CLUSTERING ANALYSES OF 300,000 PHOTOMETRICALLY CLASSIFIED QUASARS. I. LUMINOSITY AND REDSHIFT EVOLUTION IN QUASAR BIAS

ADAM D. MYERS,1,2 ROBERT J. BRUNNER,1,2 ROBERT C. NICHL,3 GORDON T. RICHARDS,4,5 DONALD P. SCHNEIDER,6 AND NETA A. BAHCALL7

Received 2006 September 7; accepted 2006 December 5

ABSTRACT

Using ~300,000 photometrically classified quasars, by far the largest quasar sample ever used for such analyses, we study the redshift and luminosity evolution of quasar clustering on scales of ~50 h⁻¹ kpc to ~20 h⁻¹ Mpc from redshifts of z ~ 0.75–2.28. We parameterize our clustering amplitudes using realistic dark matter models and find that a ΛCDM power spectrum provides a superb fit to our data with a redshift-averaged quasar bias of bQ(z=1.40) = 2.41 ± 0.08 (P_a2 = 0.847) for α = 0.9. This represents a better fit than the best-fit power-law model [ω = (0.0493 ± 0.0064)z−0.928 ± 0.055, P_a2 = 0.482]. We find bQ increases with redshift. This evolution is significant at >99.6% using our data set alone, increasing to >99.999% if stellar contamination is not explicitly parameterized. We measure the quasar classification efficiency across our full sample as α = 95.6 ± 1.9%, a star-quasar separation comparable to the star-galaxy separation in many photometric studies of galaxy clustering. We derive the mean mass of the dark matter halos hosting quasars as M_{DMH} = (5.2 ± 0.6) × 10^{12} h M_☉. At z ~ 1.9 we find a 1.5 σ deviation from luminosity-independent quasar clustering; this suggests that increasing our sample size by a factor of ~1.8 could begin to constrain any luminosity dependence in quasar bias at z ~ 2. Our results agree with recent studies of quasar environments at z < 0.4, which detected little luminosity dependence to quasar clustering on proper scales ≥50 h⁻¹ kpc. At z < 1.6, our analysis suggests that bQ is constant with luminosity to within ΔbQ ~ 0.6, and that, for g < 21, angular quasar autocorrelation measurements are unlikely to have sufficient statistical power at z ≤ 1.6 to detect any luminosity dependence in quasars’ clustering.

Subject headings: cosmology: observations — large-scale structure of universe — quasars: general — surveys

Online material: color figures

1. INTRODUCTION

As the form of the nonbaryonic, cold dark matter that underpins mass in the cosmos becomes increasingly accurately described (e.g., Cole et al. 2005), understanding the baryonic processes that fuel quasars and trigger galaxy formation becomes an increasingly realistic endeavor. It is now established that most, if not all, local galaxies harbor a supermassive black hole (Kormendy & Richstone 1995; Richstone et al. 1998) and that the mass of these black holes correlates with several properties of their host galaxy’s bulge (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000; Graham et al. 2002; Tremaine et al. 2002; Wyithe 2006), implying a causal link between black holes and star formation in galaxy spheroids (e.g., Silk & Rees 1998). Observational evidence is accumulating to suggest that such a causal link remains at higher redshift (Shields et al. 2003).

It has long been suspected that accretion of baryons onto supermassive black holes is responsible for the powerful UV-excess (UVX) emission seen in quasars (see, e.g., Rees 1984 for a review), so the role of supermassive black holes in galaxy formation (UVX) emission seen in quasars (see, e.g., Rees 1984 for a review) has provided useful broad constraints on gravitationally driven aspects of quasar evolution. However, the relevance of gas physics to quasar evolution means that measurements of the luminosity function of quasars (see Richards et al. 2006 for a review) have mergers drive the formation of quasars and supermassive black holes, which in turn seed new galaxies (e.g., Heckman et al. 1986; Carlborg 1990; Barnes & Hernquist 1992; Lacey & Cole 1993; Di Matteo et al. 2005). Given that only merging systems of a certain minimum mass can trigger a UVX quasar phase visible against background star formation in the host galaxy (e.g., Hopkins et al. 2006), this picture is consistent with emerging evidence that at z ~ 2.5 quasar bias evolves with redshift but that quasars inhabit dark matter halos of similar average mass at every redshift (Porciani et al. 2004; Croom et al. 2005; Myers et al. 2006). However, the simplicity of this picture belies a rich complexity in the important physical processes that entwine quasar, galaxy, and star formation (see Hopkins et al. 2006 for a review). Some of the many theoretical insights into this complexity have included the importance of galaxy mergers (e.g., Toomre & Toomre 1972; White 1979; Negroponte & White 1983; Barnes & Hernquist 1992; Franceschini et al. 1999), cooling flows (e.g., Ciotti & Ostriker 1997; Ciotti & Ostriker 2001), heating through various feedback mechanisms (e.g., Silk & Rees 1998; Wyithe & Loeb 2002; Springel & Hernquist 2003), and/or the eventual cutoff of accretion onto a central black hole by gas ejection (e.g., Silk & Rees 1998; Fabian 1999; Sazonov et al. 2005).

Clearly, quasar evolution is an important tracer of galaxy formation; however, the large number of components that help regulate models of quasar activity beg new constraints. Measurements of quasar clustering amplitudes, which directly correlate with the average mass of the halos that harbor quasars, have provided useful broad constraints on gravitationally driven aspects of quasar evolution. However, the relevance of gas physics to quasar evolution means that measurements of the luminosity function of quasars (see Richards et al. 2006 for a review) have
also proved key in constraining baryonic elements of quasar evolution, e.g., quasar lifetimes via the “duty cycle” (Haiman & Hui 2001; Martini & Weinberg 2001). Hopkins et al. (2005a) have suggested that the peak luminosity distribution of quasars is fundamental to characterizing the quasar population; this suggests that important observational constraints will emerge by considering baryons and gravity in tandem, by measuring the luminosity evolution of quasar clustering (e.g., Valageas et al. 2001; Lidz et al. 2006).

Since the first significant detections of quasar clustering, the two-point correlation function (e.g., Totsuji & Kihara 1969; Peebles 1980) has frequently been used to measure the amplitude of quasar clustering, and accuracy has improved in step with sample sizes. Recent detections, in the wake of large spectroscopic quasar surveys, have led to some confidence about the evolution of quasar clustering (e.g., Croom et al. 2005, hereafter C05), constraining a lack of evolution in the dark matter halos that host quasars to ∼50%. However, the dependence of quasar clustering on luminosity appears to be quite weak (e.g., C05; Lidz et al. 2006), and probing variations in quasar clustering as a function of luminosity or other physical properties, as well as more tightly constraining evolution in quasar clustering, is limited by quasar sample sizes.

