Strong Clustering of Lyman Break Galaxies around Luminous Quasars at $Z \sim 4$ [*][†]

Cristina García-Vergara[1,2,3] , Joseph F. Hennawi[2,4] , L. Felipe Barrientos[1], and Hans-Walter Rix[2]
[1] Instituto de Astrofísica, Pontificia Universidad Católica de Chile, Avenida Vicuña Mackenna 4860, Santiago, Chile; garcia@strw.leidenuniv.nl
[2] Max-Planck-Institut für Astronomie (MPIA), Königstuhl 17, D-69117 Heidelberg, Germany
[3] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands
[4] Department of Physics, University of California, Santa Barbara, CA 93106, USA

Abstract

In the standard picture of structure formation, the first massive galaxies are expected to form at the highest peaks of the density field, which constitute the cores of massive proto-clusters. Luminous quasars (QSOs) at $z \sim 4$ are the most strongly clustered population known, and should thus reside in massive dark matter halos surrounded by large overdensities of galaxies, implying a strong QSO–galaxy cross-correlation function. We observed six $z \sim 4$ QSO fields with VLT/FORS, exploiting a novel set of narrow-band filters custom designed to select Lyman Break Galaxies (LBGs) in a thin redshift slice of $\Delta z \sim 0.3$, mitigating the projection effects that have limited the sensitivity of previous searches for galaxies around $z \gtrsim 4$ QSOs. We find that LBGs are strongly clustered around QSOs, and present the first measurement of the QSO–LBG cross-correlation function at $z \sim 4$, on scales of $0.1 \lesssim R \lesssim 9 \ h^{-1}$ Mpc (comoving). Assuming a power-law form for the cross-correlation function $\xi = (r/r_0^{QG})^\gamma$, we measure $r_0^{QG} = 8.83^{+1.39}_{-1.51} \ h^{-1}$ Mpc for a fixed slope of $\gamma = 2.0$. This result is in agreement with the expected cross-correlation length deduced from measurements of the QSO and LBG auto-correlation function, and assuming a deterministic bias model. We also measure a strong auto-correlation of LBGs in our QSO fields, finding $r_0^{GG} = 21.59^{+1.72}_{-1.69} \ h^{-1}$ Mpc for a fixed slope of $\gamma = 1.5$, which is ~ 4 times larger than the LBG auto-correlation length in blank fields, providing further evidence that QSOs reside in overdensities of LBGs. Our results qualitatively support a picture where luminous QSOs inhabit exceptionally massive ($M_{halo} > 10^{12} \ M_\odot$) dark matter halos at $z \sim 4$.

Key words: cosmology: observations – early universe – galaxies: clusters: general – galaxies: high-redshift – large-scale structure of universe – quasars: general

1. Introduction

Our understanding of structure formation suggests that small inhomogeneities in the density field shortly after the big bang grew over cosmic time via gravitational instability (e.g., Dodelson 2003; Padmanabhan 2006; Schneider 2015) into massive dark matter halos at $z = 0$. As clusters of galaxies are the most massive, gravitationally bound structures in the universe, we expect them to form from the highest-density peaks at early times. This make them ideal laboratories for studying the formation and evolution of cosmic structure.

Because of the small areas of sky surveyed at high redshift, and the low comoving number density $\sim 10^{-7}$ Mpc^{-3} of local clusters (Gioia et al. 2001; Vikhlinin et al. 2009), the evolutionary link between these low-redshift clusters and high-redshift galaxies has been challenging to make. The progenitors of clusters are extremely difficult to identify when the density contrast between the forming cluster and its surroundings is small (for a review see Overzier 2016). Efforts have been made to search for these so-called proto-clusters in large galaxy surveys with subsequent spectroscopic follow-up, successfully detecting some structures (e.g., Steidel et al. 2000, 2005; Ouchi et al. 2005; Capak et al. 2011; Wang et al. 2016). However, given the small volume of such high-redshift

surveys, a commonly adopted approach is to search for proto-clusters around known high-redshift massive galaxies.

One very fruitful technique to find high-redshift proto-clusters has been to use the presence of an active supermassive black hole (BH) as a signpost for a massive galaxy and hence massive dark matter halo in the distant universe (e.g., Kashikawa et al. 2007; Venemans et al. 2007; Overzier et al. 2008; Morselli et al. 2014). This technique is motivated by several considerations. First, the masses of supermassive BHs (M_{BH}) are known to tightly correlate with the bulge mass of their host galaxy (Magorrian et al. 1998; Ferrarese & Merritt 2000; Gebhardt et al. 2000), and possibly with the masses of their host dark halos (M_{halo}) (Ferrarese 2002, but see Kormendy & Bender 2011). Intriguingly, the most luminous quasars (QSOs) at $z > 3$ have $M_{BH} \sim 1$–$6 \times 10^9 \ M_\odot$ (Shen et al. 2011), comparable to the most massive known local BHs. If the present-day $M_{BH} - M_{halo}$ relation holds at early times, such BHs should reside in exceptionally massive halos. Second, some studies have suggested that the nuclear activity in active galactic nuclei (AGNs) is triggered by processes related to the environment where they reside. For example, galaxy mergers could trigger the AGN activity (Bahcall et al. 1997; Wyithe & Loeb 2002; Hennawi et al. 2015), and galaxy mergers occur preferentially in dense environments (Lacey & Cole 1993). This would imply that the existence of an AGN requires a dense environment around it. Finally, another line of evidence that QSOs trace the rarest environments at high redshift arises from their extremely strong clustering. Indeed, Shen et al. (2007) determined that QSOs at $z > 3.5$ have a comoving auto-correlation length of $r_0 = 24.3 \ h^{-1}$ Mpc (for a

fixed correlation function slope of $\gamma = 2.0$), making them the most strongly clustered population in the universe, and demanding that they reside in the most massive $M_{halo} > 10^{12} M_\odot$ dark matter halos at this epoch. This high clustering also implies a small scattering in the $L_{QSO}-M_{halo}$ relation (White et al. 2008) and then, for the case of luminous, high-redshift QSOs, their host halo masses are well constrained. Additionally, the Shen et al. (2007) correlation function agrees with that required to explain the abundance of binary QSOs at $z > 3.5$ (Hennawi et al. 2010; Shen et al. 2010), indicating that overdense structures around QSOs extend down to scales as small as $100\,h^{-1}$ kpc. Since in hierarchical clustering models, QSOs and galaxies trace the same underlying dark matter density distribution, the generic prediction is that galaxies should be very strongly clustered around QSOs at $z \gtrsim 3.5$. Observationally this should be reflected as a strong QSO–galaxy cross-correlation function.

The QSO–galaxy cross-correlation function has been measured at $z < 4$ in the past. At $z \lesssim 1$ it is found to be in good agreement with the auto-correlation of galaxies and QSOs, and it has been shown to be weakly dependent on the QSO luminosity, and redshift (e.g., Coil et al. 2007; Padmanabhan et al. 2009; Shen et al. 2013). Adelberger & Steidel (2005) measured the AGN–galaxy cross-correlation function at higher redshifts ($2 \lesssim z \lesssim 3$), finding a cross-correlation length of $r_0 \sim 5\,h^{-1}$ Mpc for a slope of $\gamma = 1.6$, which is similar to the auto-correlation of Lyman Break Galaxies (LBGs) at $z \sim 3$ (Adelberger et al. 2003). They also claim an independence of the cross-correlation length with the AGN luminosity, implying that both faint and bright AGNs should be found in halos with similar masses. The highest-redshift measurement of QSO environments is the work of Trainor & Steidel (2012), who quantified the clustering of LBGs around 15 hyper-luminous QSOs at $z = 2.7$. They found a QSO–LBG cross-correlation length of $r_0 = 7.3 \pm 1.3\,h^{-1}$ Mpc for a fixed slope of $\gamma = 1.5$ and claimed that this measurement is in agreement with the Adelberger & Steidel (2005) results. Additionally, they computed a halo mass for those QSOs of log $(M_{halo}/M_\odot) = 12.3 \pm 0.5$, which is in agreement with the halos masses inferred for fainter QSOs at the same redshift (White et al. 2012).

Theoretical considerations suggest that high-redshift QSOs live in massive dark matter halos, but not necessarily the most massive ones (Fanidakis et al. 2013). However, a high signal-to-noise clustering analysis is necessary to confirm this hypothesis.

In addition to these statistical clustering analyses, many studies of individual AGN environments have been conducted. The population of AGNs whose environments have been studied most intensively are the high-redshift radio galaxies (HzRGs) at $z \sim 2$–4, which have been shown to often reside in proto-cluster environments (e.g., Intema et al. 2006; Venemans et al. 2007; Overzier et al. 2008; Hennawi et al. 2015). At higher redshifts the environments of other classes of AGNs, such as optically selected QSOs, are currently less well constrained. Most previous work focuses on searching for galaxies around the most distant $z \gtrsim 5$ QSOs, and these results paint a diverse and rather confusing picture: Stiavelli et al. (2005), Zheng et al. (2006), Kashikawa et al. (2007), Utsumi et al. (2010), and Morselli et al. (2014) find quite a strong enhancement of galaxies compared to control fields around $z \sim 5$–6 QSOs, whereas Willott et al. (2005), Bañados et al.

(2013), Simpson et al. (2014), and Mazzucchelli et al. (2017) find no significant excess of galaxies around QSOs at $z \sim 6$–7. Kim et al. (2009) studied five QSO fields at $z \sim 6$ and reported a mix of overdensities and underdensities, and Husband et al. (2013) found galaxy overdensities in $z \sim 5$ QSOs environments, but they noted that even some randomly chosen patches of sky without AGN signposts ("blank fields") at the same redshift contain similar galaxy overdensities. Indeed, surveys of a few deg^2 for $z \sim 6$ LBGs or Lyman alpha emitters (LAEs) have identified comparable or even more overdense regions in blank field pointings (e.g., Ouchi et al. 2005; Ota et al. 2008; Toshikawa et al. 2012). These mixed results at $z \gtrsim 5$ do not yet provide compelling evidence that QSOs inhabit massive dark matter halos at the highest redshifts, and more work is clearly required.

One complication of these studies is that the majority of them are focused on dropout selection, which selects galaxies over a broad redshift range of $\Delta z \sim 1$ (e.g., Ouchi et al. 2004a), corresponding to $\sim 520\,h^{-1}$ cMpc at $z = 4$. A large part of such a volume is unassociated with the QSO, which introduces projection effects that dilute the overdensity around the QSO, making it much more difficult to detect. Furthermore, most works at the highest redshifts have focused their searches around a handful of individual QSOs, and given the poor statistics and large sample variance (which is typically not taken into account), this could preclude the detection of an overdensity.

In this paper we study the environs of QSOs at $z \sim 4$. There are several advantages to working at this redshift. First, it is the highest redshift at which auto-correlation measurements exist for QSOs (Shen et al. 2007), establishing that they reside in massive dark matter halos. Second, the luminosity function and clustering properties of $z \sim 4$ galaxies are also well known (e.g., Ouchi et al. 2004a, 2008; Shen et al. 2007). The well-measured luminosity function allows us to accurately determine the background number density, essential for a robust clustering analysis. Furthermore, the fact that the auto-correlation of QSOs and galaxies are both known gives us an idea of what the cross-correlation should be. In practical terms, redshift $z \sim 4$ also represents a compromise since the dark matter halos hosting QSOs are still expected to be massive (Shen et al. 2007), while at the same time the characteristic galaxy luminosity L_* can be imaged with much shorter exposure times than galaxies at $z \gtrsim 5$, allowing us to observe a larger statistical sample of QSO fields. Note that at $z \sim 4$ the universe was only ~ 1.5 Gyr old, and only 0.5 Gyr has elapsed since the end of reionization. Thus, our QSO targets are young objects residing in large-scale structures that are still forming.

Here we present VLT/FORS imaging of six $z \sim 4$ luminous QSOs fields. We use a novel narrow-band (NB) filter technique designed to select LBGs in a narrow redshift range ($\Delta z \sim 0.3$) around these QSOs. This minimizes the line-of-sight contamination, dramatically reducing the projection effects that are inherent in broad-band selection. We measure the QSO–LBG cross-correlation function at $z \sim 4$ for the first time, to determine whether luminous QSOs at $z \sim 4$ are surrounded by overdensities of LBGs. The sample of six QSOs studied allows us to beat down the noise from limited numbers of galaxies and cosmic variance.

The outline of this paper is as follows. In Section 2 we describe the QSO target selection, we explain the novel NB imaging technique used to select LBGs, and we give details of

the imaging observations, data reduction, and photometry. We present the color criteria used to select LBGs and compute the redshift selection function of the sample in Section 3. The measurements of the QSO–LBG cross-correlation function and LBG auto-correlation in QSO fields are presented in Section 4, where we also estimate the power-law correlation function $\xi(r) = (r/r_0)^{-\gamma}$ parameters r_0 and γ. We test the robustness of our results in Section 5, and summarize and conclude in Section 6.

Throughout this paper magnitudes are given in the AB system (Oke 1974; Fukugita et al. 1995) and we adopt a cosmology with $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.26$ and $\Omega_\Lambda = 0.74$, which is consistent with the nine-year *Wilkinson Microwave Anisotropy Probe* observations (Hinshaw et al. 2013). Comoving and proper Mpc are denoted as "cMpc" and "pMpc," respectively.

2. Observations and Data Reduction

The data set presented in this section was obtained from the ESO Archive (Program ID: 079.A-0644, P.I: Rix). This program was designed to search for LBGs in $z \sim 4$ QSOs environments using a novel NB filter technique. The aim was to test whether QSOs with the most massive BHs at $z \sim 4$ live in the most massive dark matter halos.

2.1. QSO Target Selection

The PI of this program designed a custom set of filters (see Section 2.2 for details) to search for LBGs in QSO environments. Using experiments with mock catalogs, they showed that this filter set allowed one to select galaxies with $z = 3.78 \pm 0.08$. Given this small redshift interval, and with the goal of stacking the galaxy number counts from several QSO fields, the QSO targets were selected to span a narrow redshift range of $\Delta z = 0.04$, centered at $z = 3.78$.

Taking advantage of the large sample of QSOs from the Sloan Digital Sky Survey (SDSS; York et al. 2000), they first selected all QSOs in this redshift range. Given the goal of studying the most massive dark matter halos at $z \sim 4$, believed to be correlated with the most massive BHs, only QSOs with $M_{BH} \gtrsim 10^9\,M_\odot$ were selected. As is typical, M_{BH} was estimated from the emission line widths and continuum luminosities using the so-called single-epoch reverberation mapping technique (Vestergaard 2002). One of the targeted QSOs was not selected from SDSS, but it was added to the sample because it belongs to the redshift and M_{BH} range of interest.[5] The final sample was comprised of six bright QSOs with $i < 20.2$ mag.

We verified that none of the QSOs had a detected radio emission counterpart at 20 cm by checking the Faint Images of the Radio Sky at Twenty-centimeters (FIRST Becker et al. 1995) catalog, since it is known that radio emission could strongly affect the galaxy clustering properties in AGN environments (e.g., Venemans et al. 2007; Shen et al. 2009). The QSO properties are summarized in Table 1, where we show more recent M_{BH} estimates taken from Shen et al. (2011).

---

[5] Properties of this QSO were measured by McLeod & Bechtold (2009) and are shown in Table 1.

2.2. A Novel Method to Select LBGs

The traditional Lyman break technique used to select high-redshift galaxies relies on the detection of the 912 Å flux break (the so-called Lyman limit break) observed in galaxies due to the absorption of photons with $\lambda < 912$ Å by neutral hydrogen in their interstellar and circumgalactic media. For this selection method, two bands are typically used to bracket the break, one located at $\lambda < 912(1 + z)$ Å, and the other at $\lambda > 912(1 + z)$ Å, where z is the redshift of the galaxies in question. Given this configuration, a non-detection is expected in the band blueward of the break, whereas a clear detection is expected redward of it, such that a very red color will be measured. Additionally, a third band is added at longer wavelengths in order to eliminate possible contaminants. This method was originally explored using the *UGR* filter system to detect galaxies at $z \sim 3$ (Steidel et al. 1995, 1996, 2003); however, it was subsequently generalized to higher redshifts ($z \sim$ 4–5) using a filter set shifted to longer wavelengths (Steidel et al. 1999; Ouchi et al. 2004a).

At higher redshifts ($z \gtrsim 4$), a second break in galaxy spectra becomes important. The Lyα opacity of the intergalactic medium (IGM) rapidly increases with redshift, such that a large fraction of photons emitted by galaxies with $\lambda < 1216$ Å are absorbed by neutral hydrogen. This implies a break at $\lambda = 1216$ Å (the so-called Lyα break), which can be used to select galaxies analogous to the traditional Lyman break technique described above. This Lyα break detection technique has been used to successfully identify galaxies and QSOs at $z \gtrsim 6$ (Fan et al. 2000; Bouwens et al. 2007, 2010; Oesch et al. 2010; Bañados et al. 2016).

In order to achieve our goal of selecting galaxies physically associated with high-redshift QSOs, we need to select LBGs within a narrow redshift range centered on the QSO. However, the Lyman break method (using either the Lyman limit or Lyα breaks) efficiently selects LBGs in a broad redshift slice of $\Delta z \sim 1$ (e.g., Ouchi et al. 2004a; Bouwens et al. 2007, 2010), corresponding to $\sim 520\,h^{-1}$ cMpc at $z = 4$. For such a broad redshift range, the overdensity signal around the QSO will be significantly diluted by the projection of galaxies at much larger distances, hundreds of comoving Mpc away.

In order to address this problem, the PI proposed a novel selection technique analogous to the Lyα break method, but with the difference that the selection of LBGs is performed using two NB filters located very close to each other, instead of using broad bands. These filter curves are compared to those used for traditional LBG selection in Figure 1. The advantage of using NB filters is that they allow one to select LBGs in a much narrower redshift range of $\Delta z \sim 0.3$ (~ 167 cMpc at $z = 3.78$) (see Section 3.4), which is ~ 3.3 times smaller than the redshift range covered when broad bands are used, allowing one to minimize line-of-sight projections from physically unassociated galaxies.

This method has never been used before to select LBGs, and the filters used to perform the observations were custom designed to select LBGs at $z \sim 3.78$ centered on the redshift of our six QSO targets. The two NB filters used in this study are NB$_{571}$ ($\lambda_{eff} = 5657$ Å, FWHM = 187 Å), and NB$_{596}$ ($\lambda_{eff} = 5947$ Å, FWHM = 116 Å), which were designed to have a gap between them to exclude the Lyα emission line at $z = 3.78$. Then the galaxy selection is not influenced by the Lyα line-strength, but rather is sensitive to the Lyα break. Additionally, data were collected in the broad-band filter r_{GUNN} ($\lambda_{eff} = 6490$ Å) to help remove low-redshift interlopers.

