider D. P., Bahcall N. A., 2007, *ApJ*, 658, 99
- Myers A. D., et al., 2006, *ApJ*, 638, 622
- Myers A. D., Outram P. J., Shanks T., Boyle B. J., Croom S. M., Loaring N. S., Miller L., Smith R. J., 2005, *MNRAS*, 359, 741
- Myers A. D., Richards G. T., Brunner R. J., Schneider D. P., Strand N. E., Hall P. B., Blomquist J. A., York D. G., 2008, *ApJ*, 678, 635
- Oke J. B., Gunn J. E., 1983, *ApJ*, 266, 713
- Osmer P. S., 1981, *ApJ*, 247, 762
- Outram P. J., Hoyle F., Shanks T., Boyle B. J., Croom S. M., Loaring N. S., Miller L., Smith R. J., 2001, *MNRAS*, 328, 174
- Padmanabhan N., et al., 2008, in prep.
- Padmanabhan N., White M., Norberg P., Porciani C., 2008, *ArXiv:0802.2105*
- Peacock J. A., 1999, *Cosmological Physics*. Cambridge University Press
- Peacock J. A., et al., 2001, *Nature*, 410, 169
- Peebles P. J. E., 1973, *ApJ*, 185, 413
- Peebles P. J. E., 1980, *The Large-Scale Structure of the Universe*. Princeton University Press.
- Peebles P. J. E., 1993, *Principles of physical cosmology*. Princeton, NJ: Princeton University Press
- Percival W. J., et al., 2007, *ApJ*, 657, 645
- Pier J. R., Munn J. A., Hindsley R. B., Hennessy G. S., Kent S. M., Lupton R. H., Ivezić Ž., 2003, *AJ*, 125, 1559
- Plionis M., Rovilos M., Basilakos S., Georgantopoulos I., Bauer F., 2008, *ApJ Lett.*, 674, L5
- Porciani C., Magliocchetti M., Norberg P., 2004, *MNRAS*, 355, 1010
- Rees M. J., 1984, *ARA&A*, 22, 471
- Richards G. T., et al., 2001, *AJ*, 122, 1151
- Richards G. T., et al., 2002, *AJ*, 123, 2945
- Richards G. T., et al., 2004, *ApJS*, 155, 257
- Richards G. T., et al., 2006, *AJ*, 131, 2766
- Rosati P., Tozzi P., Giacconi R., Gilli R., Hasinger G., Kewley L., Mainieri V., Nonino M., Norman C., Szokoly G., Wang J. X., Zirm A., Bergeron J., Borgani S., Gilmozzi R., Grogin N., Koekemoer A., Schreier E., Zheng W., 2002, *ApJ*, 566, 667
- Ross N. P., et al., 2007, *MNRAS*, 381, 573
- Salpeter E. E., 1964, *ApJ*, 140, 796
- Sánchez A. G., Baugh C. M., Percival W. J., Peacock J. A., Padilla N. D., Cole S., Frenk C. S., Norberg P., 2006, *MNRAS*, 366, 189
- Scherrer R. J., Weinberg D. H., 1998, *ApJ*, 504, 607
- Schlegel D. J., et al., 2007, in *Bulletin of the American Astronomical Society* Vol. 38 of *Bulletin of the American Astronomical Society*, *SDSS-III: The Baryon Oscillation Spectroscopic Survey (BOSS)*. pp 966+
- Schlegel D. J., Finkbeiner D. P., Davis M., 1998, *ApJ*, 500, 525

- Schneider D. P., et al., 2007, *AJ*, 134, 102
- Schulz A. E., White M., 2006, *Astroparticle Physics*, 25, 172
- Scoville N., et al., 2007, *ApJS*, 172, 1
- Scranton R., et al., 2002, *ApJ*, 579, 48
- Scranton R., et al., 2005, *ApJ*, 633, 589
- Serber W., Bahcall N., Ménard B., Richards G., 2006, *ApJ*, 643, 68
- Shankar F., Weinberg D. H., Miralda-Escude' J., 2007, *ArXiv e-prints*, 710
- Shanks T., Boyle B. J., 1994, *MNRAS*, 271, 753
- Shanks T., Fong R., Boyle B. J., Peterson B. A., 1987, *MNRAS*, 227, 739
- Shanks T., Fong R., Green M. R., Clowes R. G., Savage A., 1983, *MNRAS*, 203, 181
- Shen Y., et al., 2007, *AJ*, 133, 2222
- Sheth R. K., Mo H. J., Tormen G., 2001, *MNRAS*, 323, 1
- Smith J. A., et al., 2002, *AJ*, 123, 2121
- Smith R. E., et al., 2003, *MNRAS*, 341, 1311
- Smith R. E., Scoccimarro R., Sheth R. K., 2007, *Phys. Rev. D*, 75, 063512
- Spergel D. N., et al., 2007, *ApJS*, 170, 377
- Stoughton C., et al., 2002, *AJ*, 123, 485
- Strauss M. A., et al., 2002, *AJ*, 124, 1810
- Swanson M. E. C., Tegmark M., Blanton M., Zehavi I., 2008, *MNRAS*, 385, 1635
- Tegmark M., et al., 2004, *ApJ*, 606, 702
- Tucker D. L., et al., 2006, *Astronomische Nachrichten*, 327, 821
- Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, *ApJ*, 598, 886
- Vanden Berk D. E., et al., 2005, *AJ*, 129, 2047
- Wake D. A., et al., 2004, *ApJ Lett.*, 610, L85
- Wright E. L., 2006, *PASP*, 118, 1711
- Wyithe J. S. B., Loeb A., 2005, *ApJ*, 621, 95
- Wyithe J. S. B., Padmanabhan T., 2006, *MNRAS*, 366, 1029
- York D. G., et al., 2000, *AJ*, 120, 1579
- Zehavi I., Blanton M. R., Frieman J. A., Weinberg D. H., Waddell P., Yanny B., York D. G., 2002, *ApJ*, 571, 172
- Zehavi I., et al., 2005, *ApJ*, 630, 1