



Constraining the Size of the Circumgalactic Medium Using the Transverse Autocorrelation Function of CIV Absorbers in Paired Quasar Spectra

Abby Mintz^{1,2,3} , Marc Rafelski^{2,4} , Regina A. Jorgenson³ , Michele Fumagalli⁵ , Rajeshwari Dutta⁵ , Crystal L. Martin⁶ ,

Elisabeta Lusso^{7,8} , Kate H. R. Rubin⁹ , and John M. O’Meara¹⁰ 

¹ Department of Astronomy, Yale University, New Haven, CT, USA; abby.mintz@yale.edu

² Space Telescope Science Institute, Baltimore, MD 21218, USA

³ Maria Mitchell Observatory, Nantucket, MA, USA

⁴ Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA

⁵ Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

⁶ Department of Physics, University of California, Santa Barbara, CA 93106, USA

⁷ Dipartimento di Fisica e Astronomia, Università di Firenze, via G. Sansone 1, I-50019 Sesto Fiorentino, Firenze, Italy

⁸ INAF–Osservatorio Astrofisico di Arcetri, I-50125 Florence, Italy

⁹ San Diego State University, Department of Astronomy, San Diego, CA 92182, USA

¹⁰ W.M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA

Received 2021 November 1; revised 2022 May 26; accepted 2022 June 2; published 2022 July 14

Abstract

The circumgalactic medium (CGM) plays a vital role in the formation and evolution of galaxies, acting as a lifeline between galaxies and the surrounding intergalactic medium. In this study, we leverage a unique sample of quasar pairs to investigate the properties of the CGM with absorption line tomography. We present a new sample of medium-resolution Keck/ESI, Magellan/MagE, and VLT/XSHOOTER spectra of 29 quasar pairs at redshift $2 < z < 3$. We supplement the sample with additional spectra of 32 pairs from the literature, creating a catalog of 61 quasar pairs with angular separations between $1''.7$ and $132''.9$ and projected physical separations (r_{\perp}) between 14 kpc and 887 kpc. We construct a catalog of 906 metal-line absorption doublets of C IV ($\lambda\lambda 1548, 1550$) with equivalent widths ranging from $6 \text{ m}\text{\AA} \leq W_{r,1550} \leq 2053 \text{ m}\text{\AA}$. The best-fit linear model to the log-space equivalent width frequency distribution ($\log f(W_r) = m \log(W_r) + b$) of the sample yields coefficients of $m = -1.44 \pm 0.16$ and $b = -0.43 \pm 0.16$. To constrain the projected extent of C IV, we calculate the transverse autocorrelation function. The flattening of the autocorrelation function at low r_{\perp} provides a lower limit for the coherence length of the metal enriched CGM—on the order of $200 h^{-1}$ comoving kpc. This physical size constraint allows us to refine our understanding of the metals in the CGM, where the extent of C IV in the CGM depends on gas flows, feedback, timescale of metal injection and mixing, and the mass of the host galaxies.

Unified Astronomy Thesaurus concepts: [Circumgalactic medium \(1879\)](#); [Intergalactic medium \(813\)](#); [Quasar absorption line spectroscopy \(1317\)](#)

Supporting material: figure set, machine-readable tables

1. Introduction

Models of Big Bang nucleosynthesis and stellar nucleosynthesis provide well-accepted constraints for the expected quantity of baryonic matter in the universe (O’Meara et al. 2006; Spergel et al. 2007; Cooke et al. 2018). However, surveys of the low-redshift universe have revealed that the baryon content of galaxies is far below these expectations (Fukugita & Peebles 2004). Furthermore, surveys of the high-redshift universe ($z > 1.5$) have revealed significant metal enrichment in the intergalactic medium (IGM) as well as in the circumgalactic medium (CGM, Cowie et al. 1995; Tytler et al. 1995; Péroux & Howk 2020). These discoveries, along with more recent studies (Peebles et al. 2014; Bordoloi et al. 2014; Werk et al. 2014; Tumlinson et al. 2017; Macquart et al. 2020), indicate that there are significant deposits of diffuse, enriched gas surrounding galaxies. Cataloging these CGM deposits along with their redshift, metallicity, and column density (Rafelski et al. 2012; Jorgenson et al. 2013) and connecting

those systems back to the associated galaxies (Tumlinson et al. 2013; Schroetter et al. 2016; Fumagalli et al. 2017; Lofthouse et al. 2020; Kaur et al. 2021; Rhodin et al. 2021; Schroetter et al. 2021) provides essential insight into galaxy evolution. The diffuse nature of the CGM makes it difficult to constrain the structure of its constituent gas clouds through direct imaging. Instead, the gas is typically studied in absorption to a bright background light source. Studies have used bright background galaxies (Cooke & O’Meara 2015; Péroux et al. 2018; Mortensen et al. 2021), quasars (Wagoner 1967; Wolfe et al. 1986; Chabanier et al. 2022), and gamma-ray bursts (Prochaska et al. 2009; Cucchiara et al. 2015) as sources. However, these single-sightline observations do not provide information about the characteristic spatial extent of the gas clouds, which constrains their origins, ranging from gas outflows, inflows, or reservoirs.

In recent years, researchers have begun to examine samples of quasar pairs (or galaxy pairs or galaxy–quasar pairs; Rauch et al. 1999; D’Odorico et al. 2006; Tytler et al. 2009; Chen et al. 2014; Martin et al. 2010; Steidel et al. 2010; Rubin et al. 2015, 2018; Lee et al. 2018). Paired background quasars broaden the analysis of the CGM by probing the gas at two different locations. This information can be used to measure the physical extent of the CGM via the transverse autocorrelation



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

function, providing constraints for cosmological simulations (Rahmati et al. 2013; Fumagalli et al. 2014; Peebles et al. 2019).

An effective tool for studying CGM and IGM enrichment is C IV absorption (e.g., Cooksey et al. 2010; Hasan et al. 2020). The C IV absorption lines form a doublet with rest wavelengths of $\lambda\lambda 1548.20, 1550.77 \text{ \AA}$. The characteristic wavelength separation makes the doublet easily identifiable in spectra, even those with moderate resolution. The ratio of the transitions’ oscillator strengths is 2:1, providing another avenue for verification. These features are both commonly found redward of Ly α emission and are redshifted into optical wavelengths for high redshifts, and so they are ideally positioned in wavelength space for the redshift range we investigate ($1.96 < z < 4.82$). Since the pioneering work by Chen et al. (2001), it has been recognized that C IV absorption traces the gas around galaxies (i.e., CGM) with a high covering fraction out to $\sim 100\text{--}200$ kpc and is kinematically consistent with residing in the same dark matter halo (Adelberger et al. 2005; Borthakur et al. 2013; Bordoloi et al. 2014; Turner et al. 2017; Rudie et al. 2019; Schroetter et al. 2021).

In this work, we combine a new quasar pair sample, presented here for the first time, with additional data from the literature. Using the combined sample, we compile a catalog of C IV absorbers and use the paired nature of our sample to calculate the C IV transverse autocorrelation function to constrain the coherence length of the CGM as probed by this ion. Characterizing the spatial structure of C IV absorption constrains the origin of the gas, distinguishing between different models of gas flows and mapping the spatial distribution of gas around galaxies (Chen et al. 2014; Fumagalli et al. 2014).

The paper is organized as follows. In Section 2, we discuss the data set—a new sample of medium-resolution spectra of 29 closely spaced quasar pairs. We supplement the sample with 32 additional quasar pairs: 3 from Rubin et al. (2015) and 29 from Martin et al. (2010), respectively. In Section 3, we describe the C IV absorber doublet search, including details about the selection and evaluation of candidate absorbers. In Section 4, we conduct a Monte Carlo simulation to assess the sensitivity and completeness of our doublet search. In Section 5, we calculate and compare the equivalent width frequency distributions for the Mintz et al. sample and the sample from Martin et al. (2010). In Section 6, we calculate the transverse autocorrelation function, which describes the number of paired absorbers found in paired quasar spectra compared to random expectation. We then discuss our results, suggesting a lower limit for the physical size of the high-redshift CGM traced by C IV, and conclude with a summary in Section 7. The following cosmology is used throughout: $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$.

2. Observations and Data Reduction

The quasar sample described in this paper is based on the sample from HST program 14127 (PI Fumagalli) that observed 47 quasar pairs at proper transverse separations of $r_\perp \leq 250$ kpc in the redshift interval $z \sim 2.0\text{--}2.6$ with the WFC3/G280 grism. The program was designed to search for intervening pairs of Lyman Limit Systems (LLSs) that were followed up with higher-resolution, ground-based spectroscopy to more precisely measure the redshifts of the LLSs and quasars as done in Lusso et al. (2018). The HST data sample was drawn

from a compilation of quasar pairs with $g^* < 21$ mag (Hennawi et al. 2006, 2010; Findlay et al. 2018). The high-resolution, ground-based spectroscopy forms the basis of the new sample of quasar pairs described below. The follow-up observations prioritized the brighter quasars. Six of the quasar spectra (three pairs) were previously published in Rubin et al. (2015), but because the targets are part of the HST sample, we include them in our sample. The final sample consists of 64 quasar spectra—corresponding to 32 pairs.

2.1. ESI

The 14 quasar pairs observed with the Echelle Spectrograph and Imager (ESI; Sheinis et al. 2002) at the Keck II telescope were previously presented in Lusso et al. (2018), who provide updated spectroscopic redshifts of the quasars. ESI is a fixed-format echelle that covers the observed wavelength range $\lambda \approx 4000\text{--}11000 \text{ \AA}$. These quasar pairs were observed in January and April 2017 under programs PID 2016B_N032E and 2017A_N133E, with seeing of $\sim 1''$. The $0''.75$ slit was used and both quasars were placed on the same slit when possible, while quasars with large separations were observed independently. The $0''.75$ slit corresponds to a resolution FWHM of $\sim 44 \text{ km s}^{-1}$, or $R \approx 7000$. When possible, we dithered along the slit to reduce fringing at the longest wavelengths ($\lambda > 8000 \text{ \AA}$). The ESI data were reduced as described in Rafelski et al. (2012) using the XIDL package ESIRedux (Prochaska et al. 2003) developed by J. X. Prochaska in IDL, and the spectra were extracted with optimal extraction.

2.2. X-Shooter

A subsample of 11 pairs included in this analysis has been observed with the X-SHOOTER spectrograph (Vernet et al. 2011) mounted on the Very Large Telescope as part of the ESO programs PID 099.A–0018 and 0100.A–0084 (PI Fumagalli). X-SHOOTER is a medium-resolution, cross-dispersed echelle spectrograph able to collect data over a large wavelength range (300–2500 nm) split across three arms (UVB, VIS, and NIR). For these observations, the instrument was configured with a $0''.8$ slit in the UVB arm and a $0''.9$ slit in the VIS and NIR arm, yielding respective resolutions of $\approx 6700, 8900, \text{ and } 5600$. Observations were collected in ESO Periods 99 and 100 in good weather conditions and with seeing between $0''.7\text{--}1''$. The UVB arm was reduced with the standard ESO pipeline (Freudling et al. 2013), and the VIS and NIR arms were reduced with family PyPeIt¹¹ (Prochaska et al. 2020). First, the individual frames are processed for basic calibration (bias, dark, and flat). Next, the wavelength calibrations are computed with the aid of arc frames, and the orders are identified and traced. Last, following sky subtraction, the spectra are then extracted and flux calibrated.

