Draft Version of May 11, 2017
Preprint typeset using LaTeX style emulateapj v. 12/16/11

IMPLICATIONS OF $Z \sim 6$ QUASAR PROXIMITY ZONES FOR THE EPOCH OF REIONIZATION AND QUASAR LIFETIMES

Anna-Christina Eilers[1,2*], Frederick B. Davies[1], Joseph F. Hennawi[1,3], J. Xavier Prochaska[4], Zarija Lukić[5], Chiara Mazzucchelli[1,2]

Draft Version of May 11, 2017

ABSTRACT

We study quasar proximity zones in the redshift range $5.77 \leq z \leq 6.54$ by homogeneously analyzing 34 medium resolution spectra, encompassing both archival and newly obtained data, and exploiting recently updated systemic redshift and magnitude measurements. Whereas previous studies found strong evolution of proximity zone sizes with redshift, and argued that this provides evidence for a rapidly evolving intergalactic medium (IGM) neutral fraction during reionization, we measure a much shallower trend $\propto (1+z)^{-1.44}$. We compare our measured proximity zone sizes to predictions from hydrodynamical simulations post-processed with one-dimensional radiative transfer, and find good agreement between observations and theory irrespective of the ionization state of the ambient IGM. This insensitivity to IGM ionization state has been previously noted, and results from the fact that the definition of proximity zone size as the first drop of the smoothed quasar spectrum below the 10% flux transmission level probes locations where the ionizing radiation from the quasar is an order of magnitude larger than the expected ultraviolet ionizing background that sets the neutral fraction of the IGM. Our analysis also uncovered three objects with exceptionally small proximity zones (two have $R_p < 1$ proper Mpc), which constitute outliers from the observed distribution and are challenging to explain with our radiative transfer simulations. We consider various explanations for their origin, such as strong absorption line systems associated with the quasar or patchy reionization, but find that the most compelling scenario is that these quasars have been shining for $\lesssim 10^5$ yr.

Subject headings: intergalactic medium — epoch of reionization, dark ages — methods: data analysis — quasars: absorption lines

1. INTRODUCTION

A major goal of observational cosmology is to understand the epoch of reionization, when the universe transitioned from the cosmic "dark ages" following recombination, into the ionized universe we observe today. Despite much progress in the last decade, there are still many open questions regarding the exact timing and the morphology of the reionization process. Studies of the evolution of the (Lyman-α) Lyα absorption features in the spectra of distant quasars are one of the key observational probes of this epoch. Both the steep rise in the Lyα optical depth of the intergalactic medium (IGM) with redshift for $z \gtrsim 5.5$, as well as the increased scatter in the measurements, suggest a qualitative change in the state of the intergalactic medium (IGM), likely resulting from a rapid rise in the volume averaged neutral fraction which is expected to occur during the end stages of reionization (Fan et al. 2006b; Becker et al. 2015). Indeed, the absence of large Gunn-Peterson (GP) troughs (Gunn & Peterson 1965) in spectra at $z \lesssim 5.5$ indicates the epoch of hydrogen reionization must be completed at that time (McGreer et al. 2015). However, constraining the neutral hydrogen fraction of the IGM at $z \gtrsim 6$ with quasar absorption spectroscopy has proven extremely difficult, because Lyman series transitions are overly sensitive, and saturate already for volume averaged neutral hydrogen fractions $\langle f_{\rm HI} \rangle \gtrsim 10^{-4}$.

In this paper we focus on a different technique that has been applied in previous work to constrain the neutral gas fraction at $z \sim 6$. Near the end of the cosmic reionization epoch, luminous quasars exhibit a region of enhanced transmission immediately blueward of the Lyα emission line in the so-called quasar proximity zone, before the onset of near complete Gunn-Peterson absorption. This enhanced transmission is caused by the radiation from the quasar itself, which ionizes the surrounding IGM (see e.g. Madau & Rees 2000; Cen & Haiman 2000; Haiman & Cen 2001; Wyithe et al. 2005; Bolton & Haehnelt 2007b,a; Lidz et al. 2007; Bolton et al. 2011; Keating et al. 2015). It has been argued that the evolution of the proximity zone sizes with redshift can constrain the late stages of the reionization epoch (Fan et al. 2006b; Carilli et al. 2010; Venemans et al. 2015).

Haiman & Cen (2001) showed that in a very simplistic picture of reionization in which isolated ionized H II regions expand into a neutral uniform ambient IGM, the (proper) size $R_{\rm ion}$ of the ionized region around the quasar depends on the neutral hydrogen fraction $f_{\rm HI}$, the rate at which ionizing photons are emitted $\dot{N}_\gamma$, and the quasar

* email: eilers@mpia.de
[1] Max-Planck-Institute for Astronomy, Königstuhl 17, 69117 Heidelberg, Germany; eilers@mpia.de
[2] International Max Planck Research School for Astronomy & Cosmic Physics at the University of Heidelberg
[3] Physics Department, University of California, Santa Barbara, CA 93106-9530, USA
[4] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA
[5] Lawrence Berkeley National Laboratory, CA 94720-8139, USA

age $t_{\rm Q}$:

$$R_{\rm ion} \approx \left(\frac{3\dot{N}_\gamma t_{\rm Q}}{4\pi n_{\rm H} f_{\rm HI}} \right)^{1/3}, \quad (1)$$

where $n_{\rm H}$ is the hydrogen number density. In eqn. (1) recombinations of the ionized gas inside the H II region are neglected, which should be unimportant on the timescale that the quasar turned on, because the recombination timescale of hydrogen is comparable to the Hubble time. However, there are a number of other effects that are not taken into consideration in this relation, such as overlapping ionized H II regions, large-scale structure effects, and pre-ionization by local galaxies or the clumpiness of the IGM (see also Fan et al. 2006b; Bolton & Haehnelt 2007b,a; Lidz et al. 2007; Maselli et al. 2007, 2009; Khrykin et al. 2016).

While eqn. (1) provides a reasonable description of the expansion rate of an H II ionization front embedded in a neutral IGM, it does not predict the distribution of residual neutral hydrogen within the ionized bubble. However, the *observed* size of the proximity zone R_p will depend sensitively on the fraction and distribution of the residual neutral hydrogen, because this can cause saturated absorption well before the ionization front is reached (Bolton & Haehnelt 2007b). Thus the extent of the ionized H II bubble around the quasar can be significantly larger than the observed proximity zone, which is defined to end where the transmitted flux drops below the 10% level (see e.g. Fan et al. 2006b; Carilli et al. 2010) and hence does not necessarily probe the location of the ionization front.

Furthermore Bolton & Haehnelt (2007b) showed that a quasar embedded in a highly ionized IGM can produce a proximity zone that appears qualitatively similar to that of one in a neutral ambient IGM. They show that the observed size of the proximity zone of a quasar within an already ionized surrounding IGM is independent of the neutral gas fraction of the ambient IGM and as such, may not provide insights into the evolution of the neutral hydrogen fraction during the epoch of reionization at all.

However, previous observational studies of the sizes of quasar proximity zones found evidence for a steep decrease in proximity zone size with increasing redshift within the redshift range of $5.7 \leq z \leq 6.4$ (Fan et al. 2006b; Willott et al. 2007, 2010; Carilli et al. 2010). Although there is large scatter in the observations, this has been interpreted as a strong evolution of the neutral gas fraction of the IGM assuming that the observed proximity zones trace the extent of the ionized H II region presented in eqn. (1), i.e. $R_p \approx R_{\rm ion}$. With the discovery of the current highest redshift quasar ULASJ1120+0641 (Mortlock et al. 2011) and the analysis of its proximity zone (Bolton et al. 2011; Simcoe et al. 2012; Venemans et al. 2015; Bosman & Becker 2015), this steep decrease has become somewhat shallower, but still indicates a strong evolution in proximity zone size with redshift. However, the measurements of the proximity zone sizes contain a number of uncertainties, most importantly the large uncertainties in the systemic redshifts of the quasars and the resulting ambiguity in the beginning of the proximity zone.

The extent of the H II region around the quasar additionally depends on the age of the quasar (eqn. (1)). Even if the IGM is highly ionized prior to the quasar turning on, the gas in the ionized region responds on a finite timescale (Khrykin et al. 2016) and therefore also the observed proximity zone sizes depend on the quasar ages. Thus, assumptions about the quasar ages are required to interpret quasar proximity zones, but the age remains uncertain by several orders of magnitude (Martini 2004).

To be more precise, we will distinguish between several different timescales that govern the duration of quasar activity. The *duty cycle* $t_{\rm dc}$ refers to the total time that galaxies shine as active quasars integrated over the age of the universe. However, quasar activity could be episodic, and we refer to the duration $t_{\rm episodic}$ of a single emission episode as the *episodic lifetime*. The sizes of quasar proximity zones actually depend on the quasar *age*. If we denote by $t = 0$ the time at which light emitted by a quasar just reaches our telescopes on Earth, then the quasar age, which we will henceforth refer to as $t_{\rm Q}$, is defined such that the quasar actually turned on at a time $-t_{\rm Q}$ in the past. Strictly speaking the age $t_{\rm Q}$ is a lower limit on the quasar episodic lifetime $t_{\rm episodic}$, because the quasar episode might indeed continue, which we can only record on Earth if we could conduct observations in the future.

The quasar duty cycle can be constrained by comparing the number density of quasars to their host dark matter halo abundance inferred from their clustering strength (Haiman & Hui 2001; Martini & Weinberg 2001; Martini 2004; White et al. 2008; Furlanetto & Lidz 2011). But to date this method has yielded only weak constraints on $t_{\rm dc} \sim 10^6 - 10^9$ yr owing to uncertainties in how quasars populate dark matter halos (Shen et al. 2009; White et al. 2012; Conroy & White 2013; Cen & Safarzadeh 2015). An upper limit on the duty cycle of quasars, $t_{\rm dc} < 10^9$ yr, is set by the observed evolution of the quasar luminosity function, since the whole quasar population rises and falls over roughly this timescale (e.g. Osmer 1998).

Moreover, these constraints on the duty cycle of quasars do not give insights into the duration of the individual accretion episodes of the quasar activity. A population of quasars emitting 1000 individual bursts each with a lifetime of $t_{\rm episodic} \sim 10^5$ yr, would be indistinguishable from quasars with steady continuous emission for $t_{\rm episodic} \sim 10^8$ yr. A lower limit on the individual quasar bursts of $t_{\rm episodic} \sim 3 \times 10^4$ yr is based on the argument that quasars need to maintain their ionizing luminosity long enough to explain the observed proximity zones in the Lyα forest (e.g. Bajtlik et al. 1988; Khrykin et al. 2016).

It has been argued, that in order to grow the observed sizes of SMBHs, i.e. $M_{\rm SMBH} \sim 10^9 - 10^{10} M_\odot$ at $z \sim 6 - 7$ (Mortlock et al. 2011; Venemans et al. 2013; De Rosa et al. 2014; Wu et al. 2015), the quasars require very massive initial seeds and need to accrete matter over timescales comparable to the age of the Universe at these high redshifts (Hopkins & Hernquist 2009; Volonteri 2010, 2012). It follows that the duty cycle of these quasars needs to be of the order of the Hubble time. If quasar activity is episodic, the episodic bursts need to be very long or the quiescent time in between

the bursts needs to be short for the SMBHs to grow to their observed sizes.

The dozens of $z \sim 6$ quasars that have been uncovered over the past decade (see Bañados et al. 2016, for a recent compilation) from wide field surveys results in many new quasars for proximity zone measurements. The statistical power of these data alone motivates revisiting this type of analysis to further understand the resulting constraints on reionization and the quasar emission timescales. In this paper we re-investigate the evolution of the quasar proximity zone sizes at $z \geq 5.77$ with a homogeneous analysis of a significantly enlarged sample of 31 quasar spectra with higher quality data, including updated and more precise redshift measurements from CO and [C II] line observations, and consistently measured absolute magnitudes. We then compare our analysis with state-of-the-art radiative transfer simulations in order to better understand and interpret our measurements.

This paper is structured as follows: we describe our data set and the reduction pipeline in § 2. In § 3 we specify the methods we use for the continuum normalization and the measurement of the proximity zone size of each quasar. In § 4 we describe a suite of radiative transfer simulations that we compare to our observations. We show our measurements of the proximity zone sizes in § 5 and discuss their evolution with redshift. We highlight three exceptionally small proximity zones in our sample in § 6, and discuss possible explanations for their origin. We summarize and conclude in § 7.

2. HIGH REDSHIFT QUASAR SAMPLE

Our initial data set consists of 34 quasar spectra at $5.77 \leq z_{\rm em} \leq 6.54$, observed at optical wavelengths (4000 Å - 10000 Å) with the Echellette Spectrograph and Imager (ESI; Sheinis et al. 2002) at the Keck II Telescope in the years 2001 to 2016. We collected data from the Keck Observatory Archive[7] and supplemented it with our own observations. The slit widths used range from $0.75" - 1.0"$, giving a resolution of $R \approx 4000 - 5400$. The exposure times vary from 0.3 h $\lesssim t_{\rm exp} \lesssim$ 25 h resulting in median signal-to-noise ratios in the quasar continuum at rest-frame wavelength of 1250Å-1280Å in the range of $2 \lesssim {\rm S/N} \lesssim 108$ per pixel.

We obtained spectra of the four quasars PSOJ0226 + 0302[8], PSOJ0402 + 2452[9], SDSSJ0100 + 2802, and SDSSJ1137 + 3549 with ESI on on January 11th and 12th, 2016. We used a 1" slit and exposure times varied between $1 - 3.25$ hours.

For one object CFHQSJ0227−0605 of interest because of its small proximity zone (see § 6), we obtained an additional spectrum with the Low Resolution Imaging Spectrometer (LRIS; Oke et al. 1995) on the Keck I Telescope with an exposure time of 3600 s. The observations were conducted on September 16th, 2016 using the 600/7500 grating and a slit width of 1", resulting in a resolution of $R \approx 1800$. Note that we only use this LRIS spectrum for follow-up analysis, i.e. searching for metal absorption lines. For the main analysis of the proximity zone of this quasar we use the ESI spectrum, in order to be consistent with the remaining data sample.

The details of all of the observations can be found in Table 1.

TABLE 1
Overview of the observations of the 34 quasars in our data sample.

object	RA (J2000)	DEC (J2000)	PI	observation date	exposure time
SDSS J0002+2550	$00^{\rm h}02^{\rm m}39\overset{\rm s}{.}39$	$+25°50'34\overset{''}{.}96$	Kakazu	Nov. 2004	5400 s
			Cowie	Aug. 2005	16300 s
SDSS J0005-0006	$00^{\rm h}05^{\rm m}52\overset{\rm s}{.}34$	$-00°06'55\overset{''}{.}80$	Becker	Dec. 2002	1200 s
			Sargent	Oct. 2010	15000 s
CFHQS J0050+3445	$00^{\rm h}55^{\rm m}02\overset{\rm s}{.}91$	$+34°45'21\overset{''}{.}65$	Willott	Sep. 2008	6250 s
			Sargent	Oct. 2010	6000 s
			Sargent	Oct. 2012	9000 s
SDSS J0100+2802	$01^{\rm h}00^{\rm m}13\overset{\rm s}{.}02$	$+28°02'25\overset{''}{.}92$	White	Jan. 2016	3600 s
ULAS J0148+0600	$01^{\rm h}48^{\rm m}37\overset{\rm s}{.}64$	$+06°00'20\overset{''}{.}06$	Sargent	Oct. 2010	11200 s
ULAS J0203+0012	$02^{\rm h}03^{\rm m}32\overset{\rm s}{.}38$	$+00°12'29\overset{''}{.}27$	Sargent	Oct. 2010	6600 s
CFHQS J0210-0456	$02^{\rm h}10^{\rm m}13\overset{\rm s}{.}19$	$-04°56'20\overset{''}{.}90$	Sargent	Oct. 2010	6000 s
PSO J0226+0302	$02^{\rm h}26^{\rm m}01\overset{\rm s}{.}87$	$+03°02'59\overset{''}{.}42$	White	Jan. 2016	11700 s
CFHQS J0227-0605	$02^{\rm h}27^{\rm m}43\overset{\rm s}{.}29$	$-06°05'30\overset{''}{.}20$	Willott	Sep. 2008	10540 s
SDSS J0303-0019	$03^{\rm h}03^{\rm m}31\overset{\rm s}{.}40$	$-00°19'12\overset{''}{.}90$	Sargent	Oct. 2010	6000 s
SDSS J0353+0104	$03^{\rm h}53^{\rm m}49\overset{\rm s}{.}73$	$+01°04'04\overset{''}{.}66$	Sargent	Oct. 2010	13200 s
			Becker	Jan. 2006	3600 s
PSO J0402+2452	$04^{\rm h}02^{\rm m}12\overset{\rm s}{.}69$	$+24°51'24\overset{''}{.}43$	White	Jan. 2016	10800 s
SDSS J0818+1723	$08^{\rm h}18^{\rm m}27\overset{\rm s}{.}40$	$+17°22'52\overset{''}{.}01$	Becker	Apr. 2005	2400 s
SDSS J0836+0054	$08^{\rm h}36^{\rm m}43\overset{\rm s}{.}86$	$+00°54'53\overset{''}{.}26$	Becker	Mar. 2001	1731 s
			Madau	Feb. 2002	18900 s
			Cowie	Feb. 2002	13540 s
			Djorgovski	Mar. 2002	8923 s
			Kulkarni	Jan. 2003	5400 s
			Cowie	Jan. 2003	3600 s
			Cowie	Jan. 2004	10800 s
			Kakazu	Nov. 2004	10200 s
			Djorgovski	Dec. 2004	7200 s
SDSS J0840+5624	$08^{\rm h}40^{\rm m}35\overset{\rm s}{.}30$	$+56°24'20\overset{''}{.}22$	Djorgovski	Dec. 2004	10800 s
			Becker	Mar. 2006	1200 s

[7] `https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOAlogin`
[8] also known as PSO J036.5078+03.0498 (Venemans et al. 2015).
[9] also known as PSO J060.5529+24.8567 (Bañados et al. 2016).