Table 1
Targeted QSOs Properties

Field	R.A. (J2000)	Decl. (J2000)	Redshift	i	$\log(M_{BH}/M_\odot)$[a]
SDSSJ0124+0044	01:24:03.78	00:44:32.67	3.834	17.99	10.15 ± 0.03
SDSSJ0213−0904	02:13:18.98	−09:04:58.28	3.794	19.03	9.57 ± 0.18
J2003−3300[b]	20:03:24.12	−32:51:45.02	3.773	17.01	9.7
SDSSJ2207+0043	22:07:30.48	00:43:29.37	3.767	19.47	9.13 ± 0.16
SDSSJ2311−0844	23:11:37.05	−08:44:09.56	3.745	20.18	9.41 ± 0.24
SDSSJ2301+0112	23:01:11.23	01:12:43.34	3.788	19.44	8.55 ± 0.80

Notes.
[a] Virial BH masses from Shen et al. (2011).
[b] This QSO was not selected from SDSS, but it was targeted because it belongs to the redshift range of interest. The properties shown here are from McLeod & Bechtold (2009), who do not report the error for the BH mass measurement.
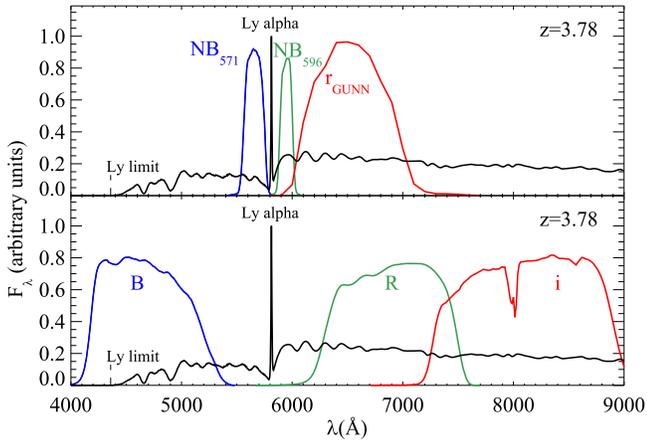
Figure 1. Upper panel: filter configuration used in this study, shown on an LBG simulated spectrum at $z = 3.78$ (see Section 3.1 for the simulated spectrum details). The NBs were designed specially for this program to identify LBGs at $z \sim 3.78$ by detecting the Lyα break. This filter configuration selects galaxies in a quiet narrow redshift slice of $\Delta z \sim 0.3$. Lower panel: example of a filter set used to identify galaxies with the standard Lyman break technique that is based in the detection of the Lyman limit break. The filter curves shown are those used by Ouchi et al. (2004a) to find LBGs at $z \sim 4$ over a redshift slice of $\Delta z \sim 1.0$.

2.3. VLT Imaging and Data Reduction

Imaging observations were acquired on three consecutive nights during 2007 September 9–11, using the FOcal Reducer and low dispersion Spectrograph 1 (FORS1; Appenzeller & Rupprecht 1992) instrument on the Very Large Telescope (VLT). The field-of-view (FOV) of FORS1 is 6.8×6.8 arcmin^2, which corresponds to $\sim 3.0 \times 3.0$ pMpc^2 at $z = 3.78$. The instrument pixel scale is 0.251 arcsec/pixel for images binned 2×2.

Each QSO field was observed in the three filters shown in Figure 1. The total exposure time for the filters was 8000s, 4000s, and 1800s for NB$_{571}$, NB$_{596}$, and r_{GUNN}, respectively. Observations were acquired in shorter individual dithered exposures, in order to fill the gap between the CCDs and to facilitate the data reduction process (cosmic ray and bad pixel rejection, building a superflat, etc.). A spectrophotometric standard star was observed only on the second and third night. The typical seeing during the three nights was 0.6–0.8 arcsec.

Science images were reduced using standard IRAF[6] tasks and our own custom codes written in the Interactive Data

---
[6] Image Reduction and Analysis Facility.

Language (IDL). The reduction process included bias subtraction and flat fielding. As our images exhibited illumination patterns, we performed the flat fielding with superflat images, created using the unregistered science frames. For that, we first masked all the objects out and then combined the science frames with an average sigma-clipping algorithm.

SExtractor (Bertin & Arnouts 1996) was used to create a source catalog for each individual image and then SCAMP (Bertin 2006) was used to compute an astrometric solution, using the SDSS-DR7 r-band star catalogs as the astrometric reference. Finally, the individual images were sky-subtracted, re-sampled, and median-combined using SWarp (Bertin et al. 2002), and then the noisy edges of the combined images were trimmed.

For the flux calibration, we only had observations of the spectrophotometric standard star SA109-949 at the beginning of the last two nights. The tabulated spectrum of this star has a coarse sampling of 25 Å (Stone 1996), which is not suitable when NB filters are used. For the first night, spectrophotometric standard stars were not observed, but we took advantage of two existing SDSS star spectra in one of the fields taken during that night. The coordinates of the stars with available SDSS spectra are R.A.$_{,star1}$ = 21.014, decl.$_{,star1}$ = 0.740872 and R.A.$_{,star2}$ = 21.057, decl.$_{,star2}$ = 0.686577 and the median signal-to-noise ratios (S/Ns) per angstrom of their spectra at the wavelengths of interest were 13.3 and 8.5, respectively.

The flux calibration process was as follows. For the first night calibration we convolved the SDSS star spectra with the three filters' curves in order to obtain standard magnitudes. These magnitudes were compared with the stars' instrumental magnitude (obtained using the MAG_AUTO of SExtractor on the combined science images) to obtain the zero points (ZPs) for each filter. A mean final ZP was computed from the two stars and the typical error for this ZP measurement was ~ 0.08 mag. For the second and third night calibration, we used the spectrum of the observed spectrophotometric star to convolve it only with the broad-band filter curve to obtain the r_{GUNN} ZP. The error in this computation was ~ 0.02 mag. After that, the differential ZPs from the first night were used to determine the NB ZPs for the second and third nights, for which we obtained a typical error of ~ 0.11 mag.

2.4. Photometric Catalogs

Object detection and photometry were performed using SExtractor in dual mode, with the r_{GUNN} image as the detection image. We set the parameters BACK_SIZE and BACK-PHOTO_THICK such that the background was calculated in

Table 2
4σ Limit Magnitudes per Field Measured in a $2''$ Diameter Aperture and
Seeing Measured on the r_{GUNN} Images

Field	NB_{571}	NB_{596}	r_{GUNN}	Seeing [$''$]
SDSSJ0124+0044	26.04	25.51	25.86	0.83
SDSSJ0213−0904	26.18	25.71	25.92	0.89
J2003−3300	26.05	25.44	25.62	0.45
SDSSJ2207+0043	26.03	25.38	25.78	0.53
SDSSJ2311−0844	26.02	25.60	25.84	0.76
SDSSJ2301+0112	26.04	25.55	25.91	0.70

regions of 64 pixels in size and then recomputed locally in an annulus area of 24 pixels of thickness centered around the object. The parameters DETECT_MINAREA and DETECT_THRESH were set such that every group of at least five contiguous pixels having a value above 1.5σ (with σ the background rms) was considered as an object.

In order to ensure an adequate color measurement we needed to carry out photometry in the same object area for the three different filters. Therefore, we convolved our images with a Gaussian kernel to degrade its point-spread function (PSF) to match it with the worst seeing image for each field. Then, the object magnitudes were estimated by the MAG_APER parameter of SExtractor using a fixed aperture of $2''$ diameter. This magnitude is not necessarily the total magnitude of the object, but is used to compute the colors of galaxies. With this choice, if galaxies at $z \sim 4$ are unresolved by the PSF, we are including the flux out to $\sim 3\sigma$ of the object's PSF (for a seeing of $0''.8$). This ensures that we measure the majority of the object's flux, as well as avoid contamination from other close sources. Magnitudes of objects not detected, or detected with S/N < 2 either in NB_{571} or NB_{596}, were assigned the value of the corresponding 2σ limiting magnitude.

Here, the S/N of each object is defined as the ratio of counts in the $2''$ aperture, given by SExtractor, to the rms sky noise in the aperture. This rms sky noise is calculated using an IDL procedure, which performs $2''$ aperture photometry in ~ 5000 different random positions in the image (avoiding the locations of objects) to compute a robust measurement of the mean sky noise. The rms sky noise is calculated as the standard deviation of the distribution of mean values.

Magnitudes were corrected for extinction due to airmass using the atmospheric extinction curve for Cerro Paranal (Patat et al. 2011), and by galactic extinction calculated using the Schlegel et al. (1998) dust maps and extinction laws of Cardelli et al. (1989) with $R_V = 3.1$. The error in the measured magnitude was computed by error propagation, with the object flux error given by the rms noise N in the aperture computed as we described above.

The mean 4σ limiting magnitude of the reduced images was 26.06 for NB_{571}, 25.53 for NB_{596}, and 25.82 for r_{GUNN} for $2''$ diameter apertures. These limiting magnitudes are listed in Table 2 for each field.

For each field, we computed the completeness of the photometric catalogs for the image detection r_{GUNN}. To do this, we linearly fitted the logarithmic magnitude distribution in the magnitude range $21.0 < r_{\mathrm{GUNN}} < 24.5$ where the photometric catalogs are assumed to be 100% complete. We extrapolated the linear fit to fainter magnitudes and measured the completeness as a function of magnitude as the ratio of the histogram relative to that linear fit. We found that at our 4σ limiting magnitude the completeness was on average $\sim 12\%$.

3. LBG Selection at $z = 3.78$

LBG candidates at $z = 3.78$ were selected using the Lyα break technique adapted to our custom filters, which target the Lyα break at $\lambda_{\mathrm{rest-frame}} = (1 + z)\,1216$ Å. Our two NB filters were chosen to bracket this break, and thus we expect that LBGs at $z = 3.78$ will have red colors in $NB_{571} - NB_{596}$. But if we used only this color criteria, we could be including some low-redshift galaxy interlopers in the sample. In order to remove them, a third filter was used to give a measurement of the LBG continuum slope using the $NB_{596} - r_{\mathrm{GUNN}}$ color.

Since the filters used in this study are not standard, the color criteria to select LBGs are unknown. We also do not know what colors low-redshift galaxy contaminants have in this filter system. For this reason, we must explore how galaxies populate the color space in order to select a complete LBG sample while avoiding low-redshift interlopers. Furthermore, in order to perform an LBG clustering analysis in QSO fields we need to know the number density of LBGs expected at random locations in the universe (i.e., in fields not specifically targeting QSOs, also here referred to as "blank fields"). When a standard filter set is used (e.g., LBG selection using broad-band filters), this number density can be computed directly from the LBG luminosity function measured from work using similar filters. However, in our case if we compute the number density from this LBG luminosity function, we have to correct this quantity to take into account the fact that our filter system is mapping a different survey volume and does not necessarily identify all of the LBGs obtained by broad-band selection. Specifically, we need to (a) determine what fraction of LBGs we are detecting at any redshift (i.e., the completeness) and (b) determine the redshift range over which we are selecting LBGs (Δz). Both of these goals can be achieved by performing an accurate computation of the redshift selection function $\phi_z(z)$, defined as the LBG completeness as a function of redshift.

In order to perform the optimal LBG selection and compute $\phi_z(z)$, we conducted detailed simulations to model the distribution of LBG colors in the color-space. In this section we detail how the color modeling was performed, we study what contaminants could be affecting our LBGs selection, and we define color criteria to select LBGs at $z = 3.78$. Finally, we present the redshift selection function providing the completeness as a function of redshift for the sample.

3.1. LBG Color Modeling

We performed a Monte Carlo simulation of 1000 LBG spectra at each redshift, which were created to have different UV continuum slopes and Lyα equivalent widths (EW$_{\mathrm{Ly}\alpha}$), such that they reproduce the space of possible LBG spectra informed by our knowledge of LBG properties.

Each simulated rest-frame spectrum was created in the following way. As a starting point, we considered a template galaxy spectrum generated from Bruzual & Charlot (2003) population synthesis models,[7] corresponding to an

---

[7] Obtained from http://bruzual.org/.
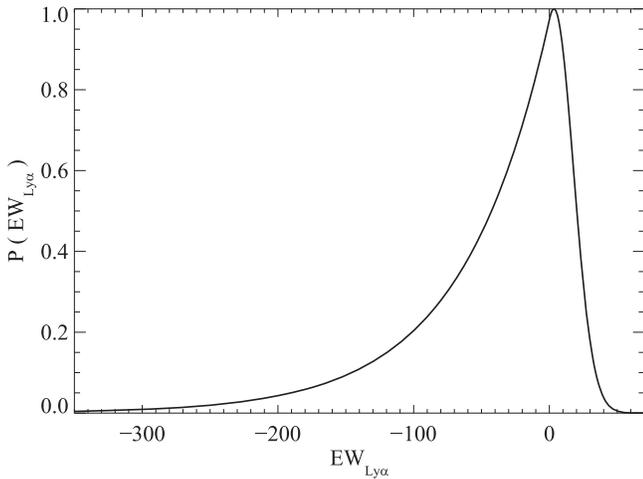
Figure 2. Normalized probability distribution function of $EW_{Ly\alpha}$ used for the simulated spectra, where negative values correspond to emission lines. $EW_{Ly\alpha}$ was chosen from a Gaussian distribution with rest-frame mean $\mu = -25$ Å and $\sigma = 40$ Å (Shapley et al. 2003) plus an exponential tail of high $EW_{Ly\alpha}$ values with scale length of $W_0 = -64$ Å (Ciardullo et al. 2012).

instantaneous burst model with an age of 70 Myr, a Chabrier (2003) IMF, and a metallicity of $0.4\,Z_\odot$, as expected for LBGs at $z \sim 4$ (Jones et al. 2012). We assumed a power-law UV continuum for this template with amplitude A and a slope α_{BC}, such that we modeled its flux as $F_{BC}(\lambda) = A\lambda^{\alpha_{BC}}$. We fit this model to the template spectrum over the UV continuum range (here defined as $1300\,\text{Å} < \lambda < 2000\,\text{Å}$) by least-squares minimization to obtain the best fit A and α_{BC} parameters.

First, we modified the UV slope of this template by multiplying its flux by $\lambda^{\alpha - \alpha_{BC}}$ in order to obtain a spectrum with a power-law UV continuum given by $A\lambda^\alpha$. The new slope α was chosen as a value taken randomly from a Gaussian distribution with mean $\mu = -1.676$ and $\sigma = 0.39$. These values are motivated by Bouwens et al. (2009), who presented the UV continuum slope distribution of LBGs at $z \sim 4$ for samples selected in different magnitude ranges.

Second, we added a Gaussian $Ly\alpha$ line with rest-frame central wavelength $\lambda_{Ly\alpha} = 1215.7$ Å, standard deviation $\sigma_{Ly\alpha}$ and amplitude B, which adjusts the intensity of the line. For all the simulated spectra we used a fixed $\sigma_{Ly\alpha} = 1$ Å, which agrees with the $\sigma_{Ly\alpha}$ of the composite spectrum of LBGs at $z \sim 4$ (Jones et al. 2012). The B value was adjusted to model a $Ly\alpha$ line with an $EW_{Ly\alpha}$ value drawn randomly from a distribution chosen to agree with observations of LBGs. The $EW_{Ly\alpha}$ distribution was given by a Gaussian core plus a tail to large negative EWs to represent strong line emitters. For the Gaussian core we adopted a mean $\mu = -25$ Å and standard deviation $\sigma = 40$ Å (rest-frame), based on the measurements of Shapley et al. (2003), who studied the spectra of 811 LBGs at $z \sim 3$. We thus assume that the Gaussian core of the LBG $EW_{Ly\alpha}$ distribution does not evolve significantly from $z \sim 3$ to $z \sim 4$. For the tail representing strong line emitters, we modified the Gaussian by adding an exponential function with rest-frame $EW_{Ly\alpha}$ scale length of $W_0 = -64$ Å, as presented in Ciardullo et al. (2012). In this way our model of line emission encompasses both LBG and LAE spectra. Figure 2 shows the $EW_{Ly\alpha}$ probability distribution function used to simulate our
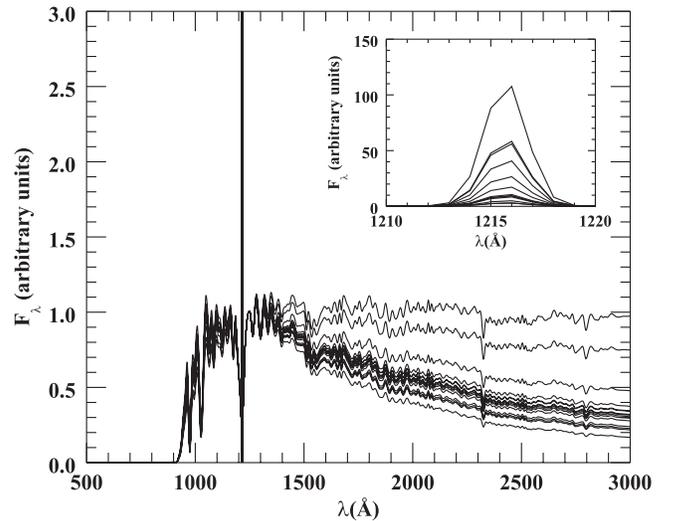
Figure 3. Example of ten rest-frame simulated spectra using our Monte Carlo simulation. The spectra have been normalized to have the same flux value at $\lambda = 1245$ Å. The subplot in the upper right corner shows a zoom-in of the region of the $Ly\alpha$ line.

spectral models. The $EW_{Ly\alpha}$ are defined as

$$EW_{Ly\alpha} = -\int \frac{F_{Ly\alpha}}{F_{cont}} d\lambda, \qquad (1)$$

where $F_{Ly\alpha}$ is the flux of the $Ly\alpha$ line (with the continuum subtracted), which is given by a Gaussian with amplitude B, as we described above, and F_{cont} is the flux of the continuum given by $A\lambda^\alpha$. Note that we defined negative values of $EW_{Ly\alpha}$ for emission lines and positive for absorption lines.

Once α and $EW_{Ly\alpha}$ were chosen for a given simulated spectrum, we dust-attenuated it using the starburst reddening curve from Calzetti et al. (2000) and adopted a color excess value of $E(B - V) = 0.16$ according to the values estimated for LBGs at $z \sim 3$ (Shapley et al. 2003).

After the dust-attenuation was applied, we modeled the fact that only a small fraction of Lyman limit photons escape LBGs with an escape fraction parameter $f_{esc}^{\lambda < 912}$. Although this value is observationally poorly constrained, studies suggest it is in the range 0.04–0.14 (Fernández-Soto et al. 2003; Ouchi et al. 2004a; Shapley et al. 2006). We assumed a fixed value of $f_{esc}^{\lambda < 912} = 0.05$, and multiplied the spectrum at $\lambda \leqslant 912$ Å by this value. We also tested our results using different values of $f_{esc}^{\lambda < 912}$, finding that the colors of simulated galaxies are relatively insensitive to the exact value of $f_{esc}^{\lambda < 912}$ used, because these wavelengths are subsequently significantly attenuated by the IGM transmission function (see below).

