ApJ, in prep February 5, 2008
Preprint typeset using LaTeX style emulateapj v. 12/01/06

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. Nichol[3], Gordon T. Richards[4,5], Donald P. Schneider[6] and Neta A. Bahcall[7]

ApJ, in prep February 5, 2008

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 $\bar{z}$ ∼0.75 to $\bar{z}$ ∼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 $b_Q^{\bar{z}=1.40} = 2.41 \pm 0.08$ ($P_{<\chi^2} = 0.847$) for $\sigma_8 = 0.9$. This represents a better fit than the best-fit power-law model ($\omega = 0.0493 \pm 0.0064\theta^{-0.928\pm0.055}$; $P_{<\chi^2} = 0.482$). We find b_Q increases with redshift. This evolution is significant at $> 99.6\%$ using our data set alone, increasing to $> 99.9999\%$ if stellar contamination is not explicitly parameterized. We measure the quasar classification efficiency across our full sample as $a = 95.6\pm^{4.4}_{1.9}\%$, a star-quasar separation comparable with the star-galaxy separation in many photometric studies of galaxy clustering. We derive the mean mass of the dark matter halos hosting quasars as $M_{DMH} = 5.2 \pm 0.6 \times 10^{12}$ h$^{-1} M_\odot$. At $\bar{z} \sim 1.9$ we find a 1.5σ 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 $z \sim 2$. Our results agree with recent studies of quasar environments at $z < 0.4$, which detected little luminosity dependence to quasar clustering on proper scales $\gtrsim$ 50 h^{-1} kpc. At $z < 1.6$, our analysis suggests that b_Q is constant with luminosity to within $\Delta b_Q \sim 0.6$, and that, for $g < 21$, angular quasar autocorrelation measurements are unlikely to have sufficient statistical power at $z \lesssim 1.6$ to detect any luminosity dependence in quasars' clustering.

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

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 correlate 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 suggests an interplay between nascent galaxies and quasar activity. A symbiotic view of galaxy formation has emerged, in which galaxy mergers drive the formation of quasars, and supermassive black holes, which in turn seed new galaxies (e.g., Heckman et al. 1986; Carlberg 1990; Barnes & Hernquist 1992; Lacey & Cole 1993; Di Matteo, Springel & Hernquist 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 $z \lesssim 2.5$ quasar bias evolves with redshift but that quasars inhabit dark matter halos of similar average mass at every redshift (Porciani, Magliocchetti & Norberg 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 1982; Barnes & Hernquist 1992; Franchesini et al. 1999), cooling flows (e.g.,

Electronic address: admyers@astro.uiuc.edu
[1] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801
[2] National Center for Supercomputing Applications, Champaign, IL 61820
[3] ICG, Mercantile House, Hampshire Terrace, University of Portsmouth, Portsmouth, P01 2EG, UK
[4] Department of Physics & Astronomy, The Johns Hopkins University, 3400 N Charles St, Baltimore, MD 21218
[5] Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104
[6] Department of Astronomy and Astrophysics, 525 Davey Laboratory, Pennsylvania State University, University Park, PA 16802
[7] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544

Ciotti & Ostriker 1997, 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; Sazanov 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. (2005b) 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; henceforth Cro05), 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., Cro05; 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 complimentary to analyses using spectroscopic data because of the larger numbers of galaxies that could be photometrically selected. Similar angular clustering analyses of complimentarily 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., 90-95% efficient in the APM survey; Maddox et al. 1990b) but star-quasar separation lagged behind (e.g.,∼60% efficient in the 2dF QSO Redshift Survey; Croom et al. 2004; henceforth 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; henceforth Mye06), 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; henceforth Paper2), we analyze quasar clustering on smaller scales.

Extensive details of our techniques, modeling, and systematics are presented in Appendixes A and B, allowing our main analysis to be presented in Section 3, after detailing our data sample in Section 2. Our main results are ordered in a concluding section (Section 5). Unless otherwise specified, we assume a ΛCDM cosmology with $(\Omega_m, \Omega_\Lambda, \sigma_8, \Gamma, h \equiv H_0/100\mathrm{km\ s^{-1}\ Mpc^{-1}}) = (0.3, 0.7, 0.9, 0.21, 0.7)$, where Γ is the shape of the matter power spectrum ($\Gamma = \Omega_m h$ for baryon-free CDM). We correct all magnitudes for Galactic extinction using the dust maps of Schlegel, Finkbeiner & Davis (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, Gunn & Szalay 1999; Hogg et al. 2001; Smith et al. 2002; Ivezic 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) $g \geq 14.5$ and (dereddened) $g < 21$. Separations, in 4-dimensional color-space, from a sample of ∼10% of point sources in SDSS Data Release 1 (DR1; Abazajian et al. 2003) and from the quasar sample of Schneider et al. (2003; henceforth 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 $z \lesssim 2.5$ quasars from stars with ∼95% efficiency (Richards et al. 2004; see also Mye06). Applying the KDE technique to SDSS DR4, results in our DR4 KDE sample of 344,431 objects[8], a sample 3.5× larger than the DR1 sample used in Mye06. 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 ($\mathrm{d}N/\mathrm{d}z$ in Equation B2). For consistency with Mye06 (see their Figure 6), we estimate $\mathrm{d}N/\mathrm{d}z$ from spectroscopic matches to DR1QSO, matches to spectra taken

[8] Available at http://sdss.ncsa.uiuc.edu/qso/nbckde

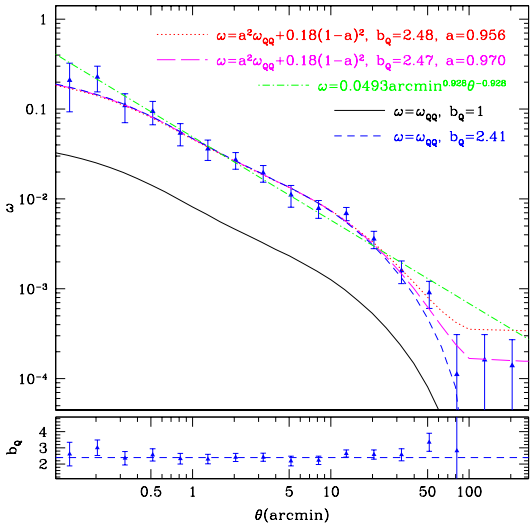

FIG. 1.— The autocorrelation of all 299,276 ($A_g < 0.21$) DR4 KDE objects in our working area. Errors are jackknifed with a resolution of 10°. The short-dashed line is the best fitting bias model (Equation B2) over the range 0.16′ to 63′ (∼55 h^{-1} kpc to 22 h^{-1} Mpc at the DR4 KDE sample's mean redshift of $z \sim 1.4$). The dotted line is the best fitting bias model (over the same range) when fitting for stellar contamination in the KDE sample (Equation C1 with $\omega_{SS} = 0.18$). The long-dashed line is the best fitting bias model with stellar contamination when the fitting range is extended out to 100′ (∼35 h^{-1} Mpc). The solid line shows the expected clustering for a linear bias model, which is easily rejected at extremely high significance. The lower panel shows the measured quasar bias relative to the linear bias model. The dot-dashed line is the best-fit power-law over 0.16′ to 63′. We estimate $\mathrm{d}N/\mathrm{d}z$ for the redshift distribution of the full DR4 KDE sample using a simple spline fit.

from the SDSS second data release (Abazajian et al. 2004; henceforth DR2) or matches to the 2QZ. As our KDE technique is currently trained on DR1QSO, obtaining $\mathrm{d}N/\mathrm{d}z$ from DR1QSO is arguably a fairer approach than using quasars from later data releases. Mye06 demonstrated that including or excluding matches with the 2QZ or DR2 has little affect on the form of $\mathrm{d}N/\mathrm{d}z$, and further showed that the methodology of using spectroscopic matches is broadly consistent with estimating $\mathrm{d}N/\mathrm{d}z$ from photometric redshifts (Weinstein et al. 2004).