2.3. MagE

For seven quasar pairs in this sample, we make use of data collected with the Magellan Echelle (MagE, Marshall et al. 2008) spectrograph at the Magellan Telescope. Similarly to ESI, MagE is a fixed-format, cross-dispersed spectrograph covering in one exposure the range 300–1000 nm at a resolution of 4100 and 5800 for typically used slits of 0.7 or $1''$. Observations have

¹¹ <https://pypeit.readthedocs.io/en/latest/>

Table 1
Quasar Sample

Name	R.A. (J2000)	Decl. (J2000)	z_{em}	g mag	S/N	FWHM km s^{-1}	Inst	θ "	r_{\perp} h^{-1} comoving kpc
J0134+2430A	01:34:58.860	+24:30:50.57	2.105	20.42	10.0	55.9	ESI	3.69	66.55
J0134+2430B	01:34:59.018	+24:30:47.57	2.095	19.85	15.6	55.9	ESI	3.69	66.55
J0735+2957A	07:35:22.429	+29:57:10.17	2.082	20.57	6.9	55.9	ESI	5.39	96.71
J0735+2957B	07:35:22.555	+29:57:04.99	2.065	20.42	9.6	55.9	ESI	5.39	96.71
J0813+1014A	08:13:29.491	+10:14:05.25	2.098	19.39	19.9	54.5	MagE	7.14	128.32
J0813+1014B	08:13:29.708	+10:14:11.62	2.071	20.09	15.1	54.5	MagE	7.14	128.32
J0846+2709A	08:46:24.330	+27:09:58.44	2.203	20.67	4.9	55.9	ESI	4.71	87.13
J0846+2710B	08:46:24.505	+27:10:02.45	2.195	20.89	6.5	55.9	ESI	4.71	87.13
J0852+3500A	08:52:30.221	+35:00:03.66	2.238	20.39	9.8	55.9	ESI	5.28	98.51
J0852+3459B	08:52:30.532	+35:00:00.03	2.235	19.98	10.3	55.9	ESI	5.28	98.51
J0920+1311A ^a	09:20:56.009	+13:11:02.66	2.427	19.20	35.5	60.0	MagE	6.23	121.58
J0920+1310B ^a	09:20:56.237	+13:10:57.42	2.446	19.29	26.3	60.0	MagE	6.23	121.58
J0937+1509A	09:37:47.249	+15:09:28.02	2.555	20.05	12.7	54.5	MagE	11.75	234.45
J0937+1509B	09:37:47.409	+15:09:39.54	2.541	19.93	14.5	54.5	MagE	11.75	234.45
J1002+3531A	10:02:33.904	+35:31:27.58	2.305	18.87	18.2	55.9	ESI	3.91	74.29
J1002+3531B	10:02:34.211	+35:31:28.68	2.320	20.04	6.2	55.9	ESI	3.91	74.29
J1045+4041A	10:45:33.318	+40:41:38.01	2.261	20.30	9.1	55.9	ESI	17.03	320.87
J1045+4041B	10:45:33.544	+40:41:21.15	2.286	20.25	9.0	55.9	ESI	17.03	320.87
J1056-0059A ^a	10:56:44.883	-00:59:33.43	2.135	20.12	14.4	54.5	MagE	7.21	131.20
J1056-0059B ^a	10:56:45.248	-00:59:38.11	2.126	20.96	8.7	54.5	MagE	7.21	131.20
J1104+2907A	11:04:30.006	+29:07:53.48	2.133	20.35	9.6	55.9	ESI	6.33	115.07
J1104+2907B	11:04:30.348	+29:07:49.02	2.127	20.48	11.4	55.9	ESI	6.33	115.07
J1139+4143A	11:39:47.063	+41:43:51.15	2.202	20.21	8.4	55.9	ESI	2.39	44.41
J1139+4143B	11:39:47.259	+41:43:52.10	2.239	19.98	11.5	55.9	ESI	2.39	44.41
J1150+0453A	11:50:31.143	+04:53:53.26	2.527	20.62	8.6	54.5	MagE	6.97	138.51
J1150+0453B	11:50:31.543	+04:53:56.85	2.517	20.52	27.0	54.5	MagE	6.97	138.51
J1236+5220A	12:36:35.145	+52:20:59.08	2.567	20.62	9.9	55.9	ESI	3.08	61.82
J1236+5220B	12:36:35.422	+52:20:57.33	2.578	20.50	10.7	55.9	ESI	3.08	61.82
J1332+2523A	13:32:09.268	+25:23:01.36	2.080	20.15	8.1	55.9	ESI	7.92	142.42
J1332+2523B	13:32:09.690	+25:23:06.83	2.093	20.11	6.2	55.9	ESI	7.92	142.42
J1345+2625A	13:45:43.640	+26:25:06.94	2.038	20.27	12.1	55.9	ESI	9.22	163.35
J1345+2625B	13:45:44.316	+26:25:05.35	2.016	19.96	14.0	55.9	ESI	9.22	163.35
J1431+2705A	14:31:04.647	+27:05:24.64	2.266	19.81	4.4	55.9	ESI	5.94	111.65
J1431+2705B	14:31:04.977	+27:05:28.63	2.263	20.26	6.2	55.9	ESI	5.94	111.65
J1549+3136A	15:49:38.172	+31:36:46.89	2.520	20.12	11.5	55.9	ESI	12.96	256.85
J1549+3136B	15:49:38.496	+31:36:34.61	2.502	19.29	21.1	55.9	ESI	12.96	256.85
J1613+0808A ^a	16:13:01.693	+08:08:06.05	2.382	19.57	20.1	60.0	MagE	9.64	186.12
J1613+0808B ^a	16:13:02.033	+08:08:14.26	2.387	18.94	32.9	60.0	MagE	9.64	186.12
J1637+2636A	16:37:00.877	+26:36:13.73	1.965	20.70	3.3	55.9	ESI	3.86	67.14
J1637+2636B	16:37:00.925	+26:36:09.92	1.961	19.39	10.4	55.9	ESI	3.86	67.14
J2103+0646A	21:03:29.249	+06:46:53.33	2.574	20.61	19.0	54.5	MagE	3.81	76.31
J2103+0646B	21:03:29.372	+06:46:49.99	2.565	20.20	12.7	54.5	MagE	3.81	76.31
J1338+0010A	13:38:31.532	+00:10:56.24	2.309	20.57	11.2	44.7	X-SHOOTER	11.65	220.93
J1338+0011B	13:38:31.962	+00:11:05.93	2.296	20.94	10.0	44.7	X-SHOOTER	11.65	220.93
J1339+1310A ^b	13:39:07.139	+13:10:39.65	2.239	18.71	46.7	44.7	X-SHOOTER	1.71	31.87
J1339+1310B ^b	13:39:07.235	+13:10:38.69	2.237	19.13	72.2	44.7	X-SHOOTER	1.71	31.87
J1421+1630A	14:21:49.008	+16:30:27.13	2.463	20.77	13.2	44.7	X-SHOOTER	10.02	196.51
J1421+1630B	14:21:48.794	+16:30:17.59	2.454	20.44	16.5	44.7	X-SHOOTER	10.02	196.51
J1515+1511A ^b	15:15:38.477	+15:11:34.84	2.052	18.60	30.0	44.7	X-SHOOTER	2.01	35.78
J1515+1511B ^b	15:15:38.592	+15:11:35.95	2.051	18.22	24.5	44.7	X-SHOOTER	2.01	35.78
J2248+0307A	22:48:56.831	+03:07:00.25	2.394	20.88	6.3	44.7	X-SHOOTER	5.94	114.88
J2248+0307B	22:48:57.226	+03:06:59.54	2.392	20.62	4.2	44.7	X-SHOOTER	5.94	114.88
J2348+0057A	23:48:19.584	+00:57:21.50	2.158	18.92	8.7	44.7	X-SHOOTER	7.07	129.42
J2348+0057B	23:48:19.191	+00:57:17.59	2.152	20.77	20.6	44.7	X-SHOOTER	7.07	129.42
J0344+1015A	03:44:07.031	+10:15:20.52	2.007	19.55	10.8	44.7	X-SHOOTER	12.13	213.55
J0344+1015B	03:44:06.644	+10:15:09.83	2.002	20.54	9.6	44.7	X-SHOOTER	12.13	213.55
J1016+2224A	10:16:52.884	+22:24:12.12	2.023	20.29	17.6	44.7	X-SHOOTER	14.83	262.20
J1016+2224B	10:16:53.946	+22:24:13.70	2.020	20.94	10.7	44.7	X-SHOOTER	14.83	262.20
J2214+1326A	22:14:27.034	+13:26:57.01	2.007	20.37	11.8	44.7	X-SHOOTER	5.84	102.77
J2214+1326B	22:14:26.793	+13:26:52.34	1.995	20.60	14.0	44.7	X-SHOOTER	5.84	102.77
J2243-0613A	22:43:25.048	-06:13:50.33	2.586	20.84	9.8	44.7	X-SHOOTER	9.44	189.61
J2243-0613B	22:43:25.681	-06:13:51.00	2.578	19.14	14.8	44.7	X-SHOOTER	9.44	189.61

Table 1
(Continued)

Name	R.A. (J2000)	Decl. (J2000)	z_{em}	g mag	S/N	FWHM km s^{-1}	Inst	θ "	r_{\perp} h^{-1} comoving kpc
J2242+0558A	22:42:04.632	+05:58:30.44	2.510	20.62	12.9	44.7	X-SHOOTER	4.27	84.44
J2242+0558B	22:42:04.373	+05:58:28.63	2.503	20.97	16.2	44.7	X-SHOOTER	4.27	84.44

Notes. The new sample of quasar pairs and the three pairs from Rubin et al. (2015). z_{em} is the emission redshift of the quasar, g mag the quasar magnitudes from Hennawi et al. (2006, 2010) and Findlay et al. (2018), S/N the signal-to-noise ratio per pixel, FWHM the full width at half maximum of the instrument, Inst the instrument, θ the angular separation of the paired quasars, and r_{\perp} the transverse separation in comoving units.

^a Spectrum published in Rubin et al. (2015).

^b Gravitational lens.

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

been collected through the years (2008–2015) with good weather conditions. Data have been reduced using the MASE pipeline (Bochanski et al. 2009), which follows a workflow very similar to that of the ESIRedux package described above. Three of the quasar pairs’ spectra were previously published in Rubin et al. 2015, as indicated in Table 1.

2.4. Complete Sample

The complete new sample of quasar pairs described above is summarized in Table 1 and will be referred to throughout as the Mintz et al. sample. It consists of 64 medium-resolution quasar spectra corresponding to 32 quasar pairs, three of which were published previously in Rubin et al. (2015). The quasars have emission redshifts ranging from $1.96 \leq z \leq 2.64$ and the pairs have projected separations of $32 \leq r_{\perp} (h^{-1} \text{ kpc}) \leq 320$ (comoving). Throughout this work, we use the quasar emission redshifts from Findlay et al. (2018), who estimated an uncertainty of $\approx 1000 \text{ km s}^{-1}$, but Lusso et al. (2018) have since calculated more precise measurements for most quasars in the sample. As our analysis deals exclusively with intervening systems, a slight difference in redshift would have no effect on our results.