Finally, we redshifted each model spectrum to different redshifts on a grid with a grid spacing of 0.02 and ranging from $z = 3.2$ to $z = 4.4$. In the redshifting process we used the IGM transmission model $T_z(\lambda)$ for the corresponding redshift z from Worseck & Prochaska (2011) to attenuate the flux blueward of the $Ly\alpha$ line.[8] Note that in principle we should attenuate both the continuum blueward of the $Ly\alpha$ line and the line itself; however, the $EW_{Ly\alpha}$ values used in this simulation are taken from the literature, which are observed values that are not corrected for IGM attenuation, such that this line emission is effectively already attenuated. In Figure 3 we show some

---

[8] Kindly provided to us by G. Worseck.
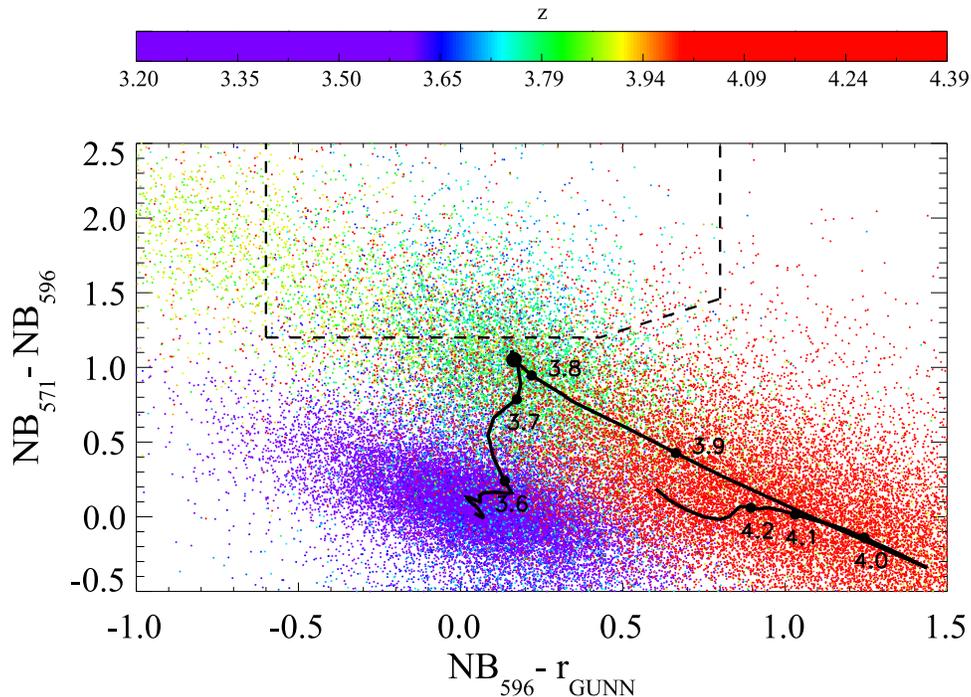
Figure 4. Color–color diagram showing the simulated colors for 1000 LBG spectra, plotted as redshift color-coded points according to the color bar. The median LBG evolutionary track is plotted as a black curve. The filled points over this curve indicate the median LBG colors at different redshift ranging from 3.6 to 4.2. The largest circle shows the exact position of the median $z = 3.78$ LBG colors. The dashed line indicates the selection region used to select LBGs according to Equation (2).

examples of our rest-frame simulated spectra, which have been normalized to have the same flux at $\lambda = 1245\,\text{Å}$.

At each redshift, we integrated the spectra against our three filter transmission curves to obtain the fluxes and then the LBG colors. In order to model the impact of noise, we added photometric errors to the simulated LBG photometry. To this end we first assigned an r_{GUNN} magnitude to each simulated object by randomly drawing a value from the $z \sim 4$ LBG luminosity function, integrated over the same magnitude range as our LBG sample ($24.0 \leqslant r_{GUNN} \leqslant 25.6$ or $0.76 \leqslant L/L_* \leqslant 3.5$; see Section 3.3).[9] We also weighted the luminosity function by the completeness of the source detection at each apparent magnitude and for each field (computed in Section 2.4), which takes into account the fact that the fraction of sources detected depends on their magnitude, such that the photometric catalog is complete for bright sources but less complete at the faint end. In this way the incompleteness of our photometry is also factored into our color modeling.

Based on the simulated LBG colors and the chosen r_{GUNN} value, we then determined the magnitude in the other two filters NB_{571} and NB_{596} for each spectrum in each redshift bin. In order to construct a noise model, we selected a galaxy sample from our photometric catalogs, and computed the median magnitude error as a function of the magnitude for each filter (with the magnitude error computed as we explained in Section 2.4). Finally, we assigned random Gaussian-distributed magnitude errors using our median relations, and then added this noise to the model photometry that defined the final photometry of the simulated spectra. The colors for the 1000 simulated spectra at each redshift are shown in Figure 4. We

also computed the median of our 1000 rest-frame Monte Carlo spectra, redshifted it, and obtained the colors at each redshift to compute the median evolutionary track of LBG colors, shown as the black solid line in Figure 4.

Figure 4 indicates that the median colors of LBGs at $z = 3.78$ are $NB_{571} - NB_{596} = 1.05$, and $NB_{596} - r_{GUNN} = 0.16$. However, if we consider the intrinsic scatter in LBG properties (continuum slope and $EW_{Ly\alpha}$) and photometric uncertainties, the $z \simeq 3.78$ LBGs (indicated by green points) span a wider color range with $NB_{571} - NB_{596} \gtrsim 0.5$ and $-0.6 \lesssim NB_{596} - r_{GUNN} \lesssim 0.8$. In principle, we should select LBGs in this broad selection region to obtain a highly complete sample; however, we also need to take into account the colors of low-redshift galaxies in our filter system to define the final selection criteria. We perform this analysis in Section 3.2, where we also test our LBG color modeling by reproducing the LBG evolutionary track presented in previous work using broad-band LBG selection.

3.2. Low-redshift Galaxy Colors

We use template galaxy spectra to develop a basic understanding of how low-redshift galaxies populate the color–color diagram in our new filters. We use a set of five commonly used templates for estimating photometric redshifts, such that they span the range of galaxy spectral energy distributions (SEDs). The templates are from the photo-z code EASY (Brammer et al. 2008), which are distilled from the PEGASE spectral synthesis models, and correspond to elliptical, Sa, Sb, Sc, and irregular galaxy spectra.

We redshifted these template spectra from $z = 0$ to $z = 3$, and integrated them over our filter transmission curves to generate their evolutionary track. Note that we need not attenuate these spectra by the IGM transmission function $T_z(\lambda)$, since our NB filters never cover rest-frame wavelengths lower than 1216 Å for the low redshifts considered. In Figure 5 we

---

[9] Given that for each field we reached slightly different limiting magnitudes, we simulated the LBG photometry field by field according to their respective r_{GUNN} limiting magnitudes. This resulted in a slightly different redshift selection function for each field.
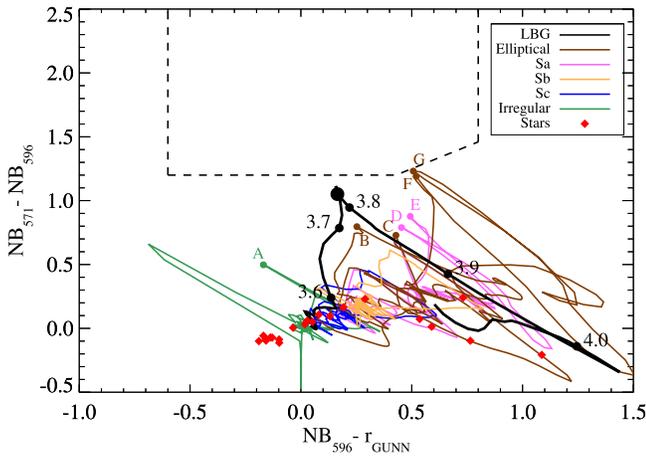
Figure 5. Evolutionary tracks of low-redshift galaxies redshifted from $z = 0$ to $z = 3$. We plot as brown, magenta, orange, blue, and green curves the evolutionary track of elliptical, Sa, Sb, Sc, and irregular galaxies, respectively. We overplot the track of LBGs computed as was explained in Section 3.1 as a black curve. Filled circles over the black curve indicate colors of LBGs from redshift 3.6 to 4.0, and the largest black point indicates the exact position of the color of LBGs at $z = 3.78$. Filled circles labeled with letters over the low-redshift galaxies evolutionary tracks indicate the colors of some contaminants that could be affecting our selection: galaxies at $z = 0.60$ (A), $z = 1.83$ (B), $z = 0.45$ (C), $z = 1.23$ (D and F), and $z = 1.04$ (E and G). Red diamonds show the mean colors of 23 different type of stars, including O, B, A, F, G, K, M, and L-type stars as well as white dwarfs, carbon, and K-type subdwarf stars. The dashed line indicates the selection region used to select LBGs according to Equation (2).

Figure 6. Same as Figure 5 but using the filter system used by Ouchi et al. (2004a) to select LBGs at $z \sim 4$ (broad-band filters B, R, i). Filled circles over the black curve indicate colors of LBGs from redshift 3.2 to 3.9. Filled circles over the curve of the low-redshift galaxies indicate colors from redshift 0.0 to 3.0. The dashed line indicates the region used Ouchi et al. (2004a) to select LBGs in their work.

show the evolutionary tracks for different galaxy types together with the median LBG evolutionary track that we computed in Section 3.1. We also show the mean colors of stars in our diagram, which were computed by using SDSS template spectra of 23 different stars,[10] including O, B, A, F, G, K, M, and L-type stars as well as white dwarfs, carbon, and K-type subdwarf stars.

In order to test our Monte Carlo simulation as well as the evolutionary tracks for low-redshift galaxies, we used our 1000 simulated spectra at each redshift to compute the median LBG evolutionary track in the standard *BRi* filter set used to select LBGs at $z \sim 4$ (see Figure 1) by Ouchi et al. (2004a). We also computed the evolutionary track of these low-redshift galaxies in the standard LBG filters in the same way as described above. These results are shown in Figure 6, where we also overplot the selection region used by Ouchi et al. (2004a) to select $z \sim 4$ LBGs. We find that the median LBG evolutionary track from our Monte Carlo model lies within the Ouchi et al. (2004a) selection region, and selects LBGs at $z \gtrsim 3.5$ as claimed. Note also that our LBG evolutionary track agrees well with the Ouchi et al. (2004a) evolutionary track (see Figure 4 of their paper) determined from a much simpler model of LBG spectra and IGM transmission. In addition we see that the evolutionary tracks of low-redshift galaxies lie comfortably outside the *BRi* LBG selection region as claimed by Ouchi et al. (2004a).

However Figure 5 shows that in our NB filter set, some of the low-redshift galaxies have similar colors as $z = 3.78$ LBGs, which suggests that our new filter configuration could make it challenging to select a sample of LBGs at $z = 3.78$ with high completeness and at the same time high purity. When we use NB filters the low-redshift galaxy colors are located in a wider

region in the color–color plot in comparison with the location of the color locus of contaminants when broad-band filters are used. We attribute this to sensitivity of the NB filters to features in the galaxy spectra such as emission or absorption lines. In the case of broad bands these features are diluted by averaging over large regions of spectra, but for NB the features result in large excursions in color with changing redshift, making the low-redshift galaxy locus extend over a larger region of color space that overlaps with the colors of $z = 3.78$ LBGs.

Given that LBG colors at $z = 3.78$ span the range $NB_{571} - NB_{596} \gtrsim 0.5$ and $-0.6 \lesssim NB_{596} - r_{GUNN} \lesssim 0.8$ (see Figure 4), there are several types of contaminants that could be affecting our LBG selection. Their colors are indicated by points on the respective low-redshift galaxy evolutionary tracks labeled by letters in Figure 5, and some examples are shown in Figure 7. The first type are red galaxies at $z \sim 0.45$ having a large $\lambda_{RF} \sim 4000$ Å Balmer break and strong calcium H and K absorption. This break is located just between our two NBs, so they present red colors (point C on the brown curve in Figure 5). The second type of interlopers are star-forming galaxies at $z \sim 0.60$ with strong [O II] 3727 Å emission lines. If NB_{596} is located just over this line, and NB_{571} over the continuum, we again detect red colors (point A on the green curve in Figure 5). The third type of interlopers are galaxies at $z \sim 1.04$ with strong Mg I and Mg II absorption lines at $\lambda_{RF} = 2852$ Å, and $\lambda_{RF} = 2799$ Å, respectively, in combination with the $\lambda_{RF} \sim 2900$ Å break. When NB_{571} is located over this absorption and NB_{596} falls on the continuum, then red colors are detected (points E and G on the magenta and brown curves, respectively, in Figure 5). Other interlopers are galaxies with strong flux breaks redshifted just between our NB filters. Examples are galaxies at $z \sim 1.23$ with a large $\lambda_{RF} \sim 2640$ Å break (points D and F on the magenta and brown curves, respectively, in Figure 5) and galaxies at $z \sim 1.83$ with a strong break at $\lambda_{RF} \sim 2085$ Å (point B on the brown curve in Figure 5).

3.3. Selection Region and LBG Sample

As we are interested in measuring the clustering properties of LBGs at $z = 3.78$, we need to select a sample with high

---

[10] Obtained from http://classic.sdss.org/dr5/algorithms/spectemplates/.
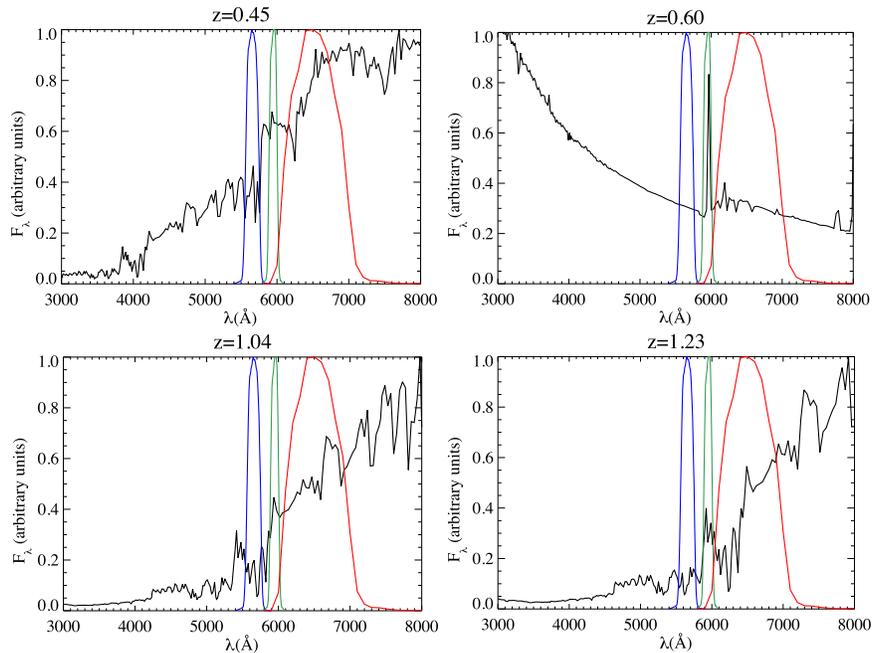
Figure 7. Examples of interlopers that could affect our LBGs selection. We show the galaxy spectra and the position of our three filters over it. Top left panel: the spectra of an elliptical galaxy at $z = 0.45$, with a strong Balmer break located at $\lambda_{obs} = 5840$ Å and intense calcium H and K absorption. Top right panel: the spectra of a galaxy at $z = 0.60$, with intense O II emission line at $\lambda_{obs} = 5925$ Å. Bottom left panel: the spectra of a galaxy at $z = 1.04$, with Mg I and Mg II absorption at $\lambda_{obs} \sim 5650$ Å. Bottom right panel: the spectra of a galaxy at $z = 1.23$, with a large break at $\lambda_{obs} = 5887$ Å.

completeness and purity. In order to avoid low-redshift contaminants, we were forced to choose a smaller selection region in the color–color diagram, which results in relatively low completeness, but it ensures that the sample is not highly contaminated.

First, we defined two vertical color cuts in Figure 4, one to the left of the median LBG colors at $z = 3.78$ and one to the right. The first cut is meant to exclude LBGs located in the upper left region of the diagram, which mostly corresponds to LBGs at $z \sim 3.9$ with strong Lyα line emission. The second cut avoids LBGs at $z > 3.9$. A third color cut defines a lower limit for $NB_{571} - NB_{596}$, which ensures we are detecting the Lyα break, while at the same time avoiding LBGs at $z \lesssim 3.7$. We used a diagonal color cut to most effectively avoid the contamination of low-redshift galaxies (see Figure 5), while at the same time including most of the LBGs at $z = 3.78$, thus maintaining the highest completeness possible.

We also tested several different color criteria to select LBGs. In Section 5.2 we will further discuss our color selection, contamination by low-redshift galaxies, and the impact that contamination can have on our clustering measurements. There we argue that the choice of color selection that we present here selects a reasonably complete LBG sample (see Section 3.4 for details about the completeness of the sample) with high purity. Our final set of color cuts is shown in Figure 4, and defined by the following relations:

$$NB_{571} - NB_{596} > 1.2$$
$$-0.6 < NB_{596} - r_{GUNN} < 0.8$$
$$NB_{571} - NB_{596} > 0.7(NB_{596} - r_{GUNN}) + 0.9. \quad (2)$$

We selected LBGs based on our galaxy photometry, but required sources to have $S/N \geqslant 4.0$ in both the NB_{596} and r_{GUNN} filters, to ensure a solid detection of the LBG continuum. In order to reduce contamination by false detections, we only considered objects that have FLAGS = 0 in SExtractor, which

excluded objects that were blended, saturated, truncated (too close to an image boundary), or affected by very bright neighboring objects. Bright stars in our images were masked in order to avoid spurious object detection due to contamination from their stellar flux. This procedure resulted in a set of masks indicating where we were able to detect galaxies, which we use later in our clustering analysis to compute the effective area of our survey.

We also imposed a lower limit on the magnitude in order to exclude bright, low-redshift interloper galaxies from our selection. Thus we only considered objects with magnitudes fainter than $r_{GUNN} = 23.97$, corresponding to LBGs with $L \sim 3.5 L_*$. We chose this value by computing the LBG luminosity function at $z \sim 4$, and finding the bright-end cut at which we would lose no more than 1% of the galaxies. In other words, 99% of the total number of LBGs have magnitudes between our bright-end cut of $r_{GUNN} = 24.0$ and the limiting magnitude $r_{GUNN} = 25.82$ (mean limit magnitude at 4σ for a $2''$ diameter aperture) of our images, which corresponds to $L = 0.76 L_*$. In this way we can safely assume we are excluding only extremely rare, bright LBGs. For the LBG luminosity function we used the Schechter parameters from Ouchi et al. (2004a) who studied the photometric properties based on a large sample of $\sim$2200 LBGs at $z \sim 4$. The values used are $\phi^* = 2.8 \times 10^{-3} h_{70}^3$ Mpc^{-3}, $M_{1700}^* = -20.6$ mag, and $\alpha = -1.6$.