3. QUASAR CLUSTERING RESULTS

3.1. *Mean Quasar Bias at $z \sim 1.4$*

In Figure 1 we show the ($A_g < 0.21$) DR4 KDE autocorrelation. We fit bias models (see Equation B2) 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 $\bar{z} = 1.4$). The fit's upper scale limit is nominally set by stellar contamination (see Appendix C) and the lower limit is set by the dark matter model we use for ω_{QQ}. Smi03 note that their models accurately reproduce Δ^2_{NL} to the limits of current simulations (∼3%) at wavenumbers of $k < 10\ h$ Mpc^{-1} ($\gtrsim 1'$); however, their models appear quite accurate even on scales several times smaller than this (see, e.g., Figure 15 of Smi03) 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 Smi03 at $k < 10\ h$ Mpc^{-1} should be easy to compare to Smi03, and thus to our results.

Our best fit bias model to the DR4 KDE autocorrelation has $b_Q^{\bar{z}=1.40} = 2.41 \pm 0.08$ ($P_{<\chi^2} = 0.847$), in good agreement with Mye06 ($b_Q^{\bar{z}=1.40} = 2.51 \pm 0.46$) but with considerably more precision. This is preferred over the best-fit power-law model ($\omega = 0.0493 \pm 0.0064\theta^{-0.928\pm0.055}$; $P_{<\chi^2} = 0.482$). A linear bias model (see Figure 1), is ruled out at an extremely high level of significance ($P_{<\chi^2} \sim 10^{-23}$). Our measured b_Q is in reasonable agreement with the value of $b_Q^{\bar{z}=1.47} = 2.42\pm^{0.20}_{0.21}$ obtained by Porciani, Magliocchetti & Norberg (2004; henceforth PMN04) in a clustering analysis of 14,000 $M_{b_J} < -22.5$ 2QZ quasars, even though PMN04 use $\sigma_8 = 0.8$, which will inflate their result by ∼13% as compared to our use of $\sigma_8 = 0.9$. Our result is slightly at odds with the value of $b_Q^{\bar{z}=1.35} = 2.02 \pm 0.07$ found by Cro05 for the full 2QZ sample after correcting their result for redshift-space distortions. It is, however, well within the error bars of their empirical fit to the evolution of quasar clustering of $b_Q = (0.53 \pm 0.19) + (0.289 \pm 0.035)(1+z)^2$. The form of this empirical fit broadly implies that our result (at $\bar{z} = 1.40$) should be ∼ 5% lower than an estimate at $\bar{z} = 1.47$ and ∼ 3% higher than at $\bar{z} = 1.35$.

To test stellar contamination effects, we also fit a model of the form $a^2\omega_{QQ} + (1-a)^2\omega_{SS}$, where a is the KDE efficiency and $\omega_{SS} = 0.18$ (as derived in Mye06) and find $b_Q = 2.48 \pm 0.15$ and $a = 0.956\pm^{0.044}_{0.019}$ ($P_{<\chi^2} = 0.920$). As we previously argued in Mye06, stellar contamination causes measurable deviation from the true ω_{QQ} only on scales of a degree or more. An appropriate approach would be to determine a at $> 1°$ and input this value to fits on scales $< 1°$. If $a \gtrsim 0.9$, this largely becomes unnecessary, as introducing an efficiency term does not significantly alter measured values of b_Q at $< 1°$.

In Figure 1, we also fit a stellar contamination model to larger scales (where stellar contamination begins to dominate fits) as 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 $a = 0.970\pm^{0.030}_{0.017}$ ($P_{<\chi^2} = 0.908$). This value of a reproduces the data very well out to at least 7° (see Figure 1); however, although altering the scale of the fit affects estimates of a, b_Q is essentially unchanged, as stellar contamination only dominates on large scales. We further note that altering ω_{SS} to 0.25, as is appropriate at ∼30′ (see Mye06) barely changes our result, giving $b_Q = 2.47 \pm 0.14$ and $a = 0.974\pm^{0.026}_{0.015}$ ($P_{<\chi^2} = 0.936$); 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 a and b_Q. However, in most cases fitting for a 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 $\mathrm{d}N/\mathrm{d}z$ than

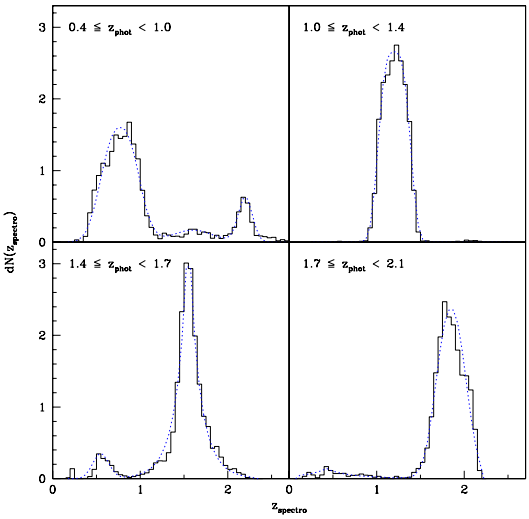

FIG. 2.— The spectroscopic redshift ($z_{spectro}$) distribution of KDE quasars for some photometric redshift bins (z_{phot}) used in our analyses. The dotted lines are functional fits of the form described by Equation 1. The plotted distributions have been normalized and the integral under the fitted functions is always within 0.1% of unity

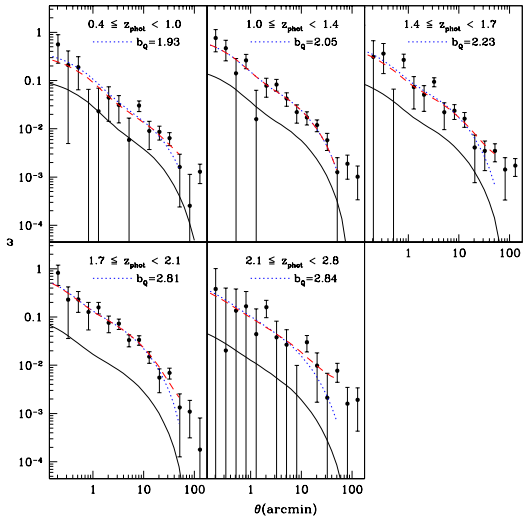

FIG. 3.— Quasar clustering evolution as a function of photometric redshift. There are ∼65,000 quasars in each bin except for the $2.1 \leq z_{phot} < 2.8$ bin, which contains ∼28,000 quasars. The solid line is the expected clustering of dark matter derived from Smi03. The dotted line is our best fit model where 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 ∼55 h^{-1} kpc to 22 h^{-1} Mpc in all bins except the lowest redshift bin, where the scales are slightly reduced to ∼50 h^{-1} kpc to 20 h^{-1} Mpc.

the simple spline fit used in Section 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

$$\mathrm{d}N = \sum_i \beta_i \exp \frac{-|z-\bar{z_i}|^{n_i}}{n_i \sigma_i^{n_i}} \mathrm{d}z \qquad (1)$$

where n (typically close to the Gaussian value of 2), $\sigma, \bar{z}$ and β are free parameters. These functions can excellently reproduce $\mathrm{d}N/\mathrm{d}z$ (see Figure 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 Mye06).

