



# Gemini GNIRS Near-infrared Spectroscopy of 50 Quasars at $z \gtrsim 5.7$

Yue Shen<sup>1,2,13</sup> , Jin Wu<sup>3,4</sup> , Linhua Jiang<sup>4</sup> , Eduardo Bañados<sup>5</sup>, Xiaohui Fan<sup>6</sup> , Luis C. Ho<sup>3,4</sup> , Dominik A. Riechers<sup>7</sup> , Michael A. Strauss<sup>8</sup> , Bram Venemans<sup>9</sup> , Marianne Vestergaard<sup>6,10</sup> , Fabian Walter<sup>9</sup> , Feige Wang<sup>11</sup> , Chris Willott<sup>12</sup> ,  
Xue-Bing Wu<sup>3,4</sup> , and Jinyi Yang<sup>6</sup>

<sup>1</sup> Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; [shenyue@illinois.edu](mailto:shenyue@illinois.edu)

<sup>2</sup> National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

<sup>3</sup> Department of Astronomy, School of Physics, Peking University, Beijing 100871, People's Republic of China

<sup>4</sup> Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China

<sup>5</sup> The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA

<sup>6</sup> Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA

<sup>7</sup> Cornell University, Space Sciences Building, Ithaca, NY 14853, USA

<sup>8</sup> Princeton University Observatory, Peyton Hall, Princeton, NJ 08544, USA

<sup>9</sup> Max Planck Institut für Astronomie, Königstuhl 17, D-69117, Heidelberg, Germany

<sup>10</sup> The Niels Bohr Institute at University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

<sup>11</sup> Department of Physics, Broida Hall, University of California, Santa Barbara, CA 93106, USA

<sup>12</sup> NRC Herzberg, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada

Received 2018 September 14; revised 2019 January 23; accepted 2019 January 28; published 2019 March 1

## Abstract

We report initial results from a large Gemini program to observe  $z \gtrsim 5.7$  quasars with GNIRS near-IR spectroscopy. Our sample includes 50 quasars with simultaneous  $\sim 0.85\text{--}2.5\ \mu\text{m}$  spectra covering the rest-frame ultraviolet and major broad emission lines from Ly $\alpha$  to Mg II. We present spectral measurements for these quasars and compare with their lower redshift counterparts at  $z = 1.5\text{--}2.3$ . We find that when quasar luminosity is matched, there are no significant differences between the rest-UV spectra of  $z \gtrsim 5.7$  quasars and the low- $z$  comparison sample. High- $z$  quasars have similar continuum and emission line properties and occupy the same region in the black hole mass and luminosity space as the comparison sample, accreting at an average Eddington ratio of  $\sim 0.3$ . There is no evidence for super-Eddington accretion or hypermassive ( $>10^{10} M_{\odot}$ ) black holes within our sample. We find a mild excess of quasars with weak C IV lines relative to the control sample. Our results, corroborating earlier studies but with better statistics, demonstrate that these high- $z$  quasars are already mature systems of accreting supermassive black holes operating with the same physical mechanisms as those at lower redshifts.

**Key words:** black hole physics – galaxies: active – line: profiles – quasars: general – surveys

**Supporting material:** figure sets, FITS file, machine-readable table

## 1. Introduction

The combination of recent progress in deep imaging sky surveys and optical/near-IR spectroscopic capabilities on large-aperture telescopes has revolutionized the study of high-redshift ( $z \gtrsim 5.7$ ) quasars (e.g., Fan et al. 2001, 2003, 2004, 2006a; Cool et al. 2006; Goto 2006; McGreer et al. 2006; Venemans et al. 2007, 2015; Willott et al. 2007, 2009, 2010b; Jiang et al. 2008, 2009, 2015, 2016; Mortlock et al. 2009, 2011; Morganson et al. 2012; Bañados et al. 2014, 2016; Matsuoka et al. 2016, 2018a, 2018b; Reed et al. 2017; Tang et al. 2017). These earliest quasars, powered by supermassive black holes (SMBHs), not only constrain the physics of black hole (BH) accretion and host galaxy assembly at cosmic dawn but also provide critical information on the physical conditions of the intergalactic medium in the early universe (e.g., Fan et al. 2006a). Since the first successful identification of  $z \gtrsim 6$  quasars (Fan et al. 2001), the searches for these high- $z$  quasars have matured and dramatically increased the inventory of these rare objects. There are now more than 250 quasars known at  $z > 5.6$ , with the most distant quasar reaching  $z \sim 7.5$  (Bañados et al. 2018b). These impressive statistics now enable a transition from individual case studies to ensemble studies of high- $z$  quasars.

The vast majority of  $z \gtrsim 6$  quasars are selected from wide-field optical+IR imaging with the dropout technique and are confirmed with optical spectroscopy (e.g., Fan et al. 2001). The confirmation optical spectroscopy provides limited information about these high- $z$  quasars themselves, and to probe the physical properties of these systems (e.g., BH mass, spectral properties, broad and narrow absorption lines, etc.), near-IR spectroscopy is necessary to cover the rest-frame UV from C IV 1549 Å to Mg II 2798 Å. Some earlier near-IR spectra of  $z \gtrsim 6$  quasars already provided a glimpse of their physical properties, such as BH masses and Eddington ratios (e.g., Jiang et al. 2007; Kurk et al. 2007; Willott et al. 2010a; De Rosa et al. 2014; Mazzucchelli et al. 2017; Kim et al. 2018) and chemical abundances in the broad-line region (e.g., Barth et al. 2003; Kurk et al. 2007; De Rosa et al. 2011). These measurements of BH properties also facilitate the study of the relations between BH growth and host galaxy assembly in the early universe, where the host galaxy properties are mostly inferred from the molecular gas emission in the millimeter regime (e.g., Walter et al. 2003, 2009; Riechers et al. 2009; Wang et al. 2013; Venemans et al. 2016; Decarli et al. 2017, 2018; Willott et al. 2017). However, near-IR spectroscopy of  $z \gtrsim 6$  quasars is expensive, and thus only a small fraction of them have existing near-IR spectra, some of which are also of low quality (e.g., insufficient spectral coverage, signal-to-noise ratio, and spectral

<sup>13</sup> Alfred P. Sloan Research Fellow.

resolution). Given the importance of understanding the physical properties and growth of these high- $z$  SMBHs, we have assembled a large sample of high-quality near-IR spectra for these objects, as statistics are the key to addressing most of the relevant questions pertaining to the early growth and evolution of these SMBHs and their hosts.

We have conducted a Gemini-GNIRS survey (Gemini Large and Long Program LLP-7) to obtain near-IR  $YJHK$  simultaneous spectroscopy for a sample of 50 quasars at  $z \gtrsim 5.7$ . The primary goal is to compile a large statistical sample of these objects with near-IR spectroscopic coverage to understand their rest-frame UV properties and derive physical properties such as BH mass and Eddington ratios with better spectral quality and sample statistics. This sample, when supplemented with existing near-IR spectroscopic data for high- $z$  quasars, can be used to measure the demography of these SMBHs in terms of their mass function (after the complicated selection function is properly quantified), as well as statistical studies such as chemical abundance in the quasar broad-line region, the prevalence of broad and narrow absorption line systems, and so on. The GNIRS sample presented here is substantially larger than earlier near-IR samples and enables a uniform analysis with the same data format and spectral fitting tools.

In this paper we describe the details of our program (Section 2) and present the spectral fitting results on the GNIRS sample (Section 3). As one application of our data, we compare the rest-frame UV spectral properties of our sample with their lower redshift counterparts (Section 4) and discuss the implications in Section 5. Additional applications of our data will be presented in successive work. Throughout this paper we adopt a flat lambda cold dark matter cosmology with  $\Omega_\Lambda = 1 - \Omega_0 = 0.7$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2. Data

Our target pool includes all known quasars at  $z \gtrsim 5.7$  compiled from the literature and unpublished works for which we are collecting near-IR spectroscopy from different sources. The majority of these quasars were discovered from dedicated high- $z$  quasar searches targeted to different depths from wide-area imaging surveys. During the 15B–17A semesters we observed a total of 51 quasars from our target pool with GNIRS on Gemini-North during band 2 allocation, although for some targets the observations were only partially executed. Most of the observations were carried out in queue mode by Gemini staff. We preferentially excluded objects with existing near-IR spectra with reasonably good quality and  $JHK$  wavelength coverage. Bright targets were assigned a higher priority over faint targets given the total time allocation. The selection of targets also depends on their visibility in each observing semester. Our GNIRS sample is by no means a complete flux-limited sample, but it covers a broad range of luminosities and samples the diversity of quasar properties at  $z \gtrsim 5.7$ . As described in Section 4, we create a comparison sample at lower  $z$  that matches the luminosity distribution of the GNIRS sample for a fair comparison. The incompleteness in our near-IR spectroscopic follow-up will be taken into account in our future work that requires the detailed selection function (such as the BH mass function).

We used the cross-dispersion (XD) mode (321/mm) on GNIRS with the short blue camera and a slit width of  $0''.675$  to balance the need for spectral resolution and throughput. This configuration provides simultaneous spectral coverage of

$\sim 0.85\text{--}2.5 \mu\text{m}$  at a spectral resolution  $R \sim 650$  with a pixel scale of  $0''.15/\text{pix}$ , sufficient to resolve the broad emission lines. We used a fixed position angle (P.A. =  $90^\circ$  east of north) to minimize the effect of differential flexure, as recommended on the GNIRS instrument page.<sup>14</sup> Given the typical low airmass of our observations ( $<1.2$ ), the atmospheric differential refraction introduces positional shifts of  $\lesssim 0''.06$  across  $JHK$  bands, which is less important than differential flexure and will not affect our relative flux calibration much. Each target was observed for a period of 30 minutes to 5 hr depending on the target brightness (with a typical single-exposure time of 300 s), in order to reach an accumulative continuum signal-to-noise ratio (S/N) of  $\sim 5$  per pixel averaged over  $H$  band. The observations were executed in the standard ABBA sequence (with an offset of  $3''$  along the slit between A and B positions), and we observed one A0 star for telluric correction and flux calibration immediately preceding or following the science observation. For absolute flux calibration we use the photometry of the targets compiled from the literature, as described in Section 2.1. The GNIRS spectra of our sample are displayed in Figure 2.

### 2.1. GNIRS Data Reduction

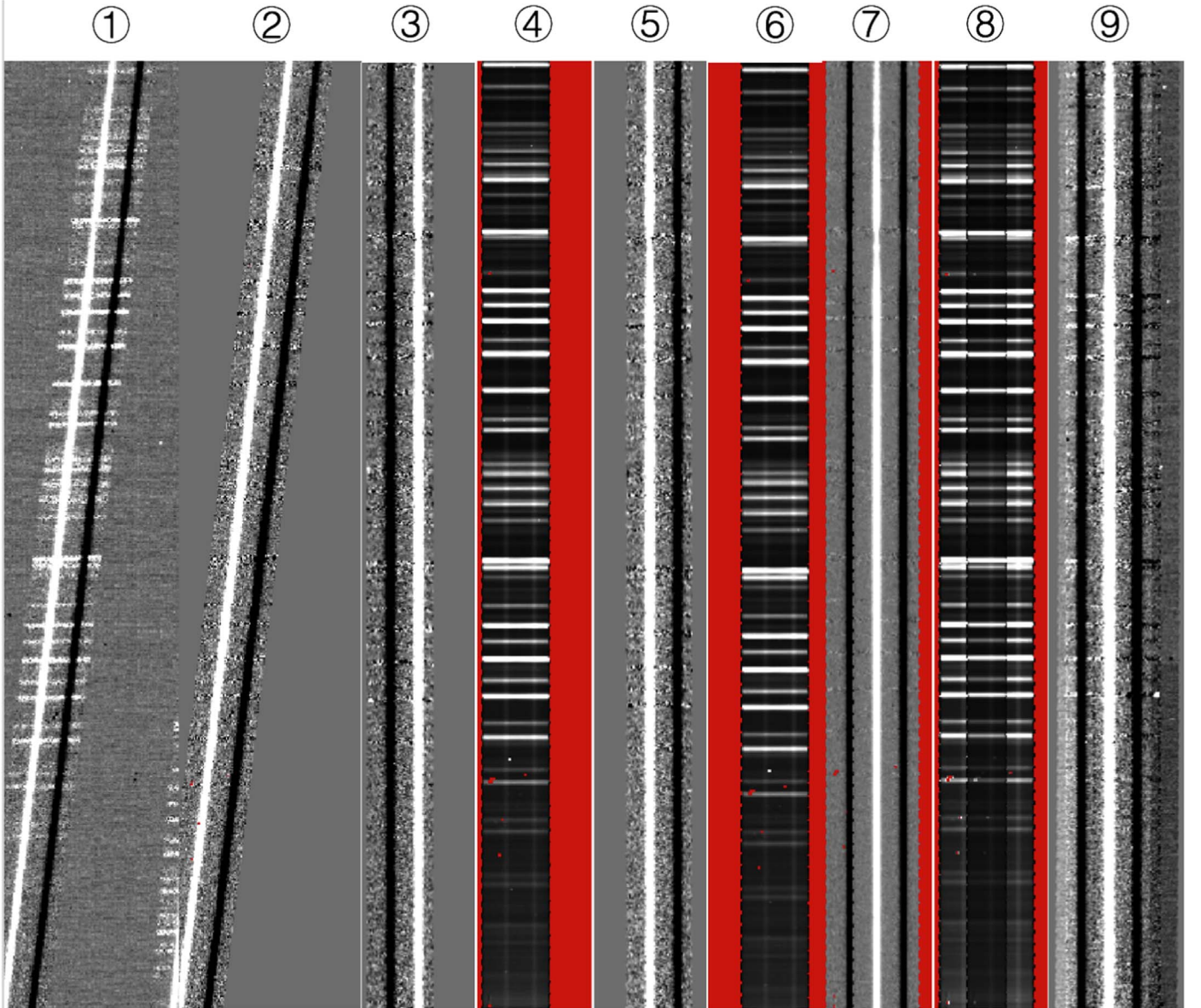
Our data reduction pipeline is a combination of two existing pipelines for GNIRS. The first one is the PyRAF-based XDNIRS (Mason et al. 2015); the other one is the IDL-based XIDL package.<sup>15</sup> XDNIRS is commonly used for bright nearby galaxy targets; it uses the standard ABBA method to perform sky subtraction and spectrum combination with 2D images before final 1D extraction, while XIDL does an extra spline fitting after the A minus B step and extracts the 1D spectrum for each single subtraction image. Both methods have their pros and cons, and we chose to combine these two methods to make our final results more accurate and robust. In brief, our data reduction consists of 3 steps, 1. preprocessing; 2. wavelength and S-distortion mapping; 3. 1D spectrum extraction and combination. The data are processed in ABBA sequence groups. All final 1D spectra are calibrated and stored in vacuum wavelengths.