Given all of these selection criteria and the color cuts defined in Equation (2), we selected LBGs in each of our fields. We compute the total area of our survey by adding the effective area of each individual field, which is defined by subtracting the masked area from the total area of the image. The the total area of our survey is 232.7 arcmin^2 corresponding to an average area per image of 38.79 arcmin^2 (recall the FOV of FORS1 is 6.8×6.8 arcmin^2 or 46.24 arcmin^2). We show color–color diagrams of objects detected in all six of our fields
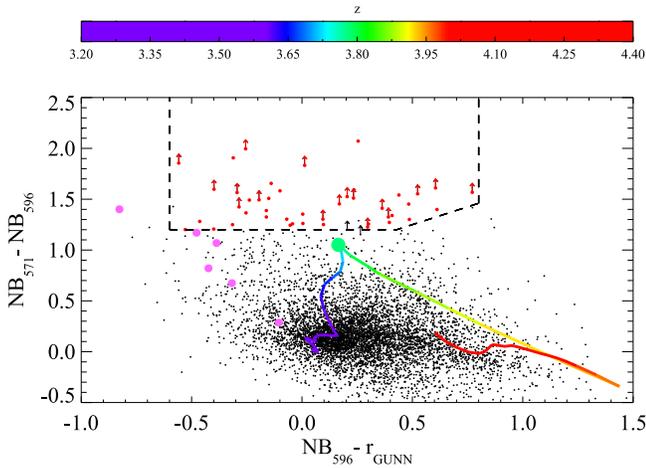
Figure 8. Color–color diagram for the six stacked QSO fields. Here the evolutionary track shown in Figure 5 is plotted as redshift color-coded track according to the color bar. We have highlighted the selected LBGs as red points. The magenta points indicate the color of each QSO in the filters. Arrows indicate lower limits for $NB_{571} - NB_{596}$ color. These are cases in which the object was not detected in NB_{571} filter at the 2σ level and magnitude was replaced by the corresponding limit magnitude.

in Figure 8. We found a total of 44 LBGs (see Table 3) corresponding to a mean number density of 0.19 LBGs arcmin^{-2}. Image cutouts in our three filters for several of our selected LBGs are shown in Figure 9. In Figure 10 we show the spatial distribution of the LBGs relative to the QSO (red dot at zero) for our six fields. We also show the individual color–color diagrams and indicate the number of LBGs found in each individual field in Figure 11. Note that the number of LBGs in the fields cannot be directly compared because each image has different limiting magnitude and different effective area (different reduced image size, masked region, etc.). In Figure 12 we show a false-color image of the field around QSO SDSS J2301+0112 with the LBG candidate positions indicated.

3.4. Redshift Selection Function

We used the LBG color modeling machinery described in Section 3.1 to compute the redshift selection function $\phi_z(z)$ of our LBG color-selection criteria. At each redshift step, we redshift the 1000 rest-frame simulated LBG spectra into the observed frame, draw luminosities from the luminosity function, compute magnitudes and colors, and add photometric errors. We then compute the completeness at each redshift by calculating the fraction of simulated LBGs that satisfy the selection criteria defined in Section 3.3, namely: fulfill the color criteria in Equation (2), and fulfill the magnitude constraints (given by the 4σ limiting magnitudes for NB_{596} and r_{GUNN} and by the bright-end cut imposed for our selection, $r_{GUNN} > 24.0$). Note that, as the limiting magnitudes of our fields are slightly different, we compute $\phi_z(z)$ for each individual field, using their corresponding NB_{596} and r_{GUNN} limiting magnitudes. The final $\phi_z(z)$ varied from field to field by a small amount, then we computed the median of $\phi_z(z)$ over the six fields, which is shown in Figure 13.

From Figure 13 we see that our selection criteria recover a peak value of $\simeq 26\%$ of LBGs at $z \sim 3.78$ over a small $\Delta z \sim 0.04$ interval. Our criteria also select $\gtrsim 10\%$ of LBGs over a wider redshift interval, ranging from $z \sim 3.65$ to $z \sim 3.95$ ($\Delta z \sim 0.3$), which corresponds to $\sim 18,800$ km s^{-1}

($\sim 167\ h^{-1}$ cMpc) at $z = 3.78$. The NB technique selects LBGs over a much narrower redshift range compared to broad-band LBG selection, which typically selects galaxies over a range $\Delta z \sim 1.0$ (e.g., Ouchi et al. 2004a; Bouwens et al. 2007, 2010), or ~ 3.3 times larger than our selection. However, as low-redshift galaxies have similar colors as the LBGs in our filters, we have to adopt relatively conservative color cuts, making our completeness relatively low.

4. Clustering Analysis

In this section we analyze the clustering of LBGs around our QSOs at $z = 3.78$. First, we present the measurement of the QSO–LBG cross-correlation function in Section 4.1. We then estimate the correlation function parameters (r_0, γ) assuming a power-law form $\xi(r) = (r/r_0)^{-\gamma}$. Our results are compared with theoretical expectations based on the auto-correlation of both LBGs and QSOs at $z \sim 4$ assuming deterministic bias. In Section 4.2 we present the LBG auto-correlation computed from our QSO fields, and finally in Section 4.3 we compare our results with previous measurements.

4.1. The QSO–LBG Cross-correlation Measurement

Following convention, we study clustering using the two-point correlation function $\xi(r)$, which measures the excess probability over a random distribution of finding an object at separation r from another randomly selected object, in a volume element dV (Peebles 1980). For the case of galaxies around QSOs this correlation function is defined by

$$dP = n_G[1 + \xi_{QG}(r)]dV, \qquad (3)$$

where $\xi_{QG}(r)$ is the QSO–galaxy cross-correlation function and n_G is the mean number density of galaxies in the universe. Here r is real space comoving distance, which is however not the observable even when redshift information is available, as peculiar velocities induce redshift space distortions along the line of sight (Sargent & Turner 1977). Typically LBG clustering studies that lack redshifts measure the angular correlation function (e.g., Giavalisco et al. 1998; Ouchi et al. 2004b; Lee et al. 2006). Although we do not have redshifts of our LBGs, our NB selection technique selects LBGs over a narrow redshift interval $\Delta z \simeq 0.3$ (see Figure 13), which allows us to measure clustering as a function of transverse comoving distance instead of angular distance (at $z = 3.78$, the angular diameter distance changes by just $\simeq 3\%$ over this redshift interval). Thus we write the real space separation r as $r^2 = R^2 + Z^2$, where R is the transverse comoving distance between the QSO and the galaxy, and Z is the radial comoving distance between them, approximately given by

$$Z = \frac{c}{H(z)}\Delta z, \qquad (4)$$

where $H(z)$ the Hubble constant evaluated at redshift $z = 3.78$, which we take to be a constant over the redshift interval considered (an approximation valid to $\simeq 5\%$).

We measure the volume-averaged projected cross-correlation function between QSOs and LBGs $\chi(R_{min}, R_{max})$, which is a dimensionless quantity defined as the real space QSO–LBG cross-correlation function $\xi_{QG}(R, Z)$ integrated over a volume

Table 3
LBG Sample

ID	R.A. (J2000)	Decl. (J2000)	r_{GUNN}	NB_{571}	NB_{596}
SDSSJ0124+0044_1	20.98541	0.73838	24.58	>26.80	25.19
SDSSJ0124+0044_2	21.06438	0.73190	25.50	>26.80	24.94
SDSSJ0124+0044_3	21.00865	0.73119	25.24	26.35	24.98
SDSSJ0124+0044_4	21.06237	0.73103	25.10	26.70	24.79
SDSSJ0124+0044_5	21.01586	0.75937	25.23	26.58	25.08
SDSSJ0124+0044_6	21.01454	0.74763	24.69	26.70	25.30
SDSSJ0213−0904_1	33.31527	−9.13219	25.09	>26.94	25.10
SDSSJ0213−0904_2	33.33407	−9.13154	24.49	26.32	24.73
SDSSJ0213−0904_3	33.29456	−9.13105	24.50	26.43	24.98
SDSSJ0213−0904_4	33.34291	−9.12850	25.16	>26.94	25.53
SDSSJ0213−0904_5	33.37631	−9.12752	24.60	25.41	24.20
SDSSJ0213−0904_6	33.29678	−9.12806	25.54	>26.94	25.63
SDSSJ0213−0904_7	33.28357	−9.12428	25.22	>26.94	25.61
SDSSJ0213−0904_8	33.33103	−9.07411	25.12	26.79	25.52
SDSSJ0213−0904_9	33.29552	−9.10705	25.39	26.61	25.35
SDSSJ0213−0904_10	33.31107	−9.05450	25.35	26.51	25.19
SDSSJ0213−0904_11	33.38310	−9.05430	24.88	26.62	25.28
J2003−3300 _1	300.85398	−32.85834	25.06	26.29	24.99
SDSSJ2207+0043_1	331.82678	0.66829	25.93	26.75	25.47
SDSSJ2207+0043_2	331.90688	0.66930	25.16	>26.78	25.33
SDSSJ2207+0043_3	331.90330	0.73064	25.04	>26.78	24.79
SDSSJ2207+0043_4	331.91995	0.71966	24.71	26.69	25.14
SDSSJ2207+0043_5	331.89738	0.69250	24.03	24.70	23.50
SDSSJ2207+0043_6	331.92999	0.68477	25.01	26.49	24.91
SDSSJ2301+0112_1	345.32769	1.16875	25.23	26.49	25.23
SDSSJ2301+0112_2	345.28149	1.21762	25.05	>26.78	25.25
SDSSJ2301+0112_3	345.34116	1.21677	25.26	26.45	25.21
SDSSJ2301+0112_4	345.28836	1.20452	25.48	>26.78	25.28
SDSSJ2301+0112_5	345.28318	1.19689	24.70	>26.78	25.22
SDSSJ2301+0112_6	345.30258	1.17353	24.49	25.74	24.25
SDSSJ2301+0112_7	345.33532	1.23626	24.91	26.26	25.01
SDSSJ2301+0112_8	345.27685	1.23673	24.16	25.51	24.16
SDSSJ2311−0844_1	347.89595	−8.70962	25.69	26.63	25.38
SDSSJ2311−0844_2	347.92391	−8.72311	24.45	>26.79	25.22
SDSSJ2311−0844_3	347.91688	−8.72487	25.65	>26.79	25.36
SDSSJ2311−0844_4	347.90256	−8.73469	25.05	>26.79	25.28
SDSSJ2311−0844_5	347.93345	−8.73878	24.72	26.24	24.59
SDSSJ2311−0844_6	347.92236	−8.74192	25.52	>26.79	25.22
SDSSJ2311−0844_7	347.94311	−8.74245	25.27	>26.79	25.56
SDSSJ2311−0844_8	347.94631	−8.75457	24.71	26.50	25.20
SDSSJ2311−0844_9	347.93114	−8.76500	25.59	>26.79	25.19
SDSSJ2311−0844_10	347.91475	−8.72482	24.32	26.65	24.58
SDSSJ2311−0844_11	347.89925	−8.72574	24.30	25.86	24.60
SDSSJ2311−0844_12	347.91793	−8.73420	25.37	26.60	25.21

Note. The magnitudes correspond to AB magnitudes measured in a 2″ diameter aperture for each band.

and then normalized by it (e.g., Hennawi et al. 2006, 2015)

$$\chi(R_{\min}, R_{\max}) = \frac{\int \xi_{QG}(R, Z) dV_{\mathrm{eff}}}{V_{\mathrm{eff}}}, \qquad (5)$$

where V_{eff} is a cylindrical volume defined by the radial bin $[R_{\min}, R_{\max}]$, the height Z probed by our filter configuration, and modulated by the selection function of our survey. We measure $\chi(R_{\min}, R_{\max})$ at scales $5.0 < \theta < 350''.0$ (corresponding to $0.124 < R < 8.706 \ h^{-1}$ cMpc) in logarithmically spaced radial bins centered on the QSO for all fields using the estimator

$$\chi(R_{\min}, R_{\max}) = \frac{\langle QG \rangle}{\langle QR \rangle} - 1, \qquad (6)$$

where $\langle QG \rangle$ and $\langle QR \rangle$ are the number of QSO–LBG and QSO–random pairs in this cylindrical volume. The quantity $\langle QG \rangle$ is directly measured by counting the QSO–LBG pairs found in our images.

The quantity $\langle QR \rangle$ is the expected random number of QSO–LBG pairs, which is given by

$$\langle QR \rangle = n_G(z, r_{\mathrm{GUNN}}^{\mathrm{lower}} < r_{\mathrm{GUNN}} < r_{\mathrm{GUNN}}^{\mathrm{limit}}) V_{\mathrm{eff}}, \qquad (7)$$

where V_{eff} is the effective volume of the radial bin in question and $n_G(z, r_{\mathrm{GUNN}}^{\mathrm{lower}} < r_{\mathrm{GUNN}} < r_{\mathrm{GUNN}}^{\mathrm{limit}})$ is the mean number density of LBGs at redshift z in the magnitude range of our survey, which will be henceforth referred to as n_G to simplify notation. Given that galaxy clustering measurements are normally performed in random locations of the universe, the
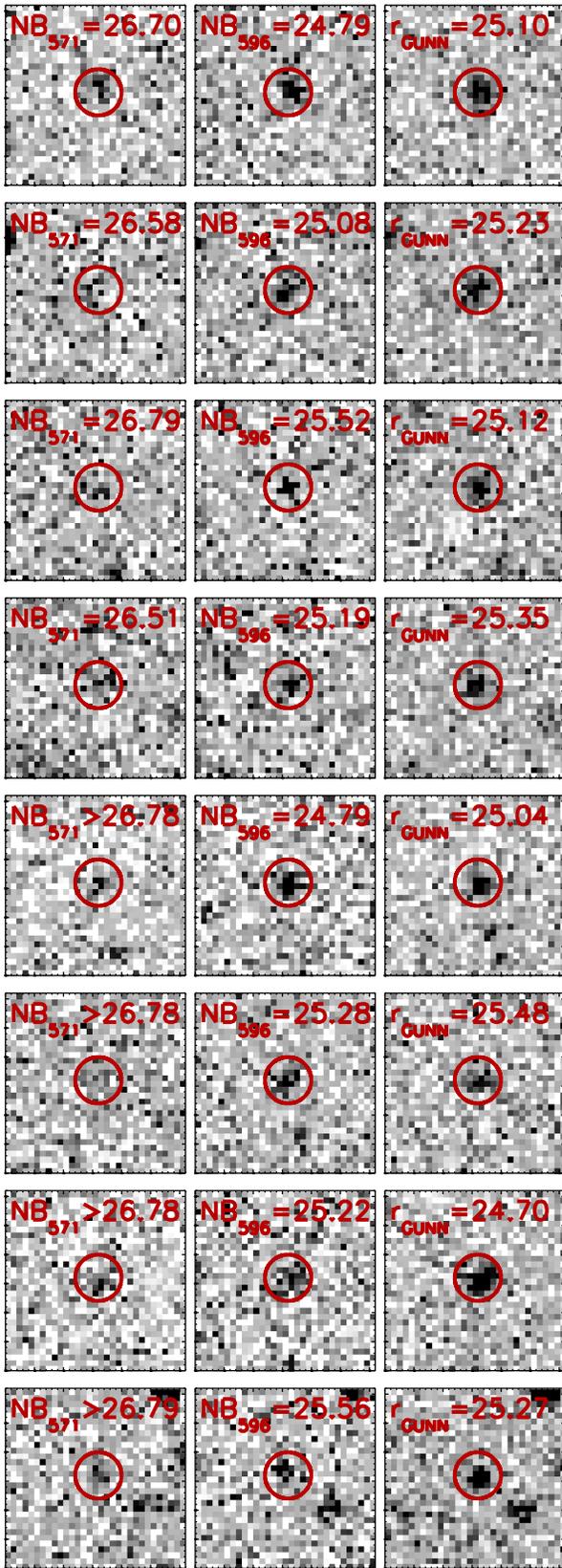
Figure 9. Images of some selected LBGs. From left to right we show the NB_{571}, NB_{596}, and r_{GUNN} images. Each panel is $7\rlap{.}{''}5$ on a side. The red circles show the position of the detected object, and its size corresponds to the region in which the photometry was done ($2''$ in diameter). The magnitudes are indicated in each panel.
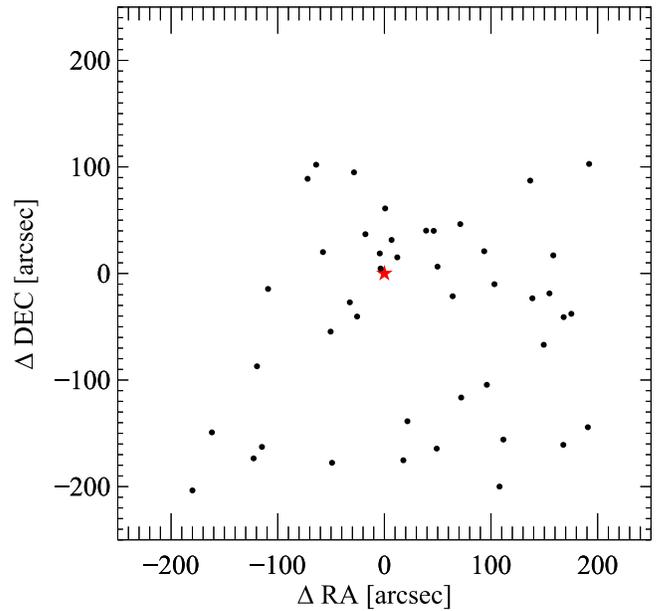
Figure 10. Distribution of LBGs around the QSO in the plane of the sky for the six stacked fields. The central QSO is located at 0.0 and is plotted by a red star.

mean number density measured from the survey is typically a good proxy for the mean number density of the universe, provided the survey volume is large enough. In such cases, the galaxy number counts for the random sample can be computed from the data themselves, and one typically constructs random catalogs with a number density determined from the survey to estimate $\langle QR \rangle$. However, in our case we are pointed toward a QSO situated in what is likely to be an overdensity. Therefore the mean number density of galaxies in our survey is not representative of the mean in random locations and we cannot follow the standard procedure for computing $\langle QR \rangle$.

If we had observations of control fields (i.e., not centered on QSOs) with our same filter configuration, then it would be possible to measure the background number density of LBGs directly and determine $\langle QR \rangle$. An alternative would be to measure this quantity directly from the outer parts of images, where the clustering becomes negligible, given a sufficiently large FOV instrument. Unfortunately, we do not have images of control fields, and the FOV of FORS1 is too small to provide a reliable measurement of the background. Thus our only alternative is to estimate $\langle QR \rangle$ from Equation (7), where n_G is calculated from the $z \sim 4$ LBG luminosity function, and the effective volume $V_{\rm eff}$ is determined from our Monte Carlo simulations of our selection function (see Section 3.4) and the effective area covered by our survey. We provide further details of these computations in what follows.