In Figure 3 we show the evolution of the angular quasar autocorrelation with photometric redshift. We derive estimates of the quasar bias using Equation B2, and also consider a two-parameter stellar-contamination model (Equation C1). In Figure 4 we display our measured quasar bias evolution. Our data alone (i.e., without assuming, $b_Q \sim 1$ at $z = 0$), rules 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 Section 3.1, not fitting for stellar contamination likely better estimates significances. Using 2QZ data, PMN04 and Cro05 have independently determined that b_Q evolves with redshift. We note that, after correcting for differing σ_8, our value of $b_Q^{\bar{z}_{phot}=1.87}$ disagrees with $b_Q^{z=1.89}$ from PMN04 but only at the 1.8σ level, dropping to 0.8σ if stellar contamination is incorporated. The $M_{b_J} < -22.5$ restriction adopted by PMN04 is immaterial in this context, as every $g < 21$ quasar at $z > 1.7$ should be brighter than $M_{b_J} = -22.5$. Bias determinations derived from cosmological models that are more like those used by

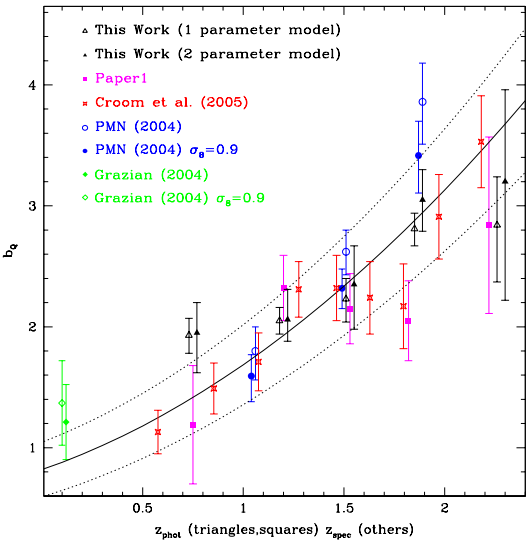

FIG. 4.— Our estimates of the evolution of quasar bias compared to other authors, and to our estimates from Mye06. The solid line plots the semi-empirical relationship $b_Q(z) = 0.53 + 0.289(1+z)^2$ derived by Cro05, and the dotted lines track the 1σ error. The open triangles show our best fit model where only the quasar bias, b_Q, is varied. The solid triangles show the results when the stellar contamination, is also allowed to vary (see Table 1 for model values). PMN04 and Grazian et al. (2004) use $\sigma_8 = 0.8$, so we also show the effect of projecting their results to match our chosen value of $\sigma_8 = 0.9$. In Section 3.1, we note that, due to the photometric nature of the redshifts we use, our points at the lowest and highest redshift are almost certainly *at least* 10% too high and 10% too low, respectively. Points at other redshifts are not at all biased by the use of photometric redshifts.

PMN04 and Cro05 are provided in Table 1 (see also the discussion in Section 4.1).

The photometric nature of our redshifts leads to problems for our $z_{phot} = 0.75$ and $z_{phot} = 2.28$ bins (i.e., our lowest and highest redshift bins); in particular, where to plot these bins on the z_{phot} axis of Figure 4. At $1 < z_{phot} < 2$, the mean redshift of our spectroscopic matches is close to $\bar{z}_{phot}$, 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 $z_{phot} < 1$ ($z_{phot} > 2$) there is a secondary solution, amounting to 12.6% (21.2%) of the area under dN/dz, at $z_{phot} > 2$ ($z_{phot} < 1$). We can examine clustering at these secondary z_{phot} solutions by adopting a similar approach to Equation C1:

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

Where, ω_1, ω_2 and ω_{1+2} are, respectively, the true clustering at the primary and secondary z_{phot} solutions, and the clustering we measure as a combination of the two z_{phot} solutions. B is the relative contribution of the primary and secondary z_{phot} solutions to dN/dz. We can then estimate the true value of the bias at the position of the primary solution in z_{phot} as

$$b_1^2 = \frac{\omega_1}{\omega_1^M} = \frac{b_{1+2}^2\omega_{1+2}^M - (1-B)^2 b_2^2 \omega_2^M}{B^2\omega_1^M} \qquad (3)$$

where the superscript M denotes the model values of ω (Equation B2 with $b_Q = 1$), and the b_i denote quasar bias. Although we don't 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 $\bar{z}_{phot} = 0.75$ ($\bar{z}_{phot} = 2.28$) 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_{phot}, rather than the (unknown) true value, to estimate the true value in the primary z_{phot} bin. If we lower our estimates of $b_Q^{\bar{z}_{phot}=0.75}$ by 10% and increase our estimates of $b_Q^{\bar{z}_{phot}=2.28}$ 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 $\bar{z}_{phot} = 1.20, 1.53, 1.87$ has no affect on b_Q values, as any secondary solutions in z_{phot} have a very small weight ($B \lesssim 0.05$).

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 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 Mye06, ideal tests of quasar evolution would attempt to break luminosity-redshift degeneracy and examine multivariate quasar properties. We now repeat our clustering analysis as a bivariate function of

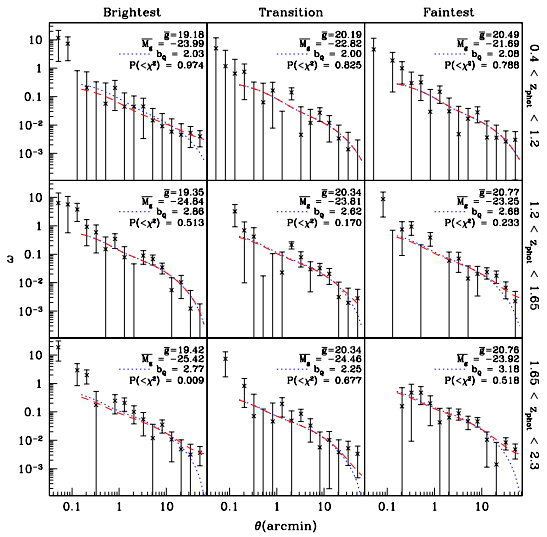

FIG. 5.— Quasar clustering evolution as a bivariate function of absolute magnitude, M_g, and photometric redshift, z_{phot}. The rows of the nine panels show three bins of roughly equal numbers in z_{phot}. The columns divide each z_{phot} bin into three of equal numbers in M_g. There are ∼30,000 quasars in each bin. Labeled in each panel are the mean g apparent and absolute magnitudes, the quasar bias, b_Q, derived as in Figure 3, and the χ^2 probability of our best fit model where only b_Q is fitted (the 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 over scales of 0.16′ to 63′ using a full covariance matrix. A scale of 0.16′ to 63′ is ∼55 h^{-1} kpc to 22 h^{-1} Mpc in all bins except the lowest redshift bin, where the scales are slightly reduced to ∼50 h^{-1} kpc to 20 h^{-1} Mpc.

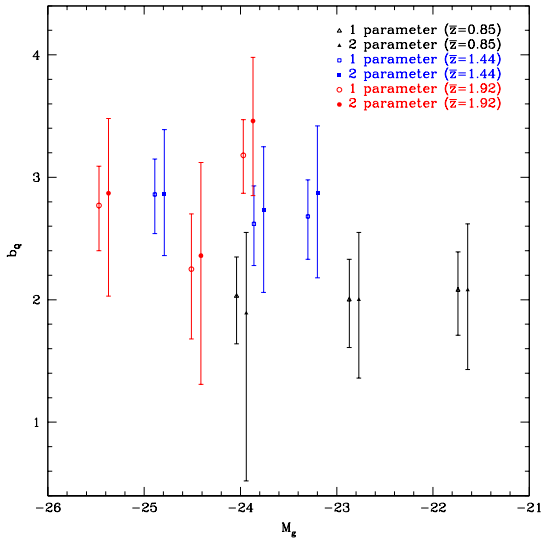

FIG. 6.— Our quasar bias estimates as a bivariate function of g-band absolute magnitude, M_g, and photometric redshift, z_{phot}. The mean photometric redshift that corresponds to each shape is labeled in the plot. The open shapes show our best fit model when only the quasar bias, b_Q, is varied. The solid shapes show estimates when a second parameter, the stellar contamination, is also allowed to vary (see Table 2 for model values).

