

PROBING LARGE GALAXY HALOS AT $Z \sim 0$ WITH AUTOMATED $\text{Ly}\alpha$ -ABSORPTION MATCHING

DAVID M. FRENCH, BART P. WAKKER

Department of Astronomy, University of Wisconsin, Madison, WI 53706, USA

ABSTRACT

We present initial results from an ongoing large-scale study of the circumgalactic medium (CGM) in the nearby Universe ($cz \leq 10,000 \text{ km s}^{-1}$), using archival Cosmic Origins Spectrograph (COS) spectra of background quasi-stellar objects (QSOs). This initial sample contains 33 sightlines chosen for their proximity to large galaxies ($D \geq 25 \text{ kpc}$) and high signal-to-noise ratio ($S/N \geq 10$), yielding 48 $\text{Ly}\alpha$ **absorption lines** which we have paired with individual galaxies (**33 unique galaxies, with 29 cases where multiple absorbers within a single sightline are paired with the same galaxy**). We introduce a likelihood parameter to facilitate the matching of galaxies to absorption lines in a reproducible manner. We find the usual anti-correlation between $\text{Ly}\alpha$ equivalent width (EW) and impact parameter (ρ) when we normalize by galaxy virial radius (R_{vir}). Galaxies associated with a $\text{Ly}\alpha$ absorber are found to be more highly-inclined than galaxies in the survey volume at a $> 99\%$ confidence level (equivalent to $\sim 3.6\sigma$ for a normal distribution). In contrast with suggestions in other recent papers of a correlation with azimuth angle for Mg II absorption, we find no such correlation for $\text{Ly}\alpha$.

Keywords: galaxies:intergalactic medium, galaxies:evolution, galaxies:halos, quasars: absorption lines

1. INTRODUCTION

It is well known that galaxies must continue to accrete gas throughout their lifetimes in order to sustain their observed levels of star formation (e.g., Erb 2008, Prochaska & Wolfe 2009, Putman et al. 2009a, 2009b, Bauermeister et al. 2010, Genzel et al. 2010). This additional gas must come from the diffuse intergalactic medium (IGM), where the majority of the baryons in the universe reside (Penton et al. 2002, 2004; Lehner et al. 2007; Danforth & Shull 2008; Shull et al. 2012). How exactly this IGM gas eventually falls into the halos and disks of galaxies is still highly uncertain, as observational constraints are hard to come by. Because of the diffuse nature of IGM gas, it is most readily and sensitively detected as absorption in the spectra of background quasi-stellar objects (QSOs). The advent of the sensitive ultraviolet (UV) *Cosmic Origins Spectrograph* (COS) on the Hubble Space Telescope (HST; Osterman et al. 2011, Green et al. 2012) has provided a wealth of information about the properties and distribution of both the ions of heavy elements as well as the Lyman series of neutral hydrogen (H I) gas around galaxies.

Individual concentrations of gas along a given sightline imprint absorption lines onto the spectrum in the direction of the QSO. The metal lines trace the star formation history within the intervening gas, and H I lines (e.g., $\text{Ly}\alpha$) indicate both the location and velocities of outflowing gas, as well as the presence of fuel for future star formation. Numerous studies using these observations have shown that many $\text{Ly}\alpha$ absorbers trace indi-

vidual galaxy halos (e.g., Lanzetta et al 1995, Tripp et al. 1998, Chen et al. 1998, 2001a, Wakker & Savage 2009, Steidel et al. 2010, Prochaska et al. 2011, Thom et al 2012, Tumlinson et al. 2011 & 2013, Stocke et al. 2013 & 2014, Liang & Chen 2014, Tejos et al. 2014, Borthakur et al