In the preprocessing step, we first clean the large-scale pattern noise by fitting a periodic function to the image and subtracting it using existing routines in the XIDL package. Cosmic rays are masked using the LACosmic method (van Dokkum 2001). The images are then flatfielded using dome flats. We split the image into different echelle orders and then apply the A-B method to do sky subtraction for each ABBA sequence. The typical exposure time for a single observation is 300 s. During this time, the sky emission may have large variations in the infrared and the A-B method will lead to large residuals near sky emission lines in such cases. After this step, the XDNIRS pipeline directly combines the sky-subtracted frames, which is reasonable for bright objects but not ideal for our faint targets. We therefore apply an extra correction to suppress these residuals.

In the wavelength and S-distorting mapping step, we use arc observations to fit the 2D  $y$ -wavelength relation and use the pinhole observations to get the 2D  $x$ -slit mapping. Having these mappings, we are able to correct the skyline residuals using the

<sup>14</sup> <https://www.gemini.edu/sciops/instruments/gnirs/spectroscopy/observing-strategies>

<sup>15</sup> <http://www.ucolick.org/~xavier/IDL/>



**Figure 1.** Sky subtraction and 2D combination results for a single order with b-spline correction by our pipeline. The nine images are, from left to right, 1. A-B; 2. A-B corrected by b-spline; 3. align B position into the center; 4. variance image of 3; 5. align A position into the center; 6. variance image of 5; 7. combined result of multiple aligned A positions and B positions; 8. variance image of 7; 9. the same combined result without b-spline correction from XDNIRS for the same order. The difference between images 7 and 9 demonstrates the improvement of our hybrid approach over using XDNIRS alone without the b-spline correction.

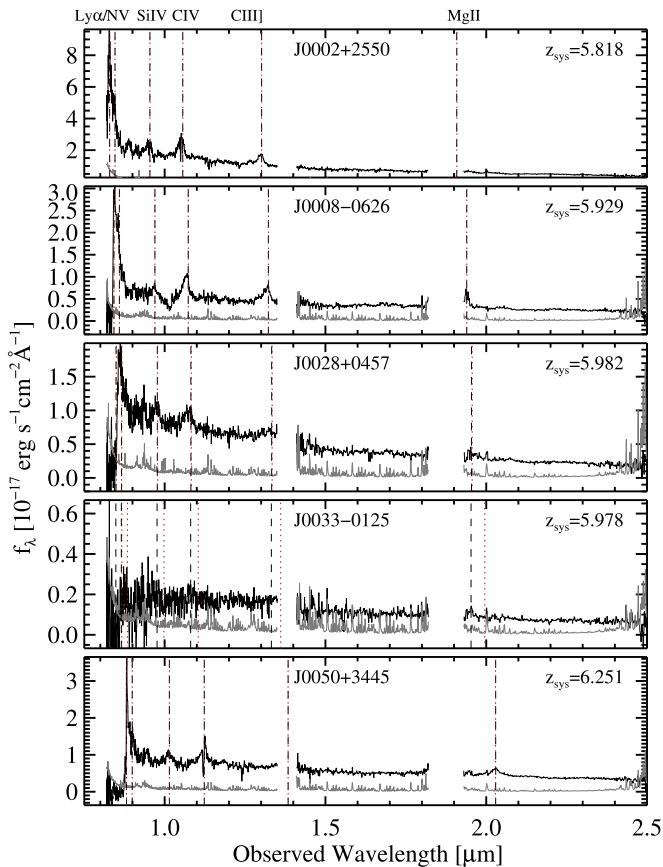
b-spline method (Kelson 2003). This correction improves the data reduction in the wavelength range with numerous sky lines. Figure 1 demonstrates the improvement using the b-spline method.

After the b-spline correction of skyline residuals, a common rectifying grid is determined for each spectrum order in each ABBA image sequence. We used our custom remapping function, which guarantees the conservation of total flux. Using the mappings above and our remapping function, all the b-spline corrected data are rectified onto the same grid. The XIDL pipeline makes 1D extraction directly from the nonrectified data and then combines the spectrum. Instead, we decided to first combine all 2D images and then use a single extraction to get the final spectrum. The main reason is that we will sometimes have undetected bad pixels after the pattern noise removal and cosmic-ray rejection process, which can be identified using sigma rejection method in the 2D combine

process. It is more robust to reject these bad pixels in the 2D image than in the extracted 1D spectrum. All the A-B images were visually checked before they were combined to make sure that all the features in the final spectrum are genuine and were not due to instrumental defects.

In the last 1D spectrum extraction and combination step, we use the boxsum (i.e., with a boxcar aperture) method to extract 1D spectra with an aperture size of 5 pixels. We also tested optimal extraction (Horne 1986) and found nearly identical results. We obtain both the quasar raw spectrum and standard star raw spectrum as well as their error spectrum. We perform absorption line removal for the standard star by fitting Voigt profiles at the position of hydrogen lines. After that, most of the absorption features in the standard star spectrum come from atmospheric absorption. We then correct for telluric absorption in our object spectrum by dividing it by the spectrum of the standard star. We always chose the standard star that has the





**Figure 2.** GNIRS spectra of our high- $z$  quasar sample. The spectrum (black; error in gray) has been smoothed with a 5-pixel median filter. The vertical black dashed lines indicate the major emission lines using our best estimate of the systemic redshifts  $z_{\text{sys}}$  (see Section 3 for details), and the red dotted lines indicate those using the original redshifts. The complete figure set (10 images) is available in the online journal.

(The complete figure set [10 images] is available.)

smallest difference in air mass and observation time for a given science target. After telluric correction, we obtain the relative flux ratio between the science object and the standard star. We then use a blackbody spectrum to model the standard star with its effective temperature. Multiplying the relative flux ratio by the model star spectrum gives a flux-calibrated spectrum. Then different orders are combined using the flux in their overlap region, and we combine all the 1D spectra for one object in different ABBA sequences with  $5\sigma$  clipping to obtain the final 1D spectrum.

As the last step, we rescale the flux (density) using the available  $J$ -band magnitude of the specific quasar target (see Table 1). We ignore the possibility of quasar variability between the  $J$  observation and the GNIRS spectrum, which is less important than the sky and seeing variations between the quasar and standard star observations in our GNIRS program. Since these high- $z$  quasars are point sources in rest-frame UV, this last flux rescaling step effectively corrects for slit losses. There is one object (J1545+6028) for which we do not have available  $J$ -band magnitudes, and we used the standard star for absolute flux calibration as well.

If we exclude one object where no signal was received at all during exposures under poor observing conditions, our final GNIRS sample includes 50 quasars. The basic target information is summarized in Table 1.

### 3. Spectral Analysis

We fit the GNIRS spectrum following the approach detailed in, e.g., Shen & Liu (2012). In short, we shift the spectrum to rest frame using the initial redshift and fit the spectrum with a global continuum+emission line model.<sup>16</sup> This differs slightly from our earlier work, which used a local continuum+line fit around each broad line (e.g., Shen et al. 2008, 2011). Several wavelength windows free of major emission lines (except for the broadband FeII emission) are used to fit the global continuum as a first step. The global continuum is modeled by a power law plus a third-order polynomial, and UV FeII emission is modeled using empirical templates from the literature (e.g., Vestergaard & Wilkes 2001) that are scaled and broadened to match our spectrum. The additional third-order polynomial component is introduced to account for any peculiar continuum curvature in the rest-frame UV that may be caused by internal reddening, as observed in a small fraction of quasars (e.g., Shen et al. 2018). The continuum and the FeII emission form a pseudocontinuum, which is subtracted from the spectrum, leaving a line-only spectrum for which we model the emission lines with multiple Gaussians. We fit the broad emission lines in individual line complexes specified in Table 2, where the main line and adjacent weak lines are fit simultaneously. We found that the number of Gaussians we use for each line is sufficient to reproduce the line profile, and using more Gaussians is unnecessary given the medium spectral quality of our sample. Figure 3 compares our model and the data around several major broad lines in one of our objects, and the full set of fitting results is provided as an online figure set. This figure set can also be used to reject certain line fits, e.g., if the line largely falls within one of the telluric gaps in the spectrum.

In some cases, the original redshift from the discovery paper is inaccurate, leading to small coverages of the broad emission lines in the deredshifted near-IR spectrum and potential biases in the emission line measurements. Therefore we perform a second fit to the spectrum with the updated redshift from the first fit and update the fitting results for all objects in our sample. One such iteration is sufficient for the fit to converge.

We measure the continuum and emission line properties from the best-fit models. To estimate the uncertainties of these spectral measurements, we use a Monte Carlo approach (e.g., Shen et al. 2011): for each original spectrum, we create a trial sample of 50 mock spectra; each is generated by randomly shuffling the flux (density) values in the original spectrum by adding a Gaussian random deviate at each pixel with its dispersion equal to the spectral error at that pixel; the same fitting approach was applied to the mock spectra and the measured quantities recorded; the nominal uncertainties of the measured spectral properties are then estimated as the semiquantile of the range enclosing the 16th and 84th percentiles of the distribution. Adding flux perturbations to the original spectrum instead of our model spectrum preserves details in the spectral features that are not captured or well fit by our model (such as absorption lines). On the other hand, the original spectrum is already a perturbed version of the noise-free true spectrum, hence the mock spectra are slightly noisier than the original spectrum, and therefore our approach will produce overly conservative measurement errors in the spectral

<sup>16</sup> The full technical details of the spectral fitting are described in Shen et al. (2018), and the associated code is made public along with that paper.