Large samples of photometrically selected objects have long been used to probe angular galaxy clustering (Groth & Peebles 1977), leading to important cosmological constraints such as early detections (Maddox et al. 1990a) of deviations from standard cold dark matter (CDM) models or early detections of dark energy (e.g., Scranton et al. 2003). Such angular analyses were complementary to analyses using spectroscopic data because of the larger numbers of galaxies that could be photometrically selected. Similar angular clustering analyses of complementarily large samples of photometrically selected quasars were impossible; however; star-galaxy separation became efficient in galaxy surveys with the advent of automatic plate measurements (e.g., Maddox et al. 1990b) but star-quasar separation lagged behind (e.g., ∼60% efficient in the 2dF QSO Redshift Survey [Croom et al. 2004, hereafter 2QZ]). With the recent advent of sophisticated photometric classification of quasars (Richards et al. 2004) star-quasar separation at many redshifts is now highly efficient (∼95%; Richards et al. 2004; Myers et al. 2006), meaning that large quasar samples can be used to measure angular quasar clustering as a function of physical properties (Myers et al. 2006) and to use angular quasar clustering to probe dark matter (Scranton et al. 2005) and dark energy (Giannantonio et al. 2006).

In Myers et al. (2006, hereafter M06), we presented a first, proof-of-concept analysis of the clustering of ∼80,000 photometrically classified quasars. In this series of papers, we extend this work, improving our modeling techniques and presenting measurements of the two-point correlation function of ∼300,000 photometrically classified quasars drawn from the fourth data release (DR4) of the Sloan Digital Sky Survey (SDSS; e.g., Stoughton et al. 2002; Abazajian et al. 2003, 2004, 2005). Our goal in this paper is to study the dependence of quasar clustering on redshift and luminosity, focusing on linear and quasi-linear scales. In a companion paper (Myers et al. 2007, hereafter Paper II), we analyze quasar clustering on smaller scales.

Extensive details of our techniques, modeling, and systematics are presented in Appendices A and B, allowing our main analysis to be presented in § 3, after detailing our data sample in § 2. Our main results are ordered in a concluding section (§ 5). Unless otherwise specified, we assume a ΛCDM cosmology with (Ω_m, Ω_Λ, σ_8, Γ, h ∼ H_0/100 km s^{-1} Mpc^{-1}) = (0.3, 0.7, 0.9, 0.21, 0.7), where Γ is the shape of the matter power spectrum (Γ = Ω_m h for baryon-free CDM). We correct all magnitudes for Galactic extinction using the dust maps of Schlegel et al. (1998).

2. THE DR4 KDE SAMPLE

The quasar sample that we analyze is constructed using the kernel density estimation (KDE) technique of Richards et al. (2004), which draws on many unique technical aspects of the SDSS (e.g., York et al. 2000), including superior photometry (e.g., Fukugita et al. 1996; Gunn et al. 1998; Lupton et al. 1999; Hogg et al. 2001; Smith et al. 2002; Ivezić et al. 2004), astrometry (e.g., Pier et al. 2003), and data acquisition (e.g., Gunn et al. 2006; Tucker et al. 2006). As in Richards et al. (2004) the sample is restricted to SDSS point sources with u − g < 1, (observed) g ≥ 14.5 and (dereddened) g < 21. Separations, in four-dimensional color-space, from a sample of ∼10% of point sources in SDSS Data Release 1 (Abazajian et al. 2003, hereafter DR1) and from the quasar sample of Schneider et al. (2003, hereafter DR1QSO) are determined for each object to be classified. A Bayesian classifier then assigns each object a probability of being a “quasar” or “star.” Taken in logarithmic ratio, the distribution of these probabilities is sufficiently bimodal to separate z ≤ 2.5 quasars from stars with ∼95% efficiency (Richards et al. 2004; see also M06). Applying the KDE technique to SDSS DR4 results in our DR4 KDE sample of 344,431 objects,8 a sample 3.5 times larger than the DR1 sample used in M06. Each KDE object is assigned a photometric redshift estimate as described in Weinstein et al. (2004).

To meaningfully model quasar clustering we must know the normalized redshift distribution of our sources (dN/dz in eq. [A6]). For consistency with M06 (see their Fig. 6), we estimate dN/dz from spectroscopic matches to DR1QSO, matches to spectra taken from the SDSS second data release (Abazajian et al. 2004, hereafter DR2), or matches to the 2QZ. As our KDE technique is currently trained on DR1QSO, obtaining dN/dz from that sample is arguably a fairer approach than using quasars from later data releases. M06 demonstrated that including or excluding matches with the 2QZ or DR2 has little effect on the form of dN/dz and further showed that the methodology of using spectroscopic matches is broadly consistent with estimating dN/dz from photometric redshifts (Weinstein et al. 2004).

3. QUASAR CLUSTERING RESULTS

3.1. Mean Quasar Bias at z ∼ 1.4

In Figure 1 we show the (A_q < 0.21) DR4 KDE autocorrelation. We fit bias models (see eq. [A6]) over scales of 0.16′ to 63′ (∼55 h^{-1} kpc to 22 h^{-1} Mpc at the DR4 KDE sample’s mean redshift of z = 1.4). The fit’s upper scale limit is nominally set by stellar contamination (see § B1) and the lower limit is set by the dark matter model we use for Ψ_{QG}. Smith et al. (2003) note that their models accurately reproduce Δ_{QG}^2 to the limits of current simulations (∼3%) at wavenumbers of k < 10 h {Mpc}^{-1} (∼1); however, their models appear quite accurate even on scales several times smaller than this (e.g., Fig. 15 of Smith et al. 2003), and in any case their models remain useful as a phenomenological description of dark matter clustering on all our scales of interest. In particular, any models that augment the approach of Smith et al. (2003) at k < 10 h {Mpc}^{-1} should be easy to compare to Smith et al. (2003) and thus to our results.

Our best-fit bias model to the DR4 KDE autocorrelation has b^{Q}_{Q} = 2.41 ± 0.08 (P_{< χ^2} = 0.847), in good agreement with

8 Available at http://sdss.ncsa.uiuc.edu/qso/nbkde.
such a model should still be valid on these scales. When a stellar contamination model is fit out to 100", we obtain $b_Q = 2.47 \pm 0.15$ and $\alpha = 0.970 \pm 0.017$ ($P_{<\chi^2} = 0.908$). This value of $\alpha$ reproduces the data very well out to at least $7'$ (see Fig. 1); however, although altering the scale of the fit affects estimates of $\alpha$, $b_Q$ is essentially unchanged, as stellar contamination only dominates on large scales. We further note that altering $\omega_{SS}$ to 0.25, as is appropriate at $\sim 30'$ (see M06), barely changes our result, giving $b_Q = 2.47 \pm 0.14$ and $\alpha = 0.974 \pm 0.025$ ($P_{<\chi^2} = 0.936$)—again, simply because stellar contamination is only influential on scales larger than those we typically fit. In general, throughout this paper we will quote results for models that simultaneously fit for $\alpha$ and $b_Q$. However, in most cases fitting for $\alpha$ at $<1'$, although illustrative, is overkill and falsely reduces the significance of our $b_Q$ measurements. We will therefore trust the significance estimates of those fits that ignore stellar contamination.