To calculate n_G we used the Schechter parameters from Ouchi et al. (2004a) $z \sim 4$ LBG luminosity function. We integrated this luminosity function over the magnitude limits given by our LBG selection, and this magnitude integral was weighted by the photometric completeness fraction of our source detection following the same procedure described in Section 3.1. Given that our fields all have slightly different limiting magnitudes and different source completeness, we computed the expected n_G for each field. We assumed that n_G is constant over the redshift ranged considered, which is a good

12
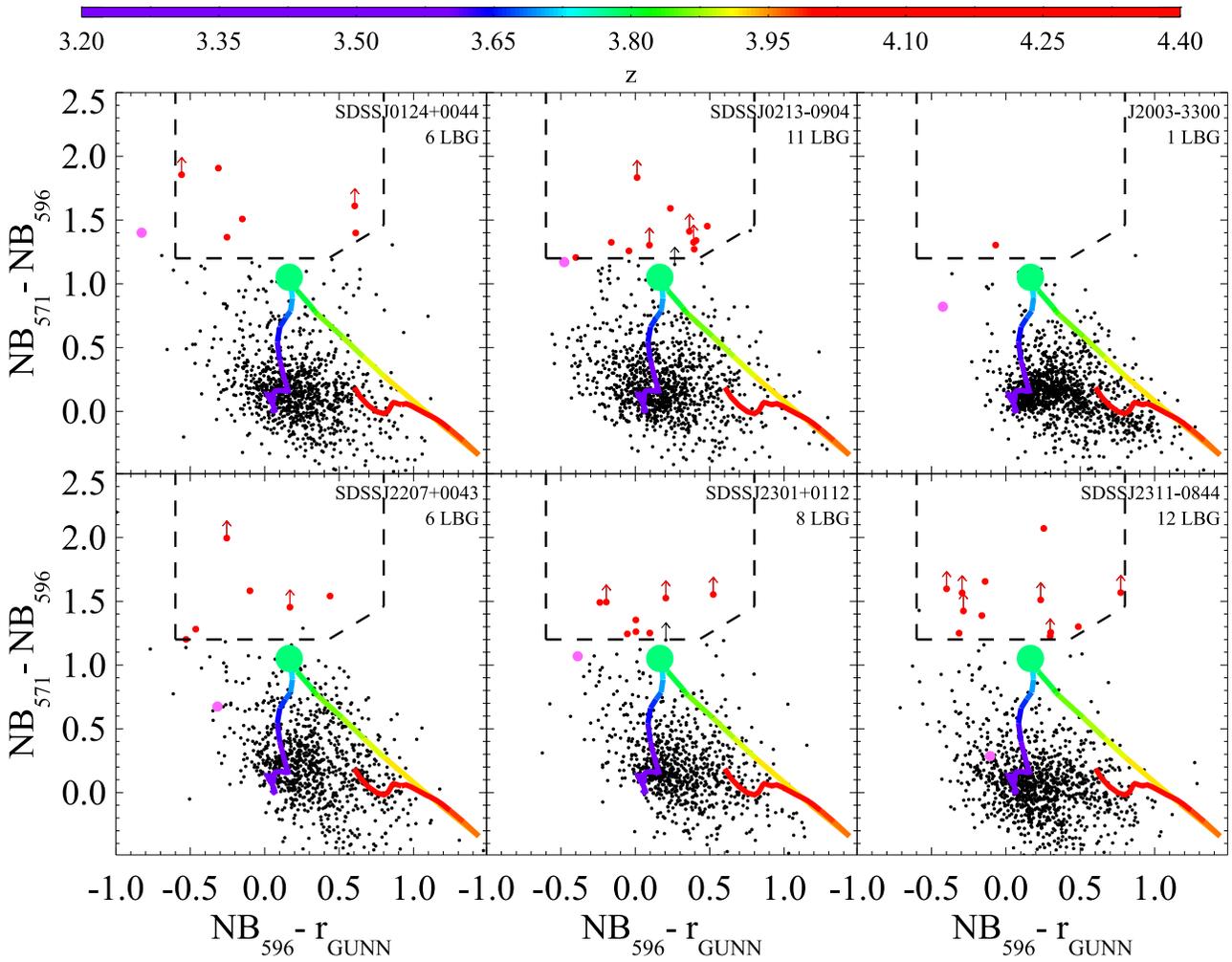
Figure 11. Same as in Figure 8 but for the six individual QSO fields. At the top right of each plot the number of LBGs found is shown.
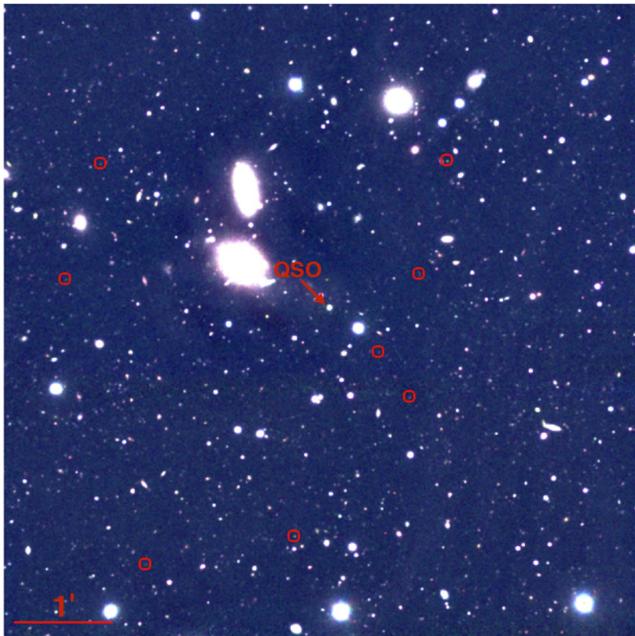
Figure 12. False-color image of 42 arcmin^2 of the field SDSSJ2301+0112. Red circles indicate the LBG candidates' positions.
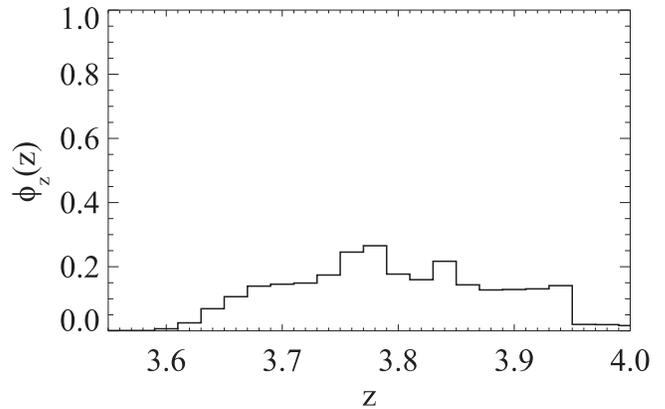
Figure 13. Redshift selection function of the LBG selection. The completeness per redshift bin was determined from 1000 simulated LBG spectra with different $EW_{Ly\alpha}$, UV continuum slope, and r_{GUNN} magnitudes. This is calculated by computing the fraction of the LBG simulated colors, per redshift bin, which was obtained from the selection region. This defines the redshift selection function, which is used for the clustering measurements in Section 4.

approximation given the narrow redshift range $\Delta z \simeq 0.3$ that we probe. The expected mean number density of LBGs in blank fields, n_G, computed for the magnitude range of each field, is given in Table 4.

13

Table 4
LBG Overdensity in Each Individual Field

Field (1)	n_G (2)	V_{field} (3)	$\langle QR \rangle_{field}$ (4)	$\langle QG \rangle_{field}$ (5)	Overdensity (6)
SDSSJ0124+0044	2.15	2600.20	5.60	6	1.07
SDSSJ0213−0904	1.79	2860.50	5.12	11	2.15
J2003−3300	1.71	2303.21	3.94	1	0.25
SDSSJ2207+0043	1.93	2032.63	3.92	6	1.53
SDSSJ2311−0844	2.13	2504.83	5.34	8	1.50
SDSSJ2301+0112	1.88	2480.15	4.66	12	2.58

Note. (1) Field ID; (2) The mean number density of $z \sim 4$ LBGs in units of (10^{-3} h^3 cMpc^{-3}), in the magnitude range of the survey $r_{GUNN}^{lower} < r_{GUNN} < r_{GUNN}^{limit}$. Given that r_{GUNN}^{limit} and completeness in the source detection are different for each field, we obtain a number density slightly different for each one; (3) total volume of the field in units of (h^{-3} cMpc^3), computed as $V_{field} = \sum_{i=1}^{N_{bins}} V_{eff,i}$; (4) total number of expected LBGs on the whole field computed as $\langle QR \rangle_{field} = n_G V_{field}$; (5) total number of observed QSO–LBG pairs on the whole field; (6) total overdensity per field, computed as $\langle QG \rangle_{field} / \langle QR \rangle_{field}$.

We define the effective volume of a radial bin as

$$V_{eff} = \int_{Z_{min}}^{Z_{max}} \int_{R_{min}}^{R_{max}} \phi(R, Z) 2\pi R dR dZ, \qquad (8)$$

where $\phi(R, Z)$ encodes the geometry of the survey, which can be separated into the radial R and the redshift (line-of-sight) Z selection function as $\phi(R, Z) = \phi_Z(Z)\phi_R(R)$. The redshift selection function of our survey $\phi_z(z)$ was modeled in Section 3.4 and we convert it to a redshift selection function in comoving units $\phi_Z(Z)$ using Equation (4). Then, we integrate it over the redshift range covered by our Monte Carlo modeling (corresponding to $3.2 \leqslant z \leqslant 4.4$).

The radial selection function $\phi_R(R)$ is easily calculated using the detection masks for our images. We created catalogs with randomly distributed galaxies with number density n_{ran} such that we had ∼100,000 sources in the entire image. Then we computed $\phi_R(R)$ in radial bins as the ratio between the number of randomly distributed galaxies and the expected number without masking $n_{ran}\pi(R_{max}^2 - R_{min}^2)$. The resulting $\phi_R(R)$ then quantified the fraction of the bin area where we could have detected LBGs. We computed the value of V_{eff} for each radial bin in each field using Equation (8). Summing the V_{eff} over the radial bins, we obtained the total volume covered by each of our six fields V_{field}, given in Table 4. We obtained that the total volume of our survey is 14,782 h^{-3} cMpc^3.

To obtain a rough estimate of the LBG overdensity in our QSO fields, we calculated the expected number of random QSO–LBG pairs, $\langle QR \rangle_{field}$, for each of our fields and compared to the number we found, $\langle QG \rangle_{field}$. These results are given in Table 4, where we also show the overdensity per field, $\langle QG \rangle_{field} / \langle QR \rangle_{field}$. We see that five of our six fields exhibit an LBG overdensity of LBGs, while one appears to be underdense. Adding up the results for all six fields, we find that the random expectation is $\langle QR \rangle = 28.6$ LBGs, whereas we detected a total of $\langle QG \rangle = 44$ LBGs, giving an overall overdensity of 1.5, and indicating that our fields are on average overdense.

To explore the profile of this overdensity around QSOs, we computed $\langle QG \rangle$ and $\langle QR \rangle$ in bins of transverse distance for each of our six fields, and then summed them to determine the binned volume averaged cross-correlation function $\chi(R_{min}, R_{max})$ according to Equation (6). These results are given in Table 5 and plotted in Figure 14. We estimate errors on $\chi(R_{min}, R_{max})$ assuming that shot-noise dominates the error budget (see Section 4.1.1 for a discussion about this

assumption), and use the one-sided Poisson confidence intervals for small number statistics from Gehrels (1986).

Given that the auto-correlation functions of both LBGs and QSOs at $z \sim 4$ have been previously measured, we can compute the expected volume-averaged QSO–LBG cross-correlation function $\chi(R_{min}, R_{max})$ assuming deterministic bias and compare it to our measurements. If we assume that both LBGs and QSOs trace the same underlying dark matter distribution, and assume deterministic bias (i.e., not stochastic), then we can write $\xi_{QG} = \sqrt{\xi_{QQ}\xi_{GG}}$. Assuming a power-law form $\xi = (r/r_0)^\gamma$ for the respective auto-correlations of QSOs and LBGs, and that they have identical slopes γ, then the cross-correlation length can be written as $r_0^{QG} = \sqrt{r_0^{QQ} r_0^{GG}}$. To compute ξ_{QG} we use respective measurements of the auto-correlation lengths of LBGs and QSOs at $z \sim 4$ from the literature. For LBGs Ouchi et al. (2004b) measured $r_0^{GG} = 4.1$ h^{-1} cMpc and $\gamma = 1.8$, whereas for QSOs we adopt $r_0^{QQ} = 22.3$ h^{-1} cMpc, which was measured by Shen et al. (2007) for $z > 3.5$ QSOs assuming a fixed $\gamma = 1.8$. Combining these implies $r_0^{QG} = 9.6$ h^{-1} cMpc for $\gamma = 1.8$. Plugging this power-law LBG–QSO cross-correlation function into Equation (5) and integrating over the effective survey volume gives us the expected value of $\chi(R_{min}, R_{max})$, which is shown as a dashed line in Figure 14. One sees that our QSO–LBG cross-correlation measurement is in reasonable agreement with the expected value of $\chi(R_{min}, R_{max})$ combining auto-correlation measurements and assuming deterministic bias. In Section 4.1.1 we quantify this agreement by fitting our cross-correlation function.

4.1.1. Fitting the Cross-correlation Function

Given the projected cross-correlation function measurement, we now determine the real-space cross-correlation parameters r_0^{QG} and γ that best fit our data. To this end we use a maximum likelihood estimator (MLE), and fit for the parameters, which maximize the probability of the data we observe. Since we are dealing with a counting process with a small number of counts in each bin (see Table 5), we can assume that Poisson error dominates the error budget. Adopting the Poisson distribution for the counts in our cross-correlation function bins, we can write the likelihood of our data as

$$\mathcal{L} = \prod_{i=1}^{N_{bins}} \frac{e^{-\lambda_i}\lambda_i^{x_i}}{x_i!}, \qquad (9)$$

Table 5
QSO–LBG Cross-correlation Function

R_{min} (h^{-1} cMpc)	R_{max} (h^{-1} cMpc)	$\langle QG \rangle$	$\langle QR \rangle$	$\chi(R_{min}, R_{max})$	$V_{eff,total}$ (h^{-3} cMpc ^{3})
0.124	0.252	1	0.039	$24.362^{+58.332}_{-20.974}$	20.84
0.252	0.513	2	0.168	$10.883^{+15.674}_{-7.676}$	88.48
0.513	1.041	2	0.771	$1.594^{+3.422}_{-1.676}$	400.18
1.041	2.112	10	3.110	$2.216^{+1.373}_{-1.000}$	1609.04
2.112	4.288	16	12.868	$0.243^{+0.395}_{-0.308}$	6644.21
4.288	8.706	13	11.637	$0.117^{+0.404}_{-0.306}$	6018.75

Note. We present the data for the volume-averaged projected cross-correlation function between QSOs and LBGs $\chi(R_{min}, R_{max})$ shown in Figure 14. This is measured in radial bins defined by R_{min} and R_{max}. $\langle QG \rangle$ is the observed number of QSO–LBG pairs per bin, and $\langle QR \rangle$ is the expected number of QSO–random pairs per bin, computed from Equation (7). We also show the total volume of the bin added over the fields, computed as $V_{eff,total} = \sum_{i=0}^{N_{fields}} V_{eff,i}$.
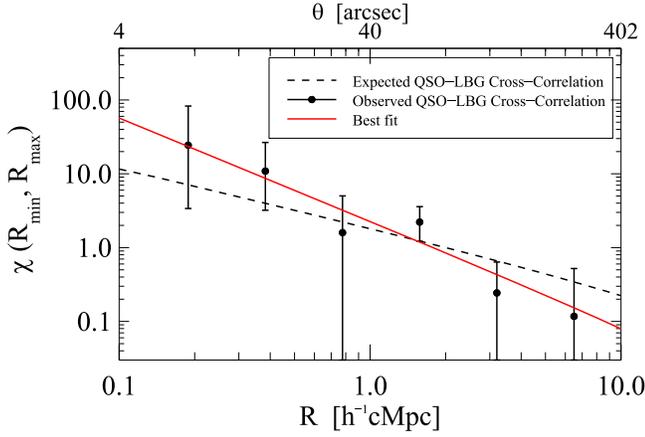
Figure 14. QSO–LBG cross-correlation function and its maximum likelihood model. The filled circles show our measurement described in Section 4.1 with 1σ Poisson error bars. The solid red curve shows the best maximum likelihood estimator for both r_0^{QG} and γ as free parameters. We obtain $r_0^{QG} = 6.93\ h^{-1}$ cMpc and $\gamma = 2.4$. The dashed black line shows the theoretical expectation of $\chi(R_{min}, R_{max})$ for the six stacked fields calculated from the independently determined QSO and LBG auto-correlation functions, assuming a deterministic bias model.

where the product is over the N_{bins} radial cross-correlation function bins, x_i is the number counts measured in the ith bin and λ_i is the expected number counts in the ith bin for a given set of model parameters. In our case we have defined $x = \langle QG \rangle$ and $\lambda = \langle QG \rangle^{exp}$, where

$$\langle QG \rangle^{exp}$$
$$= n_G \int_{Z_{min}}^{Z_{max}} \int_{R_{min}}^{R_{max}} \phi(R, Z)[1 + \xi_{QG}(R, Z)]2\pi R dR dZ$$
(10)

Here $\xi_{QG}(R, Z) = \left(\frac{\sqrt{R^2 + Z^2}}{r_0^{QG}}\right)^{-\gamma}$ and is determined by the model parameters r_0^{QG} and γ. Taking the natural logarithm of both sides of Equation (9), we obtain

$$\ln \mathcal{L} \propto \sum_{i=1}^{N_{bins}} [\langle QG \rangle_i \ln(\langle QG \rangle_i^{exp}) - \langle QG \rangle_i^{exp}], \quad (11)$$

where model-independent terms have been dropped. We calculated the log-likelihood for a grid of (r_0^{QG}, γ) values, which defines a uniform prior, ranging from $1.0 \leqslant \gamma \leqslant 5.0$ and $1.0 \leqslant r_0^{QG} \leqslant 15.0$ and maximized the likelihood to obtain
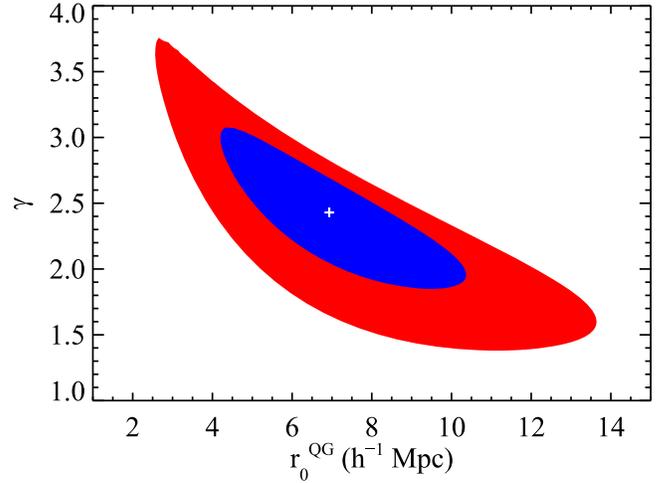
Figure 15. 1σ and 2σ confidence regions of r_0^{QG} and γ parameters (in blue and red, respectively), determined using a maximum likelihood estimator. The best estimation is shown as a white cross.