TABLE 1
Estimates of the quasar bias, b_Q and the quasar host halo mass M_{DMH} as a function of photometric redshift z_{phot}

		$\Gamma = 0.21, \sigma_8 = 0.9$				$\Gamma = 0.15, \sigma_8 = 0.8$	
z_{phot}		1 parameter model		2 parameter model[a]		1 parameter model	
range	mean	b_Q	$M_{DMH}(\mathrm{h}^{-1}M_\odot)$	a	b_Q	b_Q	$M_{DMH}(\mathrm{h}^{-1}M_\odot)$
0.4, 2.3	1.40	$2.41\pm^{0.08}_{0.09}$	$12.1\pm^{1.3}_{1.4}\times10^{12}$	$0.956\pm^{0.044}_{0.019}$	$2.48\pm^{0.13}_{0.15}$	$2.57\pm^{0.10}_{0.09}$	$5.72\pm^{0.84}_{0.70}\times10^{12}$
0.4, 1.0[c]	0.75	$1.93\pm^{0.14}_{0.15}$	$8.90\pm^{2.71}_{2.46}\times10^{12}$	$0.903\pm^{0.097}_{0.041}$	$1.95\pm^{0.25}_{0.33}$	$2.06\pm^{0.14}_{0.16}$	$4.08\pm^{1.32}_{1.25}\times10^{12}$
1.0, 1.4	1.20	$2.05\pm^{0.11}_{0.11}$	$12.5\pm^{2.3}_{2.1}\times10^{12}$	$0.997\pm^{0.003}_{0.092}$	$2.06\pm^{0.25}_{0.18}$	$2.16\pm^{0.11}_{0.12}$	$5.70\pm^{1.15}_{1.11}\times10^{12}$
1.4, 1.7	1.53	$2.23\pm^{0.17}_{0.19}$	$9.22\pm^{2.49}_{2.40}\times10^{12}$	$0.884\pm^{0.057}_{0.036}$	$2.35\pm^{0.32}_{0.37}$	$2.39\pm^{0.18}_{0.20}$	$4.30\pm^{1.33}_{1.23}\times10^{12}$
1.7, 2.1	1.87	$2.81\pm^{0.13}_{0.14}$	$11.3\pm^{1.7}_{1.7}\times10^{12}$	$0.907\pm^{0.077}_{0.037}$	$3.05\pm^{0.25}_{0.26}$	$3.00\pm^{0.14}_{0.15}$	$5.32\pm^{0.94}_{0.90}\times10^{12}$
2.1, 2.8[c]	2.28	$2.84\pm^{0.40}_{0.47}$	$11.3\pm^{5.0}_{4.6}\times10^{12}$	$0.840\pm^{0.091}_{0.039}$	$3.20\pm^{0.76}_{0.98}$	$3.13\pm^{0.43}_{0.50}$	$5.95\pm^{2.92}_{2.56}\times10^{12}$

[a] With stellar contamination $1-a$.
[b] Due to catastrophic failures, values of b_Q in this row need lowered at least 10%. The derived M_{DMH} use the reduced values. See Equation 3 and the associated discussion.
[c] Due to catastrophic failures, values of b_Q in this row need raised at least 10%. The derived M_{DMH} use the increased values. See Equation 3 and the associated discussion.

redshift and luminosity. We derive g-band absolute magnitudes (M_g) for KDE objects by assuming that each photoz is a reasonable ensemble estimate of redshift. We incorporate the K-correction from Wisotzki (2000; see, e.g., Mye06). Consistent M_g binning at every redshift is impractical for quasars (which span ∼8 magnitudes in M_g), so, as in Mye06, 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 Mye06, but other than the general increase with redshift discussed in Section 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 instead take b_Q values from models that allow a stellar contamination component, a constant b_Q model is more acceptable ($b_Q = 2.55 \pm 0.21, P_{<\chi^2} = 0.662$).

To test for luminosity-dependent bias we combine measurements for each set of two magnitudes within each redshift bin plotted in Figure 6 into a single inverse-variance-weighted estimate and determine whether this estimate would be rejected by the measurement in the third magnitude bin. Quoting the maximum rejection in each case, we find: (1) Incorporating stellar contamination, a maximum rejection of 0.2σ for $\bar{z} = 0.85$, 0.2σ for $\bar{z} = 1.44$ and 1.0σ for $\bar{z} = 1.92$, and; (2) ignoring stellar contamination, a maximum rejection of 0.1σ for $\bar{z} = 0.85$, 0.5σ for $\bar{z} = 1.44$ and 1.5σ for $\bar{z} = 1.92$.

4. DISCUSSION

4.1. *Quasar Host Masses*

Following Cro05 we can use the ellipsoidal collapse model of Sheth, Mo & Tormen (2001) to convert quasar biases into masses for the halos hosting UVX quasars. We weight this model across our (normalized) redshift distributions

$$b_Q(M_{DMH}, \bar{z}) = 1 + \int_{z=z_{min}}^{z=z_{max}} \mathrm{d}z \frac{\mathrm{d}N}{\mathrm{d}z} \frac{1}{\sqrt{a}\delta_{sc}(z)} \times \left[\sqrt{a}(a\nu^2) + \sqrt{a}b(a\nu^2)^{1-c} - \frac{(a\nu^2)^c}{(a\nu^2)^c + b(1-c)(1-c/2)}\right] \quad (4)$$

where $a = 0.707$, $b = 0.5$, $c = 0.6$ and $\delta_{sc}(z)$, the critical density contrast for spherical collapse, is given by Navarro, Frenk & White (1997) as $0.15(12\pi)^{2/3}\Omega_{mz}^{0.0055}$ (for flat cosmologies), where $\Omega_{mz} \equiv \Omega_m(z)$ is given, e.g., in Mye06.

Masses can be derived via $\nu = \delta_{sc}(z)/\sigma_r(M, z)$, where $\sigma_r(M, z) = \sigma_r(M)D(z)$, and $D(z)$ is the linear growth factor, which we approximate using the formula of Carroll, Press & Turner (1992; see, e.g., Mye06). The mass variance for a halo, $\sigma_r^2(M_{DMH})$ can be determined from the radius of a halo of mean mass M_{DMH}

$$r = \left(\frac{3M_{DMH}}{4\pi\rho_0}\right)^{1/3} \quad (5)$$

where $\rho_0 = 2.78 \times 10^{11}\Omega_m h^2 M_\odot \mathrm{Mpc}^{-3}$ is the present mean density of the universe. This mass scale implies a mass variance of

$$\sigma_r^2(M_{DMH}) = \frac{V}{2\pi^2}\int_0^\infty k^3 P(k)\left[\frac{3j_1(kr)}{kr}\right]^2 \frac{\mathrm{d}k}{k} \quad (6)$$

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 σ_8 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 0.055–22 h^{-1} Mpc), and adopt a form for the (adiabatic, CDM) linear power spectrum of $P(k) = T^2(k)k^n$ with $n = 1$ (the Harrison-Zeldovich-Peebles scale-invariant case). We adopt the transfer function, $T(k) = T_0(q)$ given by Equation 29 of Eisenstein & Hu (1998), who, following Bardeen et al. (1986; see also Efstathiou, Bond & White 1992), showed that the transfer function can be characterized by its shape (Γ) via

$$q = \frac{k}{h^{-1}\ \mathrm{Mpc}}\Theta_{2.7}^2/\Gamma \quad (7)$$

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 $\Gamma \rightarrow \Gamma_{\rm eff}$ (e.g., Sugiyama 1995). Note that, as we analyze scales far smaller than the sound horizon, $\Gamma_{\rm eff}$ can be derived from the baryon fraction via $\Gamma_{\rm eff} = \alpha_r \Omega_m h$ with α_r 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 σ_8 and Γ. However, the conversion from b_Q to M_{DMH} is somewhat dependent on σ_8 and Γ. As such, we also now analyze our results in the context of a cosmological model with (Ω_m, Ω_Λ, σ_8, Γ, 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 $\Gamma = \alpha_r \Omega_m h = 0.15$ 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 Equation 3). A Spearman rank test shows no significant correlation between redshift and M_{DMH}; however such a test is not compelling for only 5 redshift bins, particularly as we don't 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 Cro05, 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$1.3σ from the value of $M_{DMH} = 3.0 \pm 1.6 \times 10^{12}$ h$^{-1}M_\odot$ obtained by Cro05 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 section 3.3, we detected no significant luminosity dependence to quasar bias in any redshift bin, our detection of 1.5σ in our highest redshift bin ($\bar{z} = 1.92$) 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 2σ detection in our $\bar{z} = 1.92$ bin would require a sample 3.8 times larger (including stellar contamination) or 1.8 times larger (ignoring stellar contamination). Similarly, a 3σ 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 1σ errors suggest that at $z < 1.6$ quasar bias changes with luminosity by less than ± 0.6 (± 0.3 if we ignore stellar contamination).