3.2. Evolution in Quasar Bias

The spectroscopic redshift distribution of quasars at a given photometric redshift can become increasingly complex as the photometric redshift bandwidth is reduced. Therefore, to model the evolution of quasar clustering in a number of photometric redshift bins, it is convenient to have a better mechanism for modeling $dN/dz$ than the simple spline fit used in §3.1. Typically, the photometric redshift distributions we study have a primary peak where the photometric solution agrees with the spectroscopy and, in some cases, a minor, secondary peak where the photometric solution is inaccurate (due to so-called catastrophic failures). We therefore adopt the approach of summing a number of functions of the form

$$dN = \sum_{i} \beta_{i} \exp \left[ -\frac{|z - \bar{z}|^{\sigma_{i}}}{n_{i} \sigma_{i}^{n_{i}}} \right] dz,$$

where $n$ (typically close to the Gaussian value of 2), $\sigma$, $\bar{z}$, and $\beta$ are free parameters. These functions can excellently reproduce $dN/dz$ (see Fig. 2). Our schema for binning in photometric redshift is chosen to maximize object numbers in each bin while limiting discrepancies between the photometric and spectroscopic redshift estimates (see M06).

In Figure 3 we show the evolution of the angular quasar autocorrelation with photometric redshift. We derive estimates of the quasar bias using equation (A6), and also consider a two-parameter stellar-contamination model (eq. [B1]). In Figure 4 we display our measured quasar bias evolution. Our data alone (i.e., without assuming $b_Q \sim 1$ at $z = 0$) rule out constant $b_Q$ at all redshifts at $>99.99\%$, dropping to $>97.9\%$ if stellar contamination is allowed to freely vary. As discussed in §3.1, not fitting for stellar contamination likely better estimates significances. Using 2QZ data, PMN04 and C05 have independently determined that $b_Q$ evolves with redshift. We note that after correcting for differing $\sigma_{Q}$, our value of $b_{Q}^{\omega_{Q0}} = 1.87$ disagrees with $b_{Q}^{\omega_{Q0}} = 1.89$ from PMN04 but only at the 1.8 $\sigma$ level, dropping to 0.8 $\sigma$ if stellar contamination is incorporated. The $M_{b} < -22.5$ restriction adopted by PMN04 is immaterial in this context, as every $g < 21$ quasar at $z > 1.7$ should be brighter than $M_{b} = -22.5$. Bias determinations derived from cosmological models that are more like those used by PMN04 and C05 are provided in Table 1 (see also the discussion in §4.1).

The photometric nature of our redshifts leads to problems for our $z_{\text{phot}} = 0.75$ and 2.28 bins (i.e., our lowest and highest redshift bins)—in particular, where to plot these bins on the $z_{\text{phot}}$ axis of Figure 4. At $1 < z_{\text{phot}} < 2$, the mean redshift of our
spectroscopic matches is close to $z_{\text{phot}}$, but beyond this range the true redshift is further from the photometric estimate, due to catastrophic failures in the photometric redshift estimation. Figure 2 demonstrates that at $z_{\text{phot}} < 1$ ($z_{\text{phot}} > 2$) there is a secondary solution, amounting to 12.6% (21.2%) of the area under $dN/dz$, at $z_{\text{phot}} > 2$ ($z_{\text{phot}} < 1$). We can examine clustering at these secondary $z_{\text{phot}}$ solutions by adopting a similar approach to equation (B1):

$$B^2 \omega_1 = \omega_{1+2} - (1 - B)^2 \omega_2,$$

where $\omega_1$, $\omega_2$, and $\omega_{1+2}$ are, respectively, the true clustering at the primary and secondary $z_{\text{phot}}$ solutions, and the clustering we measure as a combination of the two $z_{\text{phot}}$ solutions. Here $B$ is the relative contribution of the primary and secondary $z_{\text{phot}}$ solutions to $dN/dz$. We can then estimate the true value of the bias at the position of the primary solution in $z_{\text{phot}}$ as

$$b_1^2 = \frac{\omega_1}{\omega_1^M} = \frac{b_{1+2}^2 \omega_{1+2}^M - (1 - B)^2 b_2^2 \omega_2^M}{B^2 \omega_1^M},$$

where the superscript $M$ denotes the model values of $\omega$ (eq. [A6] with $b_Q = 1$) and the $b_i$ denote quasar bias. Although we do not know the true values of $b_1$ and $b_2$, they can be estimated from measurements of $b_Q$.

If we follow this analysis, our value of $b_Q$ at $z_{\text{phot}} = 0.75$ ($z_{\text{phot}} = 2.28$) should be reduced (increased) by at least 10%. The true values of $b_Q$ must lie beyond even these 10% offsets, because we must use our measured values of $b_Q$ at the secondary solution in $z_{\text{phot}}$, rather than the (unknown) true value, to estimate the true value in the primary $z_{\text{phot}}$ bin. If we lower our estimates of $b_{Q, z_{\text{phot}}=0.75}$ by 10% and increase our estimates of $b_{Q, z_{\text{phot}}=2.28}$ by 10%, we find that our data rule out a constant $b_Q$ at all redshifts with a significance of $>99.9999\%$, dropping to $>99.6\%$ if stellar contamination is allowed to freely vary. We note that applying equation (3) to our bins at $z_{\text{phot}} = 1.20, 1.53$, and 1.87 has no affect on $b_Q$ values, as any secondary solutions in $z_{\text{phot}}$ have a very small weight ($B \leq 0.05$).

3.3. The Luminosity Evolution of Quasar Clustering

While the luminosity of UVX quasars depends somewhat on the mass of the underlying black hole, the mechanisms that drive baryons onto the accretion disk feeding the black hole are

![Fig. 2.—Spectroscopic redshift ($z_{\text{spectro}}$) distribution of KDE quasars for some photometric redshift bins ($z_{\text{phot}}$) used in our analyses. The dotted lines are functional fits of the form described by eq. (1). The plotted distributions have been normalized and the integral under the fitted functions is always within 0.1% of unity. [See the electronic edition of the Journal for a color version of this figure.]

Vol. 658
also important. As such, models of quasar formation and evolution (e.g., Hopkins et al. 2005a, 2005b, 2006) can be degenerate between mass and luminosity. It is therefore useful to examine constraints on quasar bias as a function of luminosity. As discussed in M06, ideal tests of quasar evolution would attempt to break the luminosity-redshift degeneracy and examine multivariate quasar properties. We now repeat our clustering analysis as a bivariate function of redshift and luminosity. We derive $g$-band absolute magnitudes ($M_g$) for KDE objects by assuming that each photo-$z$ is a reasonable ensemble estimate of redshift. We incorporate the $K$-correction from Wisotzki (2000; see, e.g., M06). Consistent $M_g$ binning at every redshift is impractical for quasars (which span $\sim8$ mag in $M_g$), so as in M06 we split the KDE sample into three photometric redshift bins, then subdivide these into three $M_g$ bins. We then measure the autocorrelation of each of these nine subsamples.