**Table 1**  
Sample Summary

ObjID (1)	R.A. (J2000) (2)	Decl. (J2000) (3)	$z_{\text{old}}$ (4)	$z_{\text{sys}}$ (5)	$z_{\text{sys, err}}$ (6)	$J$ (7)	$J$ err (8)	$H$ (9)	$H$ err (10)	$K_s$ (11)	$K_s$ err (12)	$J$ ref (13)	Dis. Ref (14)	Comment (15)
P000+26	00:01:21.63	+26:50:09.2	5.75	5.733	0.007	18.36	0.08	...	...	...	...	UHS	1	...
J0002+2550	00:02:39.39	+25:50:34.9	5.82	5.818	0.007	18.46	0.09	...	...	...	...	UHS	2	...
J0008-0626	00:08:25.77	-06:26:04.6	5.93	5.929	0.006	19.43	0.13	...	...	...	...	Jiang2016	3	BAL
J0028+0457	00:28:06.56	+04:57:25.7	5.99	5.982	0.012	19.16	0.12	19.05	0.20	18.32	0.18	UKIDSS LAS	4	...
J0033-0125	00:33:11.40	-01:25:24.9	6.13	5.978	0.010	20.64	0.20	...	...	...	...	Willott2007	5	...
J0050+3445	00:50:06.67	+34:45:22.6	6.25	6.251	0.006	19.05	0.15	...	...	...	...	UHS	6	...
J0055+0146	00:55:02.91	+01:46:18.3	6.02	6.017	0.054	20.93	0.15	...	...	...	...	Willott2009	7	...
J0136+0226	01:36:03.17	+02:26:05.7	6.21	6.206	0.009	21.15	0.22	...	...	...	...	Willott2010	6	...
J0203+0012	02:03:32.38	+00:12:29.2	5.86	5.709	0.010	19.05	0.10	17.75	0.07	17.32	0.09	UKIDSS LAS	8	BAL
J0221-0802	02:21:22.71	-08:02:51.5	6.16	6.161	0.054	21.09	0.14	...	...	...	...	Willott2010	6	...
J0227-0605	02:27:43.29	-06:05:30.2	6.2	6.212	0.007	21.07	0.22	...	...	...	...	VIK	7	...
J0300-2232	03:00:44.18	-22:32:27.2	5.7	5.684	0.008	18.71	0.08	19.42	0.08	...	...	Banados2014	4	...
J0353+0104	03:53:49.72	+01:04:04.4	6.049	6.057	0.005	19.45	0.16	18.53	0.17	18.16	0.22	UKIDSS LAS	9	BAL
P060+24	04:02:12.69	+24:51:24.4	6.18	6.170	0.006	19.05	0.10	18.69	0.16	17.81	0.14	UKIDSS GCS	1	...
J0810+5105	08:10:54.32	+51:05:40.1	5.80	5.805	0.010	18.77	0.06	...	...	...	...	Jiang2016	10	...
J0835+3217	08:35:25.76	+32:17:52.6	5.89	5.902	0.009	20.50	0.20	...	...	...	...	Jiang2016	10	...
J0836+0054	08:36:43.85	+00:54:53.3	5.82	5.834	0.007	17.70	0.03	17.02	0.03	16.18	0.03	UKIDSS GCS	11	radio-loud
J0840+5624	08:40:35.09	+56:24:19.9	5.85	5.816	0.010	19.00	0.01	...	...	...	...	Jiang2016	12	...
J0841+2905	08:41:19.52	+29:05:04.4	5.96	5.954	0.005	19.17	0.09	18.62	0.18	17.84	0.15	UKIDSS LAS	13	BAL
J0842+1218	08:42:29.43	+12:18:50.5	6.055	6.069	0.009	18.78	0.11	...	...	...	...	UHS	3	...
J0850+3246	08:50:48.25	+32:46:47.9	5.867	5.730	0.013	18.74	0.08	...	...	...	...	UHS	3	...
J0927+2001	09:27:21.82	+20:01:23.7	5.79	5.770	0.013	19.12	0.17	...	...	...	...	UHS	12	...
J1044-0125	10:44:33.04	-01:25:02.2	5.8	5.780	0.007	18.31	0.05	17.92	0.12	17.03	0.07	UKIDSS LAS	14	BAL
J1137+3549	11:37:17.73	+35:49:56.9	6.01	6.009	0.010	18.48	0.07	...	...	...	...	UHS	12	...
J1143+3808	11:43:38.33	+38:08:28.7	5.81	5.800	0.010	19.00	0.14	...	...	...	...	UHS	10	...
J1148+0702	11:48:03.29	+07:02:08.3	6.339	6.344	0.006	19.36	0.11	18.39	0.12	17.51	0.11	UKIDSS LAS	10	...
J1148+5251	11:48:16.64	+52:51:50.3	6.43	6.416	0.006	18.17	0.06	...	...	...	...	UHS	15	...
J1207+0630	12:07:37.43	+06:30:10.1	6.04	6.028	0.013	19.35	0.14	...	...	17.50	0.12	UKIDSS LAS	3	...
J1243+2529	12:43:40.81	+25:29:23.9	5.85	5.842	0.006	19.21	0.13	18.29	0.10	17.54	0.10	UKIDSS LAS	10	...
J1250+3130	12:50:51.93	+31:30:22.9	6.13	6.138	0.005	19.22	0.12	18.17	0.14	17.40	0.09	UKIDSS LAS	12	BAL
J1257+6349	12:57:57.47	+63:49:37.2	6.02	5.992	0.010	19.78	0.08	...	...	...	...	Jiang2016	3	...
J1335+3533	13:35:50.81	+35:33:15.8	5.93	5.870	0.020	18.90	0.13	...	...	17.61	0.14	UKIDSS LAS	12	...
P210+27	14:01:47.34	+27:49:35.0	6.14	6.166	0.007	19.56	0.21	...	...	...	...	UHS	1	...
J1403+0902	14:03:19.13	+09:02:50.9	5.86	5.787	0.013	19.17	0.10	18.59	0.10	17.93	0.11	UKIDSS LAS	3	...
J1425+3254	14:25:16.34	+32:54:09.6	5.85	5.862	0.006	19.22	0.17	...	...	...	...	UHS	16	...
J1427+3312	14:27:38.59	+33:12:41.0	6.12	6.118	0.005	19.68	0.05	...	...	...	...	McGreer2006	17	radio-loud; BAL
J1429+5447	14:29:52.17	+54:47:17.7	6.21	6.119	0.008	19.70	0.07	...	...	...	...	Willott2010	6	radio-loud
J1436+5007	14:36:11.74	+50:07:06.9	5.83	5.809	0.010	18.99	0.14	...	...	...	...	UHS	12	...
P228+21	15:14:44.91	+21:14:19.8	5.92	5.893	0.015	19.00	0.02	...	...	...	...	Banados2016	1	...
J1545+6028	15:45:52.08	+60:28:24.0	5.78	5.794	0.007	...	...	...	...	...	...	...	18	...
J1602+4228	16:02:53.98	+42:28:24.9	6.07	6.083	0.005	18.82	0.12	...	...	...	...	UHS	2	...
J1609+3041	16:09:37.27	+30:41:47.6	6.16	6.146	0.006	19.39	0.15	18.72	0.18	18.15	0.22	UKIDSS LAS	10	radio-loud
J1621+5155	16:21:00.92	+51:55:48.9	5.71	5.637	0.008	18.52	0.10	...	...	...	...	UHS	10	...
J1623+3112	16:23:31.81	+31:12:00.5	6.22	6.254	0.006	19.16	0.11	18.45	0.12	17.86	0.13	UKIDSS LAS	2	...
J1630+4012	16:30:33.90	+40:12:09.6	6.05	6.066	0.007	19.73	0.22	...	...	...	...	UHS	15	...
P333+26	22:15:56.63	+26:06:29.4	6.03	6.027	0.007	19.52	0.14	...	...	...	...	UHS	1	...
J2307+0031	23:07:35.35	+00:31:49.4	5.87	5.900	0.010	20.43	0.11	...	...	...	...	VHS	9	...

**Table 1**  
(Continued)

ObjID (1)	R.A. (J2000) (2)	Decl. (J2000) (3)	$z_{\text{old}}$ (4)	$z_{\text{sys}}$ (5)	$z_{\text{sys, err}}$ (6)	$J$ (7)	$J$ err (8)	$H$ (9)	$H$ err (10)	$K_s$ (11)	$K_s$ err (12)	$J$ ref (13)	Dis. Ref (14)	Comment (15)
J2310+1855	23:10:38.88	+18:55:19.7	6.04	5.956	0.011	17.94	0.05	...	...	...	...	UHS	10	...
J2329-0403	23:29:14.46	-04:03:24.1	5.90	5.883	0.007	21.06	0.19	...	...	...	...	Willott2009	7	BAL
J2356+0023	23:56:51.58	+00:23:33.3	6	5.987	0.014	21.18	0.07	...	...	...	...	Jiang2016	9	...

**Note.** Column (1) lists the abbreviated object IDs we assigned to each object in our GNIRS sample, which are adopted throughout this work. The original redshifts from the discovery paper for each object are compiled in Column (4), and our improved systemic redshifts (see Section 3) are compiled in Column (5). Note that some of these original redshifts may have been updated in other works.  $JHK$  magnitudes are in the Vega system. The  $J$  magnitudes are used to normalize the spectra, which are taken from different references and converted from AB magnitude (if needed) using  $J_{\text{Vega}} = J_{\text{AB}} - 0.94$ . Reference keys for the  $J$  magnitudes (and  $H$ ,  $K_s$  magnitudes if available): UHS—Dye et al. (2018); VIK—VIKING DR4 (Edge et al. 2013); VHS—McMahon et al. (2013); UKIDSS GCS—(Lawrence et al. 2007); UKIDSS LAS—UKIDSS LAS DR10 (Lawrence et al. 2007); Jiang2016—Jiang et al. (2016); Banados2016—Bañados et al. (2016); Banados2014—Bañados et al. (2014); Willott2009—Willott et al. (2009); Willott2007—Willott et al. (2007); McGreer2006—McGreer et al. (2006). Reference keys for the discovery papers in Column (14) are 1. Bañados et al. (2016), 2. Fan et al. (2004), 3. Jiang et al. (2015), 4. Bañados et al. (2014), 5. Willott et al. (2007), 6. Willott et al. (2010b), 7. Willott et al. (2009), 8. Venemans et al. (2007), 9. Jiang et al. (2008), 10. Jiang et al. (2016), 11. Fan et al. (2001), 12. Fan et al. (2006b), 13. Goto (2006), 14. Fan et al. (2000), 15. Fan et al. (2003), 16. Cool et al. (2006), 17. McGreer et al. (2006), and 18. Wang et al. (2016). Radio-loud identification is either from the discovery paper or from Bañados et al. (2015).

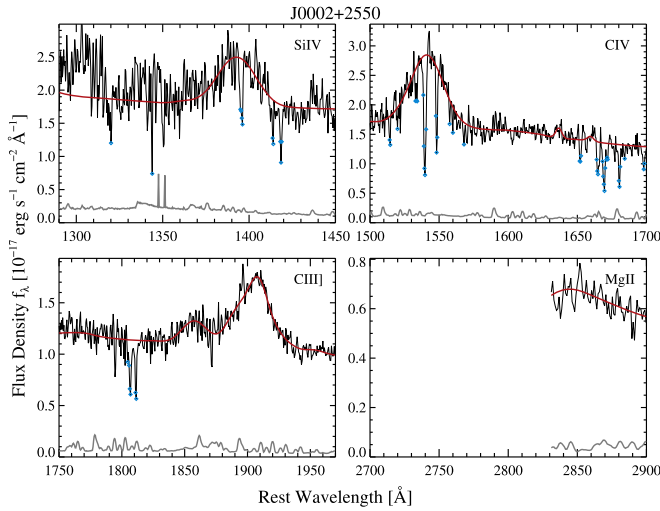
(This table is available in machine-readable form.)

**Table 2**  
Line Fitting Parameters

Line Name	Vacuum Rest Wavelength (Å)	$n_{\text{Gauss}}$	Complex
(1)	(2)	(3)	(4)
Mg II	2798.75	2	Mg II
C III]	1908.73	2	C III]
Si III]	1892.03	1	C III]
Al III	1857.40	1	C III]
C IV	1549.06	2	C IV
He II	1640.42	1	C IV
O III]	1663.48	1	C IV
Si IV	1399.41 <sup>a</sup>	2	Si IV/O IV]

**Notes.** Fitting parameters for the lines considered in this work. The third column lists the total number of Gaussians used for each line. Multiple lines in the same line complex as specified by name are fit simultaneously. The rest wavelengths of the lines are taken from Vanden Berk et al. (2001).

<sup>a</sup> The wavelength of the Si IV line is taken as the arithmetic mean of the central wavelengths of Si IV and O IV] (1396.76 Å and 1402.06 Å).



**Figure 3.** An example of our spectral fits around several major broad lines. The original spectrum and flux errors are shown in black and gray, respectively. The red lines are our model. The cyan points are masked absorption pixels in our fits. The complete figure set (50 images) is available in the online journal. (The complete figure set [50 images] is available.)

quantities. We have found that this approach produces very reasonable estimation of the measurement uncertainties (e.g., Shen et al. 2011). Shen et al. (2016, 2018) further studied the dependence of measured spectral quantities and their uncertainties on the S/N of the original spectrum and demonstrated that this overall fitting approach and error estimation are quite robust and do not depend on the S/N of the spectrum.

Once we have fitted the emission lines, we use the peak wavelength of the major broad emission lines measured from the best-fit model to improve the systemic redshift estimate of the quasar. Available optical spectra of these high- $z$  quasars often only cover the heavily absorbed Ly $\alpha$  line, and the resulting systemic redshift derived from Ly $\alpha$  is highly uncertain. Our near-IR spectroscopy covers additional broad

emission lines such as Si IV, C IV, C III], and Mg II, which provide more accurate systemic redshifts.

However, it is well known that quasar emission lines are often shifted from the systemic velocity due to various dynamical and/or radiative processes (e.g., Shen et al. 2016). We follow the approach detailed in Shen et al. (2016) to derive the best systemic redshift estimates based on a series of lines that takes into account the velocity shifts of each line relative to systemic as a function of quasar continuum luminosity. While it is difficult to measure the redshifts of these high- $z$  quasars to better than  $\sim 200 \text{ km s}^{-1}$  with broad lines only (e.g., Shen 2016; Shen et al. 2016), these new systemic redshifts are an improvement over some of the previous redshift estimates based on optical spectroscopy alone. In the catalog described in Table 3 (also in Table 1) we provide the improved systemic redshifts and their uncertainties including both measurement errors and systematic errors from intrinsic line velocity shifts as quantified in Shen et al. (2016). The median redshift uncertainty of our GNIRS sample based on the broad emission lines is  $\sim 330 \text{ km s}^{-1}$ . One broad absorption line (BAL) quasar, J0203+0012, has a systemic redshift  $z_{\text{sys}} = 5.777 \pm 0.011$  determined from C IV that is significantly lower than the redshift ( $z = 5.86$ ) based on optical spectroscopy reported in the discovery paper (Venemans et al. 2007). If we adopted the discovery redshift, then C IV and Si IV would be blueshifted by  $\sim 5000 \text{ km s}^{-1}$ , which would be extreme but still possible. Later near-IR spectroscopy of this object confirmed its BAL nature and derived revised redshifts of  $5.70 < z < 5.74$  (Mortlock et al. 2009) and  $z = 5.706$  (Ryan-Weber et al. 2009) largely based on C IV. Our derived redshift is slightly larger than the latter two redshift estimates because we took into account the typical blueshift of C IV. In any case, we are less confident about the systemic redshift determined for this object given its BAL nature and the lack of Mg II coverage.

To evaluate the overall improvement of our redshift estimation over previous results, we plot the median composite spectrum (see Section 4) around the C III] line in Figure 4. Unlike C IV, the peak of the C III] complex is known to have a modest average velocity shift of  $\sim -220 \text{ km s}^{-1}$  from systemic, with negligible luminosity dependence (Richards et al. 2011; Shen et al. 2016). The composite C III] profile using our improved redshifts is aligned with the expected location, while the composite line using previous redshifts shows a much larger blueshift. Thus we conclude that our improved redshifts are on average better than previous estimates.