$r_0^{QG} = 6.93\ h^{-1}$ cMpc and $\gamma = 2.4$. These values were used in Equation (5) to calculate the corresponding $\chi(R_{min}, R_{max})$ value shown as the red line in Figure 14. We also computed the 1σ and 2σ 2D confidence regions for these parameters, shown in the $r_0^{QG} - \gamma$ plane in Figure 15. We determined errors on the parameters by marginalization. Given that our grid of values is uniform, the normalized likelihood is the joint posterior distribution of the parameters $P(r_0^{QG}, \gamma)$. Therefore, we marginalized out r_0^{QG} and γ to obtain the probability distributions $P(\gamma)$ and $P(r_0^{QG})$, respectively. From those probability distributions we computed 68% confidence regions about our MLE to define the error on the parameters. We find $r_0^{QG} = 6.93^{+2.13}_{-1.89}\ h^{-1}$ cMpc and $\gamma = 2.4^{+0.3}_{-0.5}$.

As shown in Figure 15, our measurements are relatively noisy when we fit r_0^{QG} and γ simultaneously, and there is a clear degeneracy between these parameters. For that reason, following common practice, we also fit the correlation function with γ fixed. Independent measurements of QSO auto-correlation suggest a slope of $\gamma = 2.0$ (Shen et al. 2007), which lies within the 1σ confidence region of our measurement (see Figure 15). Thus if we choose to fix the slope to this value, the maximum likelihood and the 1σ confidence interval for the cross-correlation length is $r_0^{QG} = 8.83^{+1.39}_{-1.51}\ h^{-1}$ cMpc.

Note that in the analysis described above we have assumed that the error bars on our cross-correlation function are

dominated by Poisson counting errors. This ignores cosmic variance fluctuations, and also assumes that the positions of the LBGs around the QSO are uncorrelated. However, as galaxies are not randomly distributed in the universe, but rather have significant auto-correlations, our binned measurements of $\langle QG \rangle$ are not truly independent. Given these correlations, our results could be somewhat sensitive to our choice of binning, and our error bars could also be somewhat underestimated. We have verified that changing the binning does not significantly change our parameter constraints. In principle, we should include the correlations and cosmic variance in our likelihood, analogous to computing the non-diagonal elements of the covariance matrix for a multivariate Gaussian likelihood. However, there is no simple analytical expression for the likelihood of a correlated Poisson process and, furthermore, correctly modeling the cosmic variance would require the use of N-body simulations of massive QSO halos at $z \sim 4$. Note, however, that while the positions of LBGs in the same field will be correlated, our QSO fields are separated by Gpc distances, and hence the positions of LBGs in different fields are completely independent. Given that our correlation function comes from six distinct fields, and the relatively large Poisson error bars, we believe that ignoring correlations and cosmic variance is a reasonable approximation.

Our measurement indicates a strong cross-correlation between QSOs and LBGs at $z \sim 4$, implying that QSOs trace massive dark matter halos in the early universe, with detectable enhancements of LBGs. We expect that those halos evolve to the most massive cluster of galaxies at $z = 0$. Our results are in agreement with the expected cross-correlation function ($r_0^{QG} = 9.6\ h^{-1}$ cMpc for $\gamma = 1.8$) computed from the individual QSO and LBG auto-correlation functions assuming deterministic bias, as shown by the dashed line in Figure 14.

4.2. Auto-correlation of LBGs in QSO Fields

Another measure of the clustering of LBGs in QSO environments is the LBG auto-correlation function in our fields. If QSOs trace highly biased locations of the universe, then we expect the LBGs around them to be more highly clustered than LBGs in blank fields, resulting in an enhancement of the LBG auto-correlation function. The auto-correlation function of $z \sim 4$ LBGs in blank fields was measured by Ouchi et al. (2004b), which we compare to our results.

To measure the LBG auto-correlation function we adopt the estimator:

$$\chi(R_{\min}, R_{\max}) = \frac{\langle GG \rangle}{\langle RR \rangle} - 1 \qquad (12)$$

where $\langle GG \rangle$ is the number of observed LBG–LBG pairs, and $\langle RR \rangle$ is expected random number of LBG–LBG pairs, in a cylindrical volume defined by the radial bin $[R_{\min}, R_{\max}]$ and the height Z. We measure $\chi(R_{\min}, R_{\max})$ at scales $1.0 < \theta < 490\farcs0$ (corresponding to $0.025 < R < 12.188\ h^{-1}$ cMpc) in logarithmically spaced radial bins. We measured $\langle GG \rangle$ directly from the images by counting the LBG pairs in each radial bin. Following the same argument in Section 4.1, we use the LBG luminosity function to compute the background number density n_G, rather than estimating it from our survey images.
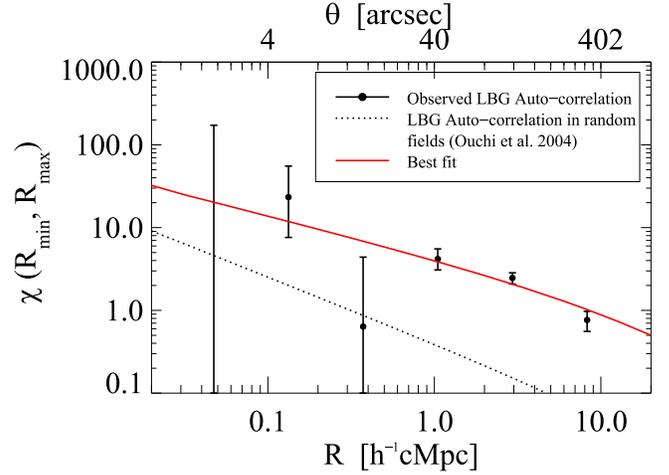
Figure 16. Data points showing the LBGs auto-correlation measurement in QSO fields as we describe in Section 4.2. The solid red curve shows the best fit for our measurements given by $r_0^{GG} = 21.59\ h^{-1}$ cMpc and $\gamma = 1.5$. The dotted black curve shows the LBGs auto-correlation in blank fields at $z \sim 4$ measured by Ouchi et al. (2004b). We find a stronger clustering in our fields in comparison with blank fields, which suggests that QSOs are located in overdense regions.

We computed the expected random number of LBG pairs as (see, e.g., Padmanabhan et al. 2007)

$$\langle RR \rangle = N_G n_G V_{\text{eff}}, \qquad (13)$$

where n_G is the same quantity defined in Section 4.1 and V_{eff} is given by Equation (8), but in this case using a different radial selection function $\phi_R(R)$, because of the different binning used. Here N_G is the expected number of LBGs for the entire volume in question in a random region of the universe, which is computed for each of our six fields as $N_G = n_G V_{\text{field}}$. The radial selection function $\phi_R(R)$ in this case is computed in an analogous way as for the cross-correlation: we create catalogs with $N_{\text{ran}} \sim 100,000$ randomly distributed galaxies on our masked images, and then we compute $\phi_R(R)$ as the ratio between the observed number of random galaxy pairs over the expected number of random galaxy pairs per radial bin. Here, the expected number of galaxy pairs per bin is computed by $N_{\text{ran}} n_{\text{ran}} \pi (R_{\max}^2 - R_{\min}^2)$. Note that, according to Equation (13), $\langle RR \rangle$ is proportional to the square of the LBG number density n_G and to the square of the redshift selection function $\phi_Z(Z)$ such that:

$$\langle RR \rangle \propto n_G^2 \left(\int_{Z_{\min}}^{Z_{\max}} \phi_Z(Z)\, dZ \right)^2. \qquad (14)$$

We computed $\langle GG \rangle$ and $\langle RR \rangle$ for each individual field and then we stacked the counts (i.e., we added the pairs in each field for a given bin) to measure the binned $\chi(R_{\min}, R_{\max})$ value as in Equation (12). We show the results in Figure 16 and the numerical values are given in Table 6. We estimate errors on $\chi(R_{\min}, R_{\max})$ using the one-sided Poisson confidence intervals for small-number statistics in the same way as in Section 4.1.

Analogous to our approach for the cross-correlation, we used an MLE to fit our auto-correlation function. In this case the

Table 6
LBG Auto-correlation Function

$R_{\min}$ (h^{-1} cMpc)	$R_{\max}$ (h^{-1} cMpc)	$\langle GG \rangle$	$\langle RR \rangle$	$\chi(R_{\min}, R_{\max})$	$V_{\mathrm{eff,total}}^{a}$ (h^{-3} cMpc^{3})
0.025	0.070	0	0.011	$-1.000^{+173.441}_{-0.000}$	1.13689
0.070	0.196	2	0.082	$23.336^{+32.099}_{-15.721}$	8.78984
0.196	0.551	1	0.611	$0.636^{+3.763}_{-1.353}$	65.3105
0.551	1.546	22	4.251	$4.175^{+1.355}_{-1.094}$	454.160
1.546	4.341	82	23.688	$2.462^{+0.382}_{-0.382}$	2528.37
4.341	12.188	72	40.779	$0.766^{+0.208}_{-0.208}$	4335.68

Note. We present the data for the LBG auto-correlation function in QSOs fields $\chi(R_{\min}, R_{\max})$ shown in Figure 16. This is measured in radial bins defined by $R_{\min}$ and $R_{\max}$. $\langle GG \rangle$ is the observed number of LBG–LBG pairs per bin, and $\langle RR \rangle$ is the expected number of random–random pairs per bin, computed from Equation (13). We also show the total volume of the bin added over the fields, computed as $V_{\mathrm{eff,total}} = \sum_{i=0}^{N_{\mathrm{fields}}} V_{\mathrm{eff,i}}$.

expected number of LBG–LBG pairs $\langle GG \rangle^{\exp}$ is modeled as

$$\langle GG \rangle^{\exp} = n_G^2 V_{\mathrm{field}}$$
$$\times \int_{Z_{\min}}^{Z_{\max}} \int_{R_{\min}}^{R_{\max}} \phi(R, Z)[1 + \xi_{GG}(R, Z)] 2\pi R \, dR \, dZ, \quad (15)$$

where ξ_{GG} is the LBG auto-correlation function assumed to have a power-law form with correlation length r_0^{GG}.

For the fitting we used a uniform prior defined by $1.0 \leqslant \gamma \leqslant 2.5$ and $5.0 \leqslant r_0^{GG} \leqslant 60.0$. We show the 1σ and 2σ 2D confidence regions for the parameters in Figure 17. We obtained that the maximum likelihood and the 1σ confidence intervals are $r_0^{GG} = 21.59^{+3.73}_{-2.96}\ h^{-1}$ cMpc and $\gamma = 1.5^{+0.1}_{-0.2}$, which are plotted as the red line in Figure 16. Following the same arguments as in Section 4.1, we also fit the auto-correlation function with γ fixed. The LBG auto-correlation function measured in blank fields at $z \sim 4$ suggests a slope of $\gamma = 1.8$ (Ouchi et al. 2004b), which lies outside the 1σ confidence region of our measurement (see Figure 17). We then prefer to fix γ to its maximum likelihood value $\gamma = 1.5$, which agrees with the LBG auto-correlation function slope measured in QSO fields at $z = 2.7$ (Trainor & Steidel 2012). After fixing γ, we obtain $r_0^{GG} = 21.59^{+1.72}_{-1.69}\ h^{-1}$ cMpc. We note that the first data point in the auto-correlation measurement is negative, indicating a smaller number of LBG–LBG pairs than expected. We also performed the fitting not including this data point and found that this does not change our results within our 1σ errors.

In order to compare this clustering signal with that computed in blank fields, we use the LBG auto-correlation at $z \sim 4$ measured by Ouchi et al. (2004b), which we assume represents the cosmic average. Plugging their best-fit values ($r_0^{GG} = 4.1\ h^{-1}$ cMpc and $\gamma = 1.8$) into Equation (5) using a power-law form for $\xi_{GG}(R, Z)$, gives the dotted line plotted in Figure 16. To better compare our auto-correlation measurement with the Ouchi et al. (2004b) blank field values, we performed a fit with fixed $\gamma = 1.8$, obtaining a maximum likelihood value and 1σ confidence interval given by $r_0^{GG} = 16.86^{+1.17}_{-1.14}\ h^{-1}$ cMpc, which is ~ 4 times higher than the correlation length in blank fields. The fact that our LBG auto-correlation measurement is higher suggests that the LBGs in our fields are more clustered than those in blank fields, which provides another indication that QSOs fields trace regions of the universe that are denser than the cosmic average, confirming our findings from the cross-correlation measurement in Section 4.1.

While it may at first seem counter-intuitive that the LBG auto-correlation is enhanced by such a large factor (~ 4) in QSO environs, this is actually exactly the expected behavior as
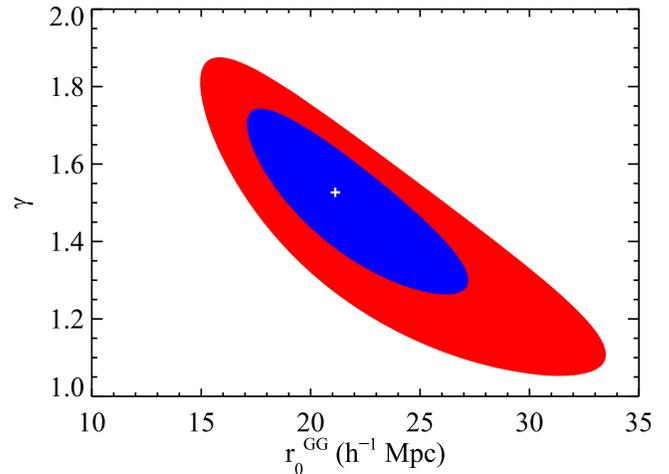
Figure 17. Same as Figure 15 but for the r_0^{GG} and γ parameters corresponding to the LBG auto-correlation function.

we clarify here. The LBG auto-correlation function measures the radially binned profile of galaxy pairs, and it is not trivial to relate the auto-correlation to the cross-correlation. In order to build intuition, we will think in terms of the total number of galaxies detected in our survey (see Table 4). On average, we found 1.5 times more galaxies in QSO fields compared with the number expected in blank fields (which is estimated from our selection function and the number density of LBGs n_G), and that means that we should detect at least 1.5^2 more galaxy pairs in our fields compared with the expectation in blank fields, simply because we are overdense by that factor. This implies that the auto-correlation function will never be less than $1.5^2 - 1 = 1.25$.

To better illustrate what happens to the auto-correlation in an overdensity, we will consider a hypothetical scenario where galaxies are randomly distributed in the universe with number density n_G, and are clustered *only* around QSOs, which are however rare objects in the universe. Now imagine that the number density of galaxies around QSOs is enhanced within a sphere of radius R_{QSO}, but that galaxies are otherwise randomly distributed within the sphere. In other words, we imagine that in QSO fields the number density of galaxies is simply increased by a factor X, but that within the sphere they are unclustered. If QSOs are rare structures, then when averaging over large volumes of the universe, we expect that the observed number of galaxy pairs $\langle GG \rangle$ will be very close to the random

expectation $\langle RR \rangle$ (computed using n_G), and then the galaxy auto-correlation function measured from blank fields will be flat and very close to zero on small scales, $r < R_{QSO}$. On larger scales, $r \sim R_{QSO}$, it would deviate more from zero, reflecting the clustering due to the top-hat overdensities around QSOs (however, if QSOs are very rare objects this positive correlation function would be diluted and could still be quite small). Note, however, that if we measure the galaxy auto-correlation around QSOs at radii $r < R_{QSO}$, then $\langle GG \rangle$ will be X^2 times larger than $\langle RR \rangle$ (which is again computed using n_G), and then we would measure an auto-correlation function of roughly $X^2 - 1$. This could be much larger than the value measured in blank fields that would be close to zero. This simple example illustrates that because $\langle RR \rangle$ is computed from the number density of galaxies in random locations n_G, overdense fields will always result in an enhanced auto-correlation relative to that in blank fields, and that these enhancements can be quite large. The situation clearly becomes more complicated if galaxies are intrinsically clustered with a power-law profile, and there is no simple analytical relationship between the cross- and auto-correlation functions. To fully quantitatively understand the relationship between the cross- and auto-correlation functions in QSO environments, one would need to analyze cosmological simulations (see, e.g., White et al. 2012). But the generic expectation is an enhancement of the auto-correlation function in QSO environs compared with blank field pointings, which is exactly what we see in Figure 16.

4.3. Comparison with Previous Measurements

The highest redshift for which the QSO–LBG cross-correlation has been measured before is at $z \sim 3$ by Trainor & Steidel (2012), who reported an overdensity of galaxies in QSO fields, and found a cross-correlation length of $r_{0,z\sim 3}^{QG} = 7.3 \pm 1.3\ h^{-1}$ cMpc for a fixed $\gamma = 1.5$. At $z \sim 4$ we find a steeper slope (we fixed $\gamma = 2.0$) than Trainor & Steidel (2012), but in order to facilitate a comparison with their results, we fit our cross-correlation measurement for their same fixed $\gamma = 1.5$. This γ value, is near the border of our 2σ confidence region (see Figure 15) and thus disfavored by our measurements, but we nevertheless proceed with this for comparison purposes. We obtain a cross-correlation length of $r_0^{QG} = 10.73_{-2.41}^{+2.20}\ h^{-1}$ cMpc, which is ~ 1.5 times higher than their cross-correlation length at $z \sim 3$, suggesting that halos hosting QSOs could be more biased and highly clustered at $z \sim 4$. This agrees with the result reported by Shen et al. (2007) who find that the QSO auto-correlation increases significantly from $z = 3$ to $z = 4$, and as such we expect to have found a larger cross-correlation. However, note that our measurements are too noisy to make strong statements about evolution, and larger LBG samples are still needed to perform more detailed comparisons.

At $z > 4$ only individual QSO fields have been studied so far. Some studies of QSO environments at $z \sim 6$–7 find no enhancements of galaxies compared with the background (e.g., Willott et al. 2005; Bañados et al. 2013; Simpson et al. 2014; Mazzucchelli et al. 2017), which could be suggesting that the strong QSO–galaxy cross-correlation breaks down at those redshifts. The lack of QSO auto-correlation measurements at these high redshifts makes it impossible to know masses of dark matter halos hosting $z \sim 6$ QSOs, but if their masses are comparable to those hosting QSOs at $z \sim 4$ (i.e., $M_{halo} \gtrsim 10^{12}\ M_\odot$ as suggested by the Shen et al. 2007 auto-correlation),

then one would generically expect a strong QSO–galaxy clustering signal as we have detected here at $z \sim 4$.

5. Testing the Robustness of Our Results

Two requirements must be fulfilled to ensure a robust clustering measurement: we need a low contamination level in the LBG sample and an accurate knowledge of the background number density of LBGs. Given that we used a novel NB technique to select LBGs, we need to carefully consider those requirements. In this section, we first discuss caveats related to the the use of this selection technique. Then we consider the effects of using a contaminated sample for clustering measurements, and finally we explore the impact of using different LBG selection criteria on our results.