In Figure 5 we plot the bivariate quasar autocorrelation as a function of photometric redshift and absolute magnitude. As before, we derive the quasar bias over the range $0.16'$ to $63'$, plotting the results in Figure 6 (see also Table 2). Our results are at least twice as precise as the equivalent results from M06, but other than the general increase with redshift discussed in §3.2, there appears to be no discernible trend in quasar clustering with absolute magnitude. However, a model with constant quasar bias as a function of absolute magnitude is only just accepted by our data ($b_Q = 2.50 \pm 0.11$, $P_{<\chi^2} = 0.092$). If we

![Figure 3: Quasar clustering evolution as a function of photometric redshift. There are $\sim65,000$ quasars in each bin except for the $2.1 \leq z_{\text{phot}} < 2.8$ bin, which contains $\sim28,000$ quasars. The solid line is the expected clustering of dark matter derived from Smith et al. (2003). The dotted line is our best-fit model in which only the quasar bias, $b_Q$, is varied. The dashed line is a two-parameter model that incorporates stellar contamination as well as quasar bias (see Table 1 for model values). Errors in this plot are jackknifed and fits are made over scales of $0.16'$ to $63'$ using a full covariance matrix. A scale of $0.16'$ to $63'$ is $\sim55$ $h^{-1}$ kpc to $22$ $h^{-1}$ Mpc in all bins except the lowest redshift bin, where the scales are slightly reduced to $\sim50$ $h^{-1}$ kpc to $20$ $h^{-1}$ Mpc. [See the electronic edition of the Journal for a color version of this figure.]
and (2) ignoring stellar contamination, a maximum rejection of 0.1 σ for \( \bar{z} = 0.85 \), 0.5 σ for \( \bar{z} = 1.44 \), and 1.5 σ for \( \bar{z} = 1.92 \).

4. DISCUSSION

4.1. Quasar Host Masses

Following C05 we can use the ellipsoidal collapse model of Sheth et al. (2001) to convert quasar biases into masses for the halos hosting UVX quasars. We weight this model across our (normalized) redshift distributions:

\[
b_Q(M_{\text{DMH}}, \bar{z}) = 1 + \frac{\int_{\bar{z} = \min}^{\bar{z} = \max} dz \frac{dN}{dz}}{\sqrt{\Delta \bar{z}(\bar{z})}} \left( \frac{1}{a(\bar{z})^2 + \sqrt{a b(\bar{z})^2}^{1-c} - \left( \frac{b(\bar{z})^2}{(1-c)(1-c/2)} \right) \right),
\]

(4)

where \( a = 0.707, b = 0.5, c = 0.6 \), and \( \delta_c(\bar{z}) \), the critical density contrast for spherical collapse, is given by Navarro et al. (1997) as 0.15(12π)^{-3/2}Ω_\text{m}^{0.665} (for flat cosmologies), where \( \Omega_\text{m} \equiv \Omega_{\text{m}}(\bar{z}) \) is given, e.g., in M06.

Masses can be derived via \( \nu = \delta_c(\bar{z}) \sigma_c(M, \bar{z}) \), where \( \sigma_c(M, \bar{z}) = \sigma_c(M) D(\bar{z}) \), and \( D(\bar{z}) \) is the linear growth factor, which we approximate using the formula of Carroll et al. (1992; see, e.g., M06). The mass variance for a halo, \( \sigma^2_c(M_{\text{DMH}}) \), can be determined from the radius of a halo of mean mass \( M_{\text{DMH}} \):

\[
r = \left( \frac{3M_{\text{DMH}}}{4\pi \rho_0} \right)^{1/3},
\]

(5)

where \( \rho_0 = 2.78 \times 10^{11} \Omega_{\text{m}} h^2 M_\odot \, \text{Mpc}^{-3} \) is the present mean density of the universe. This mass scale implies a mass variance of

\[
\sigma^2_c(M_{\text{DMH}}) = \frac{V}{2\pi^2} \int_0^\infty k^3 P(k) \left[ \frac{3j_1(kr_0)}{kr_0} \right]^2 dk
\]

(6)

where the term in square brackets represents a spherical top-hat smoothing for the density field \( j_1 \) is the spherical Bessel function of first order. We set \( V \) so that \( \sigma_c \) is tied to observations when \( r = 8 h^{-1} \) Mpc.

We assume our bias values are valid in the linear regime (as our analysis suggests \( b_Q \) is scale-independent over at least

### TABLE 1

**Estimates of the Quasar Bias \( b_Q \) and the Quasar Host Halo Mass \( M_{\text{DMH}} \) as a Function of Photometric Redshift \( \bar{z}_{\text{phot}} \)**

| \( \bar{z}_{\text{phot}} \) | \( \nu = 0.21 \), \( \sigma_c = 0.9 \) | \( \nu = 0.15 \), \( \sigma_c = 0.8 \) (one-parameter model) |
|-------------------------|---------------------------------|-------------------------------------------------|
| Range | Mean | One-Parameter Model | Mean | Two-Parameter Model \( b_Q \) | Mean | Two-Parameter Model \( b_Q \) |
| 0.4, 2.3 | 1.40 | 2.41±0.09 | 12.1±1.1×10^{12} | 0.956±0.044 | 2.48±0.13 |
| 0.4, 1.0 | 0.75 | 1.93±0.14 | 8.90±2.71×10^{12} | 0.903±0.097 | 1.92±0.33 |
| 1.0, 1.4 | 1.20 | 2.05±0.11 | 12.5±1.1×10^{12} | 0.997±0.003 | 2.06±0.33 |
| 1.4, 1.7 | 1.53 | 2.23±0.17 | 9.22±2.48×10^{12} | 0.884±0.057 | 2.35±0.33 |
| 1.7, 2.1 | 1.87 | 2.51±0.16 | 11.3±4.5×10^{12} | 0.907±0.017 | 3.05±0.25 |
| 2.1, 2.8 | 2.28 | 2.84±0.40 | 11.3±4.8×10^{12} | 0.840±0.039 | 3.20±0.29 |

* With stellar contamination \( 1 - a \).
* Due to catastrophic failures, values of \( b_Q \) in this row need lowered at least 10%. The derived \( M_{\text{DMH}} \) values use the reduced values. See eq. (3) and the associated discussion.
* Due to catastrophic failures, values of \( b_Q \) in this row need raised at least 10%. The derived \( M_{\text{DMH}} \) values use the increased values. See eq. (3) and the associated discussion.
and adopt a form for the linear power spectrum of \( P(k) = T^2(k) k^n \) with \( n = 1 \) (the Harrison-Zel'dovich-Peebles scale-invariant case). We adopt the transfer function \( T(k) \) given by equation (29) of Eisenstein & Hu (1998), who, following Bardeen et al. (1986; see also Efstathiou et al. 1992), showed that the transfer function can be characterized by its shape (\( \Gamma \)) via

\[
q = \frac{k}{h^{-1} \text{Mpc}} \Theta_{z,7}^2 / \Gamma
\]

(7)

for a CMB temperature parameterized as \( 2.7 \Theta_{2.7} \) K. In pure CDM, \( \Gamma = \Omega_m h \), however, baryons affect the power-spectrum shape, and \( \Gamma \rightarrow \Gamma_{\text{eff}} \) (e.g., Suguizawa 1995). Note that as we analyze scales far smaller than the sound horizon, \( \Gamma_{\text{eff}} \) can be derived from the baryon fraction via \( \Gamma_{\text{eff}} = \alpha_s \Omega_m h \), with \( \alpha_s \) given by equation (31) of Eisenstein & Hu (1998).