We further compare our broad-line-based redshifts with those measured for 12 quasars from [C II] and CO mm observations in Table 4. The molecular lines mostly trace the gas in the host galaxy of these quasars and may have a velocity offset with respect to the broad-line region. Figure 5 shows the velocity difference between the molecular line redshifts and our near-IR spectroscopic redshifts. There is excellent agreement between the two redshifts for half of the objects. There are two quasars for which our redshift estimates are lower than those based on molecular lines by more than  $1500 \text{ km s}^{-1}$ . These two quasars, J1429+5447 and J2310+1855, appear to have reasonably well measured broad emission lines. The UV broad lines in these objects may be more blueshifted from systemic than their lower redshift counterparts with comparable luminosities, or we underestimated the systematic uncertainties of our broad-line-based redshifts. A larger sample of high- $z$

**Table 3**  
FITS Catalog Format

Column	Format	Units	Description
OBJID	STRING	...	Object ID of the GNIRS sample
ZOLD	DOUBLE	...	Original redshift from the discovery paper
ZSYS	DOUBLE	...	Improved systemic redshift
ZSYS_ERR	DOUBLE	...	Uncertainty in ZSYS
ZMGII	DOUBLE	...	Redshift based on broad Mg II
ZMGII_ERR	DOUBLE	...	Uncertainty in ZMGII
ZCIII	DOUBLE	...	Redshift based on the C III] complex
ZCIII_ERR	DOUBLE	...	Uncertainty in ZCIII
ZCIV	DOUBLE	...	Redshift based on broad C IV
ZCIV_ERR	DOUBLE	...	Uncertainty in ZCIV
ZSIV	DOUBLE	...	Redshift based on broad Si IV/O IV]
ZSIV_ERR	DOUBLE	...	Uncertainty in ZSIV
LOGL1350	DOUBLE	(erg s <sup>-1</sup> )	Continuum luminosity at restframe 1350 Å
LOGL1350_ERR	DOUBLE	(erg s <sup>-1</sup> )	Uncertainty in LOGL1350
LOGL1700	DOUBLE	(erg s <sup>-1</sup> )	Continuum luminosity at restframe 1700 Å
LOGL1700_ERR	DOUBLE	(erg s <sup>-1</sup> )	Uncertainty in LOGL1700
LOGL3000	DOUBLE	(erg s <sup>-1</sup> )	Continuum luminosity at restframe 3000 Å
LOGL3000_ERR	DOUBLE	(erg s <sup>-1</sup> )	Uncertainty in LOGL3000
LOGLBOL	DOUBLE	(erg s <sup>-1</sup> )	Bolometric luminosity
LOGLBOL_ERR	DOUBLE	(erg s <sup>-1</sup> )	Uncertainty in LOGLBOL
MGII	DOUBLE[5]	Å, km s <sup>-1</sup> , (erg s <sup>-1</sup> ), Å, Å	peak wavelength, FWHM, log $L_{\text{line}}$ , restframe EW, top 50% flux centroid
CIII_ALL	DOUBLE[5]	...	For the entire C III] complex (C III], Si III], Al III)
CIV	DOUBLE[5]	...	For the entire C IV line
SIIV_OIV	DOUBLE[5]	...	For the 1400 Å complex
MGII_ERR	DOUBLE[5]	Å, km s <sup>-1</sup> , (erg s <sup>-1</sup> ), Å, Å	Measurement errors in MGII
CIII_ALL_ERR	DOUBLE[5]	...	Measurement errors in CIII_ALL
CIV_ERR	DOUBLE[5]	...	Measurement errors in CIV
SIIV_OIV_ERR	DOUBLE[5]	...	Measurement errors in SIIV_OIV
LOGBH_CIV_VP06	DOUBLE	( $M_{\odot}$ )	Single-epoch virial BH mass based on C IV (Vestergaard & Peterson 2006)
LOGBH_CIV_VP06_ERR	DOUBLE	...	Measurement errors in LOGBH_CIV_VP06
LOGBH_MGII_S11	DOUBLE	( $M_{\odot}$ )	Single-epoch virial BH mass based on Mg II (Shen et al. 2011)
LOGBH_MGII_S11_ERR	DOUBLE	...	Measurement errors in LOGBH_MGII_S11
LOGBH	DOUBLE	( $M_{\odot}$ )	Adopted fiducial BH mass; based on Mg II if available, otherwise based on C IV
LOGBH_ERR	DOUBLE	...	Measurement errors in LOGBH
LOGEDD_RATIO	DOUBLE	...	Eddington ratio based on the fiducial BH mass
LOGEDD_RATIO_ERR	DOUBLE	...	Measurement errors in LOGEDD_RATIO

**Note.** 1. Bolometric luminosities were computed using a bolometric correction of 5.15 (Richards et al. 2006) using the 3000 Å monochromatic luminosities; 2. uncertainties are measurement errors only; 3. null value (indicating unmeasurable) is zero for a quantity and  $-1$  for its associated error, except for LOGEDD\_RATIO, where the null value is  $-99$ . Table 3 is published in its entirety in the machine-readable FITS format.

quasars with molecular line measurements will be useful to fully address this issue.

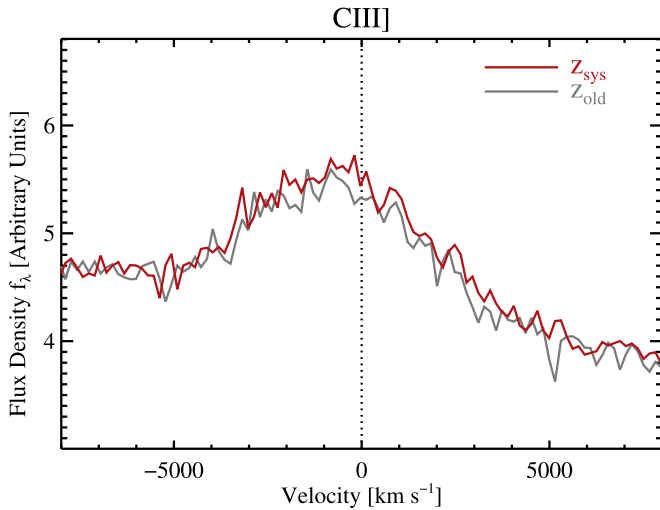
We compile our spectral measurements for our high- $z$  sample in an online FITS table, with the contents detailed in Table 3. The calibrated GNIRS spectra are displayed in Figure 2 and its online extension set.

Among the 50 quasars in our sample, we visually identified eight quasars with apparently strong broad C IV absorption lines (J0008–0626, J0203+0012, J0353+0104, J0841+2905, J1044–0125, J1250+3130, J1427+3312, J2329–0403). This fraction ( $\sim 16\%$ ) of BAL quasars is roughly consistent with the apparent fraction for lower redshift quasars (e.g., Gibson et al. 2009) but notably smaller than the intrinsic fraction of BAL quasars (e.g., Dai et al. 2008; Allen et al. 2011). As a result of our qualitative identification scheme and limited S/N of the spectra, some weak BALs may have been missed; on the other hand, we did not account for any luminosity differences between our sample and earlier samples. A more careful analysis of BALs in the high- $z$  quasar sample will be presented elsewhere.

A few quasars have low S/N in our GNIRS spectra (e.g., J0055+0146, J0221–0802, and J0227–0605), which result in low-quality spectral fits. We included these objects in our sample as the measurement uncertainties in the measured quantities are still reasonably quantified. They also contribute to the construction of the composite spectrum in Section 4. In addition, the extraction of J1427+3312 was complicated by a nearby star that falls in one of the AB positions during the observation, and at wavelengths greater than  $2.3 \mu\text{m}$  the spectrum may be compromised by this complication.

We notice that one object, J0136+0226, shows apparently narrow profiles in Ly $\alpha$  (FWHM =  $1467 \pm 25 \text{ km s}^{-1}$ ) and C IV (FWHM =  $3291 \pm 247 \text{ km s}^{-1}$ ). This may be due to emission line contamination beyond the broad-line region (such as the narrow-line region) or a combination of absorption and intrinsically narrower broad lines due to a small BH mass. Similar unobscured quasars (e.g., with detected quasar continuum) with narrow broad lines have been seen in  $z \sim 3$  SDSS quasars (Alexandroff et al. 2013) and in low-luminosity  $z > 5.7$  quasars (Matsuoka et al. 2018a).





**Figure 4.** Composite spectrum for our GNIRS sample around the C III] line. The red line shows the result using the improved redshifts, and the gray line shows the result using the old redshifts. The peak of the C III] complex is expected to be within  $\sim 200 \text{ km s}^{-1}$  of systemic (e.g., Shen et al. 2016). Therefore our redshift estimates are on average better than previous estimates.

The apparent radio-loud fraction of our GNIRS sample is  $\sim 8\%$  (4/50), similar to the fraction reported in Bañados et al. (2015).

#### 4. Rest-frame UV Properties

We now proceed to study the rest-frame UV spectral properties of our GNIRS sample. In this work we only consider our GNIRS sample because all data are processed and analyzed in a uniform manner. We are in the process of collecting other existing near-IR spectroscopic data of  $z \gtrsim 5.7$  quasars and will jointly analyze the expanded sample with a uniform spectral analysis. Nevertheless, our GNIRS sample already represents the largest statistical sample to date for the study of rest-frame UV properties of  $z \gtrsim 5.7$  quasars, in particular at the bright end.

To compare the properties of these high- $z$  quasars with their lower  $z$  counterparts, we create a control sample matched in rest frame 1350 Å continuum luminosity.<sup>17</sup> It is crucial to match the luminosity of quasars, not only because luminosity is related to the accretion power, but also because many emission line properties are functions of quasar luminosity. This is particularly true for high-ionization lines such as C IV, where the line rest-frame equivalent width (REW) decreases with luminosity (e.g., the Baldwin effect, Baldwin 1977) and the line profile (velocity shift and asymmetry) changes with luminosity in a systematic manner (e.g., Richards et al. 2002; Shen et al. 2008; Richards et al. 2011).

We select control quasars from the SDSS DR7 quasar catalog in Shen et al. (2011) with both C IV and Mg II coverage, because this range contains most of the important broad emission lines that trace the properties of quasar accretion. Using a luminosity grid (with  $\Delta \log L = 0.2$ ) and for each bin that the high- $z$  quasars reside in, we randomly select 50 times

more quasars from the DR7 catalog. Thus the resulting control sample has the same distribution in 1350 Å continuum luminosity as the high- $z$  sample. This control sample contains SDSS DR7 quasars at  $z \sim 1.5\text{--}2.3$  that cover most of the major rest-frame UV lines from Ly $\alpha$  to Mg II, providing a good reference sample to compare with our high- $z$  sample. We fit the control sample with the same global fitting recipe as for our GNIRS sample.<sup>18</sup>

We create median composite spectra for the high- $z$  sample and the control sample following Vanden Berk et al. (2001) to compare their average spectral properties. For the high- $z$  sample we adopt our best systemic redshift estimates, and for the control DR7 quasars we adopt the improved redshifts from Hewett & Wild (2010) during the coadding process. The Hewett & Wild (2010) redshifts are as good as those using our own redshift recipes in Shen et al. (2016), since both approaches take into account the velocity shifts of broad lines. As discussed in Vanden Berk et al. (2001), the median composite spectrum is more suitable for studying the relative line strength than are other spectral averaging methods. Figure 6 displays the median composite spectra for different samples. High- $z$  quasars have similar line strengths in most of the broad lines. In particular, the UV Fe II strength relative to broad Mg II is almost identical to that of low- $z$  quasars, as discovered in earlier small samples (Barth et al. 2003; De Rosa et al. 2011). This line ratio can be used to measure the most precise alpha/iron abundance ratio for these high- $z$  quasars. However, Ly $\alpha$  is apparently much weaker because of the much stronger absorption in the high- $z$  sample. Some of the narrow emission lines such as [Ne V] are also considerably weaker in the high- $z$  and control composite spectra than in the Vanden Berk et al. composite spectrum. This is likely a combination of the Baldwin effect and the fact that higher luminosity quasars (in the high- $z$  and the control samples) have more massive host galaxies, leading to broader and thus less prominent narrow emission lines such as [Ne V].

Both the high- $z$  sample and the control sample have an average C IV profile that is more blueshifted from the low-ionization lines such as Mg II than the median composite spectrum from Vanden Berk et al. (2001), as further demonstrated in Figure 7. The latter composite spectrum was generated using a sample of quasars that have lower luminosity than our high- $z$  sample and the matched control sample. The C IV blueshift relative to Mg II increases with quasar luminosity (e.g., Richards et al. 2002; Shen et al. 2016), so the C IV blueshift is smaller in that composite than in our sample and the control sample.

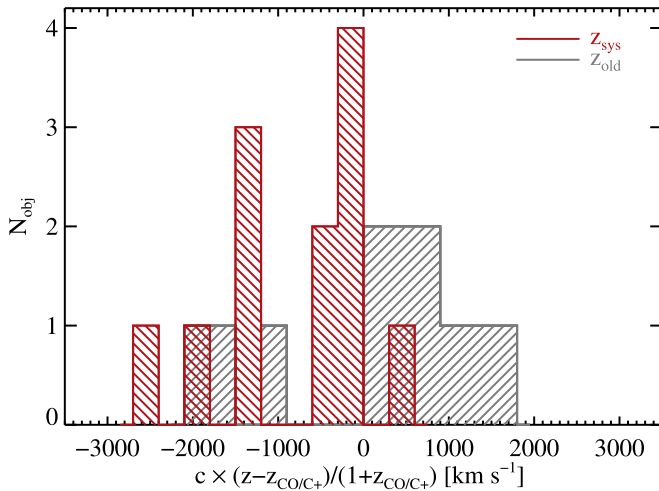
We now examine the measurements for individual objects. Figure 8 shows the distributions of the C IV REWs and C IV–Mg II blueshift for the high- $z$  sample and the control sample. Similar to our findings using the composite spectra, there is no significant difference in the typical C IV properties when quasar luminosity is matched. The C IV–Mg II velocity shift is independent of the systemic redshift estimate, and both samples

<sup>17</sup> We choose this specific monochromatic luminosity instead of a broadband luminosity because the same quantity is explicitly used in earlier work in BH mass and bolometric luminosity calculations and in the correlation between C IV properties and quasar luminosity. We do not find any significant difference in our results if we adopt the monochromatic luminosity at a different wavelength (e.g., at 3000 Å) to construct our matched control sample.

<sup>18</sup> Since we used a different continuum-fitting recipe for our high- $z$  sample than the one used in the Shen et al. (2011) catalog, we refit the control DR7 quasars with the same methodology. The new continuum fitting method mostly affected the REW of the broad C IV line. We found that the global continuum fitting method adopted in this paper produces  $\sim 18\%$  larger REWs of the C IV line on average than those reported in the Shen et al. (2011) catalog. On the other hand, other spectral properties, such as line peaks and widths, are all consistent within the errors with those compiled in the Shen et al. (2011) catalog.