We supplement our data set with additional Keck/ESI spectra of quasar pairs from Martin et al. (2010), provided by the authors. The sample from Martin et al. (2010) includes 55 quasar spectra corresponding to 29 quasar pairs (the sample includes one quasar triplet that provides three quasar pairs). The resolution of the Martin et al. sample is similar to but slightly lower than that of our sample, due to their use of a wider $1''$ slit compared to our use of a $0''.75$ one. Their sample is also different in a number of other key aspects: it probes a higher redshift range of $2.46 \leq z \leq 4.82$, it covers a higher r_{\perp} range of $88 \leq r_{\perp} (h^{-1} \text{ kpc}) \leq 3350$ (comoving), and overall has a higher signal-to-noise ratio (S/N) with a median S/N per pixel of 18.5 compared to a median S/N per pixel for our sample of 11.5. The higher S/N can be attributed in part to the lower resolution. The differences in redshift and r_{\perp} coverage are shown in Figure 1. While Martin et al. (2010) performed their own doublet search, to obtain a maximally self consistent catalog, we combine the Martin et al. quasars with our sample and search for C IV absorbers in all of the spectra using the method described in Section 3. We note that we do not include the sample by Tytler et al. (2009), as it is at lower resolution and has a significantly larger pair separation with a median separation of 1 Mpc.

Throughout this paper, unless otherwise explicitly noted, when we refer to the quasar pair sample we are referring to the

combined sample of 119 quasar spectra or 61 quasar pairs. When we wish to refer to each of the subsamples separately, for simplicity we will reference them as Mintz et al. and Martin et al. (Martin et al. 2010) respectively.

2.5. Continuum Fitting

Each of the spectra in the Mintz et al. sample are continuum fitted using the `lt_continuumfit` GUI from the `line-tools` package in Python (Prochaska et al. 2016). The GUI provided an initial guess using spline knots and then the initial fits are adjusted manually. The spectra from Martin et al. (2010) were already continuum normalized by the authors.

3. C IV Doublet Search

Each of the quasar spectra are searched automatically for candidate C IV absorbers, identified based on the characteristic wavelength separation of the C IV doublet. All of the candidates are then evaluated based on specific characteristics of C IV systems and manually assessed to account for and correct blends between C IV and other absorption lines. This section details the C IV detection process and the compilation of the absorber catalog. The line search we adopt is based on the approaches of Cooksey et al. (2008, 2010), and Mas-Ribas et al. (2018), and has been further optimized for our data and methodology.

3.1. Absorption Feature Detection

The automated doublet search first identifies all absorption features in the continuum normalized spectra. The spectra are smoothed using a Gaussian convolution with width equal to the FWHM of the instrument as provided in Table 1. Absorption candidates are identified as groups of three or more adjacent pixels with $f_G \leq 1 - \sigma_f$ where f_G is the Gaussian convolved flux and σ_f its uncertainty. Absorption features are then selected from the candidates if they have $W_{\text{obs}} \geq 3\sigma_{W_{\text{obs}}}$, where

$$W_{\text{obs}} = \sum_{\lambda_l}^{\lambda_h} (1 - f_{\lambda_l}) \Delta \lambda_l \quad (1)$$

and

$$\sigma_{W_{\text{obs}}}^2 = \sum_{\lambda_l}^{\lambda_h} \sigma_{\lambda_l}^2 \Delta \lambda_l^2, \quad (2)$$

bounded by the feature’s lower and upper wavelength limits: λ_l and λ_h . The bounds are defined by the first and last contiguous pixels that satisfy the flux requirements as described above. For

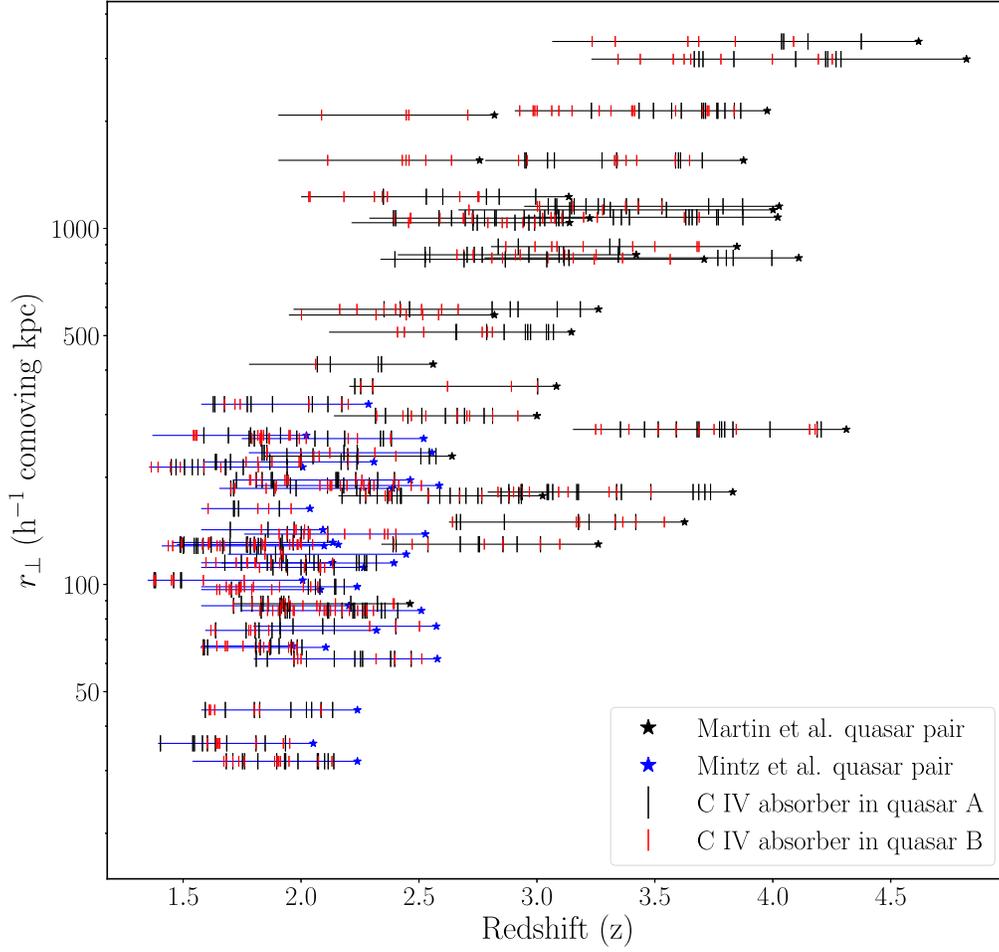


Figure 1. Quasar pairs’ search paths and detected C IV redshifts are shown as a function of pairs’ projected separations. The search path of each pair is plotted using the minimum of the two spectra’s lower limits and the maximum of the two quasar’s emission redshifts. The stars mark the higher emission redshift of the two quasars in a pair. The C IV absorbers found in the spectra of quasar A of a pair are shown in black and those found in quasar B are shown in red. The pairs from the Mintz et al. sample are plotted in blue, while those from Martin et al. (2010) are plotted in black. The differences in the two samples, including the r_{\perp} ranges, redshift ranges, and lengths of the search paths, are clearly visible in this plot.

each accepted absorption feature, we calculate its optical depth weighted C IV 1548 and C IV 1550 redshifts defined as

$$1 + z_{abs} = \frac{\sum_{\lambda_i}^{\lambda_h} \lambda_i \ln\left(\frac{1}{f_{\lambda_i}}\right)}{\lambda_r \sum_{\lambda_i}^{\lambda_h} \ln\left(\frac{1}{f_{\lambda_i}}\right)}, \quad (3)$$

with $\lambda_r = 1548.195 \text{ \AA}$ and $\lambda_r = 1550.770 \text{ \AA}$, respectively (Cooksey et al. 2008). Both the 1548 and 1550 redshifts are calculated, because each absorption feature is considered as a candidate for both the C IV 1548 and C IV 1550 line in the next step of the automatic doublet candidate detection. For the same reason, each absorption feature’s rest-frame equivalent width is calculated for both redshifts as

$$W_r = \frac{1}{1 + z_{abs}} W_{obs} \quad (4)$$

and

$$\sigma_{W_r}^2 = \frac{1}{(1 + z_{abs})^2} \sigma_{obs}^2. \quad (5)$$

The final absorption line catalog is used to identify C IV doublet candidates.

3.2. C IV Candidate Selection

The search for C IV doublets is limited to the region of the spectrum between the wavelength of the Ly α emission line (to avoid the Ly α forest, which hinders line identification) and 100 \AA redward of the wavelength of C IV 1548 at the quasar’s redshift. We extend the search slightly beyond the emission redshift, to account for high uncertainty in the quasar redshifts and for possible peculiar velocities of the probed gas. We identify 37 systems above the emission redshift in a total of 33 quasar spectra, which are included in the catalog but flagged as proximate systems. In some cases, the wavelength coverage of a spectrum does not extend to the Ly α emission line or to the wavelength of C IV 1548 at the quasar’s redshift, and in these cases the searchable path length is shorter. The search path also excludes regions with consistent contamination from skylines (i.e., $6860 \text{ \AA} < \lambda < 6930 \text{ \AA}$ and $7590 \text{ \AA} < \lambda < 7690 \text{ \AA}$).

C IV absorber candidates are selected based only on the characteristic doublet separation. Each absorption feature detected as described in Section 3.1 is assumed to be the C IV 1548 line, and the limits $\lambda_{i,1550}$ and $\lambda_{h,1550}$ of the potential

C IV 1550 line are calculated as

$$\lambda_{l,1550} = \lambda_{l,1548} \left(\frac{\lambda_{r,1550}}{\lambda_{r,1548}} \right) \quad (6)$$

and

$$\lambda_{h,1550} = \lambda_{h,1548} \left(\frac{\lambda_{r,1550}}{\lambda_{r,1548}} \right). \quad (7)$$

If a detected absorption feature falls in this range, the two features are selected as a C IV absorber candidate and added to the candidate catalog. If a single feature is sufficiently broad as to cover the C IV 1550 limits, it is included in the catalog and assessed manually as described in Section 3.3. This stage of candidate selection is intentionally designed to be generous in order to include systems that are exceedingly broad or blended with other lines.