Throughout this paper, we have used a concordance cosmological model with \( \sigma_8 = 0.9 \) and \( \Gamma = 0.21 \), as quasar bias is not greatly affected by complementarily altering \( \sigma_8 \) and \( \Gamma \). However, the conversion from \( b_Q \) to \( M_{\text{DMH}} \) is somewhat dependent on \( \sigma_8 \).

### Figure 5

Quasar clustering evolution as a bivariate function of absolute magnitude \( M_g \) and photometric redshift \( z_{\text{phot}} \). The rows show three bins of roughly equal numbers in \( z_{\text{phot}} \). The columns divide each \( z_{\text{phot}} \) bin into three of equal numbers in \( M_g \). There are \( \sim 30,000 \) quasars in each bin. Labeled in each panel are the mean \( g \) apparent and absolute magnitudes, the quasar bias \( b_Q \) derived as in Fig. 3, and the \( \chi^2 \) probability of our best-fit model in which only \( b_Q \) is fitted (dotted line). The dashed line is a two-parameter model that incorporates stellar contamination as well as quasar bias. Table 2 displays model values. Errors in this plot are jackknifed and fits are made overscalers of 0.16° to 63° using a full covariance matrix. A scale of 0.16° to 63° is \( \sim 55 \) h\(^{-1}\) kpc to 22 h\(^{-1}\) Mpc in all bins except the lowest redshift bin, where the scales are slightly reduced to \( \sim 50 \) h\(^{-1}\) kpc to 20 h\(^{-1}\) Mpc. [See the electronic edition of the Journal for a color version of this figure.]
and $\Gamma$. As such, we also now analyze our results in the context of a cosmological model with $(\Omega_m, \Omega_\Lambda, \sigma_8, \Gamma, h) = (0.28, 0.72, 0.8, 0.15, 0.7)$, motivated by recent supernovae, large-scale structure, and CMB measurements (e.g., Riess et al. 2004; Cole et al. 2005; Spergel et al. 2006). Note that equation (31) of Eisenstein & Hu (1998) suggests that $\Gamma = \alpha_h^2 \Omega_m h = 0.15$ is close to adopting a (realistic) baryon fraction of $\Omega_b/\Omega_m = 0.185$ (e.g., Cole et al. 2005).

In Table 1 we display our derived values for the mass of the dark matter halos hosting quasars. The values for $M_{DMH}$ in our highest (lowest) redshift bins have been calculated based on raising (lowering) $b_Q$ by 10% (see eq. [3]). A Spearman rank test shows no significant correlation between redshift and $M_{DMH}$. However, such a test is not compelling for only five redshift bins, particularly as we do not necessarily trust our $b_Q$ corrections at low redshift (as these corrections are dependent on the less precise values of $b_Q$ derived in our highest redshift bin). We will therefore assume, as detected in C05, that $M_{DMH}$ is constant with redshift. Across all our individual redshift bins we obtain a weighted mean of $M_{DMH} = (4.8 \pm 0.5) \times 10^{12} h^{-1} M_\odot$. If we instead only consider our “best” bins, in the range $1.0 < z < 2.1$, we obtain $M_{DMH} = (5.2 \pm 0.6) \times 10^{12} h^{-1} M_\odot$. This value of $M_{DMH}$ deviates $\sim 1.3 \sigma$ from the value of $M_{DMH} = (3.0 \pm 1.6) \times 10^{12} h^{-1} M_\odot$ obtained by C05 using a similar cosmology, and is slightly below the determination of $M_{DMH} \sim 10^{13} M_\odot$ from PMN04 (see also Porciani & Norberg 2006).

### 4.2. Luminosity-dependent Quasar Bias

Although in § 3.3 we detected no significant luminosity dependence to quasar bias in any redshift bin, our detection of $1.5 \sigma$ in our highest redshift bin ($z = 1.92$) is close to being significant. We can estimate the factor by which our data sample would have to increase in size before luminosity-dependent bias could be detected using our methodology. Assuming that the noise reduction scales as the square root of the sample size, a 2 $\sigma$ detection in our $z = 1.92$ bin would require a sample 3.8 times larger (including stellar contamination) or 1.8 times larger (ignoring stellar contamination). Similarly, a 3 $\sigma$ detection would require a sample 8.6 or 4.0 times larger, respectively.

A sample size twice as large as that used in this paper should be achievable in the near future. The necessary sample size to detect luminosity-dependent biasing will be further reduced by improved photometric techniques. For example, although we attempt to restrict the range of photometric redshift over which we analyze bivariate quasar clustering to reduce the effect of catastrophic failures on our $b_Q$ estimates, some quasars in our luminosity analysis will still be placed in entirely the wrong bin of $M_g$, diluting the significance of any comparisons we make between $M_g$ bins. Finally, we note that it is highly unlikely that luminosity-dependent quasar bias can ever be detected, via our angular analysis, to magnitudes of $g < 21$ at redshifts $z < 1.6$. However, our $1 \sigma$ errors suggest that at $z < 1.6$ quasar bias changes

---

**TABLE 2**

| $z_{ph}$ | $\tilde{M}_q$ | **ONE-PARAMETER MODEL, $b_Q$** | **TWO-PARAMETER MODEL**$^a$ |
|----------|---------------|-------------------------------|-----------------------------|
| $0.4 \leq z_{ph} < 1.2$: | | | |
| 0.85 | $-23.99$ | $2.03 \pm 0.12$ | $0.877 \pm 0.102$ \quad 1.39 \pm 0.66$ |
| 0.85 | $-22.82$ | $2.00 \pm 0.13$ | $1.00 \pm 0.00$ \quad 2.00 \pm 0.55$ |
| 0.85 | $-21.69$ | $2.08 \pm 0.14$ | $0.999 \pm 0.001$ \quad 2.08 \pm 0.54$ |
| $1.2 \leq z_{ph} < 1.65$: | | | |
| 1.44 | $-24.84$ | $2.86 \pm 0.29$ | $1.00 \pm 0.00$ \quad 2.86 \pm 0.53$ |
| 1.44 | $-23.81$ | $2.62 \pm 0.30$ | $0.92 \pm 0.07$ \quad 2.73 \pm 0.67$ |
| 1.44 | $-23.25$ | $2.68 \pm 0.30$ | $0.878 \pm 0.064$ \quad 2.87 \pm 0.69$ |
| $1.65 \leq z_{ph} < 2.3$: | | | |
| 1.92 | $-25.42$ | $2.77 \pm 0.29$ | $0.865 \pm 0.019$ \quad 2.87 \pm 0.61$ |
| 1.92 | $-24.46$ | $2.25 \pm 0.45$ | $0.947 \pm 0.053$ \quad 2.36 \pm 0.76$ |
| 1.92 | $-23.92$ | $3.18 \pm 0.51$ | $0.869 \pm 0.054$ \quad 3.46 \pm 0.68$ |

$^a$ With stellar contamination $1 - a$. 