**Table 4**  
Redshift Comparison

ObjID	$z_{\text{old}}$	$z_{\text{NIR}}$	$z_{[\text{C II}]}$	Ref	$z_{\text{CO}}$	Ref
J0055+0146	6.02	$6.017 \pm 0.054$	$6.0060 \pm 0.0008$	wil15	...	...
J0840+5624	5.85	$5.816 \pm 0.010$	...	...	$5.8441 \pm 0.0013$	wan10
J0842+1218	6.055	$6.069 \pm 0.009$	$6.0763 \pm 0.0005$	dec18	...	...
J0927+2001	5.79	$5.770 \pm 0.013$	...	...	$5.7722 \pm 0.0006$	car07
...	...	...	...	...	$5.7716 \pm 0.0012$	wan11a
J1044-0125	5.8	$5.780 \pm 0.007$	$5.7847 \pm 0.0007$	wan13	$5.7824 \pm 0.0007$	wan10
J1148+5251	6.43	$6.416 \pm 0.006$	6.419	wal09b	$6.4189 \pm 0.0006$	mai05
J1207+0630	6.04	$6.028 \pm 0.013$	$6.0366 \pm 0.0009$	dec18	...	...
J1335+3533	5.93	$5.870 \pm 0.020$	...	...	$5.9012 \pm 0.0019$	wan10
J1425+3254	5.85	$5.862 \pm 0.006$	...	...	$5.8918 \pm 0.0018$	wan10
J1429+5447	6.21	$6.119 \pm 0.008$	...	...	$6.1831 \pm 0.0007$	wan11a
J1623+3112	6.22	$6.254 \pm 0.006$	$6.2605 \pm 0.0005$	wan11a	...	...
J2310+1855	6.04	$5.956 \pm 0.011$	$6.0031 \pm 0.0002$	wan13	$6.0025 \pm 0.0007$	wan13

**Note.** Comparison between our near-IR redshifts based on broad emission lines and those derived from [C II] and CO. Reference keys are: wil15—Willott et al. (2015); wan10—Wang et al. (2010); dec18—Decarli et al. (2018); car07—Carilli et al. (2007); wan11a—Wang et al. (2011); wan13—Wang et al. (2013); wal09b—Walter et al. (2009).**Figure 5.** Comparison between the broad-line-based redshifts and those derived from CO or [C II] molecular lines for a subset of 12 objects. The red histogram is for our improved redshifts based on near-IR spectra, and the gray histogram is for the old redshifts. Our near-IR redshifts are in excellent agreement with the molecular line redshifts for half of the sample. There are two quasars for which our near-IR redshifts are lower than the molecular line redshifts by more than  $1500 \text{ km s}^{-1}$ , as discussed in the text.**Table 5**  
Composite Spectrum

Rest Wavelength (1)	$f_{\lambda}$ (2)	$\sigma_{f_{\lambda}}$ (3)	$N_{\text{obj}}$ (4)
1116.5	0.085	0.108	2
1117.5	2.846	0.988	2
...	...	...	...
3738.5	0.460	0.002	2
3739.5	0.497	0.113	2

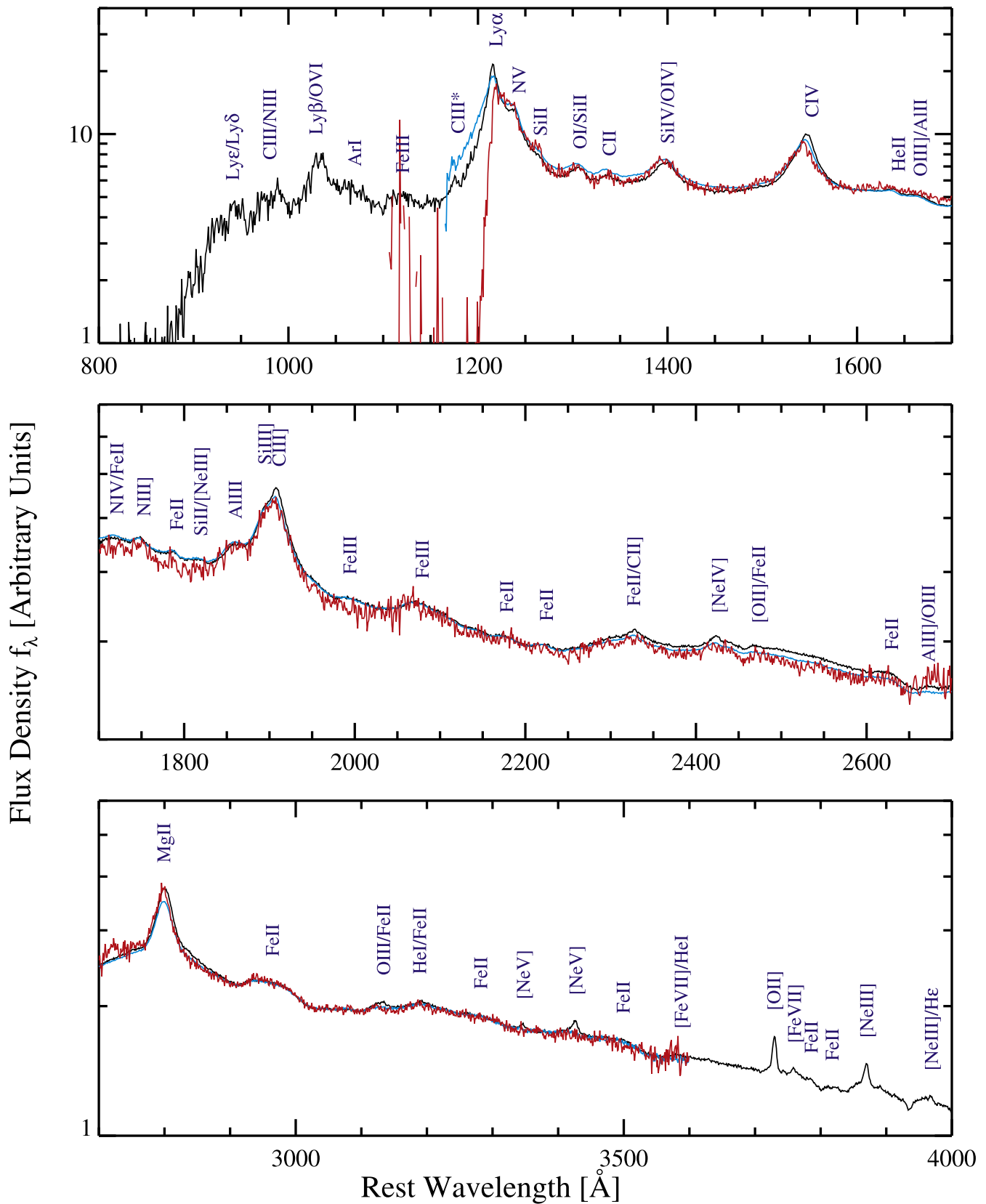
**Note.** Median composite spectrum for our near-IR quasar sample. Wavelengths are in units of Å. Flux density and flux density error units are arbitrary. The last column indicates how many objects contributed to the median composite at each wavelength pixel.

show a similar median C IV–Mg II blueshift of  $\sim 1000 \text{ km s}^{-1}$ . Examinations of other broad lines covered in our GNIRS spectra also did not result in any significant differences

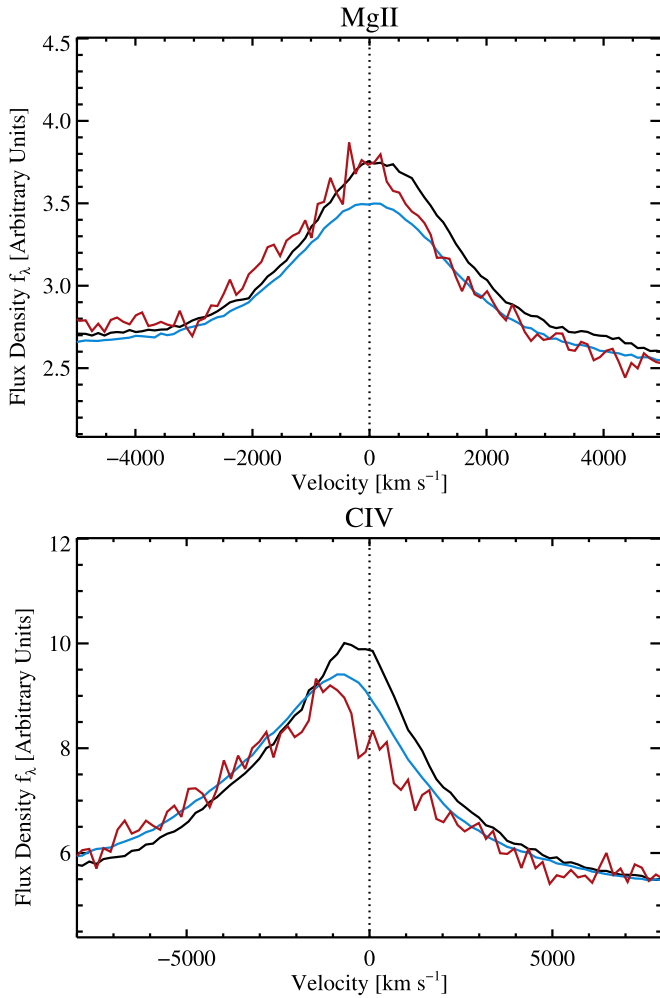
between the two samples. Interestingly, Mazzucchelli et al. (2017) reported a significantly higher average C IV blueshift relative to Mg II in nine quasars at  $z \gtrsim 6.5$  compared with low- $z$  SDSS quasars matched in luminosity. Mazzucchelli et al. fit a single Gaussian to the C IV line, which will overestimate the peak blueshift given the blue asymmetric profile (see Figure 7), while we fit multiple Gaussians consistently to both the high- $z$  sample and the low- $z$  SDSS sample. Despite the small sample statistics and the difference in the fitting details, it is possible that the observed  $z \gtrsim 6.5$  quasars have intrinsically different properties than their lower  $z$  counterpart.

Although our control sample is matched in luminosity to our high- $z$  GNIRS sample, the S/N distributions of the two samples are different. This does not affect median values of the quantities we compare or the composite spectrum but may impact the comparison of the full distribution. The median uncertainties in the measured spectral quantities in our GNIRS sample are only slightly larger than those for the SDSS control sample, but our GNIRS sample has a higher fraction of objects with large uncertainties. To test the impact of different S/N, we perturb the measured spectral quantities of the control sample using a Gaussian random variable with variance randomly assigned using the measurement uncertainties for the GNIRS sample (i.e., duplicated to match the number of the control sample). The results for the shuffled measurements for the control sample are shown as the black dotted lines in Figure 8. We found that for all the distributions studied here, the difference in S/N does not impact the results much. In fact, the shuffled distributions for the control sample better match those for the high- $z$  sample than the unshuffled distributions in the tails that are mostly contributed by noisy measurements.

We derive BH mass estimates for our high- $z$  sample using the so-called single-epoch virial BH masses that utilize the widths of the broad emission lines and the continuum luminosity measured from single-epoch spectroscopy. Despite the popularity of these virial BH mass estimates in recent years, there are substantial systematic uncertainties associated with them and differences among different broad-line estimators (for a comprehensive review, see, e.g., Shen 2013). Here we simply present these BH mass estimates, following the fiducial recipes summarized in Shen et al. (2011) based on Mg II and C IV, and caution that there is substantial systematic uncertainty



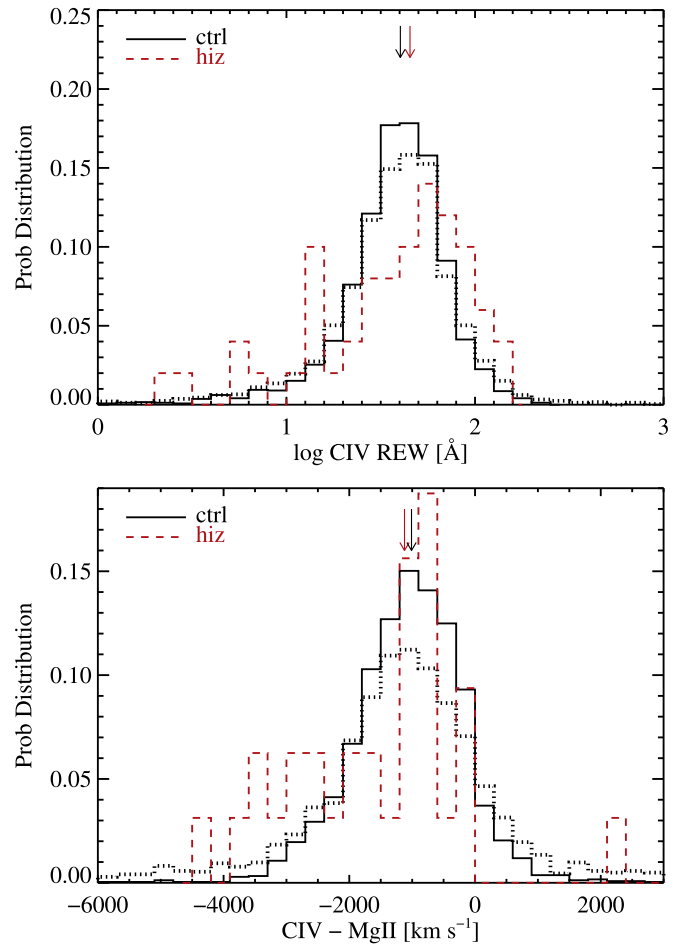
**Figure 6.** Rest-frame UV median composite spectra for our high- $z$  sample (red), the luminosity-matched control sample from SDSS DR7 (cyan), and the sample in (Vanden Berk et al. 2001, black). The spectra are normalized around different continuum wavelengths (1600, 2200, 3000 Å) in the three panels to compensate for the slight (and insignificant) differences in the continuum slopes of these samples. The high- $z$  composite around 2700 Å is not well constrained due to the small number of contributing objects as this is around the telluric absorption band between  $H$  and  $K$ . Other than the heavily absorbed Ly $\alpha$  line, the high- $z$  sample has similar average UV spectral properties as the control sample matched in quasar luminosity. The composite spectrum for the high- $z$  sample is provided in Table 5.



**Figure 7.** A detailed view of the median composite spectra as shown in Figure 6 around the Mg II (top) and C IV (bottom) lines for the high- $z$  sample (red), the luminosity-matched control sample at lower  $z$  (cyan), and the sample from Vanden Berk et al. (2001). The improved redshifts for the control sample from Hewett & Wild (2010) and our best estimates for the high- $z$  sample consistently place the Mg II peak near zero velocity, while there is a small (but negligible) Mg II redshift for the Vanden Berk et al. (2001) composite spectrum due to differences in the systemic redshift estimation in earlier work. On the other hand, C IV is notably blueshifted from Mg II, with both the high- $z$  sample and the control sample having a larger blueshift than the Vanden Berk et al. (2001) sample. This is because the C IV blueshift increases with quasar luminosity (e.g., Richards et al. 2002; Shen et al. 2016), and the Vanden Berk et al. (2001) sample has a lower average luminosity than both the high- $z$  and the control samples.