TABLE 1 — Continued

object	RA (J2000)	DEC (J2000)	PI	observation date	exposure time
SDSS J0842+1218	$08^h42^m29^s.43$	$+12°18'50''.58$	Becker	Mar. 2006	2400 s
SDSS J0927+2001	$09^h27^m21^s.82$	$+20°01'23''.64$	Becker	Mar. 2006	1200 s
SDSS J1030+0524	$10^h30^m27^s.11$	$+05°24'55''.06$	Becker	May 2001	1800 s
			Becker	Jan. 2002	12000 s
			Madau	Feb. 2002	7339 s
			Cowie	Feb. 2002	16200 s
			Djorgovski	Mar. 2002	16200 s
			Cowie	Jan. 2003	5400 s
SDSS J1048+4637	$10^h48^m45^s.07$	$+46°37'18''.55$	Becker	Dec. 2002	3600 s
			Cowie	Jan. 2003	18000 s
			Cowie	Mar. 2003	5400 s
			Sanchez	Jun. 2003	5400 s
			Djorgovski	Dec. 2004	6600 s
SDSS J1137+3549	$11^h37^m17^s.73$	$+35°49'56''.85$	White	Jan. 2016	7800 s
			Becker	Jan. 2005	2400 s
SDSS J1148+5251	$11^h48^m16^s.65$	$+52°51'50''.39$	Becker	May 2002	8100 s
			Becker	Dec. 2002	22800 s
			Djorgovski	Dec. 2002	11400 s
			Kulkarni	Jan. 2003	3600 s
			Cowie	Jan. 2003	16200 s
			Becker	Feb. 2003	16800 s
			Cowie	Jan. 2004	11700 s
SDSS J1250+3130	$12^h50^m51^s.93$	$+31°30'21''.90$	Becker	Jan. 2005	3600 s
SDSS J1306+0359	$13^h06^m08^s.27$	$+03°59'26''.36$	Becker	May 2001	900 s
			Cowie	Feb. 2002	15300 s
			Djorgovski	Mar. 2002	12600 s
			Cowie	Mar. 2003	9000 s
			Cowie	Mar. 2004	12600 s
ULAS J1319+0950	$13^h19^m11^s.30$	$+09°50'51''.52$	Steidel	Mar. 2008	2400 s
SDSS J1335+3533	$13^h35^m50^s.81$	$+35°33'15''.82$	Becker	Mar. 2006	1200 s
SDSS J1411+1217	$14^h11^m11^s.29$	$+12°17'37''.28$	Djorgovski	Apr. 2004	10200 s
			Becker	Jan. 2005	3600 s
			Cowie	May 2005	10800 s
			Cowie	Mar. 2008	12600 s
SDSS J1602+4228	$16^h02^m53^s.98$	$+42°28'24''.94$	Cowie	May. 2005	10620 s
			Cowie	Aug. 2005	3600 s
SDSS J1623+3112	$16^h23^m31^s.81$	$+31°12'00''.53$	Becker	Jan. 2004	3600 s
SDSS J1630+4012	$16^h30^m33^s.90$	$+40°12'09''.69$	Sanchez	Jun. 2003	17700 s
			Sargent	Oct. 2010	2000 s
CFHQS J1641+3755	$16^h41^m21^s.73$	$+37°55'20''.15$	Willott	Oct. 2007	2400 s
SDSS J2054-0005	$20^h54^m06^s.49$	$-00°05'14''.80$	Sargent	Oct. 2010	12000 s
CFHQS J2229+1457	$22^h29^m01^s.65$	$+14°57'09''.00$	Willott	Sep. 2008	3600 s
			Prochaska*	Sep. 2016	3600 s
SDSS J2315-0023	$23^h15^m46^s.57$	$-00°23'58''.10$	Becker	Jan. 2006	1200 s
			Sargent	Oct. 2010	27000 s
CFHQS J2329-0301	$23^h29^m08^s.28$	$-03°01'58''.80$	Willott	Oct. 2007	10800 s
			Willott	Sep. 2008	14400 s

NOTE. — Columns show the object name, its coordinates, the principal investigator of the different observing runs, the observation date and total exposure time of the object in this run.

* spectrum taken with LRIS

2.1. *Data Reduction*

We reduce all spectra uniformly using the ESIRedux pipeline[10] developed as part of the XIDL[11] suite of astronomical routines in the Interactive Data Language (IDL). This pipeline employs standard data reduction techniques which can be summarized as follows: Images are overscan subtracted, flat fielded using a normalized flat field image, and then wavelength calibrated using a wavelength image constructed from afternoon arc lamp calibration images. Objects are identified in the science frames, and then background subtracted using B-spline fits (Kelson 2003; Bernstein et al. 2015) to object free regions of the slit. Object profiles are also fit with B-splines, and optimal extraction is performed on the sky-subtracted frames. One-dimensional spectra of overlapping echelle orders are combined to produce a final spectrum for each exposure, and individual exposures are co-added into our final one-dimensional spectra. See (Bochanski et al. 2009) for a more detailed description of the algorithms used.

We further optimized the XIDL ESI pipeline to improve the data reduction for high redshift quasars. The main improvement was to remove the fringing pattern from reddest orders by differencing two images (ideally taken during the same run) with similar exposure times, analogous to the standard difference imaging techniques performed for near-infrared observations. However, this procedure only works on dithered exposures for which the trace of the science object lands at different spatial locations on the slit. Since not every observer dithered

[10] `http://www2.keck.hawaii.edu/inst/esi/ESIRedux/`
[11] `http://www.ucolick.org/~xavier/IDL/`

their object along the slit it was not possible for us to apply this procedure to ≈ 10% of the exposures that we took from the archive.

We also co-added exposures from different observing runs taken by different PIs, resulting in higher S/N data for some quasars than previous analyses of these objects. We weight each one-dimensional spectrum by its squared signal-to-noise ratio (S/N^2) that was determined in the quasar continuum region of each spectrum, i.e. at wavelengths longer than the Lyα emission line. In this way spectral regions with low or no transmitted flux, which are common in high redshift quasar spectra, are weighted by the same S/N ratio of regions with more transmitted flux.

2.2. *Quasar Properties*

Determining precise redshifts for quasars is very challenging due to the broad widths of emission lines, Gunn-Peterson absorption, and offsets between different ionization lines (Gaskell 1982; Tytler & Fan 1992; Vanden Berk et al. 2001; Richards et al. 2002; Shen et al. 2016). Most quasars show strong internal motions and winds, which displace many of the emission lines, such as the Lyα line or far-ultraviolet (far-UV) lines, far from the systemic redshift of the host galaxy. Thus the most precise determination of the location of the quasar are low ionization lines such as Mg II lines or, even better, emission lines from the molecular gas reservoir of the host galaxy itself, such as CO or [C II] lines.

TABLE 2
Overview of our data sample and the measurements of the proximity zone sizes.

object	z	Δz [km/s]	redshift line	Ref.[a]	M_{1450}	S/N[b]	R_p [pMpc]	$R_{p,\mathrm{corr}}$ [pMpc]
PSO J0226+0302	6.5412	100	[C II]	13	−27.33	10	3.64 ± 0.13	3.20 ± 0.11
CFHQS J0210-0456	6.4323	100	[C II]	12	−24.53	2	1.32 ± 0.13	3.47 ± 0.34
SDSS J1148+5251	6.4189	100	[C II]	15	−27.62	34	4.58 ± 0.13	3.59 ± 0.10
CFHQS J2329-0301	6.417	270	Mg II	9	−25.25	2	2.45 ± 0.35	4.86 ± 0.70
SDSS J0100+2802	6.3258	100	[C II]	17	−29.14	41	7.12 ± 0.13	3.09 ± 0.06
SDSS J1030+0524	6.309	270	Mg II	3	−26.99	26	5.93 ± 0.36	5.95 ± 0.36
SDSS J1623+3112	6.2572	100	[C II]	18	−26.55	9	5.05 ± 0.14	6.03 ± 0.16
CFHQS J0050+3445	6.253	270	Mg II	9	−26.70	18	4.09 ± 0.37	4.60 ± 0.41
CFHQS J0227-0605	6.20	1000	Lyα	6	−25.28	3	1.60 ± 1.37	3.15 ± 2.69
PSO J0402+2452	6.18	1000	Lyα	16	−26.95	13	4.17 ± 1.38	4.26 ± 1.40
CFHQS J2229+1457	6.1517	100	[C II]	15	−24.78	2[c]	0.45 ± 0.14	1.07 ± 0.33
SDSS J1250+3130	6.15	1000	Lyα-O I-Si IV	2	−26.53	8	6.59 ± 1.38	7.93 ± 1.66
ULAS J1319+0950	6.1330	100	[C II]	11	−27.05	10	3.84 ± 0.14	3.77 ± 0.14
SDSS J2315-0023	6.117	1000	Lyα	5	−25.66	15	3.70 ± 1.39	6.26 ± 2.36
SDSS J1602+4228	6.09	1000	Lyα-O I	1	−26.94	21	7.11 ± 1.40	7.28 ± 1.43
SDSS J0303-0019	6.078	270	Mg II	7	−25.56	2	2.21 ± 0.38	3.88 ± 0.67
SDSS J0842+1218	6.069	270	Mg II	10	−26.91	10	6.47 ± 0.38	6.71 ± 0.39
SDSS J1630+4012	6.065	270	Mg II	7	−26.19	15	4.80 ± 0.38	6.59 ± 0.52
CFHQS J1641+3755	6.047	270	Mg II	9	−25.67	4	3.98 ± 0.38	6.71 ± 0.64
SDSS J2054-0005	6.0391	100	[C II]	11	−26.21	17	3.17 ± 0.14	4.32 ± 0.19
SDSS J1137+3549	6.03	1000	Lyα-O I-Si IV	2	−27.36	24	6.98 ± 1.42	6.06 ± 1.23
SDSS J0818+1723	6.02	1000	Lyα	2	−27.52	8	5.89 ± 1.42	4.80 ± 1.16
SDSS J1306+0359	6.016	270	Mg II	4	−26.81	41	5.39 ± 0.38	5.80 ± 0.41
ULAS J0148+0600	5.98	270	Mg II	14	−27.39	30	6.03 ± 0.39	5.18 ± 0.33
SDSS J1411+1217	5.904	270	Mg II	4	−26.69	32	4.60 ± 0.39	5.19 ± 0.44
SDSS J1335+3533	5.9012	100	CO	8	−26.67	7	0.78 ± 0.15	0.89 ± 0.17
SDSS J0840+5624[d]	5.8441	100	CO	8	−27.24	28	0.88 ± 0.15	0.80 ± 0.13
SDSS J0005-0006	5.844	270	Mg II	10	−25.73	15	2.87 ± 0.40	4.73 ± 0.66
SDSS J0002+2550	5.82	1000	Lyα-O I	1	−27.31	62	5.43 ± 1.49	4.81 ± 1.31
SDSS J0836+0054	5.810	270	Mg II	4	−27.75	120	5.06 ± 0.40	3.77 ± 0.30
SDSS J0927+2001	5.7722	100	CO	8	−26.76	7	4.68 ± 0.15	5.14 ± 0.16

Note. — The columns show the object name, the redshift of the quasar and the redshift uncertainty, the lines measured to determine the redshift and the reference therefor, the quasar's magnitude M_{1450} and the S/N of the its spectrum and our measurements for the proximity zones, uncorrected and luminosity corrected.

[a] Reference for redshift. 1: Fan et al. (2004), 2: Fan et al. (2006a), 3: Jiang et al. (2007), 4: Kurk et al. (2007), 5: Jiang et al. (2008), 6: Willott et al. (2009), 7: Carilli et al. (2010), 8: Wang et al. (2010), 9: Willott et al. (2010), 10: De Rosa et al. (2011), 11: Wang et al. (2013), 12: Willott et al. (2013), 13: Bañados et al. (2015), 14: Becker et al. (2015), 15: Willott et al. (2015), 16: Bañados et al. (2016), 17: Wang et al. (2016), 18: private communication with R. Wang.

[b] Median S/N per pixel; estimated between 1250 Å$\leq \lambda_{\mathrm{rest}} \leq$ 1280 Å.

[c] The LRIS spectrum we have for this object has S/N = 7.

[d] Excluded from our analysis due to associated absorbers (see Appendix A).

We account for uncertainties in the systemic redshift of each quasar depending on the measured emission lines to determine their redshifts. For redshifts determined from the detection of a [C II] line at 158 μm or a CO line we assign a redshift error of $\Delta v = 100$ km/s. For quasars with a redshift measurement from a Mg II line we assume a redshift error of $\Delta v = 270$ km/s, in order to account for the dispersion between the redshift of the Mg II line and the redshift of the host galaxy (e.g. Richards et al. 2002; Hennawi et al. 2006; Shen & Ménard 2012; Venemans et al. 2016; Shen et al. 2016), and for the remaining quasars for which the redshift was determined by the Lyα line or far-UV lines, which usually suffer from large velocity offsets, we assume a redshift error of $\Delta v = 1000$ km/s. Note that these uncertainties are chosen to be very conservative (Shen et al. 2016), because the redshift measurements are taken from the literature, measured by various authors possibly us-

ing different data and methods. For a quasar at $z = 6$ these offsets in velocity result in a distance uncertainty of $\Delta R \approx 0.14$ proper Mpc (pMpc) for quasars with redshift measurements from CO or [C II] lines, $\Delta R \approx 0.39$ pMpc for quasars with redshift measurements from their Mg II line, and $\Delta R \approx 1.43$ pMpc for other rest-frame UV lines.

We take the absolute magnitudes M_{1450} defined at $\lambda_{\rm rest} = 1450$ Å in the rest-frame from Bañados et al. (2016), who determined M_{1450} for all sources in a consistent way. They assume a power law continuum slope $\alpha_\nu = -0.3$ and require the extrapolated $y_{\rm P1}$- or J-band magnitudes (at $\lambda_{\rm eff} = 9627.7$ Å or $\lambda_{\rm eff} = 12444.0$ Å, respectively) to be consistent with their measurements, since most optical quasar spectra at these high redshifts have limited wavelength coverage at $\lambda_{\rm rest} = 1450$Å. Table 2 summarizes the properties of all quasars in our sample.

Note that we exclude three quasars (SDSSJ1048+4637, SDSSJ0353 + 0104, and ULASJ0203 + 0012) from this analysis, because of their broad absorption line (BAL) features, which can contaminate the proximity zone and make it difficult to determine a precise redshift or the continuum level of these quasars. For a similar reason we exclude the object SDSSJ0840 + 5624 from our analysis of the proximity zones. In this spectrum we found an absorption system with associated Lyα and metal line absorption close to the quasar that attenuates enough flux such that it results in a spuriously small proximity zone. In this case, as in the case of BAL quasars, an absorption system associated with the quasar itself contaminates the proximity zone and thus we exclude this object from our analysis. Note that we do not model such intrinsic absorption line systems associated with the quasar or its host galaxy in the radiative transfer simulations (see § 4), such that any comparison would result in biases if we were not excluding these objects. We further discuss the spectrum of this excluded object SDSSJ0840 + 5624 in Appendix A.

3. METHODS

3.1. *Quasar Continuum Normalization*

We normalize all quasar spectra to unity in a region free of emission lines at $\lambda_{\rm rest} = 1280$ Å. Note that for a handful of quasars, that have very low signal-to-noise ratio data at this wavelength, we normalize their spectra to unity at a slightly different wavelength. We then estimate the quasar continua with principal component spectra (PCS) from a principal component analysis (PCA) of low redshift quasar spectra (Suzuki 2006; Pâris et al. 2011). The idea of the PCA is to represent the continuum spectrum $|q_\lambda\rangle$ of a quasar by a reconstructed spectrum that consists of a mean spectrum $|\mu_\lambda\rangle$ and a sum of m weighted PCS $|\xi_\lambda\rangle$, where the index λ denotes the wavelength. Hence

$$|q_\lambda\rangle \approx |\mu_\lambda\rangle + \sum_{i=1}^{m} \alpha_i \, |\xi_{i,\lambda}\rangle \, , \tag{2}$$

where the index i refers to the ith PCS and α_i is its weight. For the majority of our continuum fits, we use the PCS from Pâris et al. (2011) that were derived to characterize quasar continua from 78 high-quality spectra of bright $z_{\rm em} \sim 3$ quasars in the Sloan Digital Sky Survey (SDSS).

The quasars in our sample are all at very high redshift and thus experience absorption due to the intervening residual neutral hydrogen along the line of sight bluewards of the Lyα emission line, i.e. in the Lyα forest. Thus we estimate the quasar continuum solely from wavelengths redwards of rest-frame Lyα which suffer only from modest absorption by metal lines. Pâris et al. (2011) provide a set of PCS for wavelengths 1215.67 Å$\leq \lambda_{\rm rest} \leq$ 1600 Å, whose coefficients $\boldsymbol{\alpha_{red}}$ we estimate by χ^2 minimization using the noise vector from the spectra.

We use a projection matrix $\boldsymbol{P}$ to transfer the estimated coefficients for the PCS *redwards* of Lyα, $\boldsymbol{\alpha_{red}}$, onto coefficients $\boldsymbol{\alpha}$ for a set of PCS that cover the *entire* spectral region between 1020 Å $\leq \lambda_{\rm rest} \leq$ 1600 Å, i.e.

$$\boldsymbol{\alpha} = \boldsymbol{P} \cdot \boldsymbol{\alpha_{red}}. \tag{3}$$

This projection matrix $\boldsymbol{P}$ has been calibrated by Pâris et al. (2011) using the set of PCS for both the red wavelength side only and the whole spectral region covering wavelengths bluewards and redwards of Lyα.

Eqn. (2) gives us a model for the continuum of each quasar that includes the (absorbed) blue side of the spectrum. An example of a quasar spectrum from our data set and its continuum model is shown in Fig. 1. Note that in most cases we take five PCS on the red side into account to estimate the continuum model. However, a visual inspection of the continuum models reveals that for four quasars a continuum model with three or seven PCS results in a better fit.

Our PCA modeling of quasar continua is limited by the finite number of principal components and the selection of quasars that were used to create the PCS. For a few of the quasars in our sample, the PCS from Pâris et al. (2011) do not provide an acceptable fit to the continuum, particularly for the fainter quasars in our sample. For these objects we instead use the set of PCS provided by Suzuki (2006), based on fainter lower redshift quasars, which qualitatively provide better fits to this handful of objects. When using the Suzuki (2006) PCS basis, we determine the continuum in the Lyα forest region using the method they advocate. Namely, we determine the continua at blue wavelengths $\lambda_{\rm rest} < 1216.67$ Å using the best-fit continua to the red-side $\lambda_{\rm rest} > 1216.67$ Å.

The details of the continuum normalization are listed in Table 3 in Appendix B.

3.2. *Measuring the Sizes of Quasar Proximity Zones*

In order to calculate the proximity zone sizes for each quasar we adopt the standard definition used in the literature (see Fan et al. 2006b; Willott et al. 2007, 2010; Carilli et al. 2010; Venemans et al. 2015). Namely, we take the continuum normalized quasar spectra and smooth them with a 20Å-wide (observed frame) boxcar function. This smoothing scale corresponds to ≈ 1.0 pMpc or a ≈ 705 km/s window at $z = 6$. We define the proximity zone size as the distance to the first of three consecutive pixels[12] on the blue side of the Lyα emission line that show a drop of the smoothed flux below the 10% level. This is demonstrated in Fig. 2 and Fig. 3 for subsets of

[12] We checked whether it changes the measurements of the proximity zone sizes when requiring up to ten pixels to be below the 10% level, but we do not observe a significant difference.

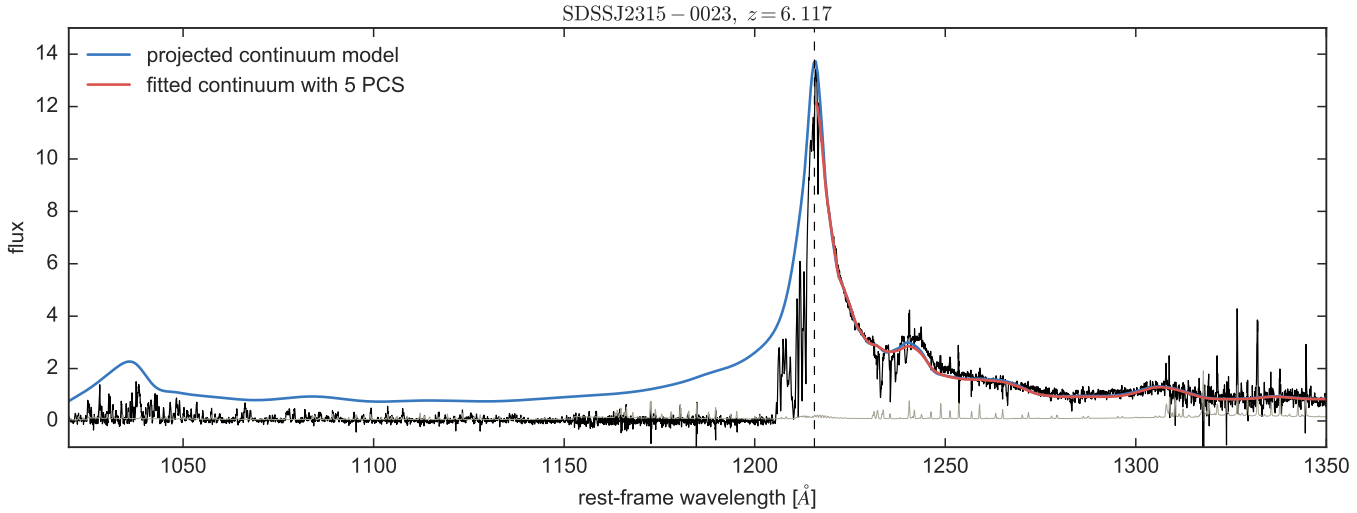

FIG. 1.— Example of a quasar spectrum and its continuum model. The continuum model is fitted to the data with five PCS from Pâris et al. (2011) on the red wavelength side of the spectrum (red curve), i.e. 1215.67 Å$\leq \lambda_{\rm rest} \leq$ 1600 Å, and then projected onto coefficients for a different set of PCS resulting in a continuum model also covering the blue wavelength side (blue curve), i.e. the whole spectral region between 1020 Å$\leq \lambda_{\rm rest} \leq$ 1600 Å, in order to predict the continuum level in the Lyα forest region.

bright ($-27.5 \leq M_{1450} \leq -26.5$) and faint ($-25.8 \leq M_{1450} \leq -24.5$) quasars from our data sample, respectively. The depicted quasars all cover a similar range in luminosity, but nevertheless show a wide range of proximity zone sizes, between 0.8 pMpc $\lesssim R_p \lesssim$ 7.1 pMpc for bright quasars and 0.5 pMpc $\lesssim R_p \lesssim$ 4.0 pMpc for fainter quasars. Similar plots for the remaining objects in our quasar sample are shown in Appendix C. All measurements of the proximity zone sizes R_p for our data set are listed in Table 2.