---

**FIG. 6.**—Our quasar bias estimates as a bivariate function of $g$-band absolute magnitude $M_g$ and photometric redshift $z_{ph}$. The mean photometric redshift that corresponds to each shape is labeled in the plot. The open shapes show our best-fit model when only the quasar bias $b_Q$ is varied. The solid shapes show estimates when a second parameter, the stellar contamination, is also allowed to vary (see Table 2 for model values). [See the electronic edition of the Journal for a color version of this figure.]
with luminosity by less than ±0.6 (±0.3 if we ignore stellar contamination).

5. SUMMARY AND FUTURE PROSPECTS

In this paper, we used a sample of ~300,000 photometrically classified quasars drawn from SDSS DR4 to study the evolution of quasar clustering. Our main results are as follows:

1. Over scales of 0.16′ to 100′ (~55 h⁻¹ kpc to 35 h⁻¹ Mpc) at our sample’s mean redshift of z ~ 1.4) quasar clustering is well described by Smith et al. (2003) fits to dark matter clustering in simulations, a ΛCDM cosmology, and a single quasar bias parameter. Quasar biasing appears to be scale-independent over this range.

2. For σ₈ = 0.9 and Γ = 0.21 the required quasar bias is b₁₄₀มวล = 2.41 ± 0.09, rising to b₁₄₀มวล = 2.57 ± 0.10 for σ₈ = 0.8 and Γ = 0.15.

3. Our sample alone is sufficient to rule out a constant quasar bias over 0.75 < z < 2.28 (for scales of ~55 h⁻¹ kpc to 22 h⁻¹ Mpc) at, conservatively, >99.6%. This significance rises to >99.9999% if stellar contamination is not explicitly fit, which is likely closer to the true significance of our detection (see § 3.1). At z ~ 2.3 we find b₀ ~ 3. Considering complementary independent constraints on redshift evolution of 99.8% (C05) and 3.6 σ (PMN04) from the 2QZ, it is certain that quasar bias is therefore evolving with cosmic time.

4. Using our best photometric redshift ranges, σ₈ = 0.8, and Γ = 0.15, we find a mean mass for the dark matter halos hosting UVX quasars of MDHM = (5.2 ± 0.6) × 10¹² h⁻¹ Mₜ⊙, approximately halfway between the values of MDHM determined from the 2QZ by PMN04 and C05.

5. We find no significant luminosity dependence to quasar clustering, but our analysis hints at a small dependence (1.5 σ) at high redshift (z = 1.92). We note that Porciani & Norberg (2006) found a similar slight dependence in a recent luminosity-segregated analysis of quasar clustering. This suggests that with improved photometric classification efficiency, a sample size of as little as 1.8 times larger (~550,000 objects) may be sufficient to detect luminosity dependence in quasar clustering at z ~ 2. This might distinguish “light bulb” accretion (where quasars are either “off” or accrete at one efficiency; see, e.g., Valageas et al. 2001) from models that allow a range of accretion efficiencies (e.g., Lidz et al. 2006).

6. Our work agrees excellently with local results from Serber et al. (2006), who studied the environments of quasars at z < 0.4 in DR3 (see their Fig. 2). We concur that there is little luminosity dependence to quasar clustering on proper scales of ~50 h⁻¹ kpc (their 100 h⁻¹ kpc comoving), and also that any weak luminosity trend is only expressed for brighter quasars (M₁₄₀ ≤ ~24; note, for comparison with Serber et al. [2006], that g − i ~ 0.3 for UVX quasars).

7. It is highly unlikely that our technique will constrain any luminosity dependence to quasar clustering to magnitudes of g < 21 at redshifts z < 1.6. The errors on our measurements suggest that quasar bias b₀ is constant at z < 1.6 to ±0.6.

Although our analyses in § 3.3 uncovered no significant pattern, they were close to favoring particular trends. In the near future, the prospects for reassessing these results are excellent. Analysis of the angular clustering of photometrically classified quasars will improve not only as photometric surveys widen and deepen, increasing total numbers of objects, but also as classification efficiency improves and is expanded to non-UVX active galactic nuclei and to higher redshifts. Thus, we expect the statistical power of clustering analyses of photometrically classified quasars to rapidly improve. Further, estimates of the evolution of quasar clustering will be enhanced as quasar photometric redshift estimates tighten and catastrophic estimates diminish (e.g., Ball et al. 2007). In combination, we expect these factors to eventually allow significant constraints on the luminosity evolution of quasar clustering, particularly at z ~ 1.6.

A. D. M. and R. J. B. wish to acknowledge support from NASA through grants NAG5-12578 and NNG06GH15G. G. T. R. was supported in part by a Gordon and Betty Moore Fellowship in Data Intensive Science. R. C. N. acknowledges the EU Marie Curie Excellence Chair for support during this work. D. P. S. acknowledges support through NSF grant AST 06-07634. The authors made extensive use of the storage and computing facilities at the National Center for Supercomputing Applications and thank the technical staff for their assistance in enabling this work.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web site is http://www.sdss.org.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, Cambridge University, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max Planck Institute for Astronomy (MPIA), the Max Planck Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

APPENDIX A

BACKGROUND METHODOLOGY

A1. CORRELATION FUNCTION AND MODEL FITTING

We estimate the two-point angular correlation function (ω) via (Landy & Szalay 1993)

$$\omega(\theta) = \frac{Q\langle(\theta) - 2Q\langle R(\theta)\rangle}{R\langle R\rangle} + 1$$  \hspace{1cm} (A1)
from counts of quasar-quasar (QQ), quasar-random (QR), and random-random (RR) pairs. We use a random catalog 100 times larger than the data catalog. The random catalog is constructed using masks from the SDSS DR4 Catalog Archive Server and cutting both the KDE data and the random catalog to the SDSS DR4 theoretical footprint (which discards ~1.7% of the data). Our approach is detailed in M06, where we also discuss how points in our random catalog are assigned values of seeing and Galactic absorption.

We estimate errors and covariance matrices using inverse-variance-weighted jackknife resampling (Scranton et al. 2002; Zehavi et al. 2002; Myers et al. 2005). The jackknife method is to divide the data into $N$ pixels, then create $N$ subsamples by neglecting each pixel in turn. Note that if we considered the contribution of each pixel, rather than neglecting its contribution, this would be an inverse-variance-weighted pixel-to-pixel (also called field-to-field or subsampled) error estimate (e.g., Myers et al. 2003). We will generally refer to the chosen length for each side of the pixels as the jackknife resolution. If we denote subsamples by the subscript $L$ and recalculate $\omega_L$ in each jackknife realization via equation (A1), then the inverse-variance-weighted covariance matrix ($C_{ij}$) can be generated as

$$C_{ij} = C(\theta_i, \theta_j) = \sum_{L=1}^{N} \frac{[RR_L(\theta_i) - \omega(\theta_i)][RR_L(\theta_j) - \omega(\theta_j)]}{[RR_L(\theta_i)][RR_L(\theta_j)]} \left[ \frac{\omega_L(\theta_i) - \omega(\theta_i)}{\omega_L(\theta_j) - \omega(\theta_j)} \right],$$

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

$$|C| = \frac{C_{ij}}{\sigma_i \sigma_j}.$$  

The $RR_L/RR$ terms in equation (A2) (Myers et al. 2005) weight by the different numbers of objects expected, due to holes, poor seeing, pixels that extend beyond the survey boundary, etc. We then estimate $\chi^2$ fits to model angular autocorrelation functions ($\omega_m$) using the inverse of the covariance matrix, and determine errors on fits from $\Delta \chi^2$,

$$\chi^2 = \sum_{ij} \left[ \omega(\theta_i) - \omega_m(\theta_i) \right] C_{ij}^{-1} \left[ \omega(\theta_j) - \omega_m(\theta_j) \right].$$

The simplest models we fit are power laws of the form $\omega_m(\theta) = A \theta^{-c}$, where the units of $A$ (as we fit it) are arcmin$^6$. In general we fit the more physical models discussed in § A2.