TABLE 2
Bivariate Estimates of the QSO bias, b_Q, as a function of photometric redshift z_{phot} and absolute magnitude M_g

		1 parameter model	2 parameter model[a]	
$\bar{z}_{phot}$	$\bar{M}_g$	b_Q	a	b_Q
$0.4 \leq z_{phot} < 1.2$				
0.85	-23.99	$2.03\pm^{0.32}_{0.39}$	$0.877\pm^{0.102}_{0.049}$	$1.89\pm^{0.66}_{1.37}$
0.85	-22.82	$2.00\pm^{0.33}_{0.39}$	$1.000\pm^{0.000}_{0.116}$	$2.00\pm^{0.64}_{0.53}$
0.85	-21.69	$2.08\pm^{0.31}_{0.37}$	$0.999\pm^{0.001}_{0.141}$	$2.08\pm^{0.54}_{0.65}$
$1.2 \leq z_{phot} < 1.65$				
1.44	-24.84	$2.86\pm^{0.29}_{0.32}$	$1.000\pm^{0.000}_{0.103}$	$2.86\pm^{0.53}_{0.50}$
1.44	-23.81	$2.62\pm^{0.31}_{0.34}$	$0.920\pm^{0.080}_{0.072}$	$2.73\pm^{0.52}_{0.67}$
1.44	-23.25	$2.68\pm^{0.30}_{0.35}$	$0.878\pm^{0.122}_{0.064}$	$2.87\pm^{0.55}_{0.69}$
$1.65 \leq z_{phot} < 2.3$				
1.92	-25.42	$2.77\pm^{0.32}_{0.37}$	$0.865\pm^{0.109}_{0.053}$	$2.87\pm^{0.61}_{0.84}$
1.92	-24.46	$2.25\pm^{0.45}_{0.57}$	$0.947\pm^{0.053}_{0.089}$	$2.36\pm^{0.76}_{1.05}$
1.92	-23.92	$3.18\pm^{0.29}_{0.31}$	$0.869\pm^{0.113}_{0.054}$	$3.46\pm^{0.52}_{0.61}$

[a] With stellar contamination $1 - a$.

5. SUMMARY AND FUTURE PROSPECTS

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

(a) Over scales of 0.16′ to 100′ ($\sim$55 h^{-1} kpc to 35 h^{-1} Mpc at our sample's mean redshift of $z \sim 1.4$) quasar clustering is well-described by Smi03 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.

(b) For $\sigma_8 = 0.9$ and $\Gamma = 0.21$ the required quasar bias is $b_Q^{\bar{z}=1.40} = 2.41 \pm 0.09$, rising to $b_Q^{\bar{z}=1.40} = 2.57 \pm 0.10$ for $\sigma_8 = 0.8$ and $\Gamma = 0.15$.

(c) Our sample alone is sufficient to rule out a constant quasar bias over $0.75 < \bar{z} < 2.28$ (for scales of $\sim$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 section 3.1). At $z \sim 2.3$ we find $b_Q \sim 3$. Considering complementary independent constraints on redshift evolution of 99.8% (Cro05) and 3.6σ (PMN04) from the 2QZ, it is certain that quasar bias is therefore evolving with cosmic time.

(d) Using our best photometric redshift ranges, σ_8=0.8, and $\Gamma = 0.15$, we find a mean mass for the dark matter halos hosting UVX quasars of $M_{DMH} = 5.2 \pm 0.6 \times 10^{12}$ h$^{-1} M_\odot$, approximately halfway between the values of M_{DMH} determined from the 2QZ by PMN04 and Cro05.

(e) We find no significant luminosity dependence to quasar clustering, but our analysis hints at a small dependence (1.5σ) at high redshift ($\bar{z} = 1.92$). This suggests that, with improved photometric classification efficiency, a sample size of as little as 1.8 times larger ($\sim$550,000 objects) may be sufficient to detect luminosity dependence in quasar clustering at $z \sim 2$. 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).

(f) 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 Figure 2). We concur that there is little luminosity dependence to quasar clustering on proper scales of $\gtrsim 50\ h^{-1}$ kpc (their $100\ h_{70}^{-1}$ kpc comoving), and, also, that any weak luminosity trend is only expressed for brighter quasars ($M_g \lesssim -24$; note, for comparison with Serber et al. 2006, that $g-i \sim 0.3$ for UVX quasars).

(g) 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 b_Q is constant at $z < 1.6$ to $\lesssim \pm 0.6$.

Although our analyses in section 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 redshift. 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 $z \gtrsim 1.6$.

ADM and RJB wish to acknowledge support from NASA through grants NAG5-12578 and NNG06GH15G. GTR was supported in part by a Gordon and Betty Moore Fellowship in Data Intensive Science. RCN acknowledges the EU Marie Curie Excellence Chair for support during this work. DPS acknowledges support through NSF grant AST-0607634. 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 U.S. 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

APPENDIX A. BACKGROUND METHODOLOGY

CORRELATION FUNCTION AND MODEL FITTING

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

$$\omega(\theta) = \frac{QQ(\theta) - 2QR(\theta)}{RR(\theta)} + 1 \tag{A1}$$

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 $\sim$1.7% of the data). Our approach is detailed in Mye06, 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 ω_L in each jackknife realization via Equation A1, then the inverse-variance-weighted covariance matrix (C_{ij}) can be generated as

$$C_{ij} = C(\theta_i, \theta_j) = \sum_{L=1}^{N} \sqrt{\frac{RR_L(\theta_i)}{RR(\theta_i)}} \left[\omega_L(\theta_i) - \omega(\theta_i)\right] \sqrt{\frac{RR_L(\theta_j)}{RR(\theta_j)}} \left[\omega_L(\theta_j) - \omega(\theta_j)\right] \tag{A2}$$

where, ω denotes the correlation function for all data and ω_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

$$|C| = \frac{C_{ij}}{\sigma_i \sigma_j} \tag{A3}$$

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 χ^2 fits to model angular autocorrelation functions (ω_m) using the inverse of the covariance matrix, and determine errors on fits from $\Delta\chi^2$, where

$$\chi^2 = \sum_{i,j} \left[\omega(\theta_i) - \omega_m(\theta_i)\right] C_{ij}^{-1} \left[\omega(\theta_j) - \omega_m(\theta_j)\right] \tag{A4}$$

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

MODELING PROJECTED QUASAR CLUSTERING

Since the seminal scaling relations of Hamilton et al. (1991), many authors (e.g., Peacock & Dodds 1994; Jain, Mo & White 1995; Peacock & Dodds 1996) have worked to obtain precise analytical descriptions of the clustering of dark matter particles. Smith et al. (2003; henceforth Smi03) married traditional approaches with a simple halo model (e.g., Cooray & Sheth 2002) to obtain fitting formulae that better approximate the non-linear clustering behavior of simulated CDM. The Smi03 formulae reproduce clustering in dark matter simulations to better than 3% at (redshift) $z < 3$ for (wavenumber) $k < 10\ h\ \mathrm{Mpc}^{-1}$.

The models of Smi03 directly predict the non-linear, dimensionless power spectrum of dark matter $\Delta^2_{\rm NL}(k, z)$ 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 $\Delta^2_{\rm Q}(k, z) = b_Q^2(k, z)\Delta^2_{\rm NL}(k, z)$, where b_Q is the quasar bias.