4. RADIATIVE TRANSFER SIMULATIONS

To interpret our measurements of R_p, we run a series of radiative transfer simulations of the effect of quasar ionizing photons on the IGM along the line of sight similar to those performed by Bolton & Haehnelt (2007b). We apply the one-dimensional ionizing radiative transfer code from Davies et al. (2016) to skewers from a 100 Mpc/h Eulerian hydrodynamical simulation run with the Nyx code (Almgren et al. 2013; Lukić et al. 2015) at $z = 6$ with 4096^3 baryonic (Eulerian) grid elements and dark matter particles. The radiative transfer computes the time-dependent ionization and recombination of six species (e^-, H I, H II, He I, He II, He III) as well as the associated photoionization heating and cooling by various processes including adiabatic cooling due to the expansion of the Universe and inverse Compton cooling off cosmic microwave background (CMB) photons (see Davies et al. 2016, for details). We use 900 skewers of density, temperature, and peculiar velocity drawn along the x, y, and z grid axes from the centers of the 150 most massive dark matter halos in the Nyx simulation, corresponding to halo masses $M_{\rm h} \gtrsim 4 \times 10^{11}$ $M_\odot$[13]. Given the relatively coarse resolution of the Nyx simulation (25 kpc/h) and the lack of metal-line cooling, star formation, and feedback processes, we do not expect to resolve or accurately model dense gas inside of galaxies. For this reason we have removed one object in our sample (SDSSJ0840 + 5624; see Appendix A) which exhibits strong associated metal line absorption at the quasar redshift that likely arises from dense gas in the vicinity of the quasar or from its host halo.

Quasar absolute magnitudes were converted to specific luminosity at the hydrogen ionizing edge ($\nu = \nu_{\rm HI}$) using the Lusso et al. (2015) spectral template, and for consistency with previous studies we assume that the spectrum at $\nu > \nu_{\rm HI}$ behaves as a power law[14] with $L_\nu \propto \nu^{-1.7}$. We assume that the quasar turns on abruptly and emits at constant luminosity for its entire age $t_{\rm Q}$, i.e. a so-called "light bulb" model. For the quasar age we assume a fiducial value of $t_{\rm Q} = 10^{7.5}$ yr, but we later investigate the dependence of R_p on $t_{\rm Q}$ in § 6.3.3. We consider two initial conditions for the ionization state of gas in the simulation: either the gas is initially highly ionized by a uniform ionizing background or the gas is initially completely neutral. In the highly ionized case, we add ionizing radiation due to the ultraviolet background (UVB) to each cell leading to an ionization rate $\Gamma_{\rm UVB} = 2 \times 10^{-13}$ s^{-1}, consistent with observations of Lyα forest opacity at $z \sim 6$ (e.g. Wyithe & Bolton 2011) and resulting in a neutral fraction of $f_{\rm HI} = 1.5 \times 10^{-4}$, with a spectrum characteristic of galactic sources (i.e. a sharp cutoff above the He II ionizing edge). For simplicity we assume that $\Gamma_{\rm UVB}$ is constant with redshift, but we note that the resulting proximity zone sizes are insensitive to changes in $\Gamma_{\rm UVB}$ of a factor of a few, a non-trivial point that we will discuss in more detail in future work (Davies et al. in prep.). We compute radiative transfer outputs for quasars of varying luminosity and redshift bracketing the properties of our observed quasar sample in in both ionized and neu-

[13] In detail we find that starting the skewers from more or less massive halos has a negligible impact on the resulting proximity zone sizes, in agreement with Bolton & Haehnelt (2007b) and Keating et al. (2015).

[14] Note that the magnitudes for the quasars in our data sample have been calibrated by Bañados et al. (2016) according to the template from Selsing et al. (2016) instead of Lusso et al. (2015). However, due to the similarities between the two templates at observed wavelengths (redward of Lyα), we expect only a very minor inconsistency in the inferred ionizing luminosity.

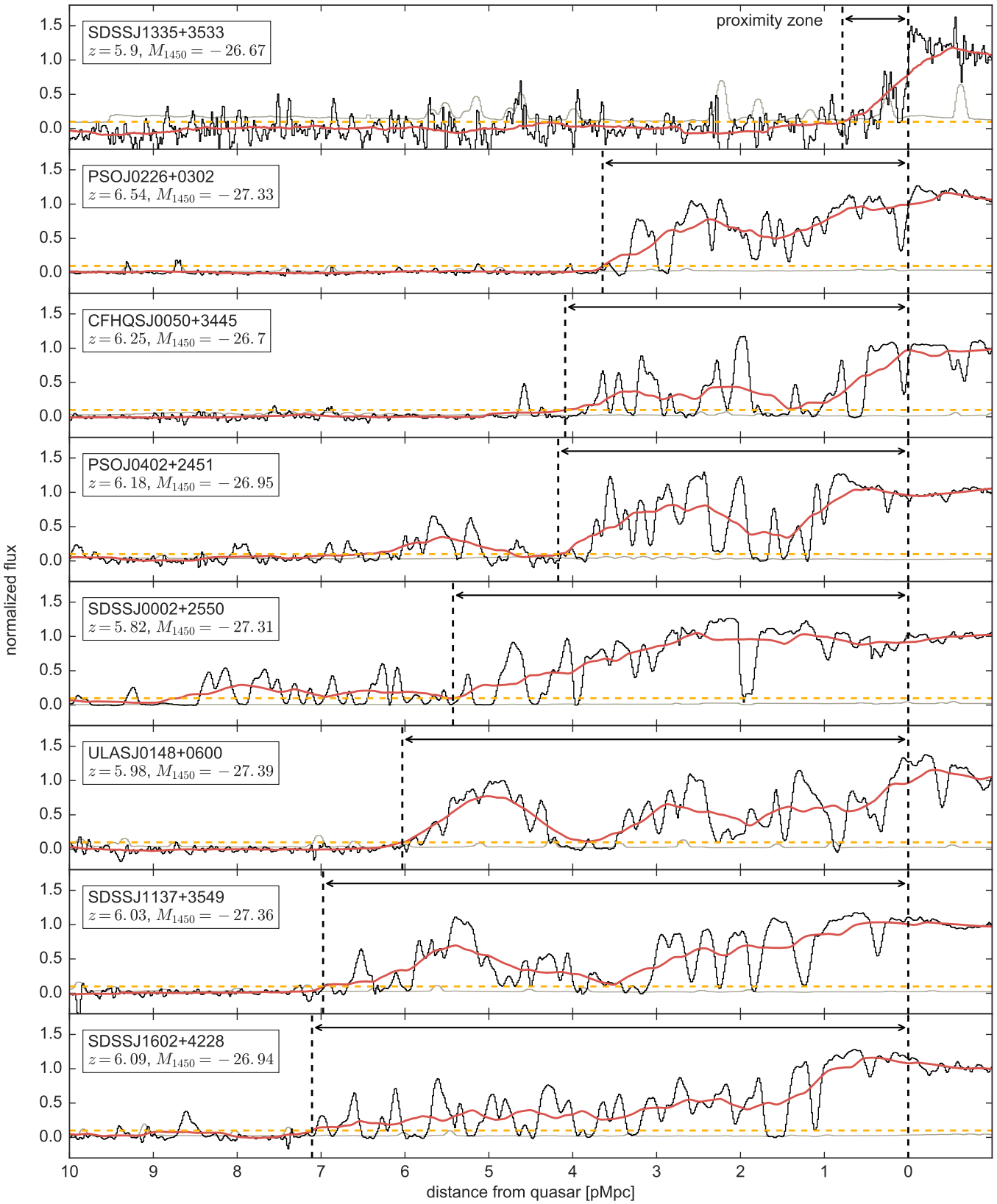

FIG. 2.— Continuum normalized spectra of a subset of bright quasars in our data set with luminosities between $-27.5 \leq M_{1450} \leq -26.5$ showing the transmission profile within 10 pMpc from the quasar. A boxcar smoothing of two pixels has been applied to the spectra (black) and their noise vectors (gray). The red curves show the quasar spectra smoothed to a resolution of 20Å in observed wavelength. The horizontal yellow dashed lines indicate a flux level of 10%. The vertical black dashed lines show the extent of the quasar proximity zone from the quasar redshift (right) to the first drop of the smoothed flux below the 10% level (left).

tral scenarios. We assume here that the *overdensity* field does not significantly evolve with redshift from our $z = 6$ output across the redshift range we study, and simply re-scale physical densities by $(1 + z)^3$.

Lyα forest spectra are computed by combining the neutral fraction $f_{\rm HI}$ and gas temperatures from the radiative transfer simulation with the velocity field from the Nyx simulation, summing the absorption from each gas element using the efficient Voigt profile approximation of Tepper-García (2006). Finally, to retrieve R_p from the Lyα forest spectra, we perform a similar analysis as is applied to the real spectroscopic data: we smooth the spectra with a boxcar filter of 20 Å in the observed frame, and then locate where the transmitted flux first drops below 10%. A few skewers through the simulation box giving mock Lyα forest spectra in a highly ionized IGM are shown in Fig. 4.

In this work we present the general trends of R_p with quasar parameters, leaving a more detailed discussion of the structure and evolution of these simulated proximity effect spectra to future work (Davies et al. in prep.).

5. THE REDSHIFT EVOLUTION OF PROXIMITY ZONE SIZES

In this section we will present our measurements of the proximity zone sizes for the ensemble of quasar spectra in our data set. Since these quasars cover a wide range of luminosities, we first put them on a common luminosity scale, in order to facilitate a study of the redshift evolution of proximity zone sizes. In the last part of this section we compare our measurements to previous work.

5.1. *Correcting for the Quasar Luminosity*

We expect the size of the proximity zone of a quasar to depend on their luminosity. Proximity zone sizes may also evolve with redshift in a way that tracks the evolution of the ionization state of the IGM driven by the underlying evolution in the UVB. To study the evolution of the proximity zone size with redshift, we have to remove the dependency on quasar luminosity, and normalize our measurements to the same absolute magnitude. The approach that has been adopted in the literature, is to re-scale all quasar luminosities to a common value assuming a particular model for the luminosity dependence, in order to remove the large scatter in proximity zone sizes driven by the wide range of quasar luminosities present in the data set. This then enables one to study the redshift evolution of proximity zone sizes. However, the scaling of the proximity zone size with luminosity depends on the physical conditions of the ambient IGM. In this section we study radiative transfer simulations to better understand the expected luminosity scaling, and its dependence on the IGM ionization state. We compute the luminosity scaling in our data set and compare to these simulations. We use these simulations to advocate for an approach to luminosity correct our measurements and search for a redshift evolution.

In general, the luminosity dependence of proximity zone sizes could depend on the ionization state of the surrounding IGM. Specifically eqn. 1 indicates that the size of an ionization front $R_{\rm ion}$ expanding into a neutral IGM scales as $R \propto \dot{N}_\gamma^{1/3}$. On the other hand, using an analytical model Bolton & Haehnelt (2007b) showed that the proximity zone size scales as $R_p \propto \dot{N}_\gamma^{1/2}$, if the IGM is in fact highly ionized by either the quasar itself or the UVB. To better understand this scaling, we turn to our radiative transfer simulations (see § 4), which were run for an ensemble of quasars that span the same range of luminosities as our data set. Note that the redshifts of the simulated quasars and the IGM are set to $z = 6$.

The simulated relation between proximity zone size and quasar magnitude for a highly ionized IGM is shown as the gray dashed line in Fig. 5, with the shaded region illustrating the 1σ scatter of simulated sizes due to cosmic variance alone (i.e. no measurement error). The relation from the simulations in a highly ionized IGM is reasonably well fit by the power-law

$$R_p \approx 5.57\,{\rm pMpc} \times 10^{-0.4 M_{1450}/2.35}. \quad (4)$$

Our simulations suggest that the proximity zone sizes in a highly ionized IGM scale as $R_p \propto \dot{N}_\gamma^{1/2.35}$, which lies in between the two theoretically expected relations for a neutral or mostly ionized ambient IGM, shown as the blue dashed-dotted and yellow dotted curves, respectively. Note that the normalization for these two analytic curves has been arbitrarily chosen such that all curves intersect at the same point. The simulated relation does not align exactly with the analytically expected relation in a highly ionized IGM, possibly due to heating effects of He II reionization.

Note however, that we obtain a very similar scaling from our radiative transfer simulations, when assuming a neutral ambient IGM:

$$R_p \approx 5.03\,{\rm pMpc} \times 10^{-0.4 M_{1450}/2.45}. \quad (5)$$

This is due to the fact that the measured proximity zone sizes R_p do not trace the extent of the ionization front $R_{\rm ion}$ described in eqn. (1) for an assumed quasar age of $t_{\rm Q} \sim 10^{7.5}$ yr, but end much earlier within the ionized H II region around the quasar. Because the IGM is indeed highly ionized in this region, one expects a scaling similar to that of a highly ionized IGM. Hence, it is no surprise to find the two relations from the radiative transfer simulations in a highly ionized ambient IGM as well as in an originally neutral ambient IGM to behave similarly. We will come back to this in more detail in § 5.3.

Given this new found intuition from radiative transfer simulations, it is now interesting to study the luminosity scaling of proximity zone sizes in our data. To this end, we show the dependence of our measured proximity zone sizes on the quasar's magnitude, which is proportional to the quasar luminosity, in Fig. 5. Measurements have been color coded by the emission redshift of the quasar. The errorbars of the measurements reflect the uncertainty on the proximity zone sizes due to the uncertainty of the quasar redshift (see § 2). We fit a power-law to our measurements and obtain

$$R_p \approx 4.71\,{\rm pMpc} \times 10^{-0.4\times(M_{1450}+27)/3.42}, \quad (6)$$

as the best fit, shown as the red dashed line. The 1σ uncertainty of this fit is given by the shaded region, which was determined by bootstrap re-sampling of our measurements with replacement, and repeating the fit 1000 times. Note that we do not weight our measurements by

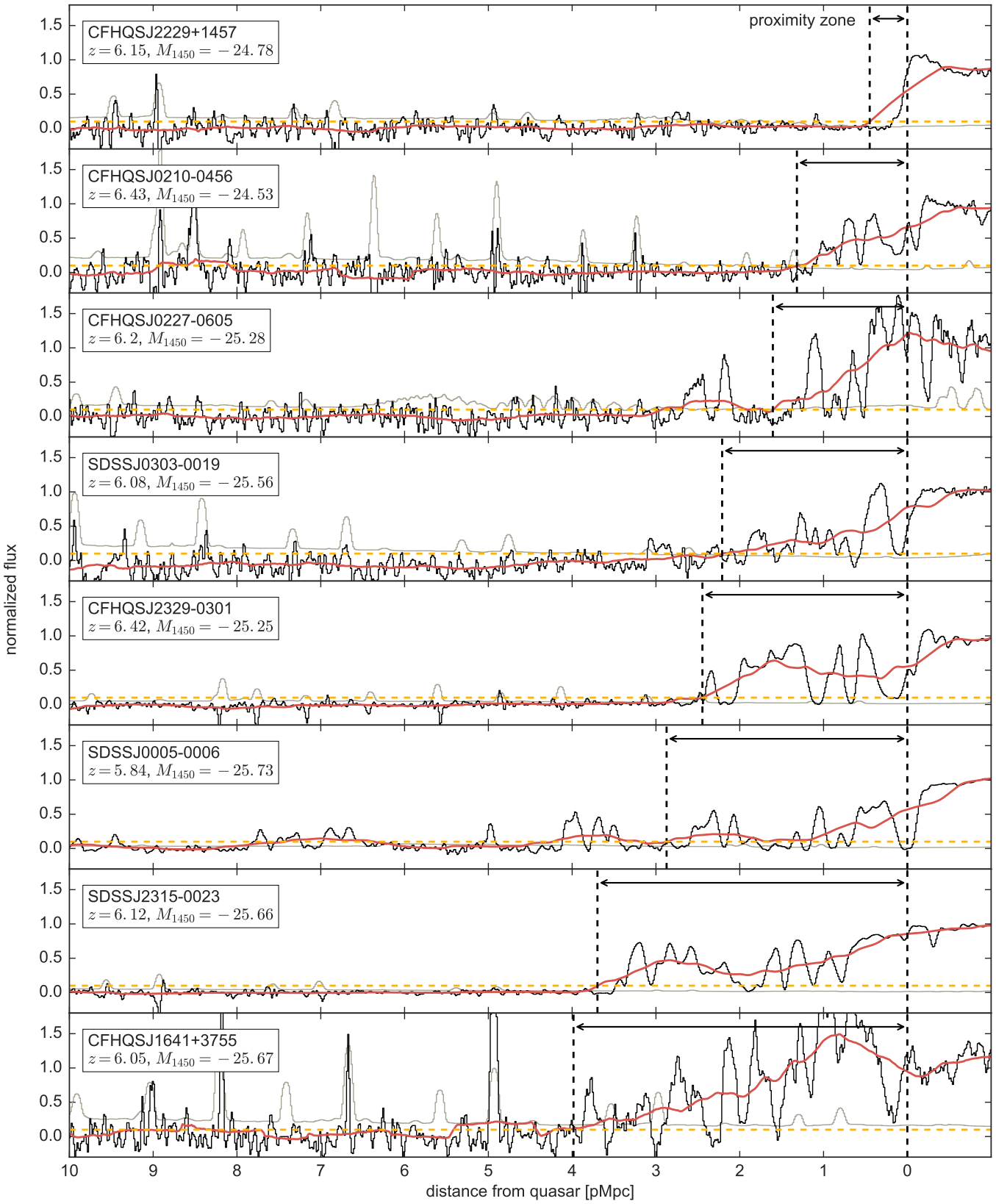

FIG. 3.— Same as Fig. 2, for a subset of faint quasars in our data set with luminosities between $-25.8 \leq M_{1450} \leq -24.5$.