A2. MODELING PROJECTED QUASAR CLUSTERING

Since the seminal scaling relations of Hamilton et al. (1991), many authors (e.g., Peacock & Dodds 1994, 1996; Jain et al. 1995) have worked to obtain precise analytical descriptions of the clustering of dark matter particles. Smith et al. (2003) married traditional approaches with a simple halo model (e.g., Cooray & Sheth 2002) to obtain fitting formulae that better approximate the nonlinear clustering behavior of simulated CDM. The formulae of Smith et al. (2003) reproduce clustering in dark matter simulations to better than 3% at (redshift) $z < 3$ for (wavenumber) $k < 10 h$ Mpc$^{-1}$.

The models of Smith et al. (2003) directly predict the nonlinear, dimensionless power spectrum of dark matter $\Delta^2_{NL}(k, z)$ for a wide range of CDM cosmologies. The clustering of objects that formed in the rare peaks of a Gaussian random field is expected to be biased relative to underlying dark matter (e.g., Kaiser 1984; Bardeen et al. 1986), in a manner that could depend on both scale and formation history. Thus the clustering of quasars relative to dark matter might be modeled as $\Delta^2_{Q}(k, z) = b^2_Q(k, z) \Delta^2_{NL}(k, z)$, where $b_Q$ is the quasar bias.

In this paper, we measure angular autocorrelations. For small angles ($\theta \ll 1$ rad), Limber’s equation can be used to project the power spectrum into the angular autocorrelation (Limber 1953; Peebles 1980; Peacock 1991; Baugh & Efstathiou 1993) via

$$\omega(\theta) = \pi \int_{z=0}^{z=\infty} \int_{k=0}^{k=\infty} \frac{\Delta^2(k, \theta)}{k} J_0(k\theta\chi(z)) \left( \frac{dn}{dz} \right)^2 \left( \frac{dz}{d\chi} \right) F(\chi) \frac{dk}{k} dz,$$

where $J_0$ is the zeroth-order Bessel function of the first kind, $\chi$ is the radial comoving distance, $dn/dz$ is the redshift selection function [normalized so that $\int_0^\infty (dn/dz) dz = 1$], and $dz/d\chi = H_0/c = H_0 [\Omega_m (1 + z)^3 + \Omega_L]^{1/2}/c$. Strictly, $\chi$ should be the angular, or transverse, comoving distance; however, in a flat cosmology radial and transverse comoving distances are equivalent and the curvature term vanishes—$F(\chi) = 1$. We ultimately model the angular quasar autocorrelation function as

$$\omega_{QQ}(\theta) = \frac{H_0 \pi}{c} \int_0^\infty \int_0^\infty b^2_Q(k, z) \Delta^2_{NL}(k, z) J_0(k\theta\chi(z)) \left( \frac{dn}{dz} \right)^2 \sqrt{\Omega_m (1 + z)^3 + \Omega_L} \frac{dk}{k} dz.$$

In theory, with sufficient data $b_Q(k, z)$ may be directly constrained by $\omega_{QQ}$, although we will generally set $b_Q$ to be a constant and constrain it by comparing the amplitudes of $\omega_{QQ}$ and the projected matter power spectrum as a function of redshift and scale.
In general, $\Delta^2_{\text{NL}}$ is not separable into individual functions of $k$ and $z$. We therefore Monte Carlo integrate under the surface described by the integrand in equation (A6) until the integration is evaluated to better than 1%. To optimize this process, two points are worth noting. First, the change of variables $dk/k = \ln(10) d \log(k)$ allows more uniform sampling in $k$-space. Second, although $\Delta^2_{\text{NL}}$ is not easily separated, we have determined that the “parameters of the spectrum” ($k_0^{-1}, n_{\text{eff}},$ and $C$; see Appendix C of Smith et al. 2003) can be approximated by splines to $\lesssim 0.3\%$ for at least $z < 4$ in a $\Lambda$CDM cosmology. Figure 7 demonstrates a typical surface we might integrate under to obtain $\omega_{\theta \theta}$. On scales $\lesssim 30'$ such surfaces are fairly smooth and the Monte Carlo integrations rapidly converge. Our integrations remain tractable at the 1% level out to the largest scales we model ($\leq 100'$).

APPENDIX B

POTENTIAL SYSTEMATICS

B1. STELLAR CONTAMINATION

We have addressed sources of systematic error in some depth in M06. In particular, we noted that the autocorrelation of KDE objects, $\omega$, combines clustering signals from a stellar component $\omega_{\text{SS}}$ and the true quasar autocorrelation $\omega_{\theta \theta}$. If $a$ is the efficiency, the fraction of genuine quasars that are classified as such by the KDE technique, then

$$
\omega(\theta) = a^2 \omega_{\theta \theta}(\theta) + (1 - a)^2 \omega_{\text{SS}}(\theta) + \epsilon(\theta),
$$

where $\epsilon$ is a tiny (theoretically zero) offset arising from QS, OR, and SR cross terms.

If the efficiency of the KDE technique is high (i.e., $a \rightarrow 1$), stellar contamination is only important as $\omega_{\theta \theta} \rightarrow 0$. In § 3.1 we fit the two-parameter model defined by equations (A6) and (B1) to the DR4 KDE autocorrelation and estimate the stellar contamination $(1 - a)$. In doing so, we derive $a = 0.956 \pm 0.044$, consistent with M06 (and with Richards et al. 2004), and find that our best analysis of the quasar bias is therefore at angles of $\theta \leq 1^\circ$. 

FIG. 7.—One example of a surface that we Monte Carlo integrate under to project the matter power spectrum into the angular autocorrelation function. The surface is plotted for $\theta = 1'$ and $\delta^2_{\text{NL}} = (\Delta^2_{\text{NL}}/k^2) J_0(kr(z)) [dN/dz]^2 dz/dy$ (see eqs. [A5] and [A6]). We obtain $dN/dz$ from spectroscopic matches (with DR1QSO, DR2, or the 2QZ) to our photometrically classified sample. Two individual contributions to power are apparent at $1'$: the nonlinear matter spectrum at low $k$ and a halo term at high $k$ (see Smith et al. 2003). [See the electronic edition of the Journal for a color version of this figure.]
B2. MISCLASSIFIED H II REGIONS

In Paper II we discuss nonstellar misclassified objects in the KDE catalog (generally H II regions in various galaxies). The fraction of such objects in the KDE sample is too small to affect clustering measurements, except on small scales where H II regions can mimic quasar pairs, and is thus generally negligible on our scales of interest. For instance, by visually inspecting pairs of KDE objects, we estimate that on scales of $12000$ (the smallest scale fit in this paper), the effect amounts to $1.0 \sigma$; by scales of $30000$, the effect is $<0.4 \sigma$; and on larger scales, where we exceed the observed angular size of most galaxies, the effect vanishes. We are engaged in determining the regions that need to be masked from KDE clustering analyses because of this small effect, but (unlike in Paper II, where our focus is small scales) we make no attempt to correct for the effect in this paper.