5.1. The Use of an NB Technique for LBG Selection: Caveats

A first complication of using our novel method for color-selecting LBGs is that the level of contamination of our sample is unknown. In principle, the purity of the sample can only be determined with follow-up spectroscopy, or detailed modeling of the population of contaminant galaxies. Both alternatives would be challenging to implement and are beyond the scope of this paper, but as a compromise we qualitatively discuss the impact of contamination on the correlation function (see Section 5.2) and we demonstrate the robustness of our results against contamination by exploring their sensitivity to the color-selection criteria (see Section 5.3). Note, however, that we could excise contamination in our LBG sample if we had additional imaging on our fields using traditional broad-band filters. This would allow us to confirm the presence of the Lyα break in our LBG candidates.

Another complication of using our novel color-selection is that we did not have an independent measurement of the background number density of LBGs required to compute the clustering. This implied that we had to rely on Monte Carlo simulations to determine the LBG selection function $\phi_Z(Z)$, and then our clustering results are sensitive to errors in this modeling. If the completeness of the sample were close to 100%, then 10% errors on $\phi_Z(Z)$ would impact our measurement at the 10% level, whereas if the completeness were $\sim 20\%$ (as is the case), then there could be 100% error on $\phi_Z(Z)$, which could strongly impact the amplitude of the measured clustering. Note that the auto-correlation is even more sensitive to this quantity compared with the cross-correlation, because while $\langle QR \rangle$ is proportional to $\phi_Z(Z)$, $\langle RR \rangle$ is proportional to the square of this quantity. In Section 5.3 we test our redshift selection function to demonstrate that it is accurate and correctly modeled.

5.2. Impact of Contamination on the Clustering Measurements

One method to qualitatively check the contamination level in the sample is by studying the shape of the measured correlation function. For example, if we measure the cross-correlation function via Equation (6) using a highly contaminated sample, the numerator in that equation would be overestimated because of the inclusion of low-redshift contaminants that are taken to be real LBGs. However, since the denominator $\langle QR \rangle$ is simply computed from the LBG luminosity function and our redshift selection function, this value does not include the extra number counts due to contamination. This implies that the measured cross-correlation will not behave like a power law, but rather it

will flatten toward larger scales. Quantitatively, for a contaminated sample, what we would actually measure is

$$\chi(R_{\min}, R_{\max}) = \frac{\langle QG \rangle + N_{\rm cont}}{\langle QR \rangle} - 1 \qquad (16)$$

where $\langle QG \rangle$ and $\langle QR \rangle$ are given by Equations (10) and (7), respectively, and $N_{\rm cont}$ is the number of contaminants in the bin. Given that the contaminants are galaxies at different redshifts, the cross-correlation between them and the $z \sim 4$ QSO is zero, then the number of contaminants will be given by $N_{\rm cont} = n_{\rm cont} V_{\rm eff, cont}$, where $n_{\rm cont}$ is the number density of contaminants and $V_{\rm eff, cont}$ is the effective volume of the bin, which is given by Equation (8), but with the redshift selection function of the contaminants $\phi_{Z,\rm cont}(Z)$. Then the Equation (16) reduces to

$$\chi(R_{\min}, R_{\max}) = \chi^{\rm true}(R_{\min}, R_{\max})$$
$$+ \frac{n_{\rm cont} \int_{Z_{\min,\rm cont}}^{Z_{\max,\rm cont}} \phi_{Z,\rm cont}(Z) dZ}{n_G \int_{Z_{\min}}^{Z_{\max}} \phi_Z(Z) dZ} \frac{D_C^2(z_{\rm cont})}{D_C^2(z_{\rm LBG})}, \qquad (17)$$

where $\chi^{\rm true}(R_{\min}, R_{\max})$ is the correlation function that we would measure from a non-contaminated sample (i.e., here $\chi^{\rm true}(R_{\min}, R_{\max}) = \langle QG \rangle / \langle QR \rangle - 1$), and $D_C(z)$ is the transverse comoving distance at redshift z. In the absence of contaminants, the second term in this equation would be zero, and we recover the correlation function defined in Equation (5), which can be approximated by a power-law function. However, if a large number of contaminants are included that span a large range in redshift, the second term becomes important, and given that it does not depend on radius, the same constant is added everywhere to the cross-correlation function, flattening its shape, with the degree of flattening dependent on the level of contamination.

This flattening effect will be even stronger for the auto-correlation function since it is proportional to the square of both the number density of contaminants and redshift range they cover. Then for a contaminated sample one obtains

$$\chi(R_{\min}, R_{\max}) = \chi^{\rm true}(R_{\min}, R_{\max})$$
$$+ \frac{n_{\rm cont}^2 \left(\int_{Z_{\min,\rm cont}}^{Z_{\max,\rm cont}} \phi_{Z,\rm cont}(Z) dZ \right)^2}{n_G^2 \left(\int_{Z_{\min}}^{Z_{\max}} \phi_Z(Z) dZ \right)^2} \frac{D_C^4(z_{\rm cont})}{D_C^4(z_{\rm LBG})} \qquad (18)$$

ignoring the clustering of the contaminants, which should be greatly diluted in projection if they span a large range of redshifts. To take clustering of contaminants into account, an additional term should be added to this equation to account for their auto-correlation. Therefore, the smoking gun of high contamination in the LBG sample would be a flat cross-correlation and auto-correlation function in Figures 14 and 16, respectively. Given that we measured a power-law shape for both correlations, we believe that our LBG sample is not strongly affected by contamination.

We have explored a third way to check contamination, which is also independent of our estimate of n_G and $\phi_Z(Z)$. For a highly contaminated sample that includes galaxies over a wide range of redshifts, it would be more appropriate to measure angular distances instead of transverse comoving distances. We thus compute the angular correlation function $\omega(\theta)$ using the standard procedure, where $\langle RR \rangle$ is determined from the angular number density of the data themselves, and we do not assume anything about the number density or selection function. In this case we only measure how clustered our sample is in comparison to a random distribution with the same number density. This angular correlation function calculation thus differs from our LBG auto-correlation function in Section 4.2, where $\langle RR \rangle$ was computed from n_G and our selection function $\phi_Z(Z)$. For a highly contaminated sample we expect the angular correlation function to be close to zero on all angular scales, because the inclusion of uncorrelated galaxies over a broad redshift range would dilute any real clustering signal. On the other hand, for a relatively pure sample composed primarily of LBGs at $z = 3.78 \pm 0.3$, we expect to measure a power-law angular auto-correlation because we would be selecting only highly biased galaxies in a small volume. Note, however, that even for a pure LBG sample, $\omega(\theta)$ computed in this way is not the true angular correlation function of LBGs, because we are pointing toward overdense regions around QSOs.

We estimate the angular auto-correlation function of the LBGs as

$$\omega(\theta) = \frac{\langle GG(\theta) \rangle}{\langle RR(\theta) \rangle} - 1, \qquad (19)$$

where $\langle GG(\theta) \rangle$ is the number of LBG–LBG pairs per angular bin, which is directly measured from our images, and $\langle RR(\theta) \rangle$ is the number of random–random pairs per angular bin. The $\langle RR(\theta) \rangle$ quantity was estimated using a random catalog of sources created as follows. First, we computed the total number of LBG candidates in all the fields, then we divided that by the total unmasked area to get the average angular number density of LBGs. Second, we multiplied the unmasked area per image by this average number density to determine the number of galaxies expected in each field. Finally, we increased the number of galaxies by a large factor F in order to decrease the noise in the measurement, and we randomly distributed those sources on the image and then measured $\langle RR(\theta) \rangle$ by counting the pairs of simulated galaxies per angular bin. We then rescaled $\langle RR(\theta) \rangle$ down by F^2.

Our measurement of the angular correlation function is shown in Figure 18. We see a non-flat correlation function, which suggests that our LBG sample is not highly contaminated. Assuming a power-law form given by $\omega(\theta) = A\theta^{-\beta}$ we performed a Levenberg–Marquardt least-squares fit to these data to quantify how consistent the measurement is with a flat shape (where $\beta = 0$). We obtained best-fit parameters of $A = 21.56 \pm 39.54$ and $\beta = 1.07 \pm 0.49$. Given the large error bars in the measurement we were unable to discard a correlation function consistent with zero; however, as we show in the next subsection, if the LBG sample were highly contaminated then the angular correlation function would be much flatter. The fact that we measure a signal in Figure 18 suggests that we are measuring real LBG clustering.
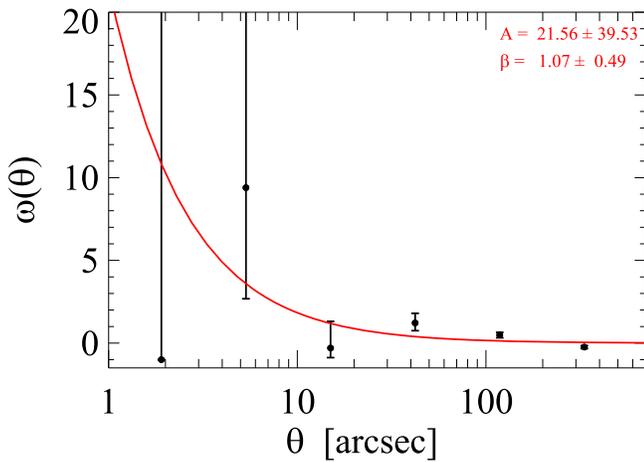
Figure 18. Angular auto-correlation function measurement for the LBG sample. This measurement is used to test if the sample is contaminated. If the sample were highly contaminated, $\omega(\theta) = 0$ at every scale and a power-law fit with slope β would be consistent with zero. We find that the LBG sample is not highly contaminated, given their power-law shape in this plot, which is well fitted by $\omega(\theta) = 21.56 \, \theta^{-1.07}$ (red line).

5.3. Robustness of Clustering Measurements against Changes in Color Selection

Here, we study the impact of using different color selections on our clustering measurements to demonstrate that our results are not significantly impacted by contamination, and to show that our Monte Carlo simulation of the completeness is robust if we change the color-selection criteria. To this end we have defined different LBG selection criteria, and for each one we compute the cross-correlation, auto-correlation, and angular auto-correlation function. The cross-correlation and auto-correlation functions were fitted in each case using an MLE following the same procedure as described in Section 4. We consider a progression of seven different selections, from the most permissive Case 1, which selects the majority of $z \sim 4$ LBGs, but also likely incurs a large fraction of low-redshift contaminants, to the most conservative Case 7, which results in a low completeness for $z \sim 4$ LBGs, but ensures low contamination. These results are shown in Figure 19, and we give the best-fit values for each case in Table 7. There we also tabulate the best-fit correlation lengths for a fixed $\gamma = 2.0$ for the QSO–LBG cross-correlation, and $\gamma = 1.5$ for the LBG auto-correlation, in order to study how r_0 varies for the different cases. Note that for the three most permissive Cases 1–3, we measure a flat correlation function and hence do not quote fits for fixed γ. Additionally for those cases we had to use a different prior for the MLE fit, since flat correlation functions result in small values for the slope and large values for the correlation length, which are not covered by the prior used for the other cases. We only quote the best-fit parameters for Cases 1–3 because the 1σ confidence region extends beyond the prior, precluding reliable error estimates.

We find that the cross-correlation function flattens and its amplitude increases for more permissive selections that increase the level of contamination, and the auto-correlation function shows a similar but even stronger tendency. This is the behavior that we expected as we describe in Section 5.2 and according to Equations (17) and (18). As for the angular correlation function, we find that the more conservative the selection, the steeper the slope of $\omega(\theta)$ and the more significant its departure from zero. These are again the trends we expect

because reduced contamination results in a more strongly clustered sample of $z \sim 4$ galaxies, selected from a narrow redshift slice, reducing the amount that the clustering is diluted by projection. Note, however, that for the less conservative cases (i.e., Case 1 and Case 2), where the sample is dominated by contaminants, the angular correlation function is close to zero, but not perfectly consistent with $\beta = 0$. We believe that the measurement of a weak clustering signal in these cases results from the actual clustering of foreground contaminants, which is diluted by the line-of-sight projection, but nevertheless remains strong enough to not be perfectly consistent with zero.

The takeaway message from Figure 19 is that we observe convergence of both the cross-correlation and auto-correlation functions for the more conservative selections. Specifically, we find stable results for Cases 5–7, with the only significant difference being the S/N of the clustering measurements, resulting from the smaller sample of LBGs selected in the more conservative cases. In Figures 20 and 21 we plot the values of the cross-correlation length r_0^{QG} (for fixed $\gamma = 2.0$) and auto-correlation length r_0^{GG} (for fixed $\gamma = 1.5$), respectively, for the four most conservative selections. The convergence of the correlation lengths demonstrates that (1) we do not suffer large contamination and hence our results are robust against contamination, (2) our Monte Carlo simulation of the selection function is reliable, since it results in consistent measurements as the color selection and selection function are varied, and (3) our results are largely independent of the exact color-selection region adopted. For these reasons we simply adopt Case 5 to present the final results in this paper.

Finally, we performed one last test to establish that the redshift selection function modeled from our Monte Carlo is essentially correct. We compared the total observed QSO–LBGs pairs in all the fields $\langle QG \rangle^{obs}$ for each selection with the expected value $\langle QG \rangle^{exp}$ based on our clustering measurements, and our Monte Carlo determination of the redshift selection function. Specifically, for each selection the $\langle QG \rangle^{obs}$ was measured by summing the observed QSO–LBGs pairs over the fields up to scales of $R \sim 9 \, h^{-1}$ cMpc, and Poisson errors were computed for this measurement. The expected value $\langle QG \rangle^{exp}$ for each field was computed using Equation (10), where we computed the corresponding $\phi_Z(Z)$ using our Monte Carlo simulation method described in Section 3.1 for each selection criterion. For all the cases, we used $r_0^{QG} = 6.93 \, h^{-1}$ cMpc and $\gamma = 2.4$, which are the best-fit parameters for our fiducial color selection (Case 5; see Figure 19) in the computation of $\langle QG \rangle^{exp}$. The total expected number of QSO–LBG pairs in the whole survey, $\langle QG \rangle^{exp}$, was computed by summing $\langle QG \rangle^{exp}$ over the bins and over the fields.

If the contamination is low, and the redshift selection function $\phi_Z(Z)$ is correctly computed for each case, we expect that $\langle QG \rangle^{exp}$ should be equal to $\langle QG \rangle^{obs}$. As the sample becomes more contaminated we expect that $\langle QG \rangle^{obs}$ will exceed $\langle QG \rangle^{exp}$ and increasingly deviate from it for more permissive selections. The results of this test are shown in Figure 22, where we plot $\langle QG \rangle^{obs}$ versus $\langle QG \rangle^{exp}$ for the seven color selections we considered, and compare to the line $\langle QG \rangle^{exp} = \langle QG \rangle^{obs}$ (solid line). We find that the total number of observed QSO–LBGs pairs is consistent with our expectations for the three more conservative selections Cases 5–7, but that $\langle QG \rangle^{obs}$ exceeds $\langle QG \rangle^{exp}$ for more permissive selections, with the deviations progressively increasing as more contaminants are included. Note that by construction we will have
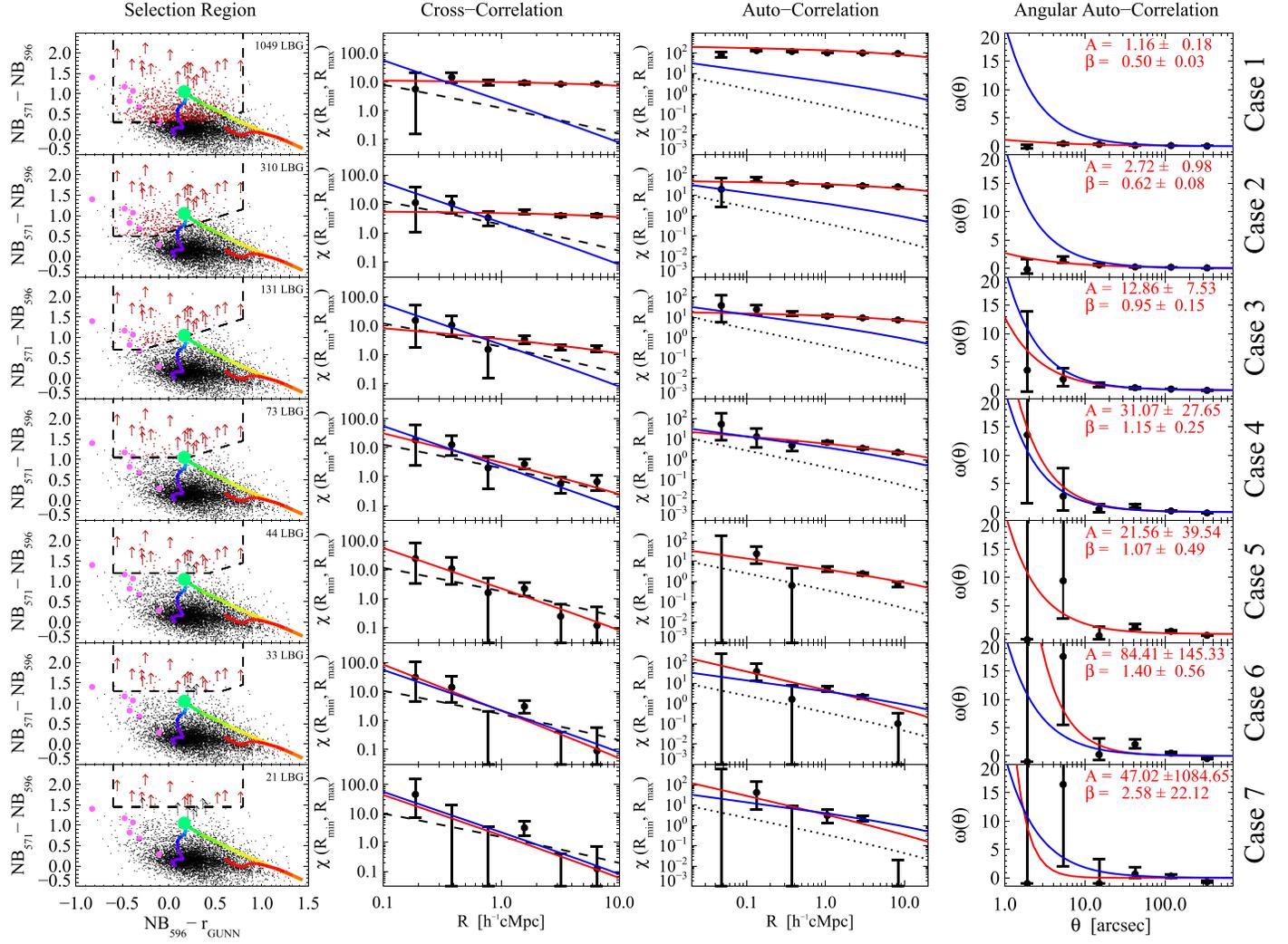
Figure 19. Impact of contaminants on the clustering measurements. We show seven different selection criteria and their respective clustering measurements. From left to right we show the color–color plot showing the color cuts used and the photometry. The cross-correlation, including the best fit according to our MLE estimator and the theoretical expectation of $\chi(R_{min}, R_{max})$ calculated from the QSO and LBGs auto-correlation functions and assuming a deterministic bias model (dashed line). The auto-correlation measurement, including the best fit according to our MLE estimator, and the LBG auto-correlation in blank fields at $z \sim 4$ measured by Ouchi et al. (2004b) (dotted line). Finally, we show the angular auto-correlation function, with a power law fit $\omega(\theta) = A\theta^{-\beta}$, with the A, β values indicated in the top right corner. In the three correlation function plots we show the best fit in each case as a red curve and the best fit for our fiducial selection (i.e., Case 5) as a blue curve. From top to bottom we show selections progressively more conservative, and thus less contaminated. The selection used in this paper to measure the clustering properties of $z \sim 4$ LBGs in QSO environments correspond to Case 5. We detected a convergence in the clustering measurements for the three last cases.