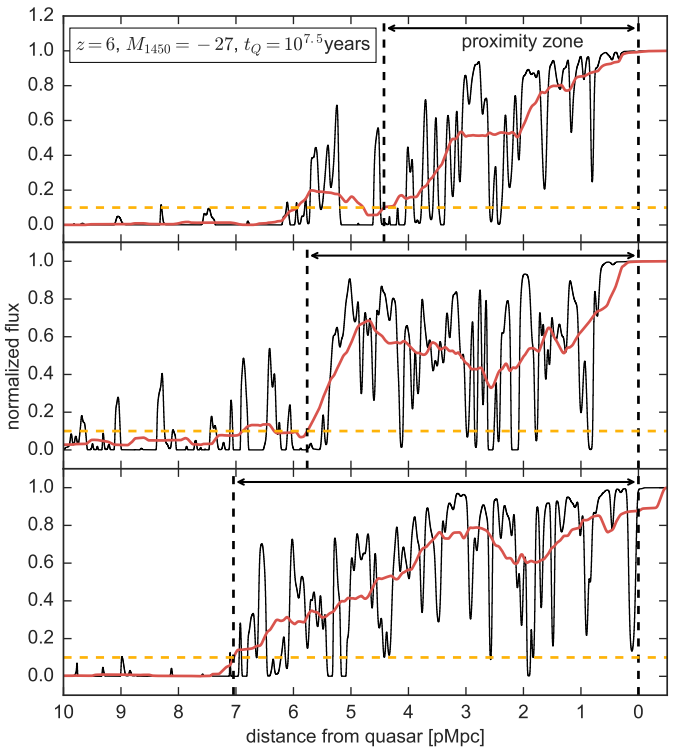

FIG. 4.— Mock spectra from three different skewers through the radiative transfer simulation box for a quasar at $z = 6$, a magnitude of $M_{1450} = -27$, and an age of $t_Q = 10^{7.5}$ yr, in a highly ionized IGM. The red curve shows the smoothed spectra and the black dashed lines indicate the extent of the proximity zones.

the redshift errors in this fit. In general, we find reasonable agreement between the scaling relations from the radiative transfer simulations and our data. The data seem to favor a shallower evolution, but this is complicated by the fact that we fit a sample over a large range of redshifts, and also, possibly, that exceptionally small proximity zones (see § 6) are pulling down the fit.

So which relation should we adopt to put the measured proximity zone sizes onto a common luminosity scale? Due to the large body of evidence for the IGM being mostly ionized at $z \sim 6$ (e.g. Wyithe & Bolton 2011; Calverley et al. 2011; Becker et al. 2015; McGreer et al. 2015) we choose the relation from our radiative transfer simulations assuming a highly ionized IGM (eqn. (4)) to eliminate the dependence on quasar luminosity and normalize our measurements to a common magnitude of $M_{1450} = -27$. However, our results are only marginally dependent on this choice, since the luminosity scaling for an ionized and neutral IGM are in fact very similar, and the main conclusions of this paper are not influenced by this choice. Thus we re-scale our measured proximity zone sizes with

$$R_{p,\mathrm{corr}} \approx R_p \cdot 10^{-0.4(-27-M_{1450})/2.35}. \quad (7)$$

Our measurements of the quasar proximity zone sizes are provided in Table 2, with the two rightmost columns showing the measured proximity zone R_p in proper Mpc and the corrected proximity zone size $R_{p,\mathrm{corr}}$, which are re-scaled to an absolute magnitude of $M_{1450} = -27$ following eqn. (7).

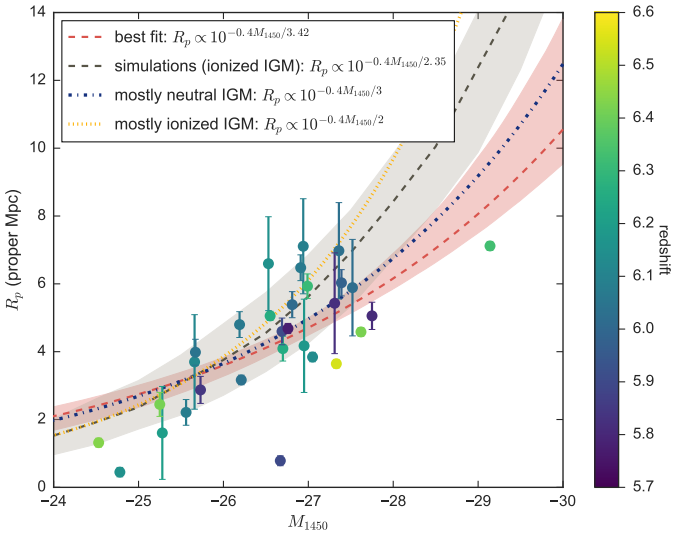

FIG. 5.— Sizes of the proximity zones shown dependent on the quasars' magnitude M_{1450}, color coded by their redshifts. The red dashed line shows the best power-law fit to the measurements with a 1σ-uncertainty level from bootstrapping errors. The gray dashed line shows the expected evolution of the proximity zones from radiative transfer simulations in a highly ionized IGM. The blue dashed-dotted and yellow dotted curves show the theoretical expectations when the IGM surrounding the quasars is neutral or highly ionized, respectively.

5.2. *The Redshift Evolution of Quasar Proximity Zone Sizes*

After correcting our measurements for the quasar luminosity, we can study the evolution of the proximity zones with redshift. In Fig. 6 we show this redshift evolution of our luminosity corrected measurements, color coded by M_{1450}, in both the left and middle panel. In the left panel, the red dashed line shows the best-fit power-law fit to the redshift evolution

$$R_{p,\mathrm{corr}} \approx 4.87\,\mathrm{pMpc} \times \left(\frac{1+z}{7}\right)^{-1.44}. \quad (8)$$

The shaded region indicates the 1σ uncertainty on the fit determined by bootstrap re-sampling of our measurements with replacement and repeating the fit 1000 times.

In the next section we use our radiative transfer simulations to better understand the redshift evolution of proximity zones, and show that the shallow redshift evolution that we measure is indeed expected. However, the black dashed and dash-dotted lines in the right panel of Fig. 6 are linear fits to the measurements of quasar proximity zones from previous analyses of similar data sets by Carilli et al. (2010) and Venemans et al. (2015), which show a much steeper trend. We compare our measurement to previous work in § 5.4, and discuss the sources of this discrepancy.

5.3. *Understanding the Shallow Redshift Evolution of Quasar Proximity Zone Sizes*

We use our radiative transfer simulations to investigate the expected redshift evolution of proximity zone sizes for both a highly ionized ambient IGM and a neutral IGM. Our results are shown in the middle panel of Fig. 6 as the gray dashed line (ionized) and blue dash-dotted line (neutral), respectively. In the case of an ionized

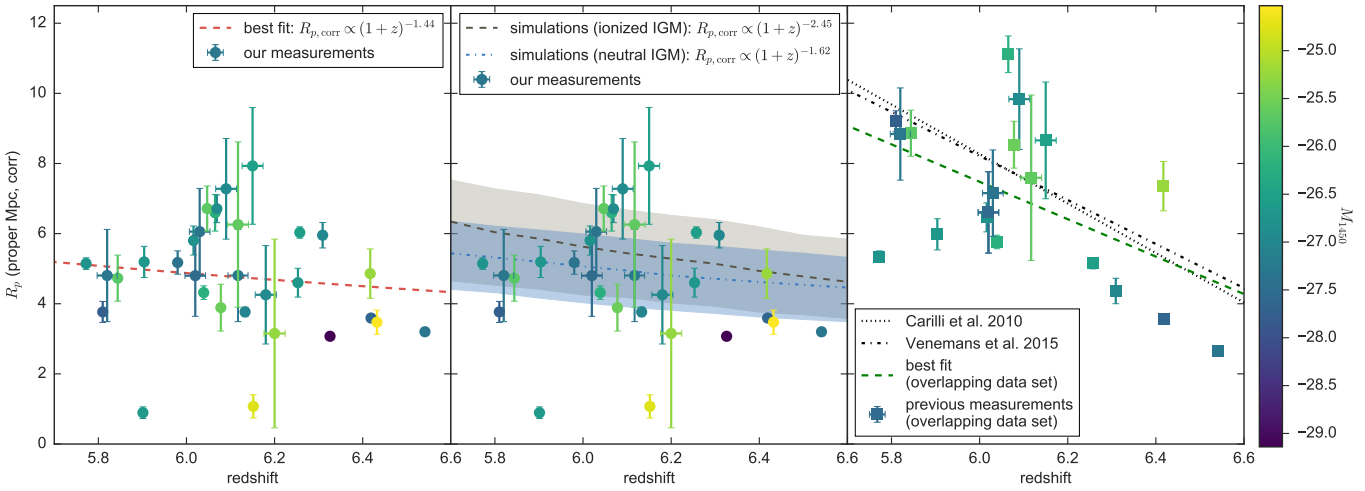

FIG. 6.— Evolution of the luminosity corrected proximity zone sizes with redshift, color coded by the quasars' actual magnitude M_{1450}. The data points showing our measurements for $R_{p,\mathrm{corr}}$ in both the right and middle panel are the same. In the left panel, the red dashed line shows the best power-law fit to the measurements with a 1σ-uncertainty level determined by bootstrapping. In the middle panel the gray dashed line shows the evolution of the proximity zones found in radiative transfer simulations when assuming a highly ionized IGM, whereas the blue dashed line is the result of the radiative transfer simulations assuming a mostly neutral ambient IGM. The shaded areas show the respective 1σ uncertainties of the relations due to cosmic variance. In the right panel the square data points show previous measurements by Carilli et al. (2010) and Venemans et al. (2015) of the quasars in common between our and their data sets, after correcting with updated redshift and magnitude estimates. The green dashed line shows the best fit to these measurements with a 1σ-uncertainty level determined by bootstrapping. The black dotted and dash-dotted curves show linear fits to the measurements of the whole data set from Carilli et al. (2010) and Venemans et al. (2015), respectively.

ambient IGM the corrected quasar proximity zones follow a power-law

$$R_{p,\mathrm{corr}} \approx 5.66\,\mathrm{pMpc} \times \left(\frac{1+z}{7}\right)^{-2.45}, \qquad (9)$$

whereas the evolution in the case of a neutral IGM is well fit by

$$R_{p,\mathrm{corr}} \approx 5.06\,\mathrm{pMpc} \times \left(\frac{1+z}{7}\right)^{-1.62}. \qquad (10)$$

The shaded regions around the curves show the scatter about these relations due to cosmic variance for quasars with magnitude $M_{1450} = -27$, but they slightly over- and underpredict the scatter for higher and lower luminosity quasars, respectively. The two scenarios of an ionized and a neutral IGM both result in a relatively shallow redshift evolution of proximity zones and are thus consistent with our measured redshift evolution shown in the left panel as the red dashed curve, but inconsistent with the steep evolution found in previous work (right panel).

Why should the redshift evolution of quasar proximity zones be so shallow? The end of the proximity zone around quasars is defined as the location where the smoothed transmitted flux drops below the 10% transmission level and a corresponding limiting optical depth of $\tau_{\mathrm{lim}} = 2.3$ is reached. Bolton & Haehnelt (2007b) determine in eqn. (8) of their paper the ionization rate Γ_{lim} that is necessary to produce a neutral gas fraction f_{lim} that results in the required limiting optical depth τ_{lim}. Assuming a highly ionized gas in ionization equilibrium and making reasonable assumptions about the gas temperature of the IGM at $z \sim 6$, one obtains an ionization rate of $\Gamma_{\mathrm{lim}} \sim 4 \times 10^{12}$ s^{-1} at the end of the proximity zone R_p, which is an order of magnitude larger than ionization rate of the UVB $\Gamma_{\mathrm{UVB}} \sim 2 \times 10^{-13}$ s^{-1} at this redshift (Wyithe & Bolton 2011). Thus the total ionization rate at the end of the proximity zone Γ_{lim}, which is the sum of the ionization rate of the background radiation Γ_{UVB} and the ionization rate of the quasar itself Γ_{QSO}, has to be totally dominated by the latter.

As such, in our simulations R_p is essentially independent of Γ_{UVB} and hence fairly insensitive to the neutral fraction f_{HI} of a highly ionized IGM, provided the quasar has been emitting light for longer than the equilibration timescale t_{eq} of the gas, which denotes the timescale on which the gas reaches ionization equilibrium (see § 6.3.3). The quasar age thus has to be $t_{\mathrm{Q}} \gtrsim t_{\mathrm{eq}} \sim 1/\Gamma_{\mathrm{QSO}}(R_p) \sim 10^5$ yr, in order for the ambient gas to have reached ionization equilibrium.

If the IGM surrounding the quasars is instead very neutral, i.e. $f_{\mathrm{HI}} \sim 0.1 - 1.0$, eqn. (1) for the location of the ionization front indicates that the size of the ionized H II region around the quasar scales as $R_{\mathrm{ion}} \propto (t_{\mathrm{Q}}/f_{\mathrm{HI}})^{1/3}$. This suggests that the proximity zones could be sensitive to the neutral gas fraction (subject to a degeneracy with the quasar age). Indeed for short ages the measured proximity zone R_p will trace the expanding ionization front R_{ion} around the quasar, i.e. $R_p \approx R_{\mathrm{ion}}$, which increases with the quasar age. However, R_p will cease to grow further once it reaches a distance given by the 10% transmission level of the flux, according to the definition of R_p. Thus even as R_{ion} continues to grow with age, the proximity zone size saturates, and will be insensitive to both the quasar age and neutral fraction. This implies that the measured proximity zone size provides a lower limit on the location of the ionization front R_{ion}. The maximum size of the proximity zone $R_{p,\mathrm{max}}$ is given by the distance at which a 10% flux transmission level corresponding to a limiting optical depth of $\tau_{\mathrm{lim}} = 2.3$ is reached. The age of the quasar at this distance is $t_{\mathrm{Q}} > t_{\mathrm{Q}}(R_{p,\mathrm{max}}) \sim 10^6$ yr under the assumption of a

homogeneous IGM (see eqn. (13) in Bolton & Haehnelt 2007b).

Thus for a highly plausible quasar age of $t_{\rm Q} \sim 10^{7.5}$ yr (see e.g. Martini 2004) the proximity zones R_p in a neutral IGM will look the same as for the highly ionized case, because the proximity zones by definition cease to grow once a flux transmission level of 10% is reached, as has been previously pointed out by (Bolton & Haehnelt 2007b). This demonstrates the insensitivity of R_p to the neutral fraction $f_{\rm HI}$ of the IGM for an appropriate choice of quasar age (see also Khrykin et al. 2016, for an analogous argument in the context of He II proximity zones).

Regarding the redshift evolution of the proximity zone sizes, eqn. (11) in Bolton & Haehnelt (2007b) predicts a very shallow scaling of $R_p \propto (1+z)^{-2.25}$ due to the density evolution in the universe. The simulations in Fig. 6 reveal a similar scaling, i.e. $R_p \propto (1+z)^{-2.45}$ for a neutral ambient IGM and $R_p \propto (1+z)^{-1.62}$ for a highly ionized IGM with a quasar age of $t_{\rm Q} \sim 10^{7.5}$ yr. The true scaling does not match exactly the analytically expected scaling due to different heating effects.

Hence we conclude that the redshift evolution of R_p is not a very useful probe of the ionization state of the IGM. In a highly ionized IGM due to the definition of proximity zones, the ionization rate at the end of the proximity zone is totally dominated by the ionization rate of the quasar itself and thus essentially independent of the background radiation and the ionization state of the surrounding IGM. In a neutral IGM, R_p ceases to grow once a maximum size $R_{p,\rm max}$ corresponding to a limiting optical depth $\tau_{\rm lim}$ at the 10% flux transmission level is reached. The observed shallow redshift evolution is thus perfectly consistent with our models for both a highly ionized and a neutral IGM for a quasar age of $t_{\rm Q} > 10^6$ yr, in particular for a fiducial quasar age of $t_{\rm Q} \sim 10^{7.5}$ yr. None of our simulations reveal the steep evolution that was measured by previous analyses (Carilli et al. 2010; Venemans et al. 2015). Our models do not predict any large proximity zones of $R_p \gtrsim 10$ pMpc that would be consistent with their best-fit curve at $z \approx 5.8$. We will now investigate possible reasons leading to the discrepancy between ours and previous analyses.

5.4. *Comparison to Previous Analyses*

There are several possible reasons causing the differences between ours and previous analyses that result in the different measured redshift evolution of quasar proximity zones: we use a different scaling with luminosity applied to correct all proximity zone measurements and put them on the same luminosity scale, different quasar spectra[15], different approaches to fitting continua and measuring the proximity zone sizes, and our measurements also rely on updated luminosities and redshift measurements. Below we discuss each of these differences and their impact on our results in turn.

To address the issue of different luminosity scalings we facilitate a comparison to previous work by Carilli et al. (2010) and Venemans et al. (2015) by taking their measurements of R_p and correcting them with our luminosity correction ($R_{p,\rm corr} \propto \dot{N}_\gamma^{1/2.35}$ instead of their $R_{p,\rm corr} \propto \dot{N}_\gamma^{1/3}$; see eqn. (7)) using updated measurements of M_{1450} from Bañados et al. (2016) for their objects. In their work they chose a linear relation between proximity zone size and redshift, i.e. $R_p \propto z$, and thus we also show a linear fit to their corrected measurements in the right panel of Fig. 6 in order to simplify the comparison. Our best fit slope to the measurements of Carilli et al. (2010) is $m \approx -7.07$ (black dotted line) and $m \approx= -6.28$ (black dash-dotted line) for measurements by Venemans et al. (2015), which is considerably steeper than the slope $m \approx -1.12$ that we obtain when applying a linear fit to our data. Thus we conclude that the difference in luminosity scaling does not cause the discrepancy between the different analyses.

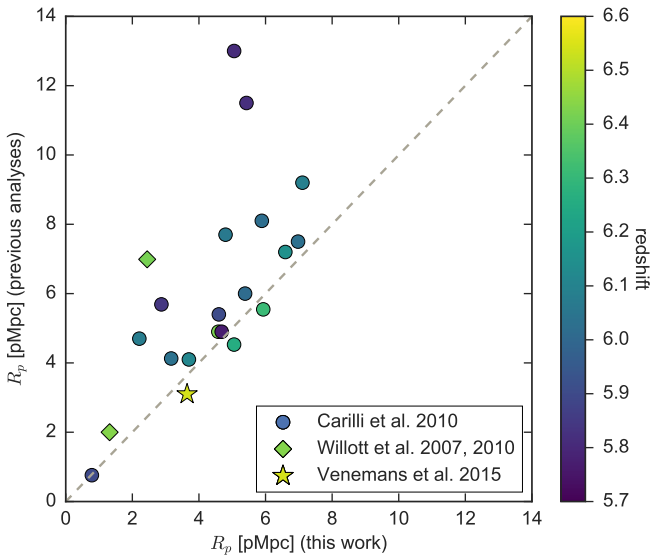

FIG. 7.— Comparison of previous measurements of quasar proximity zones R_p to the analysis presented in this paper for the overlapping objects in the various data sets. The gray dashed line shows the exact one-to-one relation. In general we recover smaller proximity zones than previous analyses, but mostly consistent within $\lesssim 2$ pMpc.

However, both previous analyses contain objects in their data sample that are not included in our sample. Although the data sets from (Carilli et al. 2010) and Venemans et al. (2015) partially overlap with our sample, they also contain distinct objects. In order to verify that differences in the slope of the redshift evolution are not being driven by differences in the quasar samples, we restrict the analysis now to measurements of the 18 quasars that are common to our sample and those of Carilli et al. (2010) and Venemans et al. (2015). For five of these quasars, our analysis uses updated redshift measurements and thus we correct the measurements of R_p from previous work for the difference between the old and the new redshift measurements in order to have consistent redshifts and afterwards use eqn. (7) to obtain $R_{p,\rm corr}$ from their measurements of R_p. Their corrected measurements of the objects overlapping in both data sets are shown in the right panel of Fig. 6 and the best linear fit for the redshift evolution is shown as the green dashed curve. It still shows a relatively steep slope of $m \approx -5.33$ although slightly shallower than before.