B3. COVARIANCE AND THE JACKKNIFE RESOLUTION

If the covariance between scales is perfectly accounted for, equation (A4) should return identical estimates of $\chi^2$ irrespective of the jackknife resolution. Given that DR4 covers close to $7000$ deg$^2$, we have adequate area with which to determine whether the jackknife resolution affects significance estimates. In Figures 8 and 9 we plot jackknife errors and covariance matrices for identical sets of data, obtained at different jackknife resolutions.

Figure 8 clearly shows that error estimates are influenced by the choice of jackknife resolution and inflate on scales similar to the resolution. However, we might expect that error discrepancies will be offset by differences in the covariance matrices plotted in Figure 9. To test this we assume a model that at every scale is equal to $\sigma_2$ (the tip of the 1 $\sigma$ upper error bar), as obtained when jackknife resampling at $10'$, and calculate $\chi^2$ for other jackknife resolutions (over the 0.1 to 100 scales we study in this work). We find that jackknife resolutions of greater than a few degrees all return highly consistent $\chi^2$ estimates (within $2\%$) but that jackknife resolutions that lie increasingly within our scales of interest give increasingly inflated estimates of the significance relative to our $10'$ control.

In this paper we use a jackknife resolution of $10'$ for several reasons. The covariance matrix is well mixed at this resolution, so that either neglecting or incorporating covariance gives similar $\chi^2$ estimates (to within $5\%$). Further, at a resolution of $10'$ pixel-to-pixel errors can be calculated (see § A1) for our scales of interest, and we find these agree with the jackknife estimates. Note that generalizing the ideal jackknife resolution for a given analysis is not our focus in this work; we simply suggest that when jackknifing errors, different jackknife resolutions should be tested, particularly when the resolution is similar to the scales being probed. It is tempting to ask, however, “if pixel-to-pixel errors are always well defined for resolutions that consistently recover significance estimates, why use the jackknife at all?”

B4. ADDITIONAL OBSERVATIONAL SYSTEMATICS

M06 discussed how observations that were made in poor seeing conditions or that trace absorption by dust in our Galaxy could contaminate the true quasar clustering signal. Using a similar analysis to M06, with the narrower binning allowed by the larger DR4 data set, we find (1) no clear pattern imposed by seeing variations, and (2) that an $A_g < 0.21$ cut on our DR4 KDE sample (and random catalog) removes clustering imprints caused by Galactic dust. In agreement with M06, we find a clear clustering pattern in the KDE sample for $A_g \geq 0.22$. We note that Yahata et al. (2006) found little change in the number density of KDE objects as a function of...
Fig. 9.—Normalized covariance matrix (see eq. [A3]) at different jackknife resolutions. The covariance grows rapidly on scales larger than the jackknife resolution. On scales far smaller than the jackknife resolution, off-diagonal elements are well mixed but still contribute to significance estimates. From top left to bottom right the panels represent jackknife resolutions of $0.3^\circ$, $1^\circ$, $3^\circ$, and $10^\circ$. The data used to estimate the covariance matrices were the DR4 KDE sample discussed in §2, with a nominal $A_g < 0.24$ Galactic absorption cut. [See the electronic edition of the Journal for a color version of this figure.]

 Galactic absorption, perhaps because their surface density analysis is insensitive to the relative numbers of stars and quasars assigned by the KDE technique (see expanded discussion in M06). We have checked that a range of cuts around $A_g > 0.21$ (in particular, our adopted cut of $A_g < 0.18$ from M06) all yield statistically similar estimates of $\omega$. For our main analyses in this paper, we adopt no seeing cut and an $A_g < 0.21$ cut, which discards $\sim 12.5\%$ of our sampled area.

REFERENCES

Abazajian, K., et al. 2003, AJ, 126, 2081 (DR1)
———. 2004, AJ, 128, 502 (DR2)
———. 2005, AJ, 129, 1755 (DR3)
Ball, N. M., et al. 2007, ApJ submitted (astro-ph/0612471)
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Barnes, J. E., & Hernquist, L. 1992, ARA&A, 30, 705
Baugh, C. M., & Efstathiou, G. 1993, MNRAS, 265, 145
Carlberg, R. G. 1990, ApJ, 350, 502
Carroll, S. M., Press, W. H., & Turner, E. L. 1992, ARA&A, 30, 499
Ciotti, L., & Ostriker, J. P. 1997, ApJ, 487, L105
———. 2001, ApJ, 551, 131
Cole, S., et al. 2005, MNRAS, 362, 505
Cooray, A., & Sheth, R. K. 2002, Phys. Rep., 372, 1
Croom, S. M., Smith, R. J., Boyle, B. J., Shanks, T., Miller, L., Outram, P. J., & Loaring, N. S. 2004, MNRAS, 349, 1397 (2QZ)
Croom, S. M., et al. 2005, MNRAS, 356, 415 (C05)
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1P
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Fabian, A. C. 1999, MNRAS, 308, L39
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Franceschini, A., Hasinger, G., Miyaji, T., & Malquori, D. 1999, MNRAS, 310, L5
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Gebhardt, K., et al. 2000, ApJ, 539, L13
Giangiantonio, T., et al. 2006, Phys. Rev. D, 74, 3520
Graham, A. W., Erwin, P., Caon, N., & Trujillo, I. 2002, in ASP Conf. Ser. 275, Disks of Galaxies: Kinematics, Dynamics, and Perturbations, ed. E. Athanassoula, A. Bosma, & R. Mujica (San Francisco: ASP), 87
Grazian, A., Negrello, M., Moscardini, L., Cristiani, S., Haehnelt, M. G., Matarrese, S., Omizzolo, A., & Vanzella, E. 2004, AJ, 127, 592
Groth, E. J., & Peebles, P. J. E. 1977, ApJ, 217, 385
Gunn, J. E., et al. 1998, AJ, 116, 3040
———. 2006, AJ, 131, 2332
Haiman, Z., & Hui, L. 2001, ApJ, 547, 27
Hamilton, A. J. S., Kumar, P., Li, E., & Matthews, A. 1991, ApJ, 374, L1
Heckman, T. M., Smith, E. P., Baum, S. A., van Breugel, W. J. M., Miley, G. K., Illingworth, G. D., Bothun, G. D., & Balick, B. 1986, ApJ, 311, 526
Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129
Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2005a, ApJ, 630, 716
———. 2006, ApJS, 163, 1
Hopkins, P. F., Hernquist, L., Martini, P., Cox, T. J., Robertson, B., Di Matteo, T., & Springel, V. 2005b, ApJ, 625, L71

No. 1, 2007 CLUSTERING OF 300,000 PHOTOCLASSIFIED QSOs. I. 97