In this paper, we measure angular autocorrelations. For small angles ($\theta \ll 1$ radian), 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

$$\omega(\theta) = \pi \int_{z=0}^{z=\infty} \int_{k=0}^{k=\infty} \frac{\Delta^2(k, z)}{k} J_0[k\theta\chi(z)] \left(\frac{\mathrm{d}N}{\mathrm{d}z}\right)^2 \left(\frac{\mathrm{d}z}{\mathrm{d}\chi}\right) F(\chi) \frac{\mathrm{d}k}{k} \mathrm{d}z \tag{B1}$$

where J_0 is the zeroth-order Bessel function of the first kind, χ is the radial comoving distance, $\mathrm{d}N/\mathrm{d}z$ is the redshift selection function (normalized so that $\int_0^\infty [\mathrm{d}N/\mathrm{d}z]\mathrm{d}z = 1$), and $\mathrm{d}z/\mathrm{d}\chi = H_z/c = H_0\sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}/c$. Strictly, χ 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

$$\omega_{QQ}(\theta) = \frac{H_0\pi}{c} \int_0^\infty \int_0^\infty b_Q^2(k, z) \frac{\Delta^2_{\rm NL}(k, z)}{k} J_0[k\theta\chi(z)] \left(\frac{\mathrm{d}N}{\mathrm{d}z}\right)^2 \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda} \frac{\mathrm{d}k}{k} \mathrm{d}z \tag{B2}$$

In theory, with sufficient data, $b_Q(k, z)$ may be directly constrained by ω_{QQ}, although we will generally set b_Q to be a constant and constrain it by comparing the amplitudes of ω_{QQ} and the projected matter power spectrum as a function of redshift and scale.

In general, $\Delta^2_{\rm NL}$ is not separable into individual functions of k and z. We therefore Monte Carlo integrate under the surface described by the integrand in Equation B2 until the integration is evaluated to better than 1%. To optimize this process, two points are worth noting. First, the change of variables $\mathrm{d}k/k = \ln(10)\mathrm{d}\log(k)$ allows more uniform sampling in k-space. Second, although $\Delta^2_{\rm NL}$ is not easily separated, we have determined that the "parameters of the spectrum" (k_σ^{-1}, n_{eff} and C; see Appendix C of Smi03) can be approximated by splines to $\lesssim 0.3\%$ for at least $z < 4$ in a ΛCDM cosmology. Figure 7 demonstrates a typical surface we might integrate under to obtain ω_{QQ}. On scales

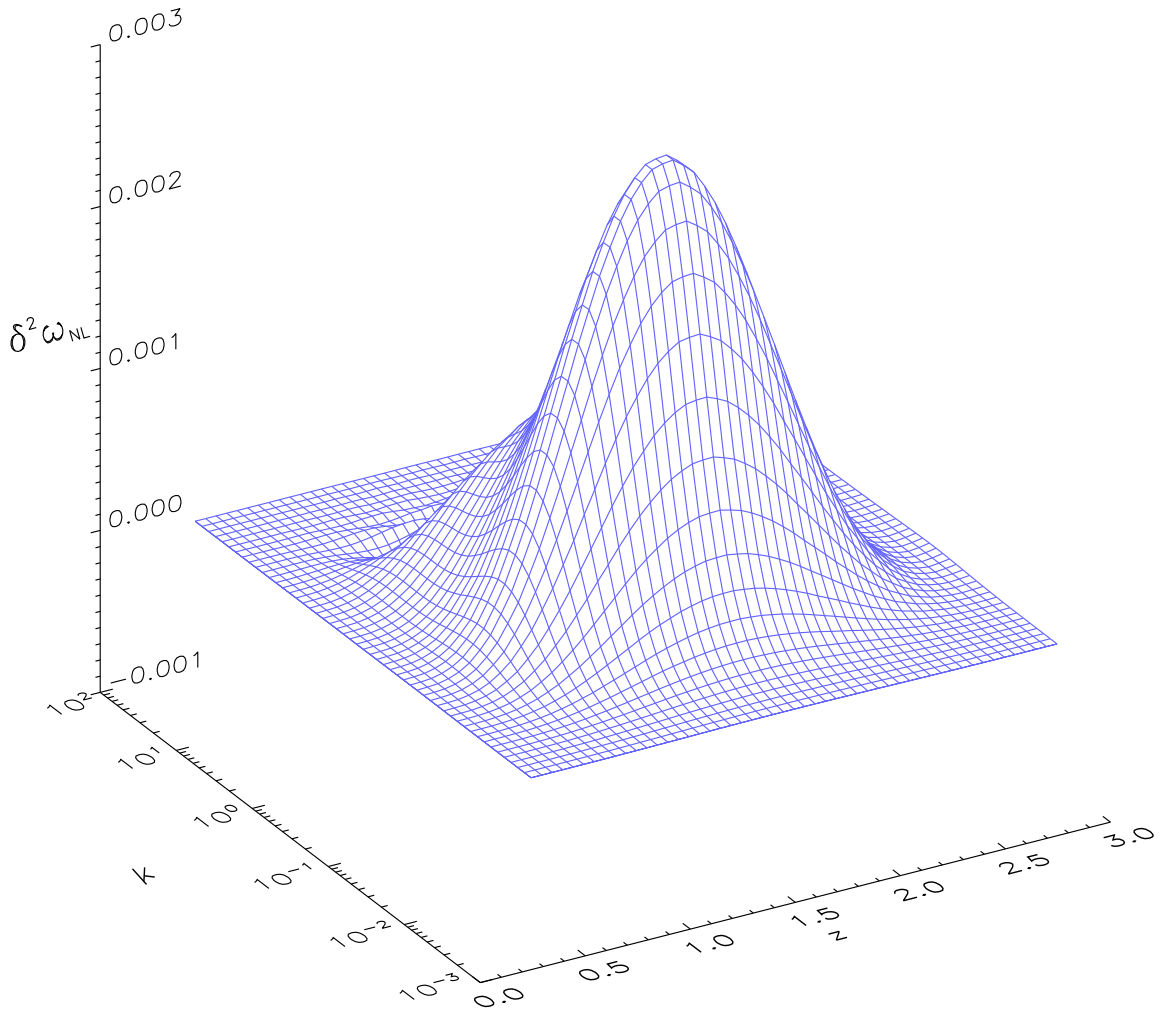

FIG. 7.— One example of a surface that we Monte Carlo integrate under to project the matter power spectrum into the angular autocorrelation function. The surface is plotted for $\theta = 1'$, and $\delta^2\omega_{NL} = \left(\Delta^2_{\rm NL}/k^2\right) J_0[k\theta\chi(z)] \left({\rm d}N/{\rm d}z\right)^2 {\rm d}z/{\rm d}\chi$ (see Equations B1 and B2). We obtain ${\rm d}N/{\rm d}z$ from spectroscopic matches (with DR1QSO, DR2 or the 2QZ) to our photometrically classified sample. Two individual contributions to power are apparent at $1'$; the non-linear matter spectrum at low k and a halo term at high k (see Smi03).

$\lesssim 30'$ 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

STELLAR CONTAMINATION

We have addressed sources of systematic error in some depth in Mye06. In particular, we noted that the autocorrelation of KDE objects, ω, combines clustering signals from a stellar component ω_{SS} and the true quasar autocorrelation ω_{QQ}. If a is the efficiency, the fraction of genuine quasars that are classified as such by the KDE technique, then

$$\omega(\theta) = a^2\omega_{QQ}(\theta) + (1-a)^2\omega_{SS}(\theta) + \epsilon(\theta) \tag{C1}$$

where ϵ is a tiny (theoretically zero) offset arising from QS, QR and SR cross-terms.

If the efficiency of the KDE technique is high (i.e., $a \to 1$), stellar contamination is only important as $\omega_{QQ} \to 0$. In Section 3.1, we fit the two-parameter model defined by Equations B2 and C1 to the DR4 KDE autocorrelation and estimate the stellar contamination $(1-a)$. In doing so, we derive $a = 0.956\pm^{0.044}_{0.019}$, consistent with Mye06 (and with Richards et al. 2004), and find that our best analysis of the quasar bias is therefore at angles of $\theta \lesssim 1^\circ$.