[15] Although many of the objects are overlapping between ours and previous data sets, we co-added different exposures from different observing runs and have thus gained higher quality spectra.

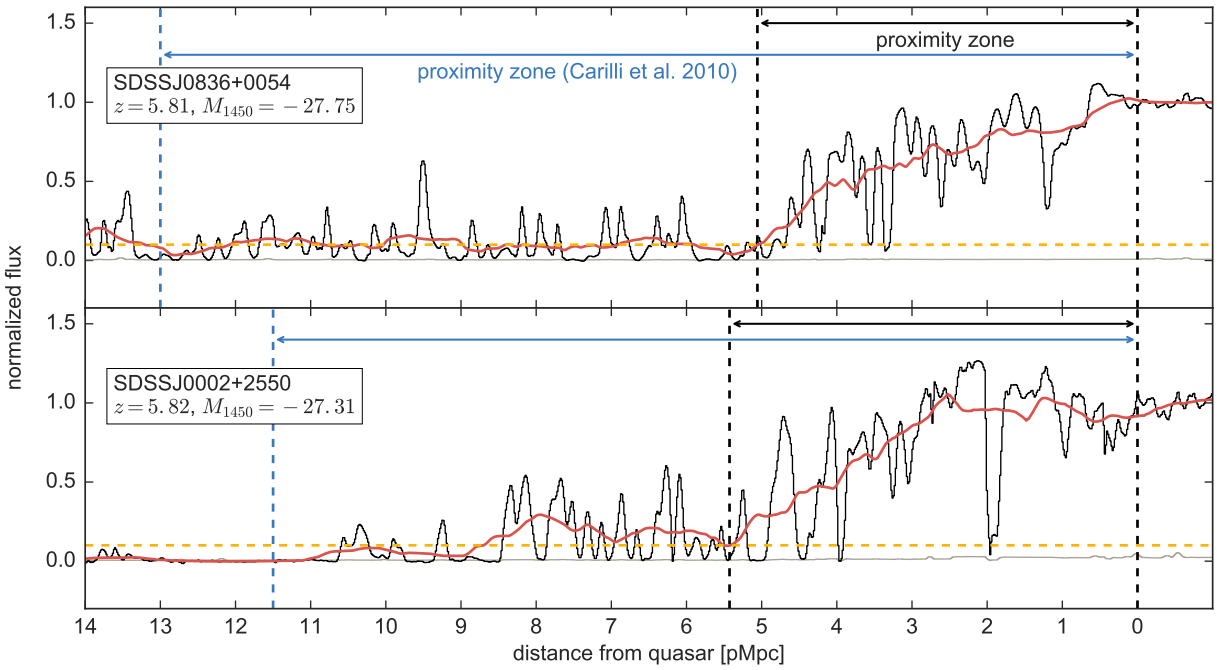

FIG. 8.— Quasar spectra from our data set with the largest discrepancy between our measurements of the proximity zones and the measurements by Carilli et al. (2010).

Thus we conclude that differences in the data sets are not driving the discrepancy in the results.

Note however, that the differences in the resulting redshift evolution of R_p appear to be driven by a handful of objects, where we measure significantly different proximity zone sizes than previous work. In Fig. 7 we compare our measurements of the (uncorrected) R_p to those measured by Willott et al. (2007), Willott et al. (2010), Carilli et al. (2010) (who updated the measurements originally performed by Fan et al. (2006b)), and Venemans et al. (2015) for all objects that are overlapping between our data sample and these previous analyses. We have adopted consistent redshifts for all objects here in order to facilitate a one-to-one comparison that is not driven by changes in redshift, i.e. for the objects for which we have updated redshift measurements, we correct the previously analyzed proximity zones for the difference.

The measurements which lie along the gray dashed line indicate agreement between the different analyses. In general, we measure smaller sizes of the proximity zones than previously obtained, but for most objects the measurements agree within $\Delta R_p \lesssim 2$ pMpc. Despite the fact the fact that for a significant fraction of these overlapping quasars the analyzed data comes from the same instrument (Keck/ESI), small differences in the measurements can be attributed to different data reduction pipelines and in many cases we have co-added data from different runs to get higher S/N spectra and thus the final quasar spectra might differ in quality. However, there are a few outliers in Fig. 7, for which our measurements differ significantly from previous work, i.e. $\Delta R_p > 2$ pMpc, which we discuss in further detail.

We measure much smaller proximity zones for three quasars in particular: SDSSJ0836 + 0054 ($R_{p,\mathrm{Carilli}} \approx 13.0$ pMpc vs. our measurement $R_p = 5.06 \pm 0.40$ pMpc) at $z = 5.810$, SDSSJ0002 + 2550, ($R_{p,\mathrm{Carilli}} \approx 11.5$ pMpc vs. our measurement $R_p = 5.43 \pm 1.49$ pMpc) at $z = 5.82$, and CFHQSJ2329 − 0301, ($R_{p,\mathrm{Willott}} \approx 7.0$ pMpc vs. our measurement $R_p = 2.45 \pm 0.35$ pMpc) at $z = 6.417$.

The spectra of the first two objects, SDSSJ0836 + 0054 and SDSSJ0002 + 2550 and our measured proximity zones (black dashed lines) as well as the measurements from Carilli et al. (2010) (blue dashed lines) are shown in Fig. 8. Both analyses use the same redshift measurements and thus no differences due to redshift errors should be causing the discrepancy we see here. The spectrum of SDSSJ0836 + 0054 in the upper panel shows transmitted flux throughout the spectrum with the smoothed flux oscillating around the 10% level. Our spectrum nevertheless shows a significant drop below the 10% level at $R_p \approx 5.06$ pMpc. This measured R_p does not change significantly when fitting the continuum model with a different set of PCA components.

For SDSSJ0002 + 2550 (lower panel) the size of the proximity zone is more sensitive to the precise placement of the quasar continuum level. A slightly different continuum normalization (i.e. taking a different set of PCA components) would increase the measurement of the proximity zone from $R_p \approx 5.43$ pMpc to $R_p \approx 8.85$ pMpc, but still falls considerably short of the $R_p = 11.5$ pMpc previous measurement by Carilli et al. (2010). Another reason for the discrepancy in the R_p measurements could be the higher signal-to-noise data in our sample. However, it remains unclear whether the differences in the continuum normalization and the better quality data could cause the whole discrepancy of the two measurements.

The third object that has a significantly different proximity zone measurement is CFHQSJ2329 − 0301, whose proximity zone was determined by Willott et al. (2007) to be $R_p \approx 6.3$ pMpc (after updating their measurements with a new redshift measurement it is now

$R_p \approx 7.0$ pMpc). However, most of this discrepancy between their measurement and ours can be contributed to the fact that for this particular object, Willott et al. (2007) decided after inspecting the spectrum by eye that they take the *second* drop of the smoothed flux below 10% level as the end of the proximity zone instead of the *first* drop according to the standard definition of proximity zones (Fan et al. 2006b). They determine the first drop in flux below the 10% transmission level to be at $R_p \approx 3.7$ pMpc (with an updated redshift measurements this would be $R_p \approx 4.4$ pMpc), which would be much more consistent with our measurement of $R_p = 2.45 \pm 0.35$ pMpc.

Another reason for the discrepancy in the redshift evolution of the proximity zone measurements is caused by the exclusion of the quasar SDSSJ1335 + 3533 at $z = 5.9012$ from previous analyses. For this particular object, which has a very small proximity zone of $R_p = 0.78 \pm 0.15$ pMpc, our measured size agrees with the one measured by Carilli et al. (2010). However, this object is a weak emission line quasar and has been somewhat arbitrarily excluded from previous analyses due to this fact (see Carilli et al. 2010; Wyithe & Bolton 2011). The reason given in the aforementioned papers for is the "fundamentally different nature" of such objects. However Diamond-Stanic et al. (2009) showed that weak emission line quasars do not show significantly different UV continuum slopes apart from the emission lines and thus do not differ in their physical properties. The original reason for Fan et al. (2006b) to exclude this object from their analysis was that the only redshift measurement they had at that time came from the onset of strong Lyα absorption. Given that we now have an accurate redshift measurement, we include SDSSJ1335+3533 into our analysis, which causes the slope of the redshift evolution of the proximity zones to become shallower due to its small zone. We will further discuss the implications of small proximity zones in § 6.

Further differences between the different analyses can be attributed to the different continuum fitting methods. Previous analyses applied a power-law fit to the quasar continuum redwards of the Lyα emission line and fitted the Lyα and N V lines with Gaussian curves, whereas we chose to model the continuum of each quasar with two different sets of PCA (i.e. Pâris et al. (2011) or Suzuki (2006)). Although differences in the continuum estimation can change the size of individual proximity zones, we do not expect them to alter the distribution of proximity zone sizes, since the scatter on R_p resulting from continuum uncertainties is much smaller than the intrinsic scatter due to density fluctuations (Kramer & Haiman 2009).

The continuum uncertainties in our analysis arising due to differences between the two sets of PCA are not very significant. The median difference between the proximity zone sizes when estimating the quasar continua with different PCA models is $\langle \Delta R_p \rangle = \langle |R_{p,\mathrm{Paris}} - R_{p,\mathrm{Suzuki}}| \rangle \approx 0.09$ pMpc. The scatter in the distribution of ΔR_p determined from the 16th and 84th percentile is $\sigma_{\Delta R_p} \approx 0.73$ pMpc.

To conclude, it is not completely clear to us, how other authors obtained some very large proximity zones, particularly for a few objects at $z < 6$ that are the main drivers for the steep redshift evolution. Some of the differences between ours and previous analyses and the resulting shallower redshift evolution of the proximity zone sizes can be attributed to a variety of reasons: first, we stick to a rigorous definition of the proximity zone sizes and continue to take the first drop of the smoothed flux below the 10% level, since we adopt the same rigorous treatment for the mock spectra from the radiative transfer simulations. Second, we analyze higher quality data, which can help to determine the drop in flux below the 10% level easier. Third, we include all quasars of the ensemble in our analysis and do not exclude the weak emission line quasar in our sample. Finally, we analyze a larger sample of quasar spectra than previous analyses. Differences in the continuum estimation, updated redshift and magnitude measurements, and the different correction for the quasar luminosity do not have a significant influence on the discrepancy.

6. EXCEPTIONALLY SMALL QUASAR PROXIMITY ZONES

Several of the quasars we studied have particularly small proximity zones, as can be seen from Fig. 6. The proximity zones of two objects, i.e. CFHQSJ2229 + 1457 and SDSSJ1335 + 3533, are $R_p < 1$ pMpc and $R_{p,\mathrm{corr}} \lesssim 1$ pMpc. Additionally, given its extreme brightness, the proximity zone of SDSSJ0100 + 2802, $R_p = 7.12 \pm 0.13$ pMpc and $R_{p,\mathrm{corr}} = 3.09 \pm 0.06$ pMpc, is also exceptionally small. In this section, we discuss the properties of these objects and possible explanations for their small proximity zones.

6.1. *Individual Objects with Small Proximity Zones*

CFHQS J2229+1457

This relatively faint quasar, $M_{1450} = -24.78$ (Bañados et al. 2016), was discovered in the Canada-France High-z Quasar Survey (CFHQS) and published by Willott et al. (2009). The measurements of the proximity zone are $R_p = 0.45 \pm 0.14$ pMpc and, when normalized to a magnitude of $M_{1450} = -27$, $R_{p,\mathrm{corr}} = 1.07 \pm 0.33$ pMpc. The quasar has a precise redshift measurement of $z = 6.1517$ from [C II] emission from the host galaxy (Willott et al. 2015) and hence a very small redshift uncertainty $\Delta v = 100$ km/s resulting in an uncertainty of the proximity measurement of $\Delta R_p \approx 0.14$ pMpc. The measurements of the proximity zone are independent of the choice of PCS used to model the quasar continuum, i.e. continuum uncertainties do not influence our measurements.The top panel of Fig. 3 shows the continuum normalized spectrum and the proximity zone of this object. In the upper panel of Fig. 9 we show the whole spectrum, revealing no BAL features.

SDSS J1335+3533

This quasar, that has been discovered in the Sloan Digital Sky Survey (SDSS) by Fan et al. (2004), has a precise redshift measurement of $z = 5.9012$ from CO $(6-5)$ emission from the host galaxy (Wang et al. 2010), and is fairly bright with a magnitude of $M_{1450} = -26.67$ (Bañados et al. 2016). The measured proximity zone size for this object is $R_p = 0.78 \pm 0.15$ pMpc and the luminosity corrected size of the proximity zone is $R_{p,\mathrm{corr}} = 0.89 \pm 0.17$ pMpc. The redshift uncertainty of

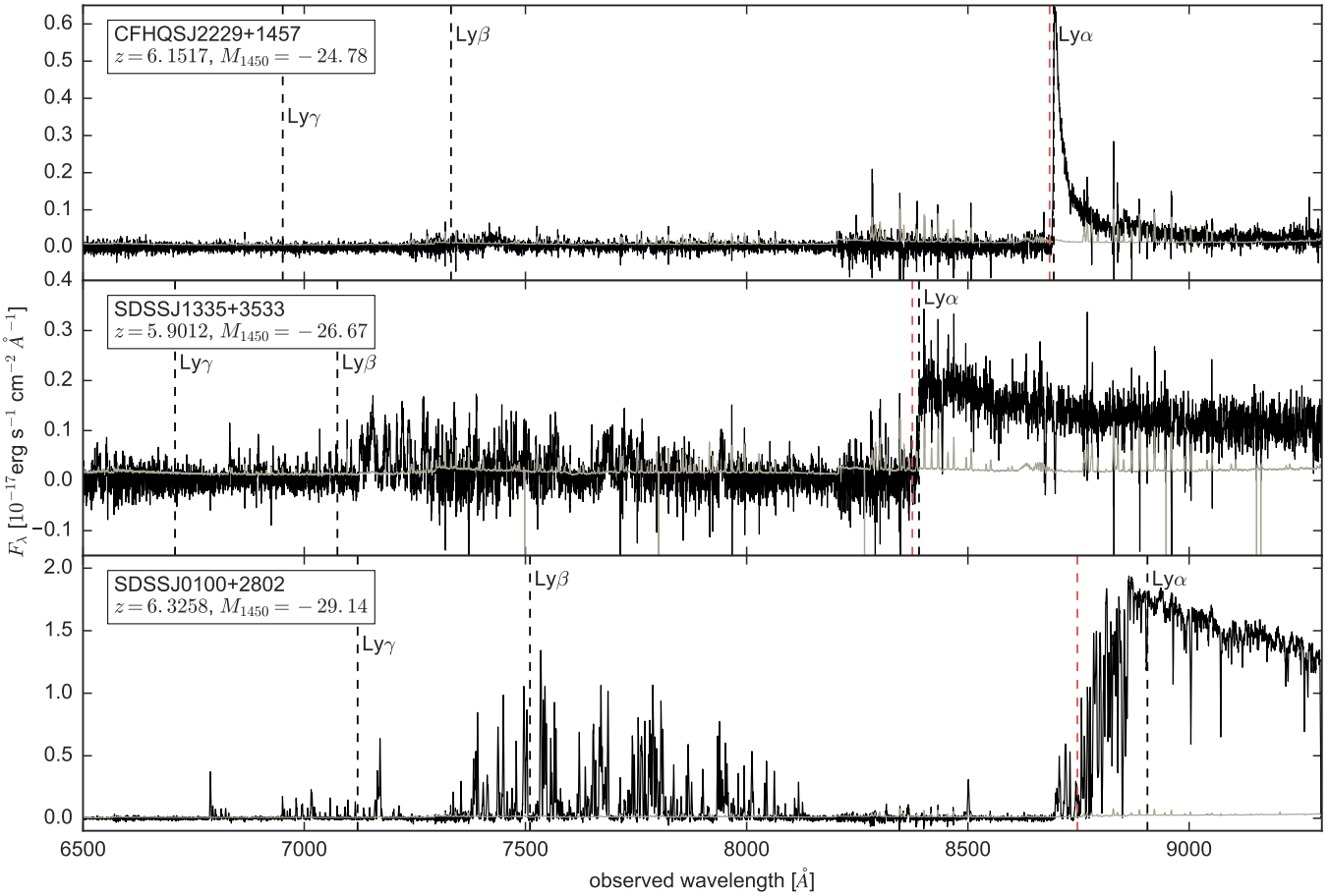

FIG. 9.— Spectra of the three quasars CFHQSJ2229 + 1457 (upper panel), SDSSJ1335 + 3533 (middle panel) and SDSSJ0100 + 2802 (lower panel), exhibiting exceptionally small proximity zones. Lyα, Lyβ and Lyγ lines are indicated by the black dashed lines, whereas the red dashed lines show the extent of the measured proximity zones.

$\Delta v = 100$ km/s results in an uncertainty of the proximity zone measurement of $\Delta R_p \approx 0.15$ pMpc. These measurements of the proximity zone are again independent on continuum modeling uncertainties. The top panel of Fig. 2 shows the continuum normalized spectrum and the proximity zone of this object. The whole spectrum of the quasar shown in the middle panel of Fig. 9 is completely devoid of broad emission lines, as previously noticed from its discovery spectrum by Fan et al. (2004). The spectrum does not show any BAL features.

SDSS J0100+2802

This object is the brightest high redshift quasar known so far (Wu et al. 2015) with an absolute magnitude of $M_{1450} = -29.14$ (Bañados et al. 2016). Its redshift, $z = 6.3258$, has been measured precisely by the detection of the [C II] line (Wang et al. 2016). Given its extreme brightness the quasar exhibits a very small proximity zone of $R_p = 7.12 \pm 0.13$ pMpc ($R_{p,\mathrm{corr}} = 3.09 \pm 0.06$ pMpc, when normalized to a luminosity of $M_{1450} = -27$) compared to a proximity zone size of $R_p = 12.0 \pm 2.0$ pMpc, that one would expect to see for an object this bright. The redshift uncertainty $\Delta v = 100$ km/s results in an uncertainty of the proximity zone measurement of $\Delta R_p \approx 0.13$ pMpc. Continuum uncertainties do not play any significant role for these measurements. The lower panel of Fig. 9 shows the spectrum of this object, which does not show any BAL features. The continuum normalized spectrum and its proximity zone is shown in the second to last panel of Fig. 16 in Appendix C.