$\langle QG \rangle^{obs} = \langle QG \rangle^{exp}$ for Case 5, since the clustering measurements ($\langle QG \rangle^{obs}$) were fit to determine the correlation function parameters, which go into the computation of $\langle QG \rangle^{exp}$. But the fact that the expected $\langle QG \rangle^{exp}$ matches the observed $\langle QG \rangle^{obs}$ for the more conservative Cases 6 and 7 demonstrates that (1) the modeling of the redshift selection function $\phi_Z(Z)$ is correct, (2) the contamination is insignificant, and (3) our clustering measurements are robust.

6. Summary and Conclusions

The strong observed auto-correlation of QSOs at $z > 3.5$ indicates that they inhabit massive dark matter halos with $M_{halo} > 10^{12} M_\odot$, which implies QSO environments should exhibit an enhancement of galaxies manifest as a strong QSO–galaxy cross-correlation function. We characterized the environments of six QSO fields at $z = 3.78$ that were chosen to host massive BHs ($\gtrsim 10^9 M_\odot$). The fields were imaged using VLT/FORS1 with two custom NB filters, and the broad-band r_{GUNN},

to identify LBGs using a novel technique that selects them in a redshift range ~ 3.3 times smaller than that typically probed when selecting LBGs with broad-band filters. This significantly reduces the line-of-sight projection effects that have hampered previous searches for overdensities around $z \gtrsim 5$ QSOs.

Since we used a non-standard filter set to select LBGs, we performed detailed Monte Carlo simulations to model LBG colors, define our selection criteria, and compute the redshift selection function and volume probed by our survey. This new method effectively selects LBGs in a narrow redshift range, but the color loci of $z \simeq 3.78$ LBGs and low-redshift galaxies overlap more than with traditional LBG selection using broader filters. Defining a pure sample free from low-redshift contaminants required adopting stricter color cuts, which decreased the completeness of the resulting LBG sample. We devised selection criteria that resulted in $\sim 26\%$ completeness at $z = 3.78$, and detected 44 LBGs in our six fields, corresponding to a number density of 0.19 LBGs arcmin^{-2}. Our survey probed $\Delta z \simeq 0.3$, and covered a volume equal to 14,782 h^{-3}

Table 7
Best-fit Parameters for the Cross-correlation and Auto-correlation Functions for the Seven Cases Shown in Figure 19

	Selection criteria	Cross-correlation			Auto-correlation		
		r_0^{QG}	γ	$r_{0,\gamma=2.0}^{QG}$	r_0^{GG}	γ	$r_{0,\gamma=1.5}^{GG}$
1	$(NB_{571} - NB_{596}) > 0.30$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.0$	873.03	0.7		5456.67	0.9	
2	$(NB_{571} - NB_{596}) > 0.50$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.6$	198.96	0.6		1100.00	0.8	
3	$(NB_{571} - NB_{596}) > 0.70$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.9$	25.67	1.3		261.08	0.8	
4	$(NB_{571} - NB_{596}) > 1.05$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.9$	$10.25^{+2.18}_{-2.08}$	$2.0^{+0.3}_{-0.3}$	$10.25^{+1.13}_{-1.19}$	$41.17^{+8.52}_{-4.19}$	$1.3^{+0.1}_{-0.1}$	$31.23^{+1.27}_{-1.32}$
5	$(NB_{571} - NB_{596}) > 1.20$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.9$	$6.93^{+2.13}_{-1.89}$	$2.4^{+0.3}_{-0.5}$	$8.83^{+1.39}_{-1.51}$	$21.59^{+3.73}_{-2.96}$	$1.5^{+0.1}_{-0.2}$	$21.59^{+1.72}_{-1.69}$
6	$(NB_{571} - NB_{596}) > 1.30$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.9$	$6.22^{+2.53}_{-1.92}$	$2.6^{+0.3}_{-0.6}$	$8.83^{+1.61}_{-1.77}$	$14.96^{+2.83}_{-1.93}$	$1.9^{+0.1}_{-0.2}$	$19.94^{+2.20}_{-2.17}$
7	$(NB_{571} - NB_{596}) > 1.45$ $-0.6 < (NB_{596} - r_{GUNN}) < 0.8$ $(NB_{571} - NB_{596}) >$ $0.7(NB_{596} - r_{GUNN}) + 0.9$	$6.46^{+2.95}_{-3.61}$	$2.4^{+0.6}_{-0.8}$	$7.88^{+2.15}_{-2.46}$	$13.06^{+4.21}_{-2.93}$	$1.9^{+0.2}_{-0.4}$	$16.62^{+3.42}_{-3.25}$

Note. r_0 is shown in $(h^{-1} \text{ cMpc})$ units.
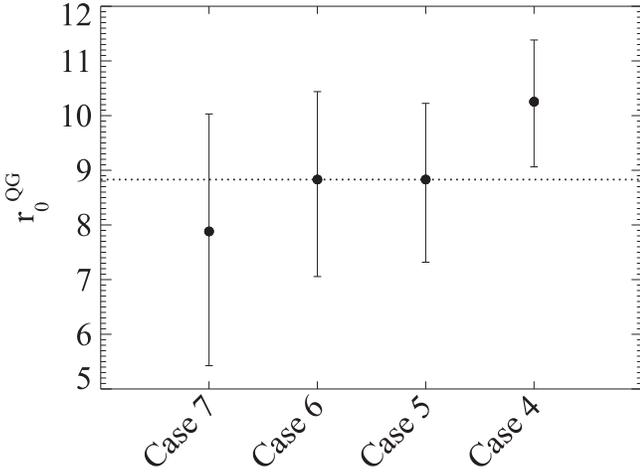
Figure 20. Best-fit r_0^{QG} values for a fixed $\gamma = 2.0$ for the four most conservative selections shown in Figure 19. We detect a convergence of the correlation length. The horizontal dashed line indicates the best-fit r_0^{QG} value for Case 5.

Figure 21. Best-fit r_0^{GG} values for a fixed $\gamma = 1.5$ for the four most conservative selections shown in Figure 19. We detected a convergence of the correlation length. The horizontal dashed line indicates the best-fit r_0^{GG} value for Case 5.

cMpc^3 within $R < 9\ h^{-1}$ cMpc from the QSO, and we found on average 1.5 times more galaxies than expected in random locations of the universe.

Our work resulted in the first volume-averaged projected QSO–LBG cross-correlation function at $z \sim 4$. We fit our measurements with a (real space) power-law cross-correlation

function, and found $r_0^{QG} = 6.93^{+2.13}_{-1.89}\ h^{-1}$ cMpc and $\gamma = 2.4^{+0.3}_{-0.5}$. When we fixed the slope at $\gamma = 2.0$ we found $r_0^{QG} = 8.83^{+1.39}_{-1.51}\ h^{-1}$ cMpc. This strong cross-correlation function is in agreement with the theoretical expectation for the cross-correlation assuming a deterministic bias model, which can be
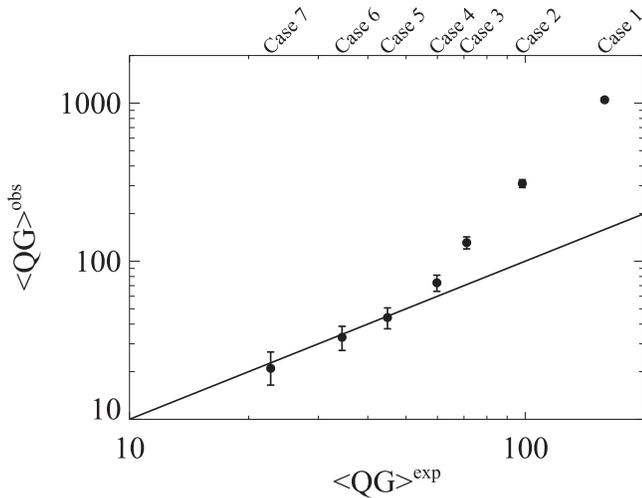
Figure 22. Total observed QSO–LBGs pairs in the whole survey for our selections with Poisson error bars compared with the expected number of QSO–LBGs pairs, assuming the cross-correlation parameters for Case 5 and using the redshift selection function $\phi_Z(Z)$ computed for each case as we described in Section 3. The solid line indicates the case where $\langle QG \rangle^{\mathrm{obs}} = \langle QG \rangle^{\mathrm{exp}}$. Our results show that the three more conservative cases (non-contaminated samples) are consistent with the expectation, which means that the modeled redshift selection function $\phi_Z(Z)$ is sensitive and that the levels of contamination in those selections are negligible.

estimated using the auto-correlation of both LBGs and QSOs at $z \sim 4$.

We also measured the auto-correlation function of LBGs near these QSOs and found an auto-correlation length of $r_0^{\mathrm{GG}} = 21.59^{+1.72}_{-1.69}\ h^{-1}$ cMpc for a fixed slope of $\gamma = 1.5$, which is ~ 4 times higher than the measured auto-correlation length of LBGs in blank fields at the same redshift. Our measurement of an enhanced LBG auto-correlation in QSO environments, and the strong QSO–LBG cross-correlation, both indicate that QSOs at $z \sim 4$ trace massive dark matter halos in the early universe, which are the likely progenitors of massive clusters of galaxies at $z = 0$.

We demonstrated that our results are robust against contamination and that our selection function modeling is reliable, by varying our color-selection criteria and showing that the cross-correlation and auto-correlation functions are converged. Spectroscopic follow-up of our LBGs candidates would provide an additional and definitive test of the reliability of our novel color-selection technique. However, because the colors of LBGs in our filters do not separate as cleanly from contaminants as for broad-band LBG selection, we had to choose relatively conservative color cuts that recovered only 26% of LBGs. As such, we believe that the preferred approach to search for overdensities around $z \sim 4$–6 QSOs using NB filters is to perform traditional LAE selection. Although LAE selection also only selects a fraction of the total population of high-redshift galaxies (Stark et al. 2010, 2011; Curtis-Lake et al. 2012), the primary advantages are: (1) contamination (from low-redshift line-emitters) is very low, and (2) the background number density is known from wide-field observations of blank fields (Hu et al. 2004; Shimasaku et al. 2006; Murayama et al. 2007; Ouchi et al. 2008).

The challenge for the future is to perform similar QSO–galaxy clustering analyses at higher redshifts. Indeed, if QSOs at $z \sim 5$–6 trace halos of similar masses as those at $z \sim 4$ (i.e., $M_{\mathrm{halo}} \gtrsim 10^{12}\ M_{\odot}$), then we expect a strong QSO–galaxy

cross-correlation function. Clustering studies based on both broad-band and narrow-band imaging, as well as follow-up spectroscopy, are now needed to search for these overdensities around QSOs, and to clarify the relationship between early supermassive BHs and the formation of structure in the early universe.

ORCID iDs

Cristina García-Vergara ⓘ https://orcid.org/0000-0002-4811-0677
Joseph F. Hennawi ⓘ https://orcid.org/0000-0002-7054-4332
Hans-Walter Rix ⓘ https://orcid.org/0000-0003-4996-9069

References

Adelberger, K. L., & Steidel, C. C. 2005, ApJ, 630, 50
Adelberger, K. L., Steidel, C. C., Shapley, A. E., & Pettini, M. 2003, ApJ, 584, 45
Appenzeller, I., & Rupprecht, G. 1992, Msngr, 67, 18
Bahcall, J. N., Kirhakos, S., Saxe, D. H., & Schneider, D. P. 1997, ApJ, 479, 642
Bañados, E., Venemans, B., Walter, F., et al. 2013, ApJ, 773, 178
Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ApJS, 227, 11
Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559
Bertin, E. 2006, in ASP Conf. Ser. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel et al. (San Francisco, CA: ASP), 112
Bertin, E., & Arnouts, S. 1996, A&As, 117, 393
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in ASP Conf. Ser. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley (San Francisco, CA: ASP), 228
Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2010, ApJL, 709, L133
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Natur, 470, 233
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chabrier, G. 2003, PASP, 115, 763
Ciardullo, R., Gronwall, C., Wolf, C., et al. 2012, ApJ, 744, 110
Coil, A. L., Hennawi, J. F., Newman, J. A., Cooper, M. C., & Davis, M. 2007, ApJ, 654, 115
Curtis-Lake, E., McLure, R. J., Pearce, H. J., et al. 2012, MNRAS, 422, 1425
Dodelson, S. 2003, Modern Cosmology (New York: Academic Press)
Fan, X., White, R. L., Davis, M., et al. 2000, AJ, 120, 1167
Fanidakis, N., Macciò, A. V., Baugh, C. M., Lacey, C. G., & Frenk, C. S. 2013, MNRAS, 436, 315
Fernández-Soto, A., Lanzetta, K. M., & Chen, H.-W. 2003, MNRAS, 342, 1215
Ferrarese, L. 2002, ApJ, 578, 90
Ferrarese, L., & Merritt, D. 2000, ApJL, 539, L9
Fukugita, M., Shimasaku, K., & Ichikawa, T. 1995, PASP, 107, 945

---

[11] http://www.mpia.de/ENIGMA/

Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJL, 539, L13
Gehrels, N. 1986, ApJ, 303, 336
Giavalisco, M., Steidel, C. C., Adelberger, K. L., et al. 1998, ApJ, 503, 543
Gioia, I. M., Henry, J. P., Mullis, C. R., et al. 2001, ApJL, 553, L105
Hennawi, J. F., Myers, A. D., Shen, Y., et al. 2010, ApJ, 719, 1672
Hennawi, J. F., Prochaska, J. X., Cantalupo, S., & Arrigoni-Battaia, F. 2015, Sci, 348, 779
Hennawi, J. F., Strauss, M. A., Oguri, M., et al. 2006, AJ, 131, 1
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hu, E. M., Cowie, L. L., Capak, P., et al. 2004, AJ, 127, 563
Husband, K., Bremer, M. N., Stanway, E. R., et al. 2013, MNRAS, 432, 2869
Intema, H. T., Venemans, B. P., Kurk, J. D., et al. 2006, A&A, 456, 433
Jones, T., Stark, D. P., & Ellis, R. S. 2012, ApJ, 751, 51
Kashikawa, N., Kitayama, T., Doi, M., et al. 2007, ApJ, 663, 765
Kim, S., Stiavelli, M., Trenti, M., et al. 2009, ApJ, 695, 809
Kormendy, J., & Bender, R. 2011, Natur, 469, 377
Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Lee, K.-S., Giavalisco, M., Gnedin, O. Y., et al. 2006, ApJ, 642, 63
Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285
Mazzucchelli, C., Bañados, E., Decarli, R., et al. 2017, ApJ, 834, 83
McLeod, K. K., & Bechtold, J. 2009, ApJ, 704, 415
Morselli, L., Mignoli, M., Gilli, R., et al. 2014, A&A, 568, A1
Murayama, T., Taniguchi, Y., Scoville, N. Z., et al. 2007, ApJS, 172, 523
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2010, ApJL, 709, L16
Oke, J. B. 1974, ApJS, 27, 21
Ota, K., Kashikawa, N., Malkan, M. A., et al. 2008, arXiv:0804.3448
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2005, ApJL, 620, L1
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004a, ApJ, 611, 660
Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004b, ApJ, 611, 685
Overzier, R. A. 2016, A&Ar, 24, 14
Overzier, R. A., Bouwens, R. J., Cross, N. J. G., et al. 2008, ApJ, 673, 143
Padmanabhan, N., White, M., & Eisenstein, D. J. 2007, MNRAS, 376, 1702
Padmanabhan, N., White, M., Norberg, P., & Porciani, C. 2009, MNRAS, 397, 1862
Padmanabhan, T. 2006, in AIP Conf. Ser. 843, Graduate School in Astronomy: X, ed. S. Daflon et al. (Melville, NY: AIP), 111
Patat, F., Moehler, S., O'Brien, K., et al. 2011, A&A, 527, A91
Peebles, P. J. E. 1980, The Large-scale Structure of the Universe (Princeton, NJ: Princeton Univ. Press)
Sargent, W. L. W., & Turner, E. L. 1977, ApJL, 212, L3

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Schneider, P. 2015, Extragalactic Astronomy and Cosmology: An Introduction (Berlin: Verlag)
Shapley, A. E., Steidel, C. C., Pettini, M., & Adelberger, K. L. 2003, ApJ, 588, 65
Shapley, A. E., Steidel, C. C., Pettini, M., Adelberger, K. L., & Erb, D. K. 2006, ApJ, 651, 688
Shen, Y., Hennawi, J. F., Shankar, F., et al. 2010, ApJ, 719, 1693
Shen, Y., McBride, C. K., White, M., et al. 2013, ApJ, 778, 98
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, AJ, 133, 2222
Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, ApJ, 697, 1656
Shimasaku, K., Kashikawa, N., Doi, M., et al. 2006, PASJ, 58, 313
Simpson, C., Mortlock, D., Warren, S., et al. 2014, MNRAS, 442, 3454
Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408, 1628
Stark, D. P., Ellis, R. S., & Ouchi, M. 2011, ApJL, 728, L2
Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2000, ApJ, 532, 170
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2005, ApJ, 626, 44
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJL, 462, L17
Steidel, C. C., Pettini, M., & Hamilton, D. 1995, AJ, 110, 2519
Stiavelli, M., Djorgovski, S. G., Pavlovsky, C., et al. 2005, ApJL, 622, L1
Stone, R. P. S. 1996, ApJS, 107, 423
Toshikawa, J., Kashikawa, N., Ota, K., et al. 2012, ApJ, 750, 137
Trainor, R. F., & Steidel, C. C. 2012, ApJ, 752, 39
Utsumi, Y., Goto, T., Kashikawa, N., et al. 2010, ApJ, 721, 1680
Venemans, B. P., Röttgering, H. J. A., Miley, G. K., et al. 2007, A&A, 461, 823
Vestergaard, M. 2002, ApJ, 571, 733
Vikhlinin, A., Kravtsov, A. V., Burenin, R. A., et al. 2009, ApJ, 692, 1060
Wang, T., Elbaz, D., Daddi, E., et al. 2016, ApJ, 828, 56
White, M., Martini, P., & Cohn, J. D. 2008, MNRAS, 390, 1179
White, M., Myers, A. D., Ross, N. P., et al. 2012, MNRAS, 424, 933
Willott, C. J., Percival, W. J., McLure, R. J., et al. 2005, ApJ, 626, 657
Worseck, G., & Prochaska, J. X. 2011, ApJ, 728, 23
Wyithe, J. S. B., & Loeb, A. 2002, ApJ, 581, 886
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579
Zheng, W., Overzier, R. A., Bouwens, R. J., et al. 2006, ApJ, 640, 574