( $\sim 0.4$  dex) for individual estimates, which is typically much larger than the measurement uncertainty from spectral fits. These adopted fiducial mass recipes from Shen et al. (2011) were tested using SDSS quasars with a spectral fitting methodology similar to the one used in this work.

For a subset of 25 quasars we have BH mass estimates from both Mg II and C IV lines. There has been extensive discussion in the literature on the reliability of C IV as a virial BH mass estimator for high-redshift quasars (e.g., Shen 2013 and references therein). Unlike the low-ionization broad lines such as  $H\beta$  and Mg II, the C IV line often displays a more asymmetric profile and a significant blueshift that correlate with the luminosity and Eddington ratio of the quasar (e.g., Sulentic et al. 2000; Richards et al. 2011; Brotherton et al. 2015). C IV line width is poorly correlated with that of Mg II or



**Figure 8.** Histograms of the C IV rest equivalent width (top) and the C IV blueshift relative to Mg II (bottom). The red dashed histograms are for the high- $z$  sample, and the black solid histograms are for the control sample at lower redshifts matched in quasar luminosity. The median value of the distribution is indicated by an arrow at the top. The black dotted lines are the results of the control sample measurements randomly shuffled by the uncertainty distribution of the high- $z$  sample to compensate for the different S/N in the two samples. The median values remain almost the same, and K-S tests on the distributions for the high- $z$  sample and the shuffled control sample suggest insignificant difference between the two distributions, with a null probability of  $\sim 0.07$  and  $\sim 0.4$  for the C IV REW and blueshift distributions, respectively. Therefore we conclude that there is no significant difference in the plotted quantities between the high- $z$  sample and the control sample.

$H\beta$  in high-luminosity quasars (e.g., Baskin & Laor 2005; Shen et al. 2008; Shen & Liu 2012; Coatman et al. 2017), and the difference correlates with quasar luminosity. It is likely that the C IV line includes a significant nonvirial component (e.g., Denney 2012), particularly in high-luminosity objects. While it is possible to calibrate the C IV estimator to yield average BH masses unbiased relative to those based on  $H\beta$  and Mg II (e.g., Shen et al. 2008; Assef et al. 2011; Shen & Liu 2012; Runnoe et al. 2013; Coatman et al. 2017; Park et al. 2017; Mejía-Restrepo et al. 2018), the scatter between the C IV mass and  $H\beta$  or Mg II mass is still considerably larger than that between Mg II and  $H\beta$ . Given that most reverberation mapping measurements to date are on the  $H\beta$  line (e.g., Peterson et al. 2004) and that the current single-epoch mass estimators are all directly or indirectly based on the  $H\beta$  reverberation mapping results, it is reasonable to assume that  $H\beta$  is the safest line from which to estimate a virial BH mass. Since Mg II width

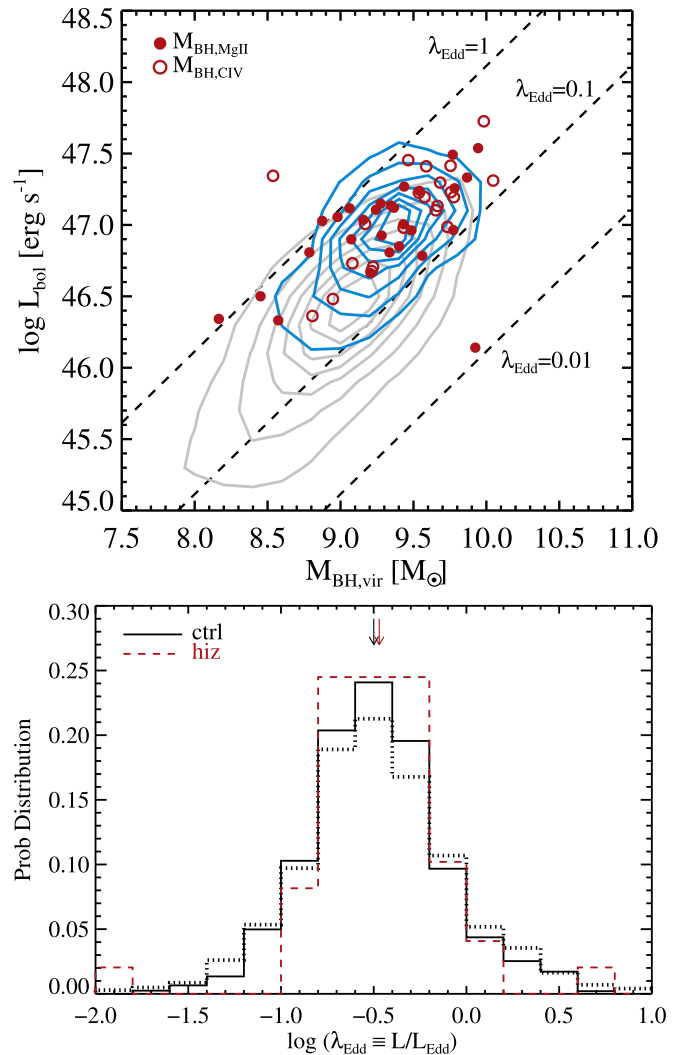


correlates with  $H\beta$  width well (e.g., Shen et al. 2008; Wang et al. 2009), it is also reasonable to assume that Mg II is relatively safe to use as a BH mass estimator. C IV can still be used because on average it provides consistent BH masses with those from Mg II or  $H\beta$ , albeit with a large intrinsic scatter.

For this work we adopt fiducial BH mass estimates based on Mg II if available; otherwise we use C IV-based masses. However, we have tested using alternative fiducial BH masses, e.g., using C IV over Mg II masses or the average of the two masses (as suggested by Vestergaard et al. 2011) and did not find any significant changes in our results. We require an additional criterion for BH mass estimation that the line must have a measured REW at  $>1\sigma$  to avoid extremely noisy measurements, which leaves one object, J0055+0146, without a virial BH mass estimate.

Figure 9 (top panel) displays the distribution of quasars in the BH mass versus bolometric luminosity plane. The gray and cyan contours show the distributions for the entire SDSS DR7 quasar sample (Shen et al. 2011) and the control sample. Our high- $z$  sample is shown in red points, where filled symbols use the Mg II-based BH masses and open symbols use the C IV-based BH masses. The bottom panel compares the distribution of the Eddington ratio,  $\lambda_{\text{Edd}} \equiv L_{\text{bol}}/L_{\text{Edd}}$ , where  $L_{\text{Edd}} = 1.3 \times 10^{38} \text{ erg s}^{-1} (M_{\text{BH}}/M_{\odot})$  is the Eddington luminosity of the BH, between the high- $z$  sample and the luminosity-matched control sample. Our high- $z$  quasars have similar BH masses and Eddington ratios as the control sample, based on the virial BH mass estimates. The median Eddington ratio for the high- $z$  sample is  $\log \lambda_{\text{Edd}} \sim -0.5$ , and we estimate a logarithmic dispersion of  $\sim 0.3$  dex by fitting a Gaussian function to the distribution of  $\log \lambda_{\text{Edd}}$ . There is one quasar (J0300–2232) with apparent super-Eddington accretion, with the BH mass estimated from C IV, which may be impacted by the absorption features adjacent to C IV. Another object, J2356+0023, has a noisy Mg II line, which resulted in a large uncertainty in the Mg II-based BH mass ( $\sim 0.5$  dex), and it appears as an apparent outlier in Figure 9 (top panel). The Eddington ratio based on C IV is  $\log \lambda_{\text{Edd}} \sim -1.2$  for this object.

We note that our Mg II-based virial BH mass recipe was calibrated to agree with both  $H\beta$  and C IV-based masses on average with the Vestergaard & Peterson (2006) recipe using SDSS quasar samples that cover two lines at low and high redshift (Shen et al. 2011). Our Mg II mass recipe is almost identical to that derived by Trakhtenbrot & Netzer (2012); both have the same slope on continuum luminosity as in McLure & Dunlop (2004) and similar zero-points. Earlier work (e.g., Jiang et al. 2007; Kurk et al. 2007; De Rosa et al. 2011) utilized the Mg II recipe from McLure & Dunlop (2004), which has the same luminosity slope but a smaller zero-point by 0.22 dex, compared with our Mg II masses. We have checked the handful of objects<sup>19</sup> in our sample that have near-IR spectroscopy in earlier work and found reasonably good agreement in line width and continuum luminosity measurements considering the different epochs and spectral fitting recipes. Therefore, the apparent difference in the reported Eddington ratios for the common objects is largely due to the difference in the adopted Mg II mass estimators. On the other hand, Willott et al. (2010a)



**Figure 9.** Top: the BH mass–luminosity plane. The gray contours show the distribution for all SDSS DR7 quasars from the Shen et al. (2011) catalog, and the cyan contours show the distribution for the control sample matched in luminosity to our high- $z$  sample. The red points show our high- $z$  sample, with the Mg II-based virial BH masses shown in filled circles and the less reliable C IV-based virial BH masses shown in open circles. We caution that individual virial BH masses could have a systematic uncertainty of  $\sim 0.4$  dex (e.g., Shen 2013), and individual bolometric luminosities could also be uncertain by a factor of up to 2 (e.g., Richards et al. 2006). The one quasar with apparent super-Eddington accretion (J0300–2232) has its virial BH mass estimated from C IV, whose fit is likely impacted by the absorption features adjacent to C IV. Another apparent outlier with the lowest Eddington ratio (J2356+0023) is caused by a poor fit to the noisy Mg II line. Bottom: the distribution of Eddington ratio  $L/L_{\text{Edd}}$  for our high- $z$  sample (red dashed line) and the luminosity-matched control quasars at lower redshifts (black solid line). The median values of the distributions are marked by the arrows. The black dotted line shows the result for the control sample measurements randomly shuffled by the uncertainty distribution of the high- $z$  sample to compensate for the different S/N in the two samples. A K-S test on the distributions for the high- $z$  sample and the shuffled control sample suggests an insignificant difference between the two distributions with a null probability of  $\sim 0.05$ .

reported Eddington ratios near unity for a sample of nine low-luminosity quasars at  $z \sim 6$ . There is negligible overlap in the Willott et al. (2010a) sample and our GNIRS sample. But the distribution at the faint luminosity end in Figure 9 shows a few quasars are closer to unity Eddington ratio than the rest of the sample, which is more in line with the typical Eddington ratios reported in Willott et al. (2010a) for low-luminosity high- $z$  quasars. Also, the Mg II BH masses in Willott et al. (2010a)

<sup>19</sup> There are three objects, J0836+0054, J1044–0125, and J1623+3112, in common with Jiang et al. (2007); one object, J0836+0054, in common with Kurk et al. (2007); and three objects, J0050+3445, J0055+0146, and J0221–0802, in common with Willott et al. (2010a). However, our GNIRS spectra for J0055+0146 and J0221–0802 are too noisy for a meaningful comparison with Willott et al. (2010a).

were based on the recipe from Vestergaard & Osmer (2009), which has a shallower luminosity dependence than our recipe and would yield on average smaller BH masses (higher Eddington ratios) than those based on our Mg II recipe for the luminosity regime of high- $z$  quasars. Willott et al. (2010a) also used a slightly larger bolometric correction, which further helps to explain the higher Eddington ratios reported in their work.

There are other existing Mg II-based mass recipes that use slightly different zero-points, luminosity, and line-width dependences (e.g., Onken & Kollmeier 2008; Wang et al. 2009; Woo et al. 2018). Using different mass recipes introduces  $\sim 0.2$ – $0.3$  dex offset in the virial BH masses, which we consider tolerable given the  $\sim 0.4$  dex systematic uncertainty in single-epoch virial BH masses (e.g., Shen 2013). Considering the large uncertainties in the BH mass estimates and bolometric corrections in individual objects, as well as the much smaller statistics in earlier work, we do not consider the factor of  $\sim 2$ – $3$  difference in the typical Eddington ratios between our work and earlier results significant.