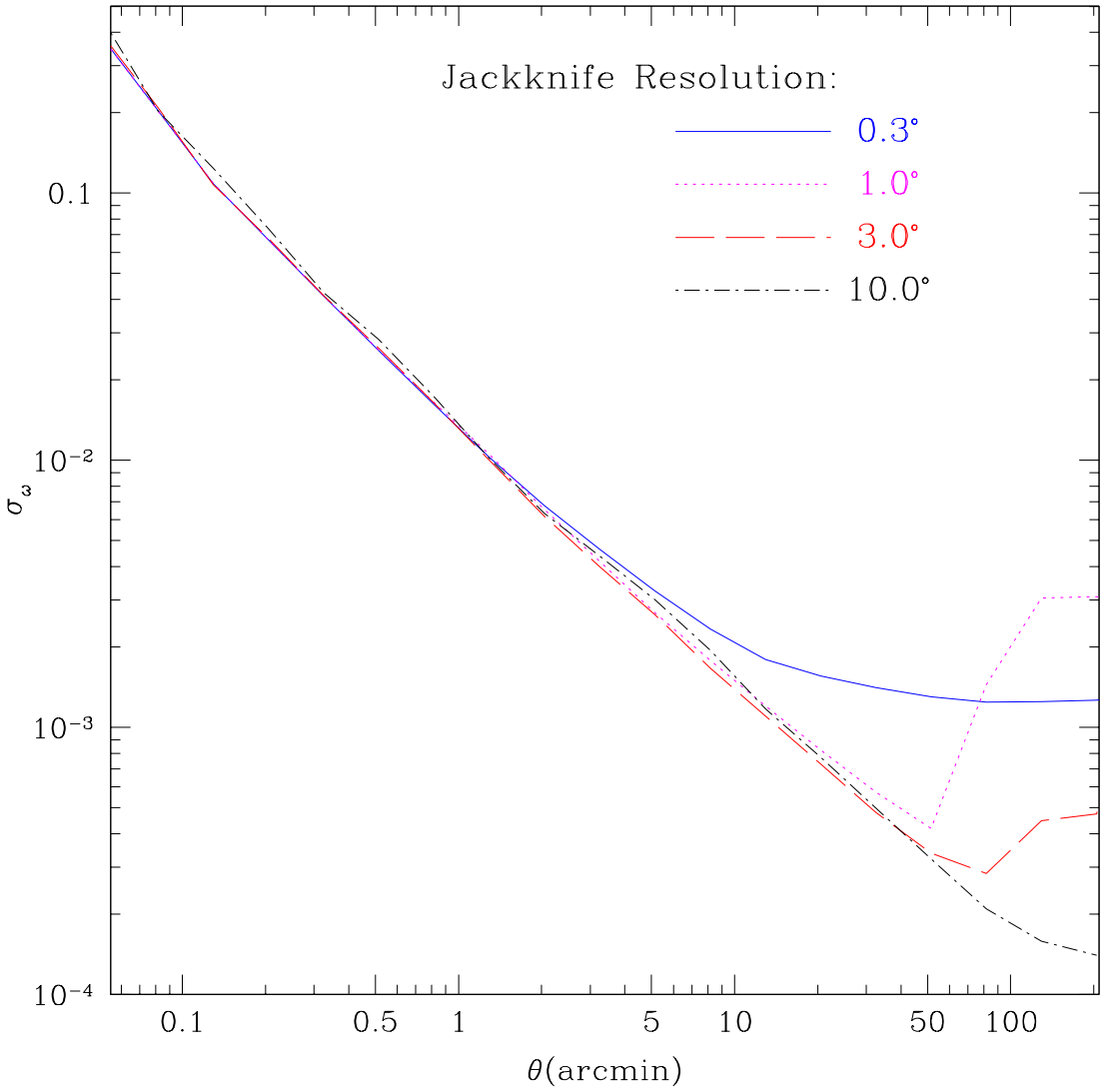

FIG. 8.— Error on the correlation function, σ_ω at different jackknife resolutions. The error appears to be poorly defined, or flatten, around scales corresponding to the resolution. Generally, this is the same scale at which a pixel-to-pixel estimate of the error would break down. The data used to estimate σ_ω was the DR4 KDE sample discussed in Section 2, with an initial dust cut $A_g < 0.24$.

MISCLASSIFIED H II REGIONS

In Paper2, we discuss non-stellar 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 $12''$ (the smallest scale fit in this paper), the effect amounts to $\sim 1.0\sigma$ (where σ is the error on ω), by scales of $30''$, 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 masked from KDE clustering analyses because of this small effect but (unlike in Paper2 where our focus is small scales) we make no attempt to correct for the effect in this paper.

COVARIANCE AND THE JACKKNIFE RESOLUTION

If the covariance between scales is perfectly accounted for, Equation A4 should return identical estimates of χ^2 irrespective of the jackknife resolution. Given that DR4 covers close to 7000 square degrees, 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 $\omega + \sigma_\omega$ (the tip of the 1σ upper error bar) as obtained when jackknife resampling at $10°$, and calculate χ^2 for other jackknife