6.2. *How Common are Small Proximity Zones?*

In order to quantify the probability of finding objects with such small proximity zones, we compare in Fig. 10 our measured proximity zones R_p from a subset of quasar spectra from our data sample to proximity zones from our radiative transfer simulations. We show the distribution of quasar proximity zone sizes for three different luminosity ranges, quasars with magnitudes $-25.8 \leq M_{1450} \leq -24.5$ are shown in the left panel, brighter objects with magnitudes $-27.5 \leq M_{1450} \leq -26.5$ are shown on the middle panel, and the right panel shows the distribution of proximity zones for very bright quasars with $M_{1450} = -29.14$. The magnitude intervals were chosen, such that each bin has a width of ~ 1 dex and includes one of the small proximity zone objects. This results in six quasars in the faintest magnitude range, eleven quasars in the brighter range of magnitudes and only one object at $M_{1450} \approx -29$ (Wu et al. 2015). Because of this split into magnitude intervals, we are comparing R_p instead of $R_{p,\mathrm{corr}}$, and model the ranges of quasar luminosities in the simulations. Note that the redshift range covered by the quasars in each bin is very broad,

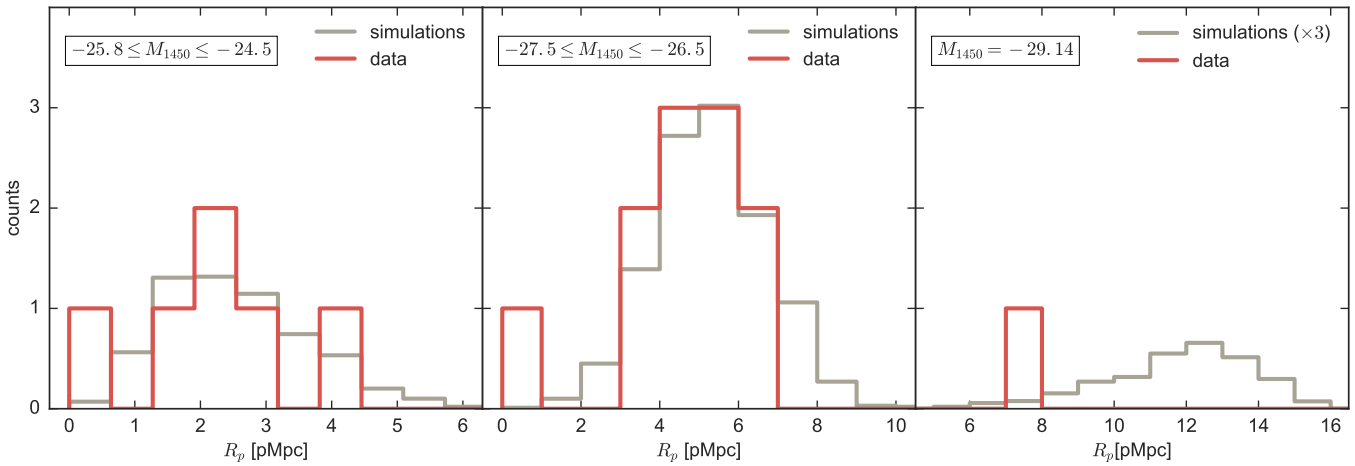

FIG. 10.— Distribution of sizes of quasar proximity zones for faint quasars with luminosities between $-25.8 \leq M_{1450} \leq -24.5$ (left panel), brighter quasars with luminosities between $-27.5 \leq M_{1450} \leq -26.5$ (middle panel) and for very bright quasars with luminosity $M_{1450} \approx -29.14$ (right panel). The red histogram shows the distribution of the measured proximity zone sizes of quasars in our data sample within the given luminosity range that have redshift measurements from Mg II, [C II] or CO lines. The gray histogram shows the expected distribution for quasars with the same redshifts and luminosity properties as in the data sample from the radiative transfer simulations. In the right panel the simulated distribution is multiplied by a factor of three for better visibility. Our data sample includes six fainter quasars with these requirements, eleven brighter ones and one very bright one.

i.e. $5.844 \leq z \leq 6.4323$ for the faintest magnitude bin, $5.7722 \leq z \leq 6.5412$ for the brighter magnitude range, and the only one bright quasar has a redshift $z \approx 6.3258$. We forward model the redshift evolution of the quasars in the radiative transfer simulations, such that the broad range of luminosities in each bin does not influence our results.

The red histograms show the distribution of measured proximity zones from our data set. In this figure we only consider quasars that have their redshifts determined by the detection of a Mg II, [C II] or CO line, i.e. with small redshift uncertainties. These histograms have been normalized to the number of objects available in our data set within the given magnitude interval.

The gray histograms show the distribution of proximity zone sizes expected from our radiative transfer runs. We simulate multiple realizations of each quasar in our sample matching the redshift, luminosity, and redshift uncertainties (which we add as a Gaussian uncertainty to the quasar position in our mock spectra; see Fig. 4). For the two leftmost panels of Fig. 10 we simulate 100 skewers for each quasar in our data sample. Thus the gray histograms show the proximity zones of 600 skewers in the left panel and 1100 skewers in the middle panel. For the single very bright quasar in the right panel of Fig. 10 we simulate 900 realizations to improve the statistics. We do not model continuum uncertainties in the simulations, which we expect to have negligible influence on the distribution of proximity zone sizes (see discussion in § 5.4, Kramer & Haiman 2009). For the radiative transfer runs we assume a highly ionized IGM, which is consistent with empirical constraints on the neutral gas fraction at $z \sim 6$ (Calverley et al. 2011; Wyithe & Bolton 2011; Becker et al. 2015; McGreer et al. 2015). Following the arguments of § 5.3 we do not expect any qualitative changes when assuming a mostly neutral IGM.

In both the faint and the brighter case (left and middle panel) the bulk of the distributions of measured and simulated proximity zones agrees very well. Note that our simulations do not reproduce the very large proximity zones of $R_p \sim 10 - 13$ pMpc that Carilli et al. (2010) measured previously.

However, the simulations also do a poor job of reproducing the frequency of very small proximity zones that we find. For the fainter quasars (left panel) our simulations indicate that the probability of finding a quasar with $R_p \approx 0.45$ pMpc, which is our measurement for CFHQSJ2229 + 1457, in a sample of six quasar spectra is $\approx 3\%$. However, the significance of this small proximity zone is limited by the definition of the proximity zone. The smoothed flux in the spectrum of CFHQSJ2229 + 1457 remains below the $< 10\%$ level within a distance of 10 pMpc to the quasar, whereas a by eye inspection of the simulated small proximity zones reveals that in most spectra the flux increases above 10% again just outside of their proximity zone. If we would adopt a different definition of the proximity zone that would also be sensitive to the length of the GP trough outside of the proximity zone, this object would be an even greater outlier.

In the middle panel showing quasars with magnitudes $-27.5 \leq M_{1450} \leq -26.5$ we have one object, SDSSJ1335 + 3533, with a measured proximity zone $R_p \approx 0.78$ pMpc. In the sample of 1100 simulated quasars with similar redshifts and magnitudes only one of them has a proximity zone size that small. Thus the probability of finding such an object in a sample of eleven quasars is $\approx 1\%$.

In the right panel we have only one object in our data set, SDSSJ0100 + 2802. Its proximity zone is relatively small given its extreme brightness. We estimate the probability of finding a quasar with a proximity zone of $R_p \lesssim 7.12$ pMpc to be $\approx 3\%$.

In summary, in all three respective magnitude ranges, the occurrence of small proximity zones is much higher in our data sample than the simulations predict. We will now investigate several possible scenarios that could explain the exceptionally small proximity zones of the three aforementioned objects.

6.3. *Possible Explanations for Small Proximity Zones*

We address three scenarios that could possibly explain the exceptionally small proximity zones of the three quasars mentioned in the previous subsection: The proximity zones could be prematurely truncated due to associated dense absorbers, such as damped Lyα systems (DLA) or Lyman limit systems (LLS), patches of remaining neutral hydrogen within the IGM could truncate the proximity zones, or the quasars could be very young.

6.3.1. *Truncation Due to Damped Lyα Systems or Lyman Limit Systems*

One possible explanation for the exceptionally small proximity zone sizes could be the truncation of the zones due to strong absorbers, such as DLAs or LLSs, which could either be intervening or physically associated with the quasar environment (Bañados et al. in prep.). The presence of such self-shielding absorbers could prematurely truncate the proximity zone at small radii by blocking the quasar's ionizing flux. These optically thick absorption line systems are not correctly modeled by our radiative transfer simulations (see § 4). We model only the gas densities in the IGM, but do not include such dense gas patches that could be coming from galaxy like overdensities resulting in these absorbers. Thus, proximity zones truncated prematurely by dense absorption systems are not reproduced in our simulations and need to be identified and eliminated from any comparison between the data and the simulations.

To this end, we search for signatures of strong absorption line systems in the quasar spectra near the end of their proximity zones by visually inspecting the spectra, and searching for evidence of damping wings which would indicate the presence of a strong absorber, as well as associated ionic metal-line transitions.

In the spectrum of SDSSJ1335 + 3533, the only strong metal line absorber we could find was a Mg II absorber at $z \approx 2.10$. However, we have to note that the signal-to-noise ratio of the spectrum is insufficient to identify very weak metal absorption features.

Searching the ESI spectrum of CFHQSJ2229 + 1457 for metal absorption lines associated with nearby dense absorption systems did not reveal any absorbers due to the very low signal-to-noise data (S/N $\approx$ 2 per pixel). To facilitate the search for metal lines, we obtained a higher quality spectrum with the Low Resolution Imaging Spectrometer (LRIS) at the Keck I telescope, which has a lower resolution ($R \sim 1800$) than the ESI spectrum ($R \sim 4000$) but a higher signal-to-noise ratio (S/N $\approx$ 7 per pixel). An excerpt of this spectrum is shown in Fig. 11 and reveals an N V doublet associated with an absorber at $z \approx 6.136$, whose H I absorption Voigt profile is shown as the red curve when assuming a column density of $N_{\rm HI} = 10^{19}$ cm^{-2} and a Doppler parameter of $b = 40$ km/s. The absorber is located at a distance of 0.91 pMpc from the quasar and thus it is highly unlikely that this absorber influences the size of the proximity zone significantly, because the proximity zone ends at a higher redshift of $z \approx 6.144$ at a distance $R_p \approx 0.45$ pMpc. Additionally, at the location of the absorber in the spectrum, there is a transmitted flux spike in the Lyα forest partly visible (at $\lambda \approx 8670$ Å), which indicates that the absorber cannot be saturated, implying $N_{\rm HI} \lesssim 10^{14}$ cm^{-2}, i.e. its column density would need to be much less than the here assumed $N_{\rm HI} = 10^{19}$ cm^{-2} in order not to violate the spectrum, and is thus unlikely to truncate the proximity zone. We also do not find evidence for any low-ion absorption lines, such as Si II at $\lambda_{\rm rest} = 1260.42$ Å, that should be present in DLAs.

In the spectrum of SDSSJ0100+2802 we find a low ionization absorption system close to the object at $z \approx 6.144$ (see Fig. 12), which has previously been identified by Wu et al. (2015). However, this absorber is at such large line-of-sight distance $R \approx 9.94$ pMpc from the quasar that it cannot truncate its proximity zone, that already ends at a distance of $R_p = 7.12 \pm 0.13$ pMpc. We show the H I Voigt absorption profile (red curve) assuming an absorber with $N_{\rm HI} = 10^{19}$ cm^{-2} and $b = 40$ km/s, and its associated lines (Si II at $\lambda_{\rm rest} = 1260.42$ Å and $\lambda_{\rm rest} = 1304.37$ Å, O I at $\lambda_{\rm rest} = 1302.16$ Å, [C II] at $\lambda_{\rm rest} = 1334.53$ Å). Note that an absorption system with a higher column density of $N_{\rm HI} \sim 10^{20.3}$ cm^{-2} characteristic of DLAs, would be too broad and would be inconsistent with the presence of flux transmission that are observed in the Lyβ forest at this redshift (see § 6.3.2).

In light of the possibility that proximate absorbers could truncate the proximity zones, it is interesting to estimate the number of those proximate absorbers that one would expect to find around quasars. Therefore, we start with an estimate of intervening absorbers along the line-of-sight to quasars by Songaila & Cowie (2010), who have estimated the number density $\mathrm{d}N/\mathrm{d}z$ of intervening LLSs in Fig. 5 and Table 3 of their paper. They determine the best power-law fit to their measurements with a maximum likelihood analysis to be

$$\frac{\mathrm{d}N}{\mathrm{d}z} \approx 2.84 \left(\frac{1+z}{4.5}\right)^{2.04}, \qquad (11)$$

which results in an estimate of $\mathrm{d}N/\mathrm{d}z \approx 6.99$ at $z = 6$.

On the one hand the abundance of LLSs could be enhanced in the proximity of quasars because of the overdense quasar environment (Hennawi et al. 2006; Prochaska et al. 2008, 2013). However, the intense radiation from the quasar will also ionize the dense gas in its surroundings making it less likely to self-shield, possibly lowering the abundance of so-called proximate LLSs (Hennawi & Prochaska 2007; Prochaska et al. 2008). Prochaska et al. (2010) found empirically that at $z \sim 4$ the number density of proximate LLSs (occurrence within $\Delta v \leq 3000$ km s^{-1} of the quasar emission redshift) roughly equals the number density of intervening LLSs (see Fig. 15 of Prochaska et al. 2010). Assuming the same approximate equality holds at $z \sim 6$, we estimate the probability p of finding a proximate LLS within 1 pMpc to a quasar at $z = 6$ to be $p \approx n(z)\Delta z \approx 0.11$. Given the size of our data set, which consists out of 30 quasars, we would expect to find a proximate LLS in ~ 3 spectra.

Where are the ~ 3 proximate LLSs that we expect to find in our data sample? We found only one in the spectrum of the quasar SDSSJ0840+5624, which we excluded from our analysis for this reason (see Appendix A). The absence of other proximate LLSs detected in our data sample could be attributed to a few different reasons: the estimate of proximate LLSs by Prochaska et al. (2010)

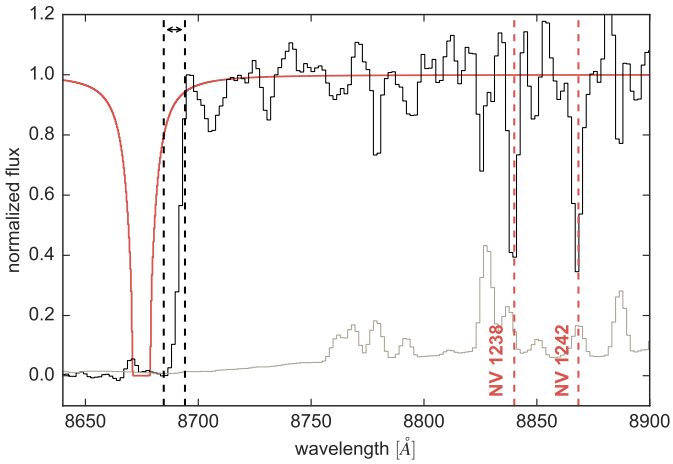

FIG. 11.— Excerpt of the spectrum of CFHQSJ2229+1457 taken with LRIS. We find an N V absorption line system (red dashed lines) associated with an absorber at $z \approx 6.136$ (red curve) with a column denisty $N_{\rm HI} = 10^{19}{\rm cm}^{-2}$ and Doppler parameter $b = 40$ km/s. A boxcar smoothing of two pixels has been applied to both the spectrum (black curve) and its noise vector (gray curve). The vertical black dashed lines indicate the extent of the proximity zone.

assumes a $\Delta v \leq 3000$ km s^{-1} window, which corresponds to a distance of ≈ 4.3 pMpc to a quasar at $z = 6$. We are just interested in the innermost 1 pMpc to the quasar, where the quasar's radiation is stronger and could thus have photoionized all possible LLSs (Hennawi & Prochaska 2007). A second possible reason for the lack of detections of more proximate LLSs in our sample, could be due to the fact that our data sample at $z \sim 6$ is brighter than the one analyzed by Prochaska et al. (2010) at $z \sim 3.5$ and thus the higher radiation could have photoionized more LLSs. Hence, it might be not surprising, that we are only detecting one proximate LLS in our data sample instead of the expected ~ 3.

Although we cannot rule out the presence of absorbers with weak metal lines or rare metal-free absorbers (Fumagalli et al. 2011; Simcoe et al. 2012; Cooper et al. 2015; Crighton et al. 2016; Cooke et al. 2017) in all objects, which would require higher quality data, we do not see evidence for proximate LLSs in the three spectra showing small proximity zones, particularly in the spectrum of the very bright quasar SDSSJ0100 + 2802, where we would most likely be able to detect it. Follow up higher S/N and resolution observations are required to confirm the lack of associated dense absorbers in these spectra.

6.3.2. *Neutral Islands of Gas in a Patchily Reionized IGM*

Another possible explanation for the existence of very small proximity zones could be the truncation of the proximity zones due to islands of very neutral gas in an otherwise highly ionized IGM, which can occur in patchy reionization models (e.g. Lidz et al. 2007; Mesinger 2010). This scenario implies that reionization is not yet complete at $z \sim 6$ and these quasars lie in (presumably rare) neutral environments that could absorb all incoming flux and thus truncate the proximity zones. Opposite to the previously described scenario where LLSs, i.e. dense gas on galactic scales, truncate the proximity zones, we assume here neutral patches of ~ 10 pMpc scales that are associated with reionization and do not have any low redshift counter parts. This explanation requires the presence of large patches of highly neutral gas just outside the proximity zone, which would also absorb flux in the Lyman-series (Lyα, Lyβ, etc.) forests. In other words, in a patchy reionization scenarios one expects to see long GP troughs that are line black in both Lyα and Lyβ, beginning right at the edge of the proximity zone.

In Fig. 13 we show continuum normalized spectra of the Lyα and the Lyβ forest (top and bottom panels, respectively) near the three quasars in question. The black dashed lines in all panels indicate the extent of the quasar proximity zone. The left panels show the Lyα and the Lyβ forest close to the quasar CFHQSJ2229 + 1457. While we do not detect any transmitted flux in the Lyβ forest, we do see a few transmission spikes in the Lyα forest at $R \approx 1.1$ pMpc and $R \approx 2.8$ pMpc, indicated by the red dashed lines. The transmission in the Lyβ forest at the same distances must have been absorbed due to foreground Lyα absorption at $z \approx 5.02$ and $z \approx 4.99$.

The middle panels show the Lyα and Lyβ forest of SDSSJ1335 + 3533, where we see transmitted flux in the Lyα forest just outside of the proximity zone, i.e. at $R \approx 0.9$ pMpc and $R \approx 1.0$ pMpc, and a bit further away at $R \approx 2.9$ pMpc and $R \approx 3.6$ pMpc. In the Lyβ forest we can see clear transmission spikes just outside of the proximity zone at $R \approx 1.4$ pMpc, $R \approx 2.8$ pMpc and $R \approx 3.3$ pMpc.

The right panels show excerpts of the Lyα and Lyβ forest for SDSSJ0100 + 2802. One can clearly see several transmitted flux spikes in the Lyα forest right outside the proximity zone. Also the Lyβ forest shows a few flux spikes close to the proximity zone, although less prominent ones.

These detections of transmitted flux just outside the proximity zones in the Lyα or Lyβ forest suggest that if there are patches of neutral gas present close to the quasar, they need to be $\lesssim 1$ pMpc in size, corresponding to the size of the regions that are both line black in the Lyα and Lyβ forest. Although we cannot completely rule out a patchy reionization scenario with such small islands of neutral gas, our constraints on their sizes already suggest that they would be difficult to distinguish from metal poor LLSs. Deeper spectra (at least for the quasars CFHQSJ2229 + 1457 and SDSSJ1335 + 3533) covering their proximity zones further down the Lyman series could reveal additional Lyβ or Lyγ transmission spikes that would further constrain this scenario, as well as enabling more sensitive searches for metal lines to rule out an LLS scenario.