3.3. C IV Candidate Evaluation

The C IV absorber candidates identified in Section 3.2 are then evaluated based on two additional characteristics of the C IV doublet: the equivalent width ratio and the velocity separation of the doublet lines. The equivalent width ratio and associated uncertainty are calculated as

$$R_W = \frac{W_{r,1548}}{W_{r,1550}} \quad (8)$$

and

$$\sigma_{R_W}^2 = R_W^2 \left[\left(\frac{\sigma_{W_{r,1548}}}{W_{r,1548}} \right)^2 + \left(\frac{\sigma_{W_{r,1550}}}{W_{r,1550}} \right)^2 \right]. \quad (9)$$

The C IV 1548 line is expected to be twice as strong as the C IV 1550 line in the unsaturated regime, and so the C IV candidates that do not satisfy the following criteria are not included as likely C IV absorbers:

$$1 - 3\sigma_{R_W} \leq R_W \leq 2 + 3\sigma_{R_W}. \quad (10)$$

The lower limit of this criteria accounts for the possibility that the two lines are saturated and therefore could have nearly equal equivalent widths. We note that our 3σ tolerance is slightly higher than those used in Cooksey et al. (2010) and Mas-Ribas et al. (2018). These studies did not perform a comprehensive visual inspection of the sightlines and so required more stringent selection criteria than our search. We inspect each sightline after the automatic search, manually correcting for blends and including multicomponent systems that were not previously identified. The tolerance levels were chosen to maximize the number of true systems found, while minimizing false positives.

The two lines of each doublet are expected to have the same redshift. To further select based on the lines' redshift difference, the velocity separation $\delta v_{\text{C IV}}$ is calculated as

$$\Delta v_{\text{C IV}} = c \left(\frac{z_{1550} - z_{1548}}{1 + z_{1548}} \right). \quad (11)$$

Systems with $|\Delta v_{\text{C IV}}| > 20 \text{ km s}^{-1}$ are not included as likely C IV absorbers. Again, this tolerance is higher than those used

in Cooksey et al. (2010) and Mas-Ribas et al. (2018), for the reasons described above.

In addition to the automatic evaluation based on R_W and $\delta v_{\text{C IV}}$, the entire search path is also visually assessed. Occasionally, the presence of blends or multiple velocity components leads to an incorrect evaluation of a candidate doublet. Such systems are manually flagged and the wavelength limits of their equivalent width integration are adjusted where appropriate. We link C IV absorber components separated by less than 250 km s^{-1} based on the distribution of damped Ly α C IV velocity widths reported in Prochaska et al. (2019). An example of a full C IV doublet search is shown in Figure 2.

We compile a catalog of 555 likely C IV absorbers in the Mintz et al. sample, from a candidate list of 2567. As our primary area of interest in this study regards the intervening C IV absorbers, we flag any systems within 5000 km s^{-1} of the quasar redshift as proximate systems and remove them from subsequent analysis. We find 85 proximate systems, leaving 470 intervening C IV absorbers with redshifts in the range $1.36 \leq z \leq 2.58$ and a median redshift of 1.91.

In the Martin et al. quasar sample, we find 524 likely C IV absorbers from a candidate list of 1224. Of the likely absorbers, 88 are proximate systems, leaving a total of 436 intervening absorbers with redshifts of $1.74 \leq z \leq 4.38$. The median redshift is 2.96. The acceptance fraction of likely absorbers from the candidate list is much higher in the Martin et al. sample, due to the higher S/N of the quasars as compared to the Mintz et al. sample. This is discussed further in Section 4. We also note that our absorber catalog compiled from the Martin et al. sample is $\sim 30\%$ larger than the catalog found in Martin et al. (2010). Our methods for detecting C IV absorbers are not identical, and so it is expected that the resulting catalogs would differ as well. For example, we do not require similar velocity structure in the two doublet lines as in Martin et al. (2010), we do not separate multiple velocity components of absorbers, and we use a different linking length chosen as described above based on the results of Prochaska et al. (2019). We refer the reader to Martin et al. (2010) for more details on their method. The variations in methodology and definition of C IV absorbers resulting in different sample sizes do not affect the results of the equivalent width distribution function or the autocorrelation as discussed in Sections 5 and 6, because the same methodological differences are applied in the sensitivity assessment. In Section 5, we find that the completeness-corrected distribution agrees with the Martin et al. (2010) results.

Altogether, we compile a catalog of 906 intervening C IV absorbers. In all analysis that follows, the C IV absorber catalog refers to this combined catalog described in Table 2, whose $W_{r,1550}$ distribution is shown in Figure 3.

4. Sensitivity and Completeness

In order to conduct a proper analysis of our C IV absorber catalog, we must address the detection sensitivity of our spectra and the completeness of our absorber catalog. We follow the methods of Cooksey et al. (2010) and perform a Monte Carlo simulation to assess the sensitivity of the process described in Section 3. First, each spectrum is cleaned of absorption features. The features are identified as described in Section 3.1 and replaced with flux randomly selected from the region adjacent to the feature so that the cleaned spectra accurately

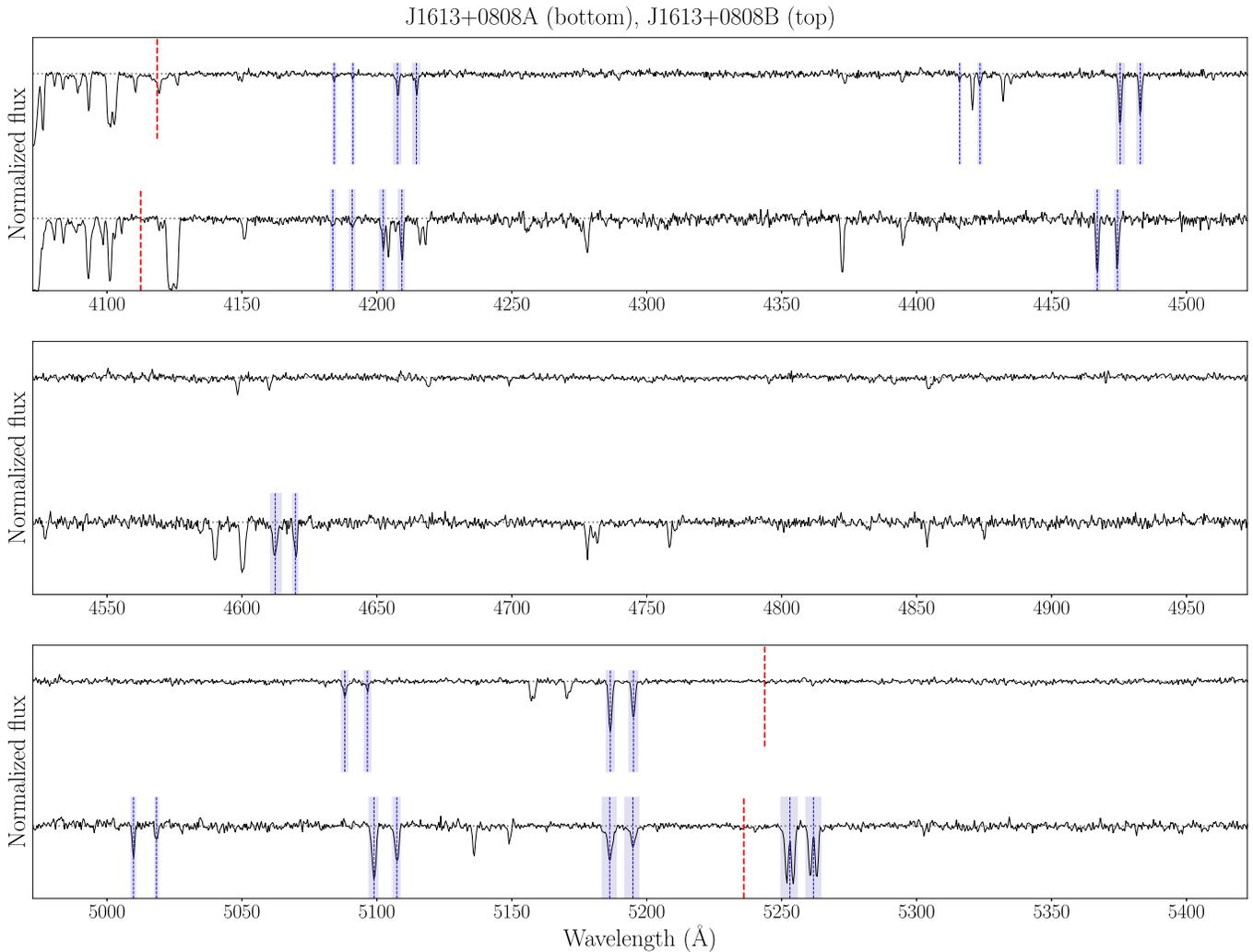


Figure 2. An example of a C IV doublet search. The normalized flux is plotted in black, for the spectrum of J1613+0808A on the bottom and that of its pair, J1613+0808B, on top. The normalized continuum is shown as a dashed gray line. The lower limit of the search and the wavelength of C IV in the quasar’s rest frame are marked by vertical, dashed red lines. The lower limit is determined by the location of the quasar’s Ly α emission line, and the upper limit extends 100 Å beyond the C IV wavelength at the quasar’s redshift. This extension accounts for uncertainty in the quasars’ emission redshifts and possible redshifting of C IV absorbers due to peculiar velocities. The likely C IV absorbers are indicated by the shaded blue regions, which span the lines’ equivalent width integration limits. The absorber redshift of the 1548 and 1550 Å lines are shown with dotted, vertical blue lines. We see several examples of paired absorbers in this quasar pair—around 4190 and 5190 Å. The absorbers around 4200, 4470, and 5100 Å demonstrate the importance of the velocity width tolerance δv_{pair} discussed in Section 6.1. (The complete set (61 elements) of doublet search plots are available in the figure set online.)

(The complete figure set (61 images) is available.)

represent the noise of the original data. We choose to remove all absorption lines because we handled blends by eye in the construction of our absorber catalog.

Each spectrum is then injected with simulated C IV doublets. The redshifts and column densities of the doublets for a given spectrum are selected so that they fill a grid of equally spaced redshift bins of width 0.01 covering the search-path range and equally spaced log column density bins ranging from $12.5 < \log N(\text{C IV}) [\text{cm}^{-2}] < 15.5$. For each spectrum, 100 simulated doublets are drawn from each redshift-column density bin. For example, in one such simulation, there are 100 doublets with $13.5 < \log N(\text{C IV}) [\text{cm}^{-2}] < 13.6$ and $2.00 < z_{1548} < 2.01$. The Doppler parameters of the systems are drawn randomly and range from 10 km s^{-1} to 35 km s^{-1} , reflecting the width of the real data systems. Each doublet is also randomly assigned up to five velocity components with randomly selected fractions of the total column density and

velocity offsets up to 50 km s^{-1} . The equivalent width of each simulated system is calculated by integrating over the doublet’s flux prior to adding the spectral noise.