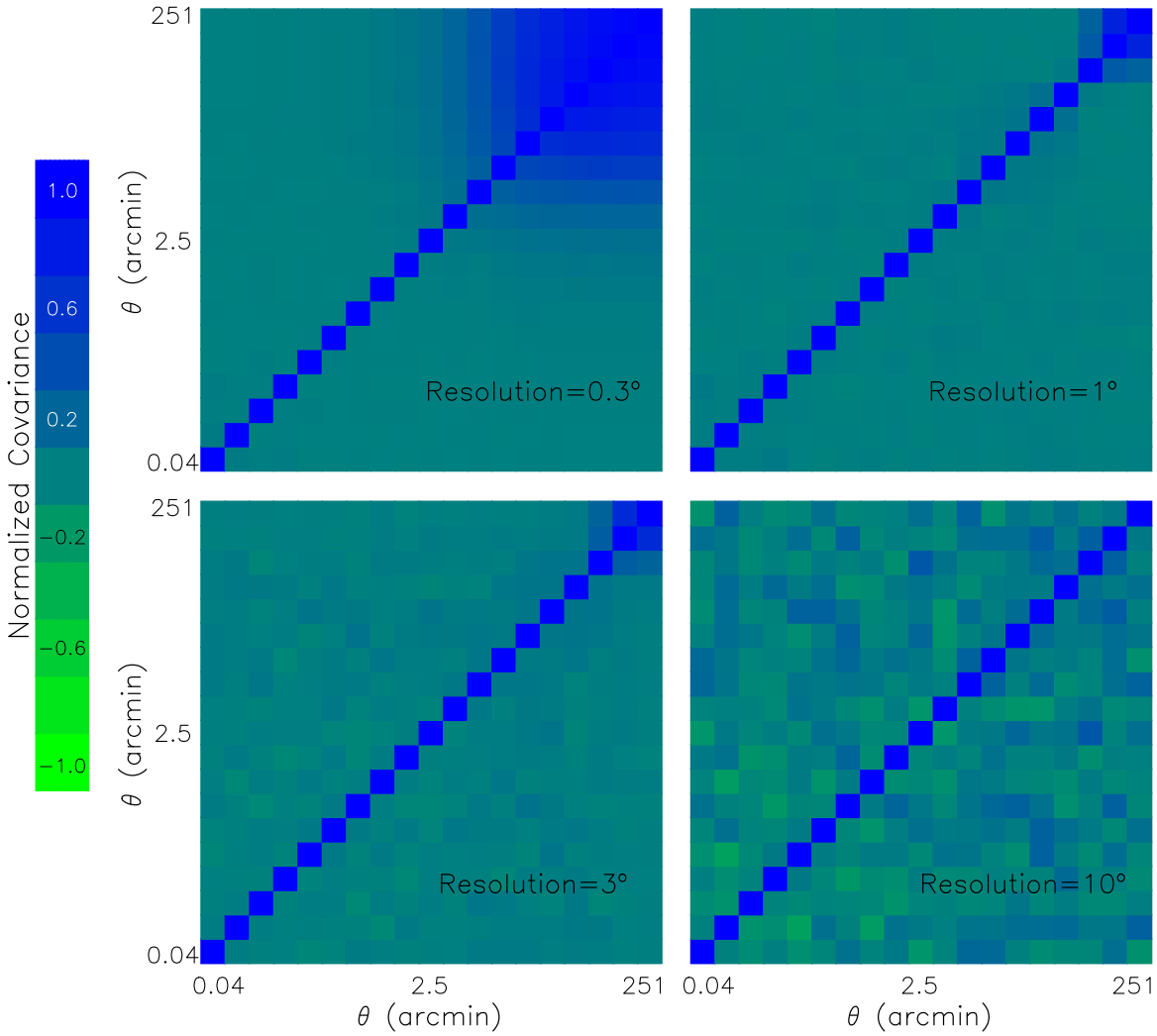

FIG. 9.— Normalized covariance matrix (see Equation A3) at different jackknife resolutions. The covariance grows rapidly on scales larger than the jackknife resolution. On scales far smaller than the jackknife resolution, off-diagonal elements are well-mixed but still contribute to significance estimates. From top-left to bottom-right the panels represent jackknife resolutions of $0.3°, 1°, 3°$ and $10°$. The data used to estimate the covariance matrices was the DR4 KDE sample discussed in Section 2, with a nominal $A_g < 0.24$ Galactic absorption cut.

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 χ^2 estimates (within $\sim$2%) but 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 will 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 χ^2 estimates (to within 5%). Further, at a resolution of $10°$ pixel-to-pixel errors can be calculated (see Appendix A) 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?

ADDITIONAL OBSERVATIONAL SYSTEMATICS

Mye06 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 Mye06, 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 Mye06, we find a clear clustering pattern in the KDE sample for $A_g \gtrsim 0.22$. We note that Yahata et al. (2006) found little change in the number density of KDE objects as a function of 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 Mye06). We have checked that a range of cuts around $A_g \lesssim 0.21$ (in particular, our adopted cut of $A_g < 0.18$ from Mye06) all yield statistically similar estimates of ω. For our main

analyses in this paper, we adopt no seeing cut and an $A_g < 0.21$ cut, which discards ∼12.5% of our sampled area.

REFERENCES

Abazajian, K., et al. 2003, AJ, 126, 2081 (DR1)
Abazajian, K., et al. 2004, AJ, 128, 502 (DR2)
Abazajian, K., et al. 2005, AJ, 129, 1755 (DR3)
Ball, N. M., et al. 2007, in preparation
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
Carroll, S. M., Press, W. H., & Turner, E. L. 1992, ARA&A, 30, 499
Carlberg, R. G. 1990 ApJ, 350, 505
Ciotti, L., & Ostriker, J. P. 1997, ApJ, 487, 105
Ciotti, L., & Ostriker, J. P. 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 (Cro05)
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Efstathiou, G., Bond, J. R., & White, S. D. M. 1992, MNRAS, 258, 1
Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605
Fabian, A. C. 1999, MNRAS, 308, 39
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, 9
Franceschini, A., Hasinger, G., Miyaji, T., & Malquori, D. 1999, MNRAS, 310, 5
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, 13
Giannantonio, T., et al. 2006, in preparation
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
Gunn, J. E., et al. 2006, AJ, 131, 2332
Haiman, Z., & Hui, L. 2001, ApJ, 547, 27
Hamilton, A. J. S., Kumar, P., Lu, E., & Matthews, A. 1991, ApJ, 374, 1
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., Martini, P., Cox, T. J., Robertson, B., Di Matteo, T., & Springel V. 2005a, ApJ, 625, 71
Hopkins, P. F, Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel V. 2005b, ApJ, 630, 716
Hopkins, P. F, Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel V. 2006, ApJS, 163, 1
Ivezic, Z., et al. 2004, AN, 325, 583
Jain, B., Mo, H. J., & White, S. D. M. 1995, MNRAS, 276, 25
Kaiser, N. 1984, ApJ, 284, 9
Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581
Lacey C., & Cole S. 1993, MNRAS, 262, 627
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Lidz, A., Hopkins, P. F, Cox, T. J., Hernquist, L., & Robertson, B. 2006, ApJ, 641, 41
Limber, D. N. 1953, ApJ, 117, 134
Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406
Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990a, MNRAS, 242, 43
Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990b, MNRAS, 243, 692
Magorrian, J., et al. 1998, AJ, 115, 2285
Martini, P., & Weinberg, D. H. 2001, ApJ, 547, 12
Myers, A. D., Outram, P. J., Shanks, T., Boyle, B. J., Croom, S. M., Loaring, N. S., Miller, L., & Smith, R. J. 2003, MNRAS, 342, 467
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., et al. 2006, ApJ, 638, 622 (Mye06)
Myers, A. D., et al. 2007, in preparation (Paper2)
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Negroponte, J., & White, S. D. M. 1983, MNRAS, 205, 1009
Peacock, J. A. 1991, MNRAS, 253, 1
Peacock, J. A., & Dodds, S. J. 1994, MNRAS, 267, 1020
Peacock, J. A., & Dodds, S. J. 1996, MNRAS, 280, 19
Peebles, P. J. E. 1980, in The Large Scale Structure of the Universe, (Princeton University Press)
Pier, J. R., Munn, J. A., Hindsley, R. B., Hennessy, G. S., Kent, S. M., Lupton, R. H., & Ivezic, Z. 2003, AJ, 125, 1559
Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010 (PMN04)
Porciani, C., & Norberg, P. 2006, MNRAS, in press (astro-ph/0607348)
Rees, M. J. 1984, ARA&A, 22, 471
Richards, G. T., et al. 2004, ApJS, 155, 257
Richards, G. T., et al. 2006, AJ, 131, 2766
Richstone, D., et al. 1998, Nature, 395, 14
Riess, A. G., et al. 2004, ApJ, 607, 665
Sazonov, S. Y., Ostriker, J. P., Ciotti, L., & Sunyaev, R. A. 2005, MNRAS, 358, 168
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schneider, D. P., et al. 2003, AJ, 126, 2579 (DR1QSO)
Scranton, R., et al. 2002, ApJ, 579, 48
Scranton, R., et al. 2003, Phys. Rev. Lett., submitted, (astro-ph/0307335)
Scranton, R., et al. 2005, ApJ, 633, 589
Serber, W., Bahcall, N. A., Ménard, B., & Richards, G. 2006, ApJ, 643, 68
Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1
Shields, G. A., Gebhardt, K., Salviander, S., Wills, B. J., Xie, B., Brotherton, M. S., Yuan, J., & Dietrich, M. 2003, ApJ, 583, 124
Silk, J., & Rees, M. J. 1998, A&A, 331, 1
Smith, J. A., et al. 2002, AJ, 123, 2121
Smith, R. E., et al. 2003, MNRAS, 341, 1311 (Smi03)
Spergel, D. N., et al. 2006, ApJ, submitted (astro-ph/0603449)
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Stoughton, C., et al. 2002, AJ, 123, 485
Sugiyama, N. 1995, ApJS, 100, 281
Toomre, A., & Toomre, J. 1972, ApJ, 178, 623
Totsuji, H., & Kihara, T. 1969, PASJ, 21, 221
Tremaine, S., et al. 2002, ApJ, 574, 740
Tucker, D., et al. 2006, AN, in press
Valageas, P., Silk, J., & Schaeffer, R. 2001, A&A, 366, 363
Weinstein, M. A., et al. 2004, ApJS, 155, 243
White, S. D. M. 1979, MNRAS, 189, 831
Wisotzki, L. 2000, A&A, 353, 861
Wyithe, J. S. B. 2006, MNRAS, 365, 1082
Wyithe, J. S. B., & Loeb, A. 2002, ApJ, 581, 886
Yahata, K., Atsunori, Y., Yasushi, S., Turner, E. L., Broadhurst, T., & Finkbeiner, D. P. 2006, PASP, submitted, (astro-ph/0607098)
York, D. G., et al. 2000, AJ, 120, 1579
Zehavi, I., et al. 2002, ApJ, 571, 172