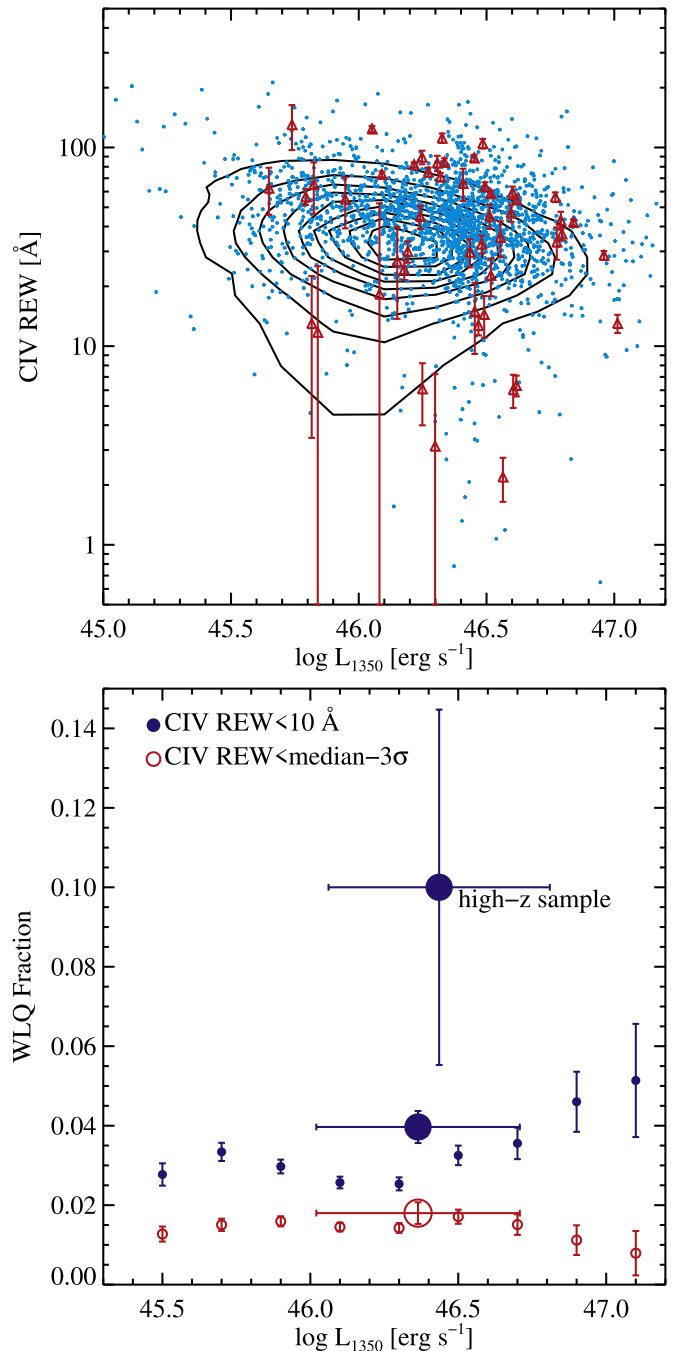
## 5. Discussion

### 5.1. Weak-line Quasars

A small fraction of radio-quiet quasars have significantly weaker (small REW) broad emission lines (in particular the UV lines), dubbed “weak-line quasars” (WLQs, e.g., Fan et al. 1999; Plotkin et al. 2008, 2010; Diamond-Stanic et al. 2009; Shemmer et al. 2010; Shemmer & Lieber 2015), where the weakness of the line is not apparently due to contamination from jet emission as in radio-loud quasars. WLQs sometimes have weaker X-ray emission than quasars with normal broad-line strength (e.g., Wu et al. 2011; Luo et al. 2015). While the exact cause of the weaker line strengths in WLQs is still unclear, one working scenario is that WLQs have a softer ionizing continuum than normal quasars such that the UV broad-line region sees fewer ionizing photons relative to the continuum underneath the line in WLQs, leading to reduced equivalent width (e.g., Luo et al. 2015). To produce a different ionizing spectrum, whether it is intrinsic or caused by filtering by some intermediate gas between the accretion disk and the broad-line region, requires that the accretion properties be different in WLQs. Thus studying the WLQ fraction at different redshifts can reveal possible changes in the accretion modes of SMBHs. Since the UV broad-line strength also decreases with quasar luminosity (Baldwin 1977), one needs to study the WLQ fraction as a function of quasar luminosity.

Since the definition of WLQs is somewhat arbitrary, we consider two definitions: 1. a quasar is a WLQ if its C IV rest frame EW is less than  $10 \text{ \AA}$  (e.g., Diamond-Stanic et al. 2009), or 2. its C IV REW is lower than  $3\sigma$  from the median of a population (e.g., quasars in a certain luminosity bin), where  $\sigma$  is the sample standard deviation. The second definition is meaningful when we consider the WLQ fraction as a function of quasar luminosity, where the median and standard deviation are calculated for each luminosity bin. We focus on the C IV line for our study on WLQs because Ly $\alpha$  is heavily absorbed and other broad lines are not as strong as C IV in our sample.

Figure 10 (top) shows the Baldwin effect of C IV for various samples. Our high- $z$  sample follows a similar trend as the



**Figure 10.** Top: C IV rest-frame EW as a function of luminosity. The black contours show all SDSS DR7 quasars from the Shen et al. (2011) catalog. The cyan points and the red triangles show the control sample and the high- $z$  sample, respectively. The general trend of decreasing C IV REW with continuum luminosity is known as the Baldwin effect (e.g., Baldwin 1977). Bottom: the WLQ fraction as a function of continuum luminosity. The small symbols are computed using all SDSS DR7 quasars in Shen et al. (2011) binned by luminosity. The large symbols are for our high- $z$  sample (top blue filled circle) and the control sample (bottom blue filled circle and red open circle), computed for the entire sample. The two types of symbols represent the two definitions of WLQ (see text). The error bars on luminosity for the high- $z$  and control samples are the standard deviations in each sample, and the vertical error bars are estimated with Poisson counting uncertainties. For the high- $z$  sample we use only the first WLQ definition (C IV REW  $< 10 \text{ \AA}$ ) because the dispersion in the C IV REW is not well defined due to both the smaller sample size and noisier measurements.

control sample. In the bottom panel of Figure 10 we show the WLQ fraction as a function of luminosity, using all SDSS DR7 quasars in Shen et al. (2011). The blue-filled symbols use the first WLQ definition with a constant threshold in C IV REW. The WLQ fraction with this definition rises toward high luminosities due to the Baldwin effect. The red open symbols use the second WLQ definition, which is roughly flat with luminosity because the average Baldwin effect is taken out. Clearly, if one uses a constant cut in C IV REW to define WLQs and compare two different samples, one must match the luminosity of these two samples for a fair comparison.

For our high- $z$  sample, the sample dispersion in C IV REW is not well defined given the smaller number of objects than the control sample, as well as substantial measurement uncertainties in individual objects. Therefore we consider only the first definition of WLQs and compare with the control sample, which is matched in luminosity to the high- $z$  sample. We identify 5 WLQ quasars in the high- $z$  sample with measured C IV REW less than  $10 \text{ \AA}$ : J0850+3246, J1257+6349, J1335+3533, J1429+5447, and J1621+5155. Inspection of their spectra suggests that these should be genuine WLQs. The WLQ fraction for the high- $z$  sample is thus  $\sim 10\% \pm 4\%$  (5 out of 50), compared with  $4.0\% \pm 0.4\%$  for the control sample. Thus the high- $z$  sample seems to have a mild excess in the WLQ fraction compared with its low- $z$  counterparts matched in luminosity, albeit with large uncertainties. This result is consistent with the distributions of C IV REW shown in Figure 8, where the distribution of the high- $z$  sample has a more prominent tail at the low REW end. Bañados et al. (2016) reported a WLQ fraction of  $\sim 14\%$  (16/117) using Ly $\alpha$  with a larger sample of  $z > 5.6$  quasars, which is fully consistent with our results here based on C IV.

There is one known radio-loud quasar (J1429+5447) among the five quasars with C IV REW  $< 10 \text{ \AA}$ , where the UV continuum could be contaminated by jet emission.

### 5.2. Early Formation of SMBHs

Our results in Section 4 and Figure 9 reveal that our high- $z$  quasars are accreting at moderately high Eddington ratios, with an average value of  $\lambda_{\text{Edd}} \approx 0.3$ , similar to their low- $z$  counterparts with matched luminosities. This typical Eddington ratio is very similar to the value of  $\sim 0.4$  for a sample of 15  $z \gtrsim 6.5$  quasars studied in Mazzucchelli et al. (2017). Interestingly, these objects appear to be well bounded by the Eddington limit, even considering the systematic uncertainties in the estimation of BH masses and bolometric luminosities.

The observed high- $z$  quasars are drawn from an underlying quasar population with an intrinsic distribution of BH masses accreting at a range of Eddington ratios. The inevitable flux limit will preferentially select the most luminous objects from the intrinsic distribution. As we naively expect that there are more abundant lower mass SMBHs than higher mass ones, those lower mass BHs with higher than average Eddington ratios may scatter above the flux limit and bias the observed Eddington ratios high. At the same time, massive BHs accreting at lower than average Eddington ratios may be missed from the sample due to the flux limit. Thus our observed apparent Eddington ratios suggest that the majority of the unobserved quasars should be accreting at even lower Eddington ratios, as argued with robust statistical approaches using lower redshift quasar samples (e.g., Kelly et al. 2010; Shen & Kelly 2012; Kelly & Shen 2013).

The  $e$ -folding time of BH growth at constant Eddington ratio is  $t_e \approx 4.5 \times 10^8 \epsilon / [\lambda_{\text{Edd}}(1 - \epsilon)] \text{ yr}$ , where  $\epsilon$  is the radiative efficiency. The typical radiative efficiency for  $z \gtrsim 5.7$  quasars is constrained to be  $\epsilon \approx 0.1$  (e.g., Trakhtenbrot et al. 2017), using the observed luminosity and BH mass, as well as scaling relations derived from accretion disk theories (e.g., Davis & Laor 2011). It would have been difficult (although not impossible, e.g., Li et al. 2007) for these SMBHs to grow to their current mass from small ( $\sim 100 M_\odot$ ) seeds at these sub-Eddington levels of accretion rates, given the limited time available for their growth from seed BHs. Episodes of higher accretion rates (e.g., Pezzulli et al. 2016) or heavier seed BHs (e.g., Latif et al. 2013) are possible solutions to this early growth problem (for recent reviews on this topic, see, e.g., Volonteri 2010; Haiman 2013; Latif & Schleicher 2019). If the early BH growth invokes much higher accretion rates, this phase is most likely obscured in the UV-optical (e.g., Sanders & Mirabel 1996; Hopkins et al. 2008) and/or has much lower radiative efficiency (e.g., Pezzulli et al. 2016), otherwise we would have observed much higher Eddington ratios among the current high- $z$  quasar sample. These heavily obscured quasars can only be probed at other wavelengths (e.g., Glikman et al. 2018; Hickox & Alexander 2018, and references therein).

We also do not find evidence for a population of excessively large ( $> 10^{10} M_\odot$ ) BHs at these highest redshifts. The one object (J0203+0012) with a C IV-only virial BH mass slightly greater than  $10^{10} M_\odot$  is a BAL quasar, where the C IV line fit is likely impacted by the broad absorption, and the individual C IV-based mass may not be as reliable as that based on the Mg II line as discussed earlier. If such a population of hypermassive SMBHs exists in large numbers, they must be accreting at substantially lower Eddington ratios, or are obscured and missed from our sample. It is possible that even more luminous quasars discovered at  $z \gtrsim 5.7$  could harbor  $> 10^{10} M_\odot$  BHs (e.g., Wu et al. 2015), but such objects are increasingly rare. In any case, we caution that individual quasars with reported virial BH mass greater than  $10^{10} M_\odot$  in some past studies could be due to systematic uncertainties in their BH mass estimation (see discussions in, e.g., Vestergaard 2004; Shen 2013). Using a robust statistical approach on the lower redshift SDSS quasar samples, Kelly et al. (2010) and Kelly & Shen (2013) constrained the maximum BH mass in quasars that can be observed in a theoretical all-sky survey without a flux limit to be several times  $10^{10} M_\odot$ , albeit with considerable systematic uncertainties. The absence of such hypermassive BHs in our sample is consistent with the constraints for lower  $z$  quasars. From a theoretical perspective, some papers have argued for an upper limit of SMBH mass of  $\sim 10^{10-11} M_\odot$  (e.g., Natarajan & Treister 2009; King 2016). But even if these hypermassive SMBHs do exist, they would be extremely difficult to find given their apparent rareness and potential low activity (or possible obscuration).

On the other hand, the distribution of high- $z$  quasars in the mass–luminosity plane of Figure 9 suffers from the flux limit of the search, lacking objects that are too faint to be included in the sample. Thus at fixed BH mass, the sample is incomplete toward lower luminosities, or lower Eddington ratios. While we observe these quasars accreting at moderately high Eddington ratios, it is possible that the majority of high- $z$  SMBHs are accreting at substantially lower Eddington ratios, as mentioned earlier. In order to constrain the intrinsic distribution of quasars in the mass–luminosity plane, one needs to take into account



the selection function, which is based on flux (not BH mass), as well as the uncertainties of the BH mass estimates. The best approach is a forward modeling as detailed in earlier work (e.g., Kelly et al. 2010; Shen & Kelly 2012; Kelly & Shen 2013), which will be one of the follow-up studies based on our sample.

## 6. Summary

We have presented near-IR (simultaneous *YJHK*) spectra for a sample of 50 quasars at  $z \gtrsim 5.7$  from our large Gemini GNIRS program. Based on these near-IR spectra, we measured the rest-frame UV spectral properties of these quasars, including broad emission line properties and derived quantities such as bolometric luminosities and virial BH mass estimates.















We compared the UV spectral properties of these high- $z$  quasars with a control sample of quasars from SDSS at  $z = 1.5\text{--}2.3$  matched in continuum luminosity. We found that our high- $z$  quasars have a mild excess of weak-line quasars (with C IV REW less than 10 Å) compared with the luminosity-matched control sample at lower redshifts. But other than that, the UV spectral properties are very similar between the high- $z$  and the low- $z$  quasars. This similarity between the high- $z$  and low- $z$  quasars also extends to X-ray properties when luminosity is matched (e.g., Nanni et al. 2017; Bañados et al. 2018a). These results suggest that the same physical mechanisms are operating for these accreting SMBHs at all redshifts. Using broad Mg II and C IV lines covered in the near-IR spectra, we estimated the BH masses and Eddington ratios of these high- $z$  quasars using the single-epoch virial mass method. We found that these high- $z$  quasars are accreting at moderately high Eddington ratios with a median value of  $L/L_{\text{Edd}} \sim 0.3$ , similar to their low- $z$  counterparts at comparable luminosities as well as quasars at even higher redshifts (e.g., Mazzucchelli et al. 2017). We did not find evidence of a population of hypermassive ( $>10^{10} M_{\odot}$ ) BHs in our sample. Such objects are either too rare to be included in our sample or are accreting at substantially lower Eddington ratios (or are heavily obscured) and are missed from current high- $z$  quasar searches. Our results are consistent with earlier studies of high- $z$  quasars with near-IR spectroscopy but with substantially better statistics.

The data and spectral measurements presented in this work can form the basis of additional follow-up studies on  $z \gtrsim 5.7$  quasars, which include the measurements of their demography in the mass–luminosity plane, detailed analysis of narrow absorption lines and BALs, chemical abundances of gas in the broad-line region, and so on. Moreover, our new data add to the increasingly larger multiwavelength database to study these earliest SMBHs and strongly motivate new spectroscopic observations to extend the wavelength coverage, e.g., mid-IR observations with *JWST* to study the rest-frame optical properties of these high- $z$  quasars (Kalirai 2018).