To assess the detection sensitivity, each simulated doublet is injected into the cleaned spectrum, which is then run through the C IV absorber candidate selection process described in Section 3. The candidates are not evaluated automatically as in Section 3.3 based on Equations (10) and (11), because these are intended to exclude false-positive detections—i.e., noise or other absorption lines that are not truly C IV doublets. Given that all of the absorption lines in the simulated spectra are by definition C IV, we do not apply these criteria. Furthermore, true systems with blends or multiple components may be discarded in the automatic evaluation. While we would recover these systems with our manual inspection in the true data, it is impossible to do so for the thousands of simulated spectra. Therefore, relaxing the selection slightly by omitting the

Table 2
C IV Absorbers

Quasar	z_{abs}	$W_{r,1548}$ m Å	$\sigma_{W_{r,1548}}$ m Å	$W_{r,1550}$ m Å	$\sigma_{W_{r,1550}}$ m Å	R_W	σ_{R_W}	$\delta v_{C\ IV}$ km s ⁻¹	Proximate
J0134+2430A	1.585	445	86	207	42	2.15	0.4	13.3	False
J0134+2430A	1.590	406	72	248	60	1.64	0.4	6.9	False
J0134+2430A	1.604	182	33	93	29	1.94	0.5	-15.3	False
J0134+2430A	1.827	113	21	173	22	0.66	0.2	-5.3	False
J0134+2430A	1.868	51	11	80	15	0.63	0.2	9.1	False
J0134+2430A	1.959	76	19	111	20	0.68	0.3	13.7	False
J0134+2430A	1.984	495	19	380	18	1.30	0.07	-12.3	False
J0134+2430A	2.004	68	13	68	12	1.00	0.3	4.1	False
J0134+2430A	2.089	56	9	66	9	0.85	0.2	-9.4	True
J0134+2430A	2.094	32	6	31	6	1.05	0.3	4.1	True

Note. The C IV absorber catalog. z_{abs} is the absorber redshift defined as the redshift of the $W_{r,1548}$ line. $W_{r,1548}$, $\sigma_{W_{r,1548}}$, $W_{r,1550}$, $\sigma_{W_{r,1550}}$ are the rest-frame equivalent widths and uncertainties of the 1548 and 1550 lines, respectively. R_W and σ_{R_W} are the equivalent width ratio and its uncertainty. $\delta v_{C\ IV}$ is the velocity difference between the 1548 and 1550 lines. The Proximate column indicates if the absorber is located within 5000 km s⁻¹ of the quasar. (We show only the first 10 absorbers, but this table is available in its entirety in machine-readable form.)

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

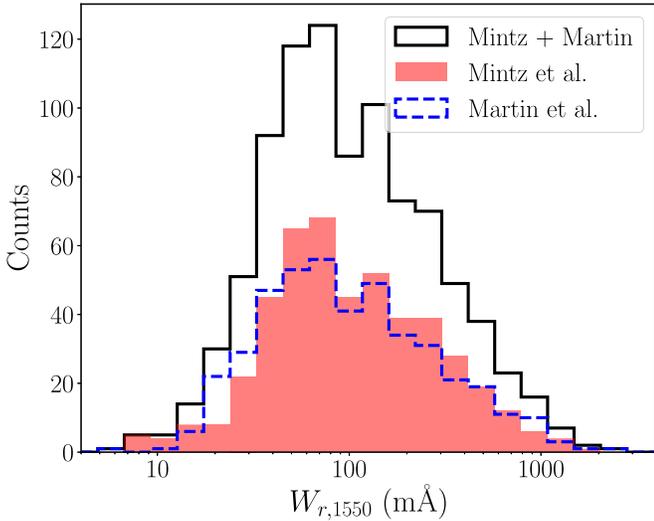


Figure 3. A histogram of the $W_{r,1550}$ of the final C IV absorber catalog. The $W_{r,1550}$ distribution for absorbers from the Mintz et al. sample is shown in a shaded red histogram. The distribution for absorbers from the Martin et al. sample is shown with a dashed blue step histogram. The combined distribution is shown with a black step histogram. Despite the higher S/N of the Martin et al. sample, the two samples have similar lower $W_{r,1550}$ detection limits due to the higher resolution of the Mintz et al. sample.

automatic evaluation more accurately represents the conditions of the search on the actual data.

To measure the searchable path as a function of column density and equivalent width, we first convert from redshift space to comoving absorption path length space (X) following Cooksey et al. (2010), using $\delta z = 0.01$ from the Monte Carlo simulation:

$$X(z) = \frac{2}{3\Omega_M} \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda}, \quad (12)$$

$$\delta X(z) = X(z + 0.5\delta z) - X(z - 0.5\delta z). \quad (13)$$

We calculate the total searchable path for a given system strength as a weighted sum over the redshift bins, multiplying the width of the bin by the fraction of doublets of that strength detected in the simulation at that redshift. We define $C_i(\log N)$

to be this fraction in redshift bin i for a system with column density of $\log N$ with equivalent definitions for $W_{r,1548}$ and $W_{r,1550}$. The total path length (δX) as a function of column density (N) or equivalent width is then calculated as a sum over redshift bins i :

$$\Delta X(\log N) = \sum_i \delta X * C_i(\log N), \quad (14)$$

$$\Delta X(W_{r,1548}) = \sum_i \delta X * C_i(W_{r,1548}), \quad (15)$$

$$\Delta X(W_{r,1550}) = \sum_i \delta X * C_i(W_{r,1550}). \quad (16)$$

The total searchable path length function for each of the quasar pair samples is calculated as the sum of all individual searchable path length functions. The results of a single spectrum's sensitivity analysis are shown in Figure 4, and the total completeness as a function of equivalent width is shown for both the Mintz et al. and the Martin et al. samples in Figure 5. This completeness function further illustrates the difference in S/N between the two samples as discussed above. The Martin et al. sample is generally less noisy than the Mintz et al. sample, due to the higher S/N (although at lower resolution), and so has higher completeness at all equivalent widths.

5. Distribution Function

Using the results from our Monte Carlo sensitivity assessment from Section 4, we calculate the equivalent width frequency distribution of C IV 1550 for the Mintz et al. sample and the Martin et al. sample, which is the number of systems per unit redshift path per unit rest equivalent width. The distribution function represents a nearly comprehensive summary of the absorption line survey. Following Martin et al. (2010), we perform our analysis in terms of equivalent width, and not column density, as converting saturated systems' equivalent widths only provides a lower limit on column density. Additionally, to be conservative we use the equivalent width of the 1550 line instead of the 1548 line, because it is the weaker of the two and detection of a doublet therefore depends on the detection of the 1550 line.

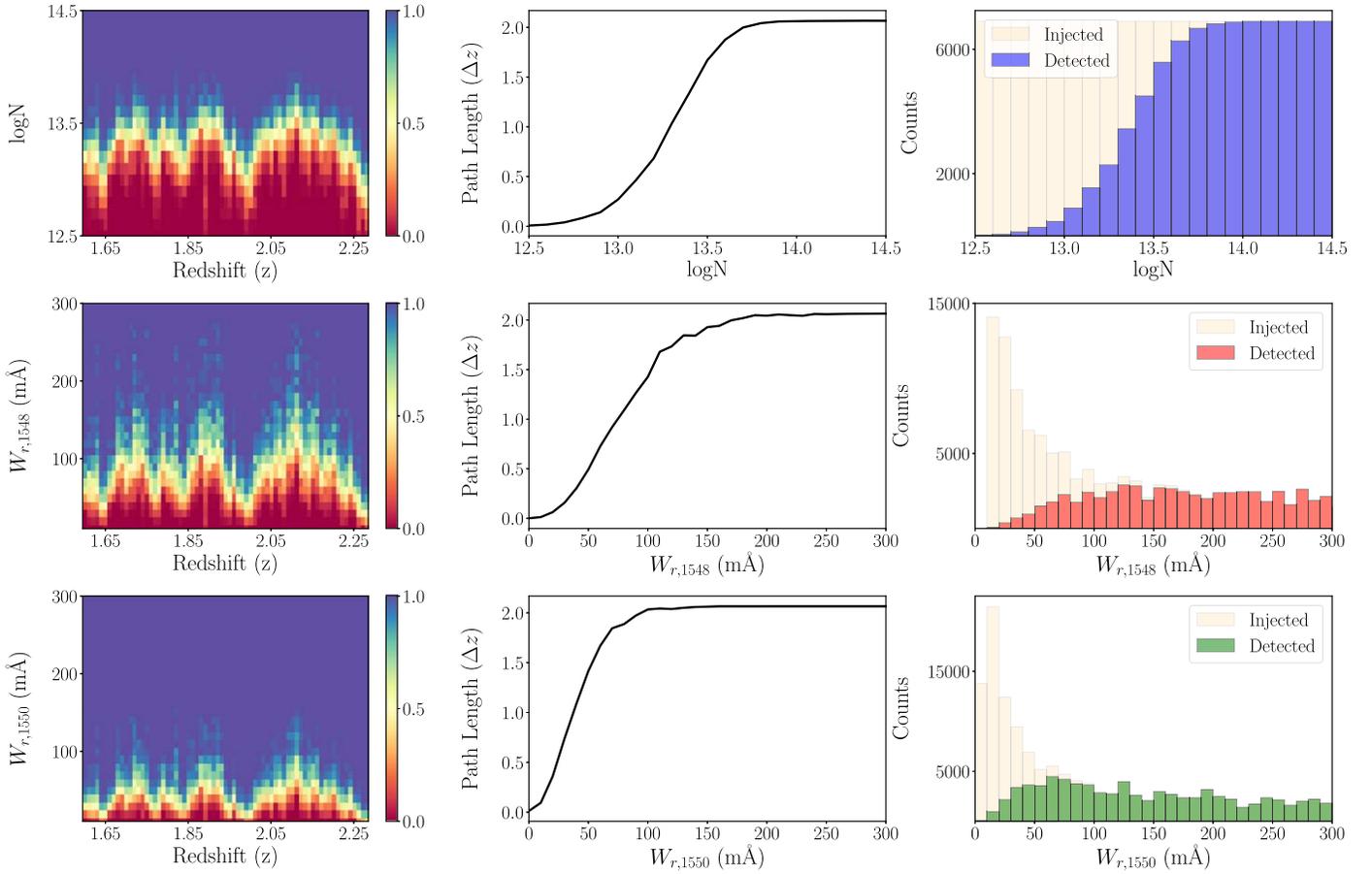


Figure 4. An example of the completeness assessment for J100233.9+353127.5. The upper row shows the completeness in terms of $\log N$, the middle row in terms of $W_{r,1548}$, and the bottom row in terms of $W_{r,1550}$. The leftmost column shows the completeness fraction as a function of redshift and system strength, with bluer values indicating higher completeness. The middle column shows the total searchable path length in units of δz as a function of $\log N$ or W_r . The rightmost column shows the number of injected vs. detected systems in the completeness simulation. Note that the number of injected systems is constant in $\log N$ but not in W_r , due to the injection method.

The equivalent width distribution $f(W_r)$ is calculated following the method of Cooksey et al. (2010):

$$f(W_r) = \frac{\Delta \mathcal{N}}{\Delta W_r \Delta X(W_r)}, \quad (17)$$

where $\Delta \mathcal{N}$ is the number of absorbers per equivalent width bin, δW_r is the width of the equivalent width bin, and $\delta X(W_r)$ is the searchable path length sensitive to absorbers of equivalent width W_r calculated as described in Section 4. We calculate the uncertainty in the distribution function using the 84.1% Poisson confidence limits (corresponding to 1σ for Gaussian statistics) for $\Delta \mathcal{N}$.

It is well-established that the distribution function can be described by a power law (Cooksey et al. 2010; Martin et al. 2010). Therefore, we use a Markov Chain Monte Carlo sampler to fit a linear model to our equivalent width distributions in log-log space using the width of the equivalent width bins as the uncertainty in the x -direction. For the fit, we use only the $f(W_r)$ bins with W_r between 30 and 500 mÅ in order to avoid bias due to incompleteness and to compare with the results in Martin et al. (2010). For the Mintz et al. sample, we obtain parameters for the fit to $\log f(W_r) = m \log(W_r) + b$ of $m = -1.5 \pm 0.2$ and $b = -0.3 \pm 0.2$ for equivalent width in Å and $b = 4.1 \pm 0.4$ for equivalent width in mÅ. For the Martin et al. sample, we found $m = -1.4 \pm 0.2$ and

$b = -0.5 \pm 0.2$ for equivalent width in Å and $b = 3.7 \pm 0.4$ for equivalent width in mÅ. For both samples together, we found $m = -1.44 \pm 0.16$ and $b = -0.43 \pm 0.16$ for equivalent width in Å and $b = 3.9 \pm 0.3$ for equivalent width in mÅ. These fit values agree with those reported in Martin et al. (2010), providing additional support for the consistency of our methodologies despite slight differences in our absorber detection strategies and catalogs. We show the distribution functions and their best fits in Figure 6. We note that we do not expect the distribution function to be biased by our paired sightlines, as systems found at similar redshifts in both sightlines of a pair do not necessarily have similar strengths. To confirm this, we repeated the analysis described above for half of the sample, excluding one sightline of every pair, and obtained consistent results.

Cooksey et al. (2013) and Hasan et al. (2020) studied the evolution of the C IV distribution function with redshift using large samples of single-sightline background quasars. Their studies used C IV catalogs with thousands of absorbers with redshifts $1.0 < z < 4.75$. Cooksey et al. (2013) used low-resolution SDSS spectra while Hasan et al. (2020) worked with high-resolution spectra from Keck/HIRES and VLT/UVES. They both found evidence of higher C IV abundance at lower redshifts, which corresponds to a larger value for the intercept of the linear model. While slightly overlapping in redshift space, the Martin et al. sample is generally at a higher redshift

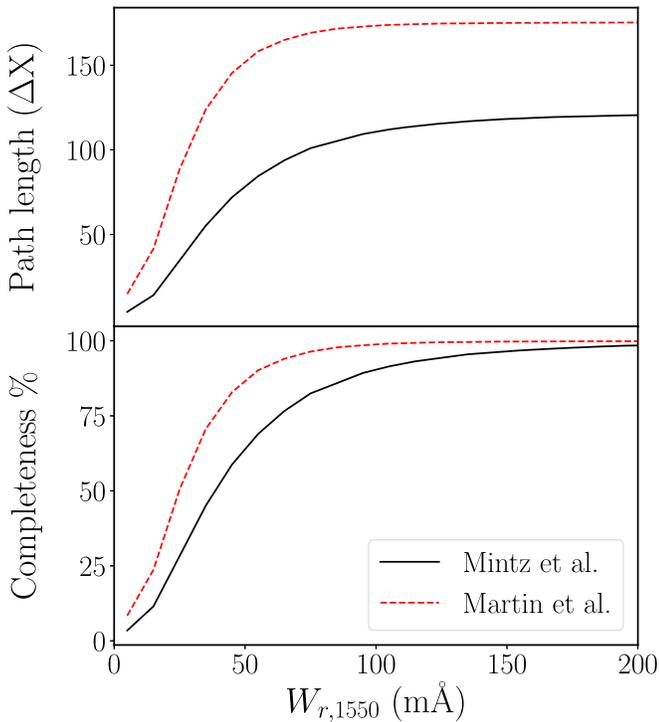


Figure 5. Total searchable path in comoving units and completeness percentage are shown as a function of $W_{r,1550}$. The functions for the Mintz et al. sample are shown as black solid lines and those for the Martin et al. sample are shown as red dashed lines. Due to the higher emission redshifts of the quasars and the higher S/N of the spectra, the Martin et al. sample has a greater total searchable path and higher completeness for every $W_{r,1550}$.

than the Mintz et al. sample, with a median C IV absorption redshift of 2.96 compared to that of the Mintz et al. sample of 1.91. Based on Cooksey et al. (2013) and Hasan et al. (2020), we would expect the Mintz et al. sample to have a larger intercept than the Martin et al. sample. We find that the slopes and intercepts of our best-fit lines to the Mintz et al. and the Martin et al. samples are statistically indistinguishable from one another, but attribute this to the small sample size. Our quasar-pair data set is not well-suited to a thorough investigation of the redshift evolution of C IV enrichment, so we cannot comment confidently on agreement or disagreement with predictions. Paired quasars are not necessary for this type of analysis—and are vastly more rare than unpaired quasars.

6. Transverse Autocorrelation Function

So far in this work, we have compiled a catalog of high-redshift C IV absorbers and computed the equivalent width distributions of our sample. We have not yet utilized the unique paired nature of our quasar sample. As described in Section 1, the transverse autocorrelation of absorbers in closely spaced quasar pairs provides additional insight into the spatial distribution of metals in the CGM and yields constraints on the coherence length of the C IV, which is related to the physical size of the CGM as traced by this ion (Adelberger et al. 2005; Bordoloi et al. 2014; Turner et al. 2017; Rudie et al. 2019; Schroetter et al. 2021). In this section, we calculate the transverse autocorrelation function for our sample following the methods of Hennawi & Prochaska (2007), Martin et al. (2010), and Fumagalli et al. (2014).

6.1. Data Absorber Pair Counts

A careful definition of paired absorbers is essential for computing and understanding the implications of the auto-correlation function. An absorber pair is a group of two absorbers found at similar redshifts but in different sightlines of a quasar pair. There are several examples of paired absorbers in Figure 2 at $\lambda\lambda 4190$ and 5190 \AA .

Whether or not two absorbers in paired quasar spectra are considered an absorber pair is dependent on the velocity width tolerance. A smaller tolerance will lead to a smaller number of paired absorbers, but may exclude absorbers that are physically associated but at slightly different redshifts. This could occur for a number of reasons, including differing gas velocities or distances at the two probed locations. A larger tolerance will include a larger number of such systems, but may also include chance coincidences of absorbers at similar redshifts that are not truly physically associated. There is no preferred method for choosing an optimized value of the velocity width tolerance δv_{pair} , and so we report our results for a number of choices for δv_{pair} as done in previous studies.

To determine reasonable values of δv_{pair} , we compare a normalized histogram of velocity separations between absorbers in sightline pairs of the true data catalog with the same normalized histogram for the simulated catalogs described below. We found that the number of data–data pairs (DD) only begins to significantly exceed the number of random–random pairs (RR) at $\delta v_{\text{pair}} \leq 500 \text{ km s}^{-1}$. In other words, there are approximately the same number of data absorbers separated by 500 km s^{-1} and above as there are random absorbers. To focus on the regime where paired absorbers are less likely to be chance coincidences, we choose to use values for δv_{pair} of 100, 200, and 500 km s^{-1} , but we note that these values differ from those used in Martin et al. (2010).

Data–data pairs compare absorbers within the true paired sightlines while random–random pairs compare absorbers in the simulated catalogs, which have no true physical association. We show a histogram of the DD absorber counts as a function of perpendicular separation of the paired quasars in Figure 7. In this figure, we normalize the counts by the total searchable path in each bin so that a bin with more quasars or with quasars with longer searchable paths will not bias the r_{\perp} dependence. This normalization removes some of the binning dependence of the distribution. We anticipate that quasar pairs with smaller r_{\perp} (those that are more closely located in the sky) will have a larger number of DD absorber pairs, as they will likely probe more of the same galaxy halos than pairs that are farther apart. We see general agreement with this expectation in Figure 7. As described above, for a given r_{\perp} value, we also see an increasing number of counts per unit searchable path length as δv_{pair} increases from 100 to 500 km s^{-1} as expected.

6.2. Random Absorber Pair Counts

We follow the approach of Martin et al. (2010) and run a Monte Carlo simulation using the EW frequency distribution to calculate the expected number of randomly paired absorbers (or RR). Generally, the simulation uses the parameters of the EW frequency distribution function from Section 5 to create 1000 fake C IV catalogs for each quasar that realistically replicate the redshift-dependent detection limitations of the actual spectra but do not contain the physical associations between paired quasars.

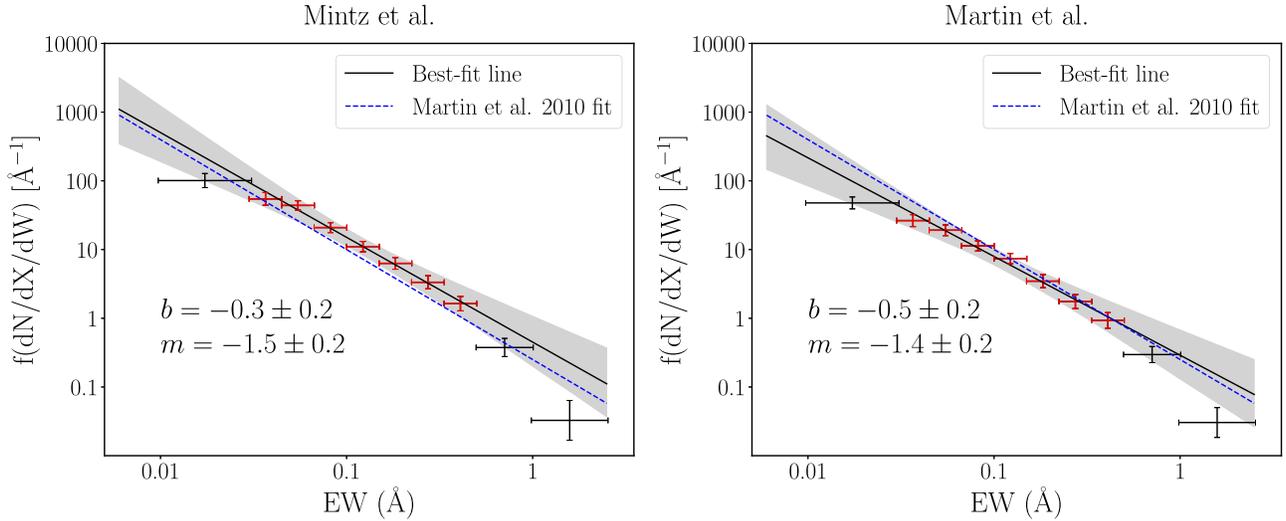


Figure 6. Equivalent width frequency distributions for $W_{r,1550}$ for the absorbers found in the Mintz et al. sample (left) and those found in the Martin et al. sample (right). The black line shows the best-fit line using only the bins plotted in red, which fall between 30 and 500 m Å, chosen to minimize bias from bins with few absorbers. The shaded gray region is the 2σ confidence interval for the best-fit line. The dashed blue line shows the best fit reported in Martin et al. (2010), shown in both plots for comparison. Our calculated frequency distributions agree well with the published results despite slight differences in our methodology.

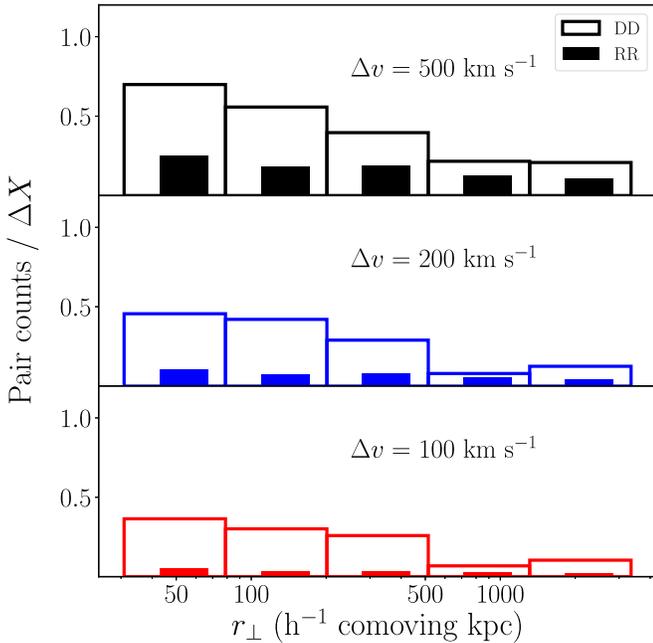


Figure 7. A histogram of DD and RR counts for absorbers in the combined sample as a function of projected separation, normalized by total searchable path length per bin. The three rows show pair counts for velocity widths of $\delta v_{\text{pair}} = 500, 200,$ and 100 km s^{-1} . The empty bars show the DD pair counts and the solid bars show the RR pair counts. There are significantly more DD pairs than RR pairs for all values of r_{\perp} and δv_{pair} . The number of pairs per bin increases as δv_{pair} increases, demonstrating the impact of the velocity width tolerance. We also see that the number of DD counts generally increases with decreasing r_{\perp} but begins to flatten at $\approx 200 h^{-1}$ comoving kpc.

We choose to use the EW frequency distribution parameters we derived for the Mintz et al. sample, which as noted above are statistically indistinguishable from those of the Martin et al. sample. In order to integrate and draw samples from the EW frequency distribution function, we must set finite EW limits for the simulation. We choose values of 1 and 1000 m Å, which represent the EW range of our actual absorber catalog.

To create the simulated catalogs, we first integrate the EW frequency distribution over EW and dX for each sightline to calculate the number of fake absorbers to simulate, setting an upper limit of 50 absorbers per sightline. We then draw an EW for each absorber from the EW frequency distribution using inverse transform sampling, and draw a redshift for each absorber from a uniform distribution covering the searchable redshift range for the given spectrum. Each simulated absorber is included in the final catalog if a randomly drawn number is greater than the detection probability for a system of that strength at that redshift in the given spectrum as determined in the sensitivity assessment described in Section 4. Finally, absorbers in a spectrum closer together than 250 km s^{-1} were linked as in the compilation of the data catalog.

On average, the resulting simulated catalogs have EW distributions that agree with the true data distribution and have a similar number of systems as the actual data. The average number of systems in each simulated catalog (for all 119 quasars) is 890, compared to 892 systems in the true catalog that have EWs between 1 and 1000 m Å. These similarities provide evidence that our simulated catalogs accurately reflect the conditions of the spectra.

We calculate the number of RR pairs per sightline pair by averaging over the 1000 runs of the simulation. We show the calculated number of RR pairs binned and normalized by search path as a function of r_{\perp} with solid bars in Figure 7. As expected, there are significantly fewer RR pairs than DD pairs at all r_{\perp} and δv_{pair} values.

6.3. Autocorrelation Function

Having calculated the number of DD pairs and the expected number of RR pairs as function of r_{\perp} , we then follow the method used in Fumagalli et al. (2014) and Scannapieco et al. (2006) and calculate the autocorrelation function ξ as

$$\xi = \frac{DD}{RR} - 1.$$

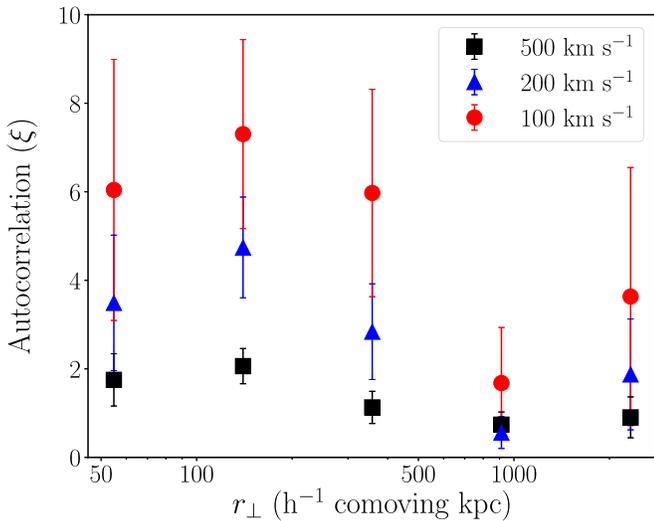


Figure 8. The transverse autocorrelation function $\xi = \frac{DD}{RR} - 1$ for the combined sample plotted as a function of perpendicular separation of the quasars pairs for three values of δv_{pair} . As expected, smaller values of δv_{pair} yield higher values of ξ , indicating higher significance of paired absorbers. Following the trend shown in Figure 7, ξ generally increases with decreasing separation. The function begins to level off at $\approx 200 h^{-1}$ comoving kpc. This indicates that the structures we are probing are no smaller than $\approx 200 h^{-1}$ comoving kpc.

If all of the DD pairs were chance coincidences, we would find $DD = RR$ and therefore $\xi = 0$. A larger value of ξ indicates a higher likelihood of true physical association of DD absorbers.

While Martin et al. (2010) used a slightly more sophisticated estimator of the autocorrelation function, introduced by Landy & Szalay (1993) in order to decrease the variance, we choose to use a simpler estimator because our uncertainty will be dominated by Poisson noise due to the small sample size. We calculate the uncertainty on our measurement of the autocorrelation function by propagating the 84.1% Poisson confidence limits on the absorber pair counts. We show the resulting autocorrelation function in Figure 8 for the same values of δv_{pair} reported in Figure 7.

The flattening of the transverse autocorrelation function shown in Figure 8 provides a constraint on the minimal extent of C IV in the CGM. A transverse autocorrelation function that is not increasing with decreasing r_{\perp} indicates that further decreasing the proper separation of the sightlines does not yield significantly more physically associated absorbers. In Figure 8, we see this flattening at $\approx 200 h^{-1}$ comoving kpc for all values of δv_{pair} . If these systems had a physical extent significantly smaller than this value, we would expect the autocorrelation function to increase more steeply for smaller r_{\perp} . The flattening we see indicates that quasar sightlines separated by less than $\approx 200 h^{-1}$ comoving kpc are equally likely to probe the same galaxy halo. Provided that the filling factor is sufficiently high, it is possible that the individual clouds of C IV are smaller than this size and we are probing the extent of the C IV gas in the halo. In fact, previous work has shown that individual C IV clouds have physical sizes of at least 300 pc and only begin to exhibit signs of clumping at kpc scales (Rauch et al. 2001; Tzanavaris & Carswell 2003).

While the absolute values of the autocorrelation function are lower than those from Martin et al. (2010) for $\delta v_{\text{pair}} = 200 \text{ km s}^{-1}$, we note that the shape of the function agrees with Martin et al. (2010) for the r_{\perp} bins covered in that study.

As described above, our approach to detect C IV doublets, assess the search sensitivity, and simulate fake catalogs is not identical to that of Martin et al. (2010). Each of these differences impacts the number of RR absorbers and therefore the absolute normalization of the autocorrelation function. We note that the absolute normalization of the autocorrelation is largely dependent on the number of RR pairs and the chosen values of δv_{pair} . Regardless, our constraint on the extent of C IV is based on the trend of the autocorrelation function, which is similar for each of the reported values of δv_{pair} , so our conclusions are not dependent on the absolute normalization.

We note that the value at which the flattening occurs is not well-specified by our data and is somewhat dependent on the binning of the autocorrelation function. Using a range of 4–7 bins, we find that ξ begins to flatten at $\approx 200 h^{-1}$ comoving kpc with a range of $\pm 50 h^{-1}$ comoving kpc. While the uncertainties in the autocorrelation are large for $\delta v_{\text{pair}} = 200 \text{ km s}^{-1}$, they are lower for the other larger δv_{pair} values, and the trend is consistent for all three values of δv_{pair} . This consistency over δv_{pair} suggests that the trend is likely real. Furthermore, this value is consistent with other estimates of the spatial extent of C IV from the impact parameters of absorbers to galaxies (Steidel et al. 2010; Borthakur et al. 2013; Bordoloi et al. 2014; Rudie et al. 2019; Schroetter et al. 2021; Dutta et al. 2021).

To better constrain the inflection point of the autocorrelation function from rising to flat, we would require more quasar pairs with projected separations at $\sim 200 h^{-1}$ comoving kpc to better sample the r_{\perp} space. To decrease the uncertainty in ξ , we would need more systems per bin, which could be achieved by including more quasar pairs or using spectra with higher S/N, higher resolution, or both. We note, however, that substantially increasing the resolution of the spectra would allow for the detection of significantly weaker C IV systems than is possible with lower-resolution data. Such a study would potentially trace a different C IV population outside of the halo. Therefore, the best approach would be to supplement the sample with additional quasar pairs at a similar resolution, none of which are currently available, and so additional data would be required.

7. Summary

We have conducted a search for C IV absorbers in a sample of high-redshift, medium-resolution spectra of closely located quasar pairs to investigate the metal distribution and extent of the high-redshift CGM.

1. We present a new sample of medium-resolution spectra of 32 high-redshift quasar pairs with redshifts $1.96 < z < 2.64$ and comoving projected separations of $32 h^{-1} < r_{\perp} < 320 h^{-1}$ comoving kpc. We supplement our sample with data from Martin et al. (2010) to cover a larger range of redshift and r_{\perp} .
2. We conduct a doublet search to identify C IV absorbers in the quasar spectra. We find 906 intervening likely absorbers, including 470 from the Mintz et al. sample. The absorbers have equivalent widths $6 \text{ m}\text{\AA} < W_{r,1550} < 2053 \text{ m}\text{\AA}$ with a median $W_{r,1550}$ of $92 \text{ m}\text{\AA}$.
3. We calculate the equivalent width frequency distribution for the C IV 1550 line using the absorber catalog and the results of the completeness assessment. We find that our results agree with the results published in Martin et al. (2010). We see potential evidence of redshift evolution in

the comparison of the two samples, in agreement with predictions from the literature. A larger sample would be necessary for a thorough investigation of the redshift evolution of CIV enrichment or extent, which is impractical with our current set of quasar pairs.

4. We compute the autocorrelation function $\xi = \frac{DD}{RR} - 1$ as a function of perpendicular separation, to assess the significance of the number of paired absorbers found in the quasar pairs. We find that the number of data–data pairs and ξ tend to increase with decreasing r_{\perp} up to $\approx 200 h^{-1}$ comoving kpc, indicating that the halos probed by CIV are likely no smaller than $\approx 200 h^{-1}$ comoving kpc.

While our results suggest the existence of a lower limit for the physical size of the CGM, we note that additional work is needed in order to better specify this value. To decrease the uncertainty in ξ , more systems per bin are required, which would be best achieved by obtaining spectra of additional quasar pairs at similar resolution. To better identify the turnover point in ξ , we would need more quasar pairs at small r_{\perp} in order to sample the parameter space more robustly.

Despite the limitations, our results provide a constraint for simulations and demonstrate the power of quasar pairs and the study of the transverse autocorrelation function as tools for investigating the nature of the CGM. Indeed, the addition of information in the transverse direction provides a valuable metric to constrain the extent of CIV enriched gas around galaxies, beyond what is possible along the line of sight (Scannapieco et al. 2006; Martin et al. 2010). As this metric depends on the implementation for feedback, the time at which metals have been injected, and the characteristic masses of the galaxies responsible for the enrichment (e.g., Booth et al. 2012), detailed comparisons of simulations and observations in quasar pairs offer a new way to refine our understanding of metal enrichment in the CGM of galaxies.

Moreover, by combining measurements on small scales (e.g., from this work) and larger-scale clustering (e.g., Gontcho et al. 2018), it will be possible to separate the contribution of the one-halo and two-halo terms, to place constraints on both the characteristic halo mass of the galaxies contributing to the CIV absorption and the spatial extent and covering factor of metals in the CGM (Fumagalli et al. 2014). Refined measurements in larger samples together with the analysis of cosmological simulations have the potential to uncover how metals spread and enrich halos at early epochs, beyond what is currently possible by studying absorbers along the line of sight.

This work was supported by NASA Keck PI Data Awards 2016B N032E, 2017A N133E, & 2019A N027E (PI Rafelski) and 2021B N087 (PI Prichard), administered by the NASA Exoplanet Science Institute. Data presented herein were obtained at the W.M. Keck Observatory from telescope time allocated to the National Aeronautics and Space Administration through the agency’s scientific partnership with the California Institute of Technology and the University of California. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. This work is based on observations collected at

the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programs PID 099.A–0018 and 0100.A–0084. This work was supported in part by NSF through AST-1817125 (C.L.M.). Support for HST Program GO-14127 was provided by NASA through grants from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS526555. This project was supported in part by the NSF REU grant AST-1757321 and by the Nantucket Maria Mitchell Association. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No 757535) and by Fondazione Cariplo (grant No 2018-2329). K.H.R.R. acknowledges partial support from NSF grants AST-1715630 and AST-2009417. This research made use of `PyPeIt`,¹² a Python package for semi-automated reduction of astronomical slit-based spectroscopy (Prochaska et al. 2020).

Facilities: HST (WFC3/UVIS), Keck:II(ESI), VLT(X-Shooter), Magellan(Mage).

Software: ESIRedux (Prochaska et al. 2003), MASE pipeline (Bochanski et al. 2009), linetools (Prochaska et al. 2016), `PyPeIt` (Prochaska et al. 2020).

ORCID iDs

Abby Mintz  <https://orcid.org/0000-0002-9816-9300>
 Marc Rafelski  <https://orcid.org/0000-0002-9946-4731>
 Regina A. Jorgenson  <https://orcid.org/0000-0003-2973-0472>
 Michele Fumagalli  <https://orcid.org/0000-0001-6676-3842>
 Rajeshwari Dutta  <https://orcid.org/0000-0002-6095-7627>
 Crystal L. Martin  <https://orcid.org/0000-0001-9189-7818>
 Elisabetta Lusso  <https://orcid.org/0000-0003-0083-1157>
 Kate H. R. Rubin  <https://orcid.org/0000-0001-6248-1864>
 John M. O’Meara  <https://orcid.org/0000-0002-7893-1054>

References

- Adelberger, K. L., Shapley, A. E., Steidel, C. C., et al. 2005, *ApJ*, 629, 636
 Bochanski, J. J., Hennawi, J. F., Simcoe, R. A., et al. 2009, *PASP*, 121, 1409
 Bordoloi, R., Tumlinson, J., Werk, J. K., et al. 2014, *ApJ*, 796, 136
 Borthakur, S., Heckman, T., Strickland, D., et al. 2013, *ApJ*, 768, 18
 Booth, C. M., Schaye, J., Delgado, J. D., et al. 2012, *MNRAS*, 420, 1053
 Chabanier, S., Etourneau, T., Le Goff, J.-M., et al. 2022, *ApJS*, 258, 18
 Chen, H.-W., Gauthier, J.-R., Sharon, K., et al. 2014, *MNRAS*, 438, 1435
 Chen, H.-W., Lanzetta, K. M., & Webb, J. K. 2001, *ApJ*, 556, 158
 Cooke, J., & O’Meara, J. M. 2015, *ApJL*, 812, L27
 Cooke, R. J., Pettini, M., & Steidel, C. C. 2018, *ApJ*, 855, 102
 Cooksey, K. L., Kao, M. M., Simcoe, R. A., et al. 2013, *ApJ*, 763, 37
 Cooksey, K. L., Prochaska, J. X., Chen, H.-W., et al. 2008, *ApJ*, 676, 262
 Cooksey, K. L., Thom, C., Prochaska, J. X., et al. 2010, *ApJ*, 708, 868
 Cowie, L. L., Songaila, A., Kim, T.-S., & Hu, E. M. 1995, *AJ*, 109, 1522
 Cucchiara, A., Fumagalli, M., Rafelski, M., et al. 2015, *ApJ*, 804, 51
 D’Odorico, V., Viel, M., Saitta, F., et al. 2006, *MNRAS*, 372, 1333
 Dutta, R., Fumagalli, M., Fossati, M., et al. 2021, *MNRAS*, 508, 4573
 Findlay, J. R., Prochaska, J. X., Hennawi, J. F., et al. 2018, *ApJS*, 236, 44
 Freudling, W., Romaniello, M., Bramich, D. M., et al. 2013, *A&A*, 559, A96
 Fukugita, M., & Peebles, P. J. E. 2004, *ApJ*, 616, 643
 Fumagalli, M., Hennawi, J. F., Prochaska, J. X., et al. 2014, *ApJ*, 780, 74
 Fumagalli, M., Mackenzie, R., Trayford, J., et al. 2017, *MNRAS*, 471, 3686
 Gontcho A Gontcho, S., Miralda-Escudé, J., Font-Ribera, A., et al. 2018, *MNRAS*, 480, 610
 Hasan, F., Churchill, C. W., Stemmock, B., et al. 2020, *ApJ*, 904, 44
 Hennawi, J. F., Myers, A. D., Shen, Y., et al. 2010, *ApJ*, 719, 1672
 Hennawi, J. F., & Prochaska, J. X. 2007, *ApJ*, 655, 735

¹² <https://pypeit.readthedocs.io/en/latest/>

- Hennawi, J. F., Prochaska, J. X., Burles, S., et al. 2006, *ApJ*, 651, 61
- Jorgenson, R. A., Murphy, M. T., & Thompson, R. 2013, *MNRAS*, 435, 482
- Kaur, B., Kanekar, N., Rafelski, M., et al. 2021, *ApJ*, 921, 68
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Lee, K.-G., Krolewski, A., White, M., et al. 2018, *ApJS*, 237, 31
- Lofthouse, E. K., Fumagalli, M., Fossati, M., et al. 2020, *MNRAS*, 491, 2057
- Lusso, E., Fumagalli, M., Rafelski, M., et al. 2018, *ApJ*, 860, 41
- Marshall, J. L., Burles, S., Thompson, I. B., et al. 2008, *Proc. SPIE*, 7014, 701454
- Martin, C. L., Scannapieco, E., Ellison, S. L., et al. 2010, *ApJ*, 721, 174
- Mas-Ribas, L., Riemer-Sørensen, S., Hennawi, J. F., et al. 2018, *ApJ*, 862, 50
- Macquart, J.-P., Prochaska, J. X., McQuinn, M., et al. 2020, *Natur*, 581, 391
- Mortensen, K., Keerthi Vasani, G. C., Jones, T., et al. 2021, *ApJ*, 914, 92
- O'Meara, J. M., Burles, S., Prochaska, J. X., et al. 2006, *ApJ*, 649, L61
- Peeples, M. S., Corlies, L., Tumlinson, J., et al. 2019, *ApJ*, 873, 129
- Peeples, M. S., Werk, J. K., Tumlinson, J., et al. 2014, *ApJ*, 786, 54
- Péroux, C., & Howk, J. C. 2020, *ARA&A*, 58, 363
- Péroux, C., Rahmani, H., Arrigoni Battaia, F., et al. 2018, *MNRAS*, 479, L50
- Prochaska, J., Hennawi, J., Westfall, K., et al. 2020, *JOSS*, 5, 2308
- Prochaska, J. X., Castro, S., & Djorgovski, S. G. 2003, *ApJS*, 148, 317
- Prochaska, J. X., Neeleman, M., Kanekar, N., et al. 2019, *ApJL*, 886, L35
- Prochaska, J. X., Sheffer, Y., Perley, D. A., et al. 2009, *ApJL*, 691, L27
- Prochaska, J. X., Tejos, N., Crighton, N., et al. 2016, *Linetools/Linetools: Second Major Release*, Zenodo, doi:10.5281/zenodo.168270
- Rafelski, M., Wolfe, A. M., Prochaska, J. X., et al. 2012, *ApJ*, 755, 89
- Rahmati, A., Pawlik, A. H., Raičević, M., et al. 2013, *MNRAS*, 430, 2427
- Rauch, M., Sargent, W. L. W., & Barlow, T. A. 1999, *ApJ*, 515, 500
- Rauch, M., Sargent, W. L. W., & Barlow, T. A. 2001, *ApJ*, 554, 823
- Rhodin, N. H. P., Krogager, J.-K., Christensen, L., et al. 2021, *MNRAS*, 506, 546
- Rubin, K. H. R., Diamond-Stanic, A. M., Coil, A. L., et al. 2018, *ApJ*, 868, 142
- Rubin, K. H. R., Hennawi, J. F., Prochaska, J. X., et al. 2015, *ApJ*, 808, 38
- Rudie, G. C., Steidel, C. C., Pettini, M., et al. 2019, *ApJ*, 885, 61
- Scannapieco, E., Pichon, C., Aracil, B., et al. 2006, *MNRAS*, 365, 615
- Schroetter, I., Bouché, N., Wendt, M., et al. 2016, *ApJ*, 833, 39
- Schroetter, I., Bouché, N. F., Zabl, J., et al. 2021, *MNRAS*, 506, 1355
- Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, *PASP*, 114, 851
- Spergel, D. N., Bean, R., Doré, O., et al. 2007, *ApJS*, 170, 377
- Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, *ApJ*, 717, 289
- Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, *ARA&A*, 55, 389
- Tumlinson, J., Thom, C., Werk, J. K., et al. 2013, *ApJ*, 777, 59
- Turner, M. L., Schaye, J., Crain, R. A., et al. 2017, *MNRAS*, 471, 690
- Tytler, D., Fan, X.-M., Burles, S., et al. 1995, in *QSO Absorption Lines*, ed. G. Meylan (New York: Springer), 289
- Tytler, D., Gleed, M., Melis, C., et al. 2009, *MNRAS*, 392, 1539
- Tzanavaris, P., & Carswell, R. F. 2003, *MNRAS*, 340, 937
- Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, 536, A105
- Wagoner, R. V. 1967, *ApJ*, 149, 465
- Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, *ApJ*, 792, 8
- Wolfe, A. M., Turnshek, D. A., Smith, H. E., et al. 1986, *ApJS*, 61, 249