6.3.3. *Young Quasar Age*

Finally, another possibility to explain small proximity zones is that the quasars may have turned on only very recently. When a quasar turns on the IGM requires a finite amount of time to adjust to the new higher photoionization rate. We will focus the discussion in this section on a highly ionized surrounding IGM, consistent with optical depth measurements at $z \sim 6$ (e.g. Wyithe & Bolton 2011; Calverley et al. 2011; Becker et al. 2015; McGreer et al. 2015), although we note similar arguments apply in the neutral case (see Khrykin et al. 2016, for fur-

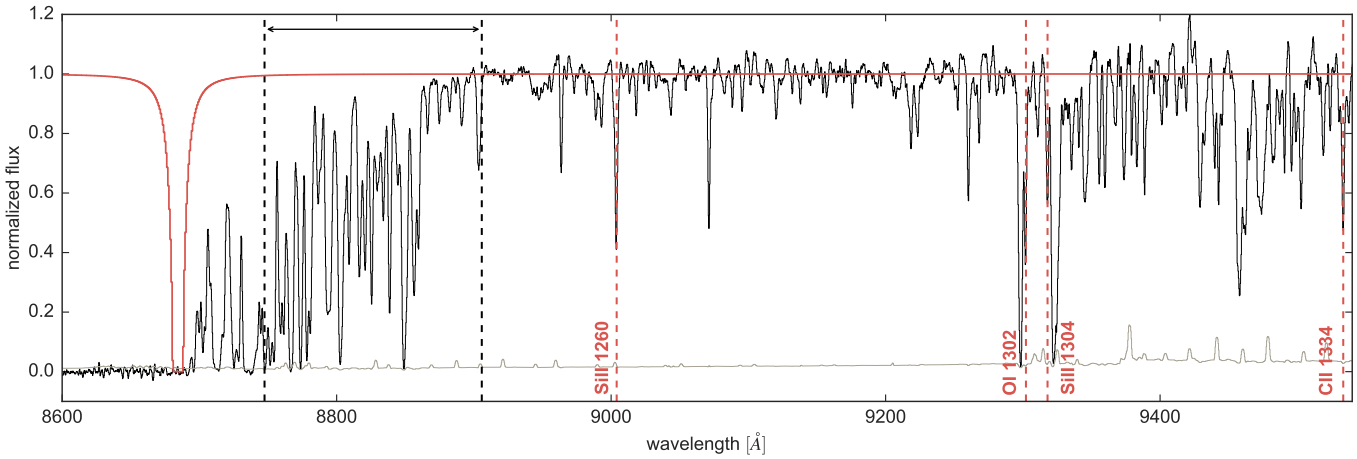

FIG. 12.— Excerpt of the spectrum SDSSJ0100 + 2802. We find a close low ionization absorption system at $z \approx 6.144$ (red curve) with column density $N_{\rm HI} = 10^{19}{\rm cm}^{-2}$ and Doppler parameter $b = 40$ km/s and associated absorption lines (red dashed lines). The continuum normalized[a] spectrum (black curve) and its noise vector (gray curve) are smoothed with a three pixel boxcar function. The vertical black dashed lines indicate the extent of the proximity zone.

[a]Note that in this figure we show the quasar spectrum normalized with a hand-fitted continuum.

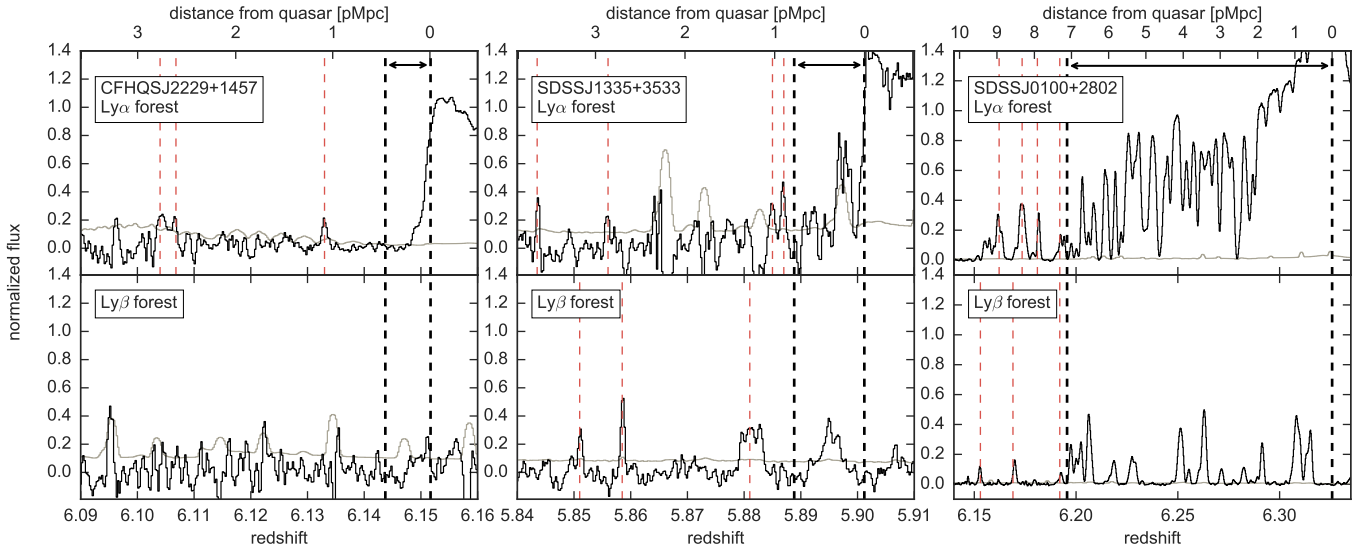

FIG. 13.— Excerpts of the Lyα and the Lyβ forests of the three spectra of CFHQSJ2229 + 1457 (left panels), SDSSJ1335 + 3533 (middle panels) and SDSSJ0100 + 2802 (right panels). The panels show the Lyα and the Lyβ forest (upper and lower panels, respectively) of the spectra at the same redshift and distance close to the quasars. A boxcar smoothing of three pixels has been applied to both the spectra (black curves) and noise vectors (gray curves). The extents of the proximity zones are indicated by the vertical black dashed lines. Red dashed lines indicate transmission spikes outside of the respective proximity zones.

ther details). Following (Khrykin et al. 2016)[16] the time evolution of the neutral fraction in the proximity zone is well described by the equation

$$x_{\rm HI}(t_{\rm Q}) = x_{\rm HI,0} + (x_{\rm HI,0} - x_{\rm HI,eq})\exp^{-t_{\rm Q}/t_{\rm eq}}, \quad (12)$$

where $t_{\rm Q}$ is the quasar age, $x_{\rm HI,0} \approx n_{\rm e}\alpha_{\rm HII}/\Gamma_{\rm UVB}$ is the neutral fraction of the IGM before the quasar turned on, and $x_{\rm HI,eq} \approx n_{\rm e}\alpha_{\rm HII}/(\Gamma_{\rm UVB} + \Gamma_{\rm QSO})$ is the new lower neutral fraction that the IGM will reach once ionization equilibrium is attained. Here $n_{\rm e}$ and $\alpha_{\rm HII}$ are the IGM electron density and recombination coefficient, respectively. The characteristic timescale to reach ionization equilibrium, the equilibration timescale $t_{\rm eq}$, is

$$t_{\rm eq} \approx \frac{1}{\Gamma_{\rm UVB} + \Gamma_{\rm QSO}}. \quad (13)$$

Eqn. (12) thus implies that a quasar has to have been shining for at least $t_{\rm Q} \gtrsim t_{\rm eq}$, in order for the surrounding IGM to have reached ionization equilibrium with the quasar and result in the maximum proximity zone size. Ionizing equilibrium is achieved more rapidly close the quasar due to the stronger radiation field ($\Gamma_{\rm QSO} \propto R^{-2}$). In order to obtain an optical depth of $\tau_{\rm lim} = 2.3$ at $z \sim 6$, which corresponds to a transmitted flux level of 10% in the Lyα forest, a value of $\Gamma_{\rm QSO} \sim 10^{-12}$ s^{-1} is implied (see § 5.3), resulting in a typical equilibration time scale

[16] Khrykin et al. (2016) wrote down the analogous equation for the He II singly ionized fraction in the context of He II Lyα proximity zones at $z \sim 3$, the exact same set of arguments applies here for neutral hydrogen around $z \sim 6$ quasars.

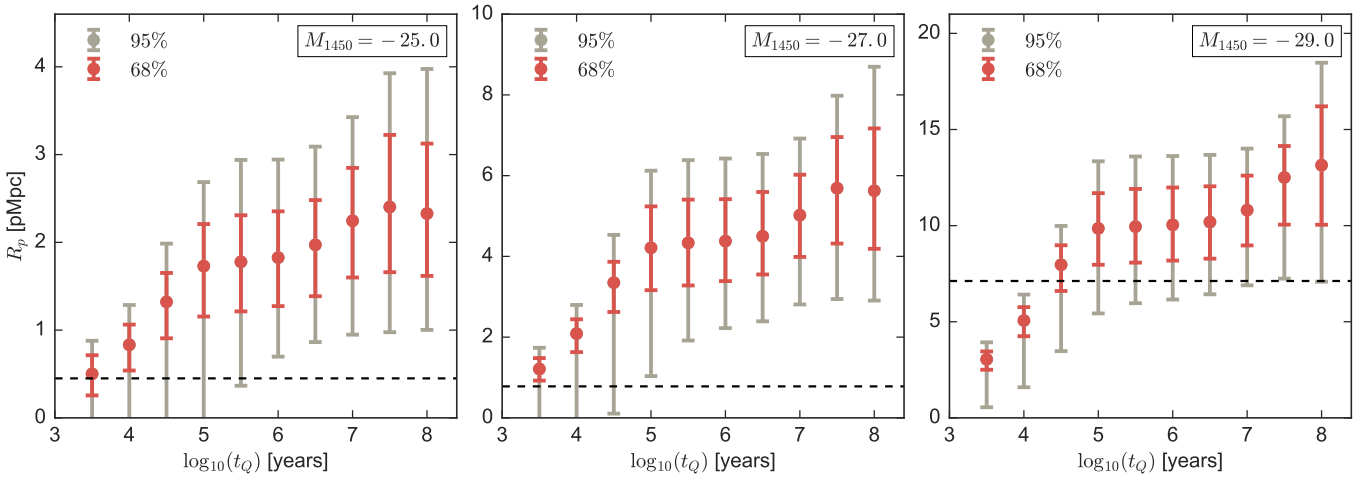

FIG. 14.— Distributions of proximity zone sizes shown as a function of the average quasar age t_Q for quasars with magnitudes of $M_{1450} = -25.0$ (left panel), magnitudes of $M_{1450} = -27.0$ (middle panel) and magnitudes of $M_{1450} = -29.0$ (right panel). The proximity zone sizes of CFHQSJ2229 + 1457, SDSSJ1335 + 3533 and SDSSJ0100 + 2802 are shown as the dashed black line in the left, middle and right panel, respectively.

of $t_{\rm eq} \sim 3 \times 10^4$ yr. Full ionization equilibrium is reached after a few equilibration times, or for $t_Q \sim 10^5$ yr. The region in which ionization equilibrium is reached, and hence also the size of the quasar's proximity zone, will grow with time until $t_Q \approx 3t_{\rm eq}$. Thus, even for a highly ionized IGM, the sizes of the proximity zones are dependent on the quasar age for $t_Q \lesssim 10^5$ yr, whereas for ages $t_Q \gtrsim 10^5$ yr the proximity zone sizes are largely independent on the precise quasar age[17].

In Fig. 14 we show the dependence of proximity zone size on quasar age determined from radiative transfer simulations in a highly ionized IGM. Quasars with magnitudes $M_{1450} = -25.0$ (left) $M_{1450} = -27.0$ (middle) and $M_{1450} = -29.0$ (right) are shown, where we have simulated 900 proximity zones at each magnitude. The black horizontal dashed lines show the measured sizes of the proximity zones of CFHQSJ2229 + 1457 (left panel), SDSSJ1335+3533 (middle panel) and SDSSJ0100+2802 (right panel).

The distribution of proximity zone sizes in the left and middle panel encompasses values of $R_p \approx 1$ pMpc that we have measured only when assuming a short quasar age of $t_Q \lesssim 10^5$ yr. Hence a plausible explanation for the exceptionally small sizes of the proximity zones could be young ages for these quasars. The proximity zone of the very bright quasar SDSSJ0100+2802 (right panel) is not as significant of an outlier as the other two objects, but nevertheless the small size of its zone would be more probable for $t_Q \sim 10^5$ yr.

At lower redshifts $z \lesssim 4$, studies of quasar clustering imply that the duty cycle of quasar activity is $t_{\rm dc} \sim 10^9$ yr (Shen et al. 2007; White et al. 2008)[18]. However, whereas clustering constraints the duty cycle, proximity zones actually probe a different timescale, which is the duration of quasar emission episodes which could be considerably shorter. Indeed, we actually obtain a lower limit on the episodic lifetime, because the quasar could continue to emit for many years after we observe it on Earth. Nevertheless, for a given episodic lifetime $t_{\rm episodic}$ the probability of measuring an age of t_Q is $p = t_Q/t_{\rm episodic}$, assuming the simplest "light-bulb" lightcurve for the quasars. If we assume an average episodic lifetime of $t_{\rm episodic} \sim 10^8$ yr, the probability of observing a quasar that has only been shining for $\sim 10^5$ yr is $p \sim 0.1\%$, or $p \sim 1\%$ for $t_{\rm episodic} \sim 10^7$ yr. We have discovered three objects suggesting ages of $t_Q \sim 10^5$ yr in a sample of 30 quasars, i.e. $p \approx 10\%$. To be consistent with finding a few of these small proximity zones, the quasar episodic lifetimes would need to be $t_{\rm episodic} \sim 10^6$ yr. This would leave the sizes of the proximity zones of the vast majority of quasars unchanged, which we typically observe much later in their evolution, but could explain the very small zones we find.

7. SUMMARY

In this paper we analyze a sample of 34 high redshift quasar spectra taken with the ESI instrument on the Keck II telescope. We reduce the spectra in a homogeneous way, and analyze the sizes of their proximity zones for a subset of 30 quasars which do not exhibit BAL features or have obvious signatures of nearby associated absorption line systems that could prematurely truncate their proximity zones. Our analysis uses updated redshift measurements, fully consistent values for the quasar absolute magnitudes M_{1450}, and carefully fits to the quasar continua based on principal component spectra.

Previous work found a strong evolution of proximity zone sizes with redshift, and it was argued that this provided evidence for rapid evolution of the IGM neutral fraction during the epoch of reionization. We instead find a much shallower redshift evolution, which is however consistent with the evolution predicted by our radiative transfer simulations, irrespective of assumptions about the ionization state of the IGM. The size of the proximity zone ends at a distance corresponding to a flux transmission level of 10% according to the definition of quasar proximity zones (Fan et al. 2006b) and, in a

[17] A mild increase in proximity zone sizes for long quasar ages, i.e. $t_Q \gtrsim 10^7$ yr, can be attributed to heating effects due to He II reionization (Bolton et al. 2010, 2012; Davies et al. in prep.)

[18] Whether this also applies at redshifts $z \sim 6$ still remains to be determined.

highly ionized IGM, this distance is reached at a location where the ionization rate of the quasar dominates the ionization rate of the UVB by at least an order of magnitude. As such, the size of the quasar proximity zone is essentially independent of the UVB and neutral gas fraction. Assuming a neutral ambient IGM the observed proximity zone sizes R_p track the growth of the ionized H II region $R_{\rm ion}$, but then cease to grow when the distance with a 10% flux transmission level is reached. For a highly plausible quasar age of $t_{\rm Q} \sim 10^{7.5}$ yr, R_p is thus independent of the ionization state of the IGM for a mostly neutral IGM as well. Thus contrary to previous claims, both the observed shallow redshift evolution and the results from our simulations imply that the redshift evolution of proximity zone size R_p does not significantly constrain the evolution of the neutral gas fraction during the epoch of reionization.

Our analysis uncovered three quasars with exceptionally small proximity zones, two of them with $R_p \lesssim 1$ pMpc. The occurrence of zones this small is extremely rare in our simulations. We estimate the probability of finding these objects in a sample of 30 quasar spectra to be $p \approx 1\%$ for SDSSJ1335 + 3533, and $p \approx 3\%$ for CFHQSJ2229 + 1457 and SDSSJ0100 + 2802. We search for evidence of proximate absorption systems such as DLAs or LLSs, or islands of remaining neutral hydrogen associated with patchy reionization, both of which could result in truncation of the proximity zones closer to the quasars. However the absence of strong metal absorption lines or large GP troughs (in both the Lyα and the Lyβ forest) appears to disfavor these truncation scenarios. Nevertheless we cannot completely rule out the presence of low metallicity or metal-free absorbers due to the low signal-to-noise ratio of the data of CFHQSJ2229 + 1457 and SDSSJ1335 + 3533.

The most plausible explanation for the small proximity zones that we observe is that these quasars are relatively young. It takes the IGM roughly $\sim 10^5$ yr to reach ionization equilibrium with the quasar radiation. Proximity zones of quasars with ages longer than this timescale are essentially independent of the exact age of the quasar. However, for shorter quasar ages the surrounding IGM has not yet reached ionization equilibrium resulting in proximity zones comparable to the small sizes we observe. If the duration of quasar emission episodes is $t_{\rm episodic} \sim 10^8$ yr, the detection of these small zones would be very unlikely, i.e. $p \sim 0.1\%$. This discrepancy can be resolved if one assumes a shorter duration of $t_{\rm episodic} \sim 10^6$ yr, resulting in a probability of $p \sim 10\%$ for finding these objects with ages of $t_{\rm Q} \sim 10^5$ yr.

However, an average episodic lifetime of $t_{\rm episodic} \sim 10^6$ yr causes significant tension with the inferred sizes of SMBHs at these redshifts, since the presence of $\sim 10^9 M_\odot$ SMBHs at $z \sim 6$ requires that quasars accrete continuously at around the Eddington limit for nearly the entire Hubble time of $\sim 10^9$yr (Volonteri 2012). Thus, although the proximity zone measurements constrain the episodic lifetime $t_{\rm episodic}$, whereas the sizes of SMBH depend on the integral over all emission episodes, i.e. the duty cycle, the presence of large SMBHs at $z \sim 6$ either requires long episodic lifetimes comparable to the duty cycle $\sim 10^9$ yr, or that the quiescent time between episodic emission bursts must be very short, implying that SMBHs grow via many short episodic phases (Schawinski et al. 2015; Mao & Kim 2016; Davies et al. in prep.).

It would clearly be interesting to uncover more quasars with small proximity zones, particularly if they are indeed young quasars that have only been shining for $t_{\rm Q} \sim 10^5$ yr. In the future, we plan to use large quasar surveys, such as Pan-STARRS or SDSS, to identify candidate small-zone quasars and measure accurate redshifts from [C II] or CO emission lines with the Atacama Large Millimeter Array (ALMA), in order to confirm the potentially small sizes of their proximity zones. Deep optical and near-IR follow-up observations will be helpful to rule out a premature truncation due to associated absorption systems. Further theoretical work is also required to investigate the causes and implications of small proximity zones, and their relationship to the distribution of quasar ages, which is a subject we address in future work (Davies et al. in prep.).

ACKNOWLEDGMENT

The authors would like to thank J. Bolton and M. Haehnelt for their questions about the small proximity zones in our sample at the "Dark Ages" meeting in Heidelberg, which provided part of the motivation for this work. Additionally, we wish to thank E. Bañados, X. Fan, B. Venemans, and F. Walter for valuable input and discussion.

The data presented in this paper were obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation.

This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

ZL was supported by the Scientific Discovery through Advanced Computing (SciDAC) program funded by U.S. Department of Energy Office of Advanced Scientific Computing Research (ASCR) and the Office of High Energy Physics. Calculations presented in this paper used resources of the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors would like to thank Dmitriy Morozov and Gunther H. Weber for the halo finder work, which was funded by the ASCR project "Scalable Analysis Methods and In Situ Infrastructure for Extreme Scale Knowledge Discovery," program manager Lucy Nowell.

REFERENCES

Almgren, A. S., Bell, J. B., Lijewski, M. J., Lukić, Z., & Van Andel, E. 2013, Astrophysical Journal, 765, 39

Bañados, E., Decarli, R., Walter, F., et al. 2015, The Astrophysical Journal Letters, 805, L8

Bañados, E., Venemans, B. P., Decarli, R., et al. 2016, ArXiv e-prints, arXiv:1608.03279
Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, The Astrophysical Journal, 327, 570
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, Monthly Notices of the Royal Astronomical Society, 447, 3402
Bernstein, R. M., Burles, S. M., & Prochaska, J. X. 2015, Publications of the Astronomical Society of Pacific, 127, 911
Bochanski, J. J., Hennawi, J. F., Simcoe, R. A., et al. 2009, The Publications of the Astronomical Society of the Pacific, 121, 1409
Bolton, J. S., Becker, G. D., Raskutti, S., et al. 2012, Monthly Notices of the Royal Astronomical Society, 419, 2880
Bolton, J. S., Becker, G. D., Wyithe, J. S. B., Haehnelt, M. G., & Sargent, W. L. W. 2010, Monthly Notices of the Royal Astronomical Society, 406, 612
Bolton, J. S., & Haehnelt, M. G. 2007a, Monthly Notices of the Royal Astronomical Society, 381, L35
—. 2007b, Monthly Notices of the Royal Astronomical Society, 374, 493
Bolton, J. S., Haehnelt, M. G., Warren, S. J., et al. 2011, Monthly Notices of the Royal Astronomical Society, 416, L70
Bosman, S. E. I., & Becker, G. D. 2015, Monthly Notices of the Royal Astronomical Society, 452, 1105
Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, Monthly Notices of the Royal Astronomical Society, 412, 2543
Carilli, C. L., Wang, R., Fan, X., et al. 2010, The Astrophysical Journal, 714, 834
Cen, R., & Haiman, Z. 2000, The Astrophysical Journal, 542, L75
Cen, R., & Safarzadeh, M. 2015, Astrophysical Journal Letters, 798, L38
Conroy, C., & White, M. 2013, The Astrophysical Journal, 762, 70
Cooke, R. J., Pettini, M., & Steidel, C. C. 2017, Monthly Notices of the Royal Astronomical Society, arXiv:1701.03103
Cooper, T. J., Simcoe, R. A., Cooksey, K. L., O'Meara, J. M., & Torrey, P. 2015, The Astrophysical Journal, 812, 58
Crighton, N. H. M., O'Meara, J. M., & Murphy, M. T. 2016, Monthly Notices of the Royal Astronomical Society, 457, L44
Davies, F. B., Eilers, A.-C., & Hennawi, J. F. in prep.
Davies, F. B., Furlanetto, S. R., & McQuinn, M. 2016, Monthly Notices of the Royal Astronomical Society, 457, 3006
De Rosa, G., Decarli, R., Walter, F., et al. 2011, The Astrophysical Journal, 739, 56
De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, The Astrophysical Journal, 790, 145
Diamond-Stanic, A. M., Fan, X., Brandt, W. N., et al. 2009, The Astrophysical Journal, 699, 782
Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, The Astronomical Journal, 128, 515
Fan, X., Strauss, M. A., Richards, G. T., et al. 2006a, The Astronomical Journal, 131, 1203
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006b, The Astronomical Journal, 132, 117
Fumagalli, M., O'Meara, J. M., & Prochaska, J. X. 2011, Science, 334, 1245
Furlanetto, S. R., & Lidz, A. 2011, The Astrophysical Journal, 735, 117
Gaskell, C. M. 1982, The Astrophysical Journal, 263, 79
Gunn, J. E., & Peterson, B. A. 1965, Astrophysical Journal, 142, 1633
Haiman, Z., & Cen, R. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 222, The Physics of Galaxy Formation, ed. M. Umemura & H. Susa, 101
Haiman, Z., & Hui, L. 2001, The Astrophysical Journal, 547, 27
Hennawi, J. F., & Prochaska, J. X. 2007, The Astrophysical Journal, 655, 735
Hennawi, J. F., Prochaska, J. X., Burles, S., et al. 2006, The Astrophysical Journal, 651, 61
Hopkins, P. F., & Hernquist, L. 2009, The Astrophysical Journal, 698, 1550
Jiang, L., Fan, X., Vestergaard, M., et al. 2007, The Astronomical Journal, 134, 1150
Jiang, L., Fan, X., Annis, J., et al. 2008, The Astronomical Journal, 135, 1057
Keating, L. C., Haehnelt, M. G., Cantalupo, S., & Puchwein, E. 2015, Monthly Notices of the Royal Astronomical Society, 454, 681
Kelson, D. D. 2003, The Publications of the Astronomical Society of the Pacific, 115, 688
Khrykin, I. S., Hennawi, J. F., McQuinn, M., & Worseck, G. 2016, The Astrophysical Journal, 824, 133
Kramer, R. H., & Haiman, Z. 2009, Monthly Notices of the Royal Astronomical Society, 400, 1493
Kurk, J. D., Walter, F., Fan, X., et al. 2007, The Astrophysical Journal, 669, 32
Lidz, A., McQuinn, M., Zaldarriaga, M., Hernquist, L., & Dutta, S. 2007, The Astrophysical Journal, 670, 39
Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, Monthly Notices of the Royal Astronomical Society, 446, 3697
Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, Monthly Notices of the Royal Astronomical Society, 449, 4204
Madau, P., & Rees, M. J. 2000, The Astrophysical Journal, 542, L69
Mao, J., & Kim, M. 2016, The Astrophysical Journal, 828, 96
Martini, P. 2004, Coevolution of Black Holes and Galaxies, 169
Martini, P., & Weinberg, D. H. 2001, The Astrophysical Journal, 547, 12
Maselli, A., Ferrara, A., & Gallerani, S. 2009, Monthly Notices of the Royal Astronomical Society, 395, 1925
Maselli, A., Gallerani, S., Ferrara, A., & Choudhury, T. R. 2007, Monthly Notices of the Royal Astronomical Society, 376, L34
McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, Monthly Notices of the Royal Astronomical Society, 447, 499
Mesinger, A. 2010, Monthly Notices of the Royal Astronomical Society, 407, 1328
Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616
Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, The Publications of the Astronomical Society of the Pacific, 107, 375
Osmer, P. S. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 146, The Young Universe: Galaxy Formation and Evolution at Intermediate and High Redshift, ed. S. D'Odorico, A. Fontana, & E. Giallongo, 1
Pâris, I., Petitjean, P., Rollinde, E., et al. 2011, Astronomy & Astrophysics, 530, A50
Prochaska, J. X., Hennawi, J. F., & Herbert-Fort, S. 2008, The Astrophysical Journal, 675, 1002
Prochaska, J. X., O'Meara, J. M., & Worseck, G. 2010, The Astrophysical Journal, 718, 392
Prochaska, J. X., Hennawi, J. F., Lee, K.-G., et al. 2013, The Astrophysical Journal, 776, 136
Richards, G. T., Vanden Berk, D. E., Reichard, T. A., et al. 2002, The Astronomical Journal, 124, 1
Schawinski, K., Koss, M., Berney, S., & Sartori, L. F. 2015, Monthly Notices of the Royal Astronomical Society, 451, 2517
Selsing, J., Fynbo, J. P. U., Christensen, L., & Krogager, J.-K. 2016, Astronomy & Astrophysics, 585, A87
Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, The Publications of the Astronomical Society of the Pacific, 114, 851
Shen, Y., & Ménard, B. 2012, The Astrophysical Journal, 748, 131
Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, The Astronomical Journal, 133, 2222
Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, The Astrophysical Journal, 697, 1656
Shen, Y., Brandt, W. N., Richards, G. T., et al. 2016, The Astrophysical Journal, 831, 7
Simcoe, R. A., Sullivan, P. W., Cooksey, K. L., et al. 2012, Nature, 492, 79
Songaila, A., & Cowie, L. L. 2010, The Astrophysical Journal, 721, 1448
Suzuki, N. 2006, The Astrophysical Journal Supplement Series, 163, 110
Tepper-García, T. 2006, Monthly Notices of the Royal Astronomical Society, 369, 2025
Tytler, D., & Fan, X.-M. 1992, Astrophysical Journal Supplement Series, 79, 1
Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, The Astronomical Journal, 122, 549
Venemans, B. P., Walter, F., Zschaechner, L., et al. 2016, The Astrophysical Journal, 816, 37
Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, The Astrophysical Journal, 779, 24
Venemans, B. P., Bañados, E., Decarli, R., et al. 2015, The Astrophysical Journal Letters, 801, L11
Volonteri, M. 2010, The Astronomy and Astrophysics Review, 18, 279
—. 2012, Science, 337, 544
Wang, R., Carilli, C. L., Neri, R., et al. 2010, The Astrophysical Journal, 714, 699
Wang, R., Wagg, J., Carilli, C. L., et al. 2013, The Astrophysical Journal, 773, 44
Wang, R., Wu, X.-B., Neri, R., et al. 2016, ArXiv e-prints, arXiv:1606.09634
White, M., Martini, P., & Cohn, J. D. 2008, Monthly Notices of the Royal Astronomical Society, 390, 1179
White, M., Myers, A. D., Ross, N. P., et al. 2012, Monthly Notices of the Royal Astronomical Society, 424, 933
Willott, C. J., Bergeron, J., & Omont, A. 2015, The Astrophysical Journal, 801, 123
Willott, C. J., Omont, A., & Bergeron, J. 2013, The Astrophysical Journal, 770, 13

Willott, C. J., Delorme, P., Omont, A., et al. 2007, The Astronomical Journal, 134, 2435
Willott, C. J., Delorme, P., Reylé, C., et al. 2009, The Astronomical Journal, 137, 3541
Willott, C. J., Albert, L., Arzoumanian, D., et al. 2010, The Astronomical Journal, 140, 546
Wu, X.-B., Wang, F., Fan, X., et al. 2015, Nature, 518, 512
Wyithe, J. S. B., & Bolton, J. S. 2011, Monthly Notices of the Royal Astronomical Society, 412, 1926
Wyithe, J. S. B., Loeb, A., & Carilli, C. 2005, The Astrophysical Journal, 628, 575

APPENDIX

A. ASSOCIATED ABSORBERS IN SDSSJ0840+5624

This object has been excluded from our analysis because we found several absorption systems at high redshift associated with the quasar itself. These absorbers cause additional absorption within the proximity zone of the quasar, such that our analysis recovers a spuriously small proximity zone of $R_p \approx 0.88$ pMpc. However, inspecting the spectrum by eye reveals that the actual proximity zone of this object extends much further out to $R_p \gtrsim 5.0$ pMpc.

We find two N V absorption line systems ($\lambda_{\rm rest} = 1238.82$ Å and $\lambda_{\rm rest} = 1242.80$ Å) associated with close absorbers at $z \approx 5.835$ and $z \approx 5.825$, which are shown in Fig. 15. We also find another absorption system in the spectrum at $z \approx 5.594$ with associated Si II absorption ($\lambda_{\rm rest} = 1260.42$ Å) that falls within the proximity zone. Thus there is a lot of absorption within the proximity zone that cannot be contributed to residual neutral hydrogen only and an exact determination of the proximity zone for this object without contamination from other absorption systems is impossible.

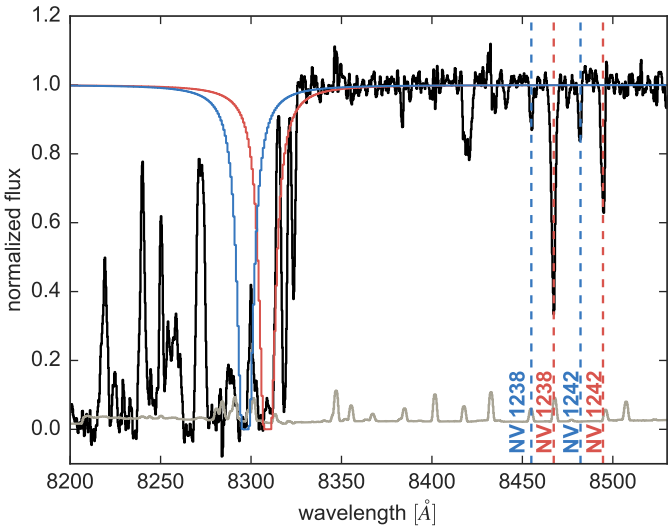

FIG. 15.— Two N V absorption line systems shown as the blue and red dashed lines asscoiated with close absorbers (blue and red curves) at $z_{\rm blue} \approx 5.825$ and $z_{\rm red} \approx 5.835$ in the spectrum of the quasar SDSSJ0840 + 5624. The continuum normalized[a] spectrum (black curve) and its noise vector (gray curve) are smoothed with a three pixel boxcar function. The depicted absorbers have a column density $N_{\rm HI} = 10^{19}{\rm cm}^{-2}$ and a Doppler parameter $b = 20$ km/s.

[a]Note that in this figure we show the quasar spectrum normalized with a hand-fitted continuum.

B. DETAILS OF THE CONTINUUM MODELING OF EACH QUASAR

Table 3 shows the details of modeling the continuum of each quasar spectrum.

TABLE 3
Details of the continuum modeling of each quasar.

object	normalization to unity at rest-frame	set of PCS	number of PCS
SDSS J0002+2550	1280 Å	Suzuki (2006)	5
SDSS J0005-0006	1280 Å	Pâris et al. (2011)	5
CFHQS J0050+3445	1280 Å	Pâris et al. (2011)	5
SDSS J0100+2802	1280 Å	Pâris et al. (2011)	5
ULAS J0148+0600	1280 Å	Pâris et al. (2011)	5
CFHQS J0210-0456	1245 Å	Pâris et al. (2011)	5
PSO J0226+0302	1280 Å	Pâris et al. (2011)	5
CFHQS J0227-0605	1265 Å	Pâris et al. (2011)	5
SDSS J0303-0019	1280 Å	Pâris et al. (2011)	7
PSO J0402+2452	1280 Å	Suzuki (2006)	5
SDSS J0818+1723	1280 Å	Pâris et al. (2011)	5
SDSS J0836+0054	1280 Å	Suzuki (2006)	5
SDSS J0840+5624	1280 Å	Pâris et al. (2011)	5
SDSS J0842+1218	1280 Å	Pâris et al. (2011)	5
SDSS J0927+2001	1280 Å	Pâris et al. (2011)	5
SDSS J1030+0524	1280 Å	Pâris et al. (2011)	5
SDSS J1137+3549	1280 Å	Suzuki (2006)	5
SDSS J1148+5251	1280 Å	Pâris et al. (2011)	5
SDSS J1250+3130	1280 Å	Pâris et al. (2011)	5
SDSS J1306+0359	1285 Å	Pâris et al. (2011)	5
ULAS J1319+0950	1280 Å	Suzuki (2006)	5
SDSS J1335+3533	1280 Å	Pâris et al. (2011)	5
SDSS J1411+1217	1280 Å	Pâris et al. (2011)	3
SDSS J1602+4228	1280 Å	Suzuki (2006)	5
SDSS J1623+3112	1280 Å	Pâris et al. (2011)	5
SDSS J1630+4012	1280 Å	Suzuki (2006)	5
CFHQS J1641+3755	1280 Å	Pâris et al. (2011)	3
SDSS J2054-0005	1280 Å	Pâris et al. (2011)	5
CFHQS J2229+1457	1250 Å	Pâris et al. (2011)	7
SDSS J2315-0023	1280 Å	Pâris et al. (2011)	5
CFHQS J2329-0301	1230 Å	Pâris et al. (2011)	5

Note. — The columns show the name of the object, the rest-frame wavelength at which the quasar spectra is normalized to unity, and the set and number of PCS we used to model the quasar continua.

C. ZONES OF REMAINING QUASARS FROM OUR DATA SET

Fig. 16 shows the same as Fig. 2 and Fig. 3 for the remaining quasars in our data sample sorted by the redshift of the quasar.

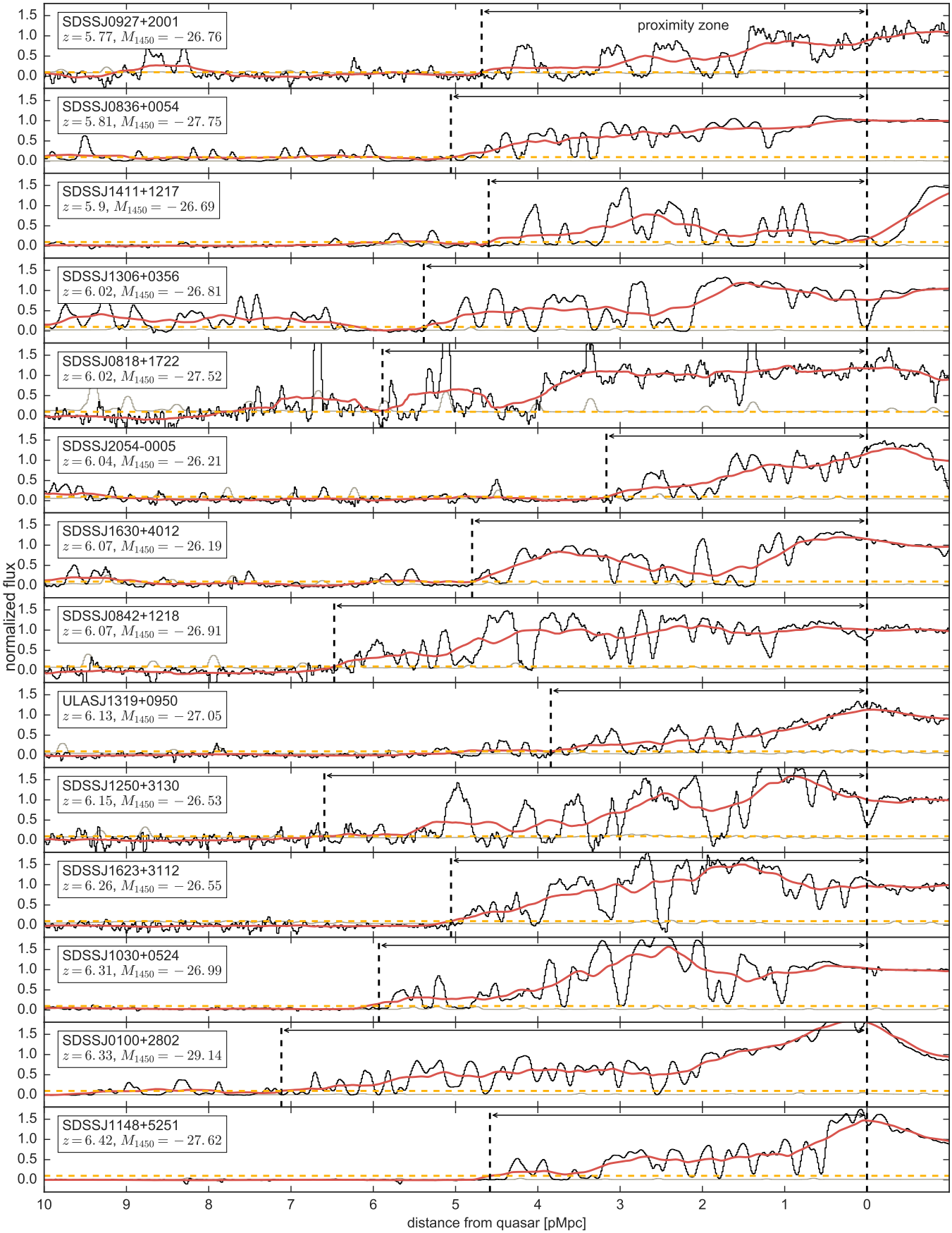

FIG. 16.— Remaining quasar spectra and their proximity zone measurements from our sample sorted by quasar redshift. The spectra (black curves) and their noise vectors (gray curves) are smoothed with a boxcar function of two pixels. The black dashed lines indicate the beginning (location of the quasar) and the end (where the smoothed flux shown as the red curve drops below the 10% level indicated by the yellow dashed lines) of the proximity zones.