We thank the Gemini Observatory staff for their help with our observing program and the anonymous referee for comments that improved the manuscript. J.W., L.J., L.C.H. and X.W. acknowledge support from the National Key R&D Program of China (2016YFA0400702, 2016YFA0400703) and from the National Science Foundation of China (grant No. 11473002, 11533001, 11721303). Y.S. acknowledges support from an Alfred P. Sloan Research Fellowship and National Science Foundation grant No. AST-1715579. D.R.

acknowledges support from the National Science Foundation under grant No. AST-1614213. M.V. gratefully acknowledges support from the Independent Research Fund Denmark via grant Nos. DFF 4002-00275 and 8021-00130.

## ORCID iDs

Yue Shen  <https://orcid.org/0000-0003-1659-7035>  
 Jin Wu  <https://orcid.org/0000-0002-6168-3867>  
 Linhua Jiang  <https://orcid.org/0000-0003-4176-6486>  
 Xiaohui Fan  <https://orcid.org/0000-0003-3310-0131>  
 Luis C. Ho  <https://orcid.org/0000-0001-6947-5846>  
 Dominik A. Riechers  <https://orcid.org/0000-0001-9585-1462>  
 Michael A. Strauss  <https://orcid.org/0000-0002-0106-7755>  
 Bram Venemans  <https://orcid.org/0000-0001-9024-8322>  
 Marianne Vestergaard  <https://orcid.org/0000-0001-9191-9837>  
 Fabian Walter  <https://orcid.org/0000-0003-4793-7880>  
 Feige Wang  <https://orcid.org/0000-0002-7633-431X>  
 Chris Willott  <https://orcid.org/0000-0002-4201-7367>  
 Xue-Bing Wu  <https://orcid.org/0000-0002-7350-6913>  
 Jinyi Yang  <https://orcid.org/0000-0001-5287-4242>

## References

- Alexandroff, R., Strauss, M. A., Greene, J. E., et al. 2013, *MNRAS*, **435**, 3306  
 Allen, J. T., Hewett, P. C., Maddox, N., Richards, G. T., & Belokurov, V. 2011, *MNRAS*, **410**, 860  
 Assef, R. J., Denney, K. D., Kochanek, C. S., et al. 2011, *ApJ*, **742**, 93  
 Baldwin, J. A. 1977, *ApJ*, **214**, 679  
 Bañados, E., Connor, T., Stern, D., et al. 2018a, *ApJL*, **856**, L25  
 Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, *ApJS*, **227**, 11  
 Bañados, E., Venemans, B. P., Mazzucchelli, C., et al. 2018b, *Natur*, **553**, 473  
 Bañados, E., Venemans, B. P., Morganson, E., et al. 2014, *AJ*, **148**, 14  
 Bañados, E., Venemans, B. P., Morganson, E., et al. 2015, *ApJ*, **804**, 118  
 Barth, A. J., Martini, P., Nelson, C. H., & Ho, L. C. 2003, *ApJL*, **594**, L95  
 Baskin, A., & Laor, A. 2005, *MNRAS*, **356**, 1029  
 Brotherton, M. S., Runnoe, J. C., Shang, Z., & DiPompeo, M. A. 2015, *MNRAS*, **451**, 1290  
 Carilli, C. L., Neri, R., Wang, R., et al. 2007, *ApJL*, **666**, L9  
 Coatman, L., Hewett, P. C., Banerji, M., et al. 2017, *MNRAS*, **465**, 2120  
 Cool, R. J., Kochanek, C. S., Eisenstein, D. J., et al. 2006, *AJ*, **132**, 823  
 Dai, X., Shankar, F., & Sivakoff, G. R. 2008, *ApJ*, **672**, 108  
 Davis, S. W., & Laor, A. 2011, *ApJ*, **728**, 98  
 De Rosa, G., Decarli, R., Walter, F., et al. 2011, *ApJ*, **739**, 56  
 De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, *ApJ*, **790**, 145  
 Decarli, R., Walter, F., Venemans, B. P., et al. 2017, *Natur*, **545**, 457  
 Decarli, R., Walter, F., Venemans, B. P., et al. 2018, *ApJ*, **854**, 97  
 Denney, K. D. 2012, *ApJ*, **759**, 44  
 Diamond-Stanic, A. M., Fan, X., Brandt, W. N., et al. 2009, *ApJ*, **699**, 782  
 Dye, S., Lawrence, A., Read, M. A., et al. 2018, *MNRAS*, **473**, 5113  
 Edge, A., Sutherland, W., Kuijken, K., et al. 2013, *Msngr*, **154**, 32  
 Fan, X., Carilli, C. L., & Keating, B. 2006a, *ARA&A*, **44**, 415  
 Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, *AJ*, **128**, 515  
 Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, *AJ*, **122**, 2833  
 Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, *AJ*, **132**, 117  
 Fan, X., Strauss, M. A., Gunn, J. E., et al. 1999, *ApJL*, **526**, L57  
 Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, *AJ*, **125**, 1649  
 Fan, X., White, R. L., Davis, M., et al. 2000, *AJ*, **120**, 1167  
 Gibson, R. R., Jiang, L., Brandt, W. N., et al. 2009, *ApJ*, **692**, 758  
 Glikman, E., Lacy, M., LaMassa, S., et al. 2018, *ApJ*, **861**, 37  
 Goto, T. 2006, *MNRAS*, **371**, 769  
 Haiman, Z. 2013, in *The First Galaxies*, 396 ed. T. Wiklund, B. Mobasher, & V. Bromm (Berlin: Springer), 293  
 Hewett, P. C., & Wild, V. 2010, *MNRAS*, **405**, 2302  
 Hickox, R. C., & Alexander, D. M. 2018, *ARA&A*, **56**, 625  
 Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, *ApJS*, **175**, 356  
 Horne, K. 1986, *PASP*, **98**, 609  
 Jiang, L., Fan, X., Annis, J., et al. 2008, *AJ*, **135**, 1057  
 Jiang, L., Fan, X., Bian, F., et al. 2009, *AJ*, **138**, 305



- Jiang, L., Fan, X., Vestergaard, M., et al. 2007, *AJ*, **134**, 1150
- Jiang, L., McGreer, I. D., Fan, X., et al. 2015, *AJ*, **149**, 188
- Jiang, L., McGreer, I. D., Fan, X., et al. 2016, *ApJ*, **833**, 222
- Kalirai, J. 2018, *ConPh*, **59**, 251
- Kelly, B. C., & Shen, Y. 2013, *ApJ*, **764**, 45
- Kelly, B. C., Vestergaard, M., Fan, X., et al. 2010, *ApJ*, **719**, 1315
- Kelson, D. D. 2003, *PASP*, **115**, 688
- Kim, Y., Im, M., Jeon, Y., et al. 2018, *ApJ*, **855**, 138
- King, A. 2016, *MNRAS*, **456**, L109
- Kurk, J. D., Walter, F., Fan, X., et al. 2007, *ApJ*, **669**, 32
- Latif, M., & Schleicher, D. R. G. 2019, in *Formation of the First Black Holes*, ed. M. Latif & D. R. G. Schleicher (Singapore: World Scientific)
- Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013, *MNRAS*, **433**, 1607
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, **379**, 1599
- Li, Y., Hernquist, L., Robertson, B., et al. 2007, *ApJ*, **665**, 187
- Luo, B., Brandt, W. N., Hall, P. B., et al. 2015, *ApJ*, **805**, 122
- Mason, R. E., Rodríguez-Ardila, A., Martins, L., et al. 2015, *ApJS*, **217**, 13
- Matsuoka, Y., Iwasawa, K., Onoue, M., et al. 2018a, *ApJS*, **237**, 5
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, *ApJ*, **828**, 26
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2018b, *PASJ*, **70**, S35
- Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017, *ApJ*, **849**, 91
- McGreer, I. D., Becker, R. H., Helfand, D. J., & White, R. L. 2006, *ApJ*, **652**, 157
- McLure, R. J., & Dunlop, J. S. 2004, *MNRAS*, **352**, 1390
- McMahon, R. G., Banerji, M., Gonzalez, E., et al. 2013, *MNRAS*, **434**, 35
- Mejía-Restrepo, J. E., Trakhtenbrot, B., Lira, P., & Netzer, H. 2018, *MNRAS*, **478**, 1929
- Morganson, E., De Rosa, G., Decarli, R., et al. 2012, *AJ*, **143**, 142
- Mortlock, D. J., Patel, M., Warren, S. J., et al. 2009, *A&A*, **505**, 97
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, *Natur*, **474**, 616
- Nanni, R., Vignali, C., Gilli, R., Moretti, A., & Brandt, W. N. 2017, *A&A*, **603**, A128
- Natarajan, P., & Treister, E. 2009, *MNRAS*, **393**, 838
- Onken, C. A., & Kollmeier, J. A. 2008, *ApJL*, **689**, L13
- Park, D., Barth, A. J., Woo, J.-H., et al. 2017, *ApJ*, **839**, 93
- Peterson, B. M., Ferrarese, L., Gilbert, K. M., et al. 2004, *ApJ*, **613**, 682
- Pezzulli, E., Valiante, R., & Schneider, R. 2016, *MNRAS*, **458**, 3047
- Plotkin, R. M., Anderson, S. F., Brandt, W. N., et al. 2010, *ApJ*, **721**, 562
- Plotkin, R. M., Anderson, S. F., Hall, P. B., et al. 2008, *AJ*, **135**, 2453
- Reed, S. L., McMahon, R. G., Martini, P., et al. 2017, *MNRAS*, **468**, 4702
- Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, *AJ*, **123**, 2945
- Richards, G. T., Kruczek, N. E., Gallagher, S. C., et al. 2011, *AJ*, **141**, 167
- Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, *ApJS*, **166**, 470
- Riechers, D. A., Walter, F., Bertoldi, F., et al. 2009, *ApJ*, **703**, 1338
- Runoe, J. C., Brotherton, M. S., Shang, Z., & DiPompeo, M. A. 2013, *MNRAS*, **434**, 848
- Ryan-Weber, E. V., Pettini, M., Madau, P., & Zych, B. J. 2009, *MNRAS*, **395**, 1476
- Sanders, D. B., & Mirabel, I. F. 1996, *ARA&A*, **34**, 749
- Shemmer, O., & Lieber, S. 2015, *ApJ*, **805**, 124
- Shemmer, O., Trakhtenbrot, B., Anderson, S. F., et al. 2010, *ApJL*, **722**, L152
- Shen, Y. 2013, *BASI*, **41**, 61
- Shen, Y. 2016, *ApJ*, **817**, 55
- Shen, Y., Brandt, W. N., Richards, G. T., et al. 2016, *ApJ*, **831**, 7
- Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008, *ApJ*, **680**, 169
- Shen, Y., Hall, P. B., Horne, K., et al. 2018, *ApJS*, in press (arXiv:1810.01447)
- Shen, Y., & Kelly, B. C. 2012, *ApJ*, **746**, 169
- Shen, Y., & Liu, X. 2012, *ApJ*, **753**, 125
- Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, *ApJS*, **194**, 45
- Sulentic, J. W., Zwitter, T., Marziani, P., & Dultzin-Hacyan, D. 2000, *ApJL*, **536**, L5
- Tang, J.-J., Goto, T., Ohya, Y., et al. 2017, *MNRAS*, **466**, 4568
- Trakhtenbrot, B., & Netzer, H. 2012, *MNRAS*, **427**, 3081
- Trakhtenbrot, B., Volonteri, M., & Natarajan, P. 2017, *ApJL*, **836**, L1
- van Dokkum, P. G. 2001, *PASP*, **113**, 1420
- Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, *AJ*, **122**, 549
- Venemans, B. P., Bañados, E., Decarli, R., et al. 2015, *ApJL*, **801**, L11
- Venemans, B. P., McMahon, R. G., Warren, S. J., et al. 2007, *MNRAS*, **376**, L76
- Venemans, B. P., Walter, F., Zschaechner, L., et al. 2016, *ApJ*, **816**, 37
- Vestergaard, M. 2004, *ApJ*, **601**, 676
- Vestergaard, M., Denney, K., Fan, X., et al. 2011, in *Narrow-Line Seyfert 1 Galaxies and their Place in the Universe*, Vol. NLS1, ed. L. Foschini et al. (Trieste: PoS), **38**
- Vestergaard, M., & Osmer, P. S. 2009, *ApJ*, **699**, 800
- Vestergaard, M., & Peterson, B. M. 2006, *ApJ*, **641**, 689
- Vestergaard, M., & Wilkes, B. J. 2001, *ApJS*, **134**, 1
- Volonteri, M. 2010, *A&ARv*, **18**, 279
- Walter, F., Bertoldi, F., Carilli, C., et al. 2003, *Natur*, **424**, 406
- Walter, F., Riechers, D., Cox, P., et al. 2009, *Natur*, **457**, 699
- Wang, F., Wu, X.-B., Fan, X., et al. 2016, *ApJ*, **819**, 24
- Wang, J.-G., Dong, X.-B., Wang, T.-G., et al. 2009, *ApJ*, **707**, 1334
- Wang, R., Carilli, C. L., Neri, R., et al. 2010, *ApJ*, **714**, 699
- Wang, R., Wagg, J., Carilli, C. L., et al. 2011, *ApJL*, **739**, L34
- Wang, R., Wagg, J., Carilli, C. L., et al. 2013, *ApJ*, **773**, 44
- Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010a, *AJ*, **140**, 546
- Willott, C. J., Bergeron, J., & Omont, A. 2015, *ApJ*, **801**, 123
- Willott, C. J., Bergeron, J., & Omont, A. 2017, *ApJ*, **850**, 108
- Willott, C. J., Delorme, P., Omont, A., et al. 2007, *AJ*, **134**, 2435
- Willott, C. J., Delorme, P., Reylé, C., et al. 2009, *AJ*, **137**, 3541
- Willott, C. J., Delorme, P., Reylé, C., et al. 2010b, *AJ*, **139**, 906
- Woo, J.-H., Le, H. A. N., Karouzos, M., et al. 2018, *ApJ*, **859**, 138
- Wu, J., Brandt, W. N., Hall, P. B., et al. 2011, *ApJ*, **736**, 28
- Wu, X.-B., Wang, F., Fan, X., et al. 2015, *Natur*, **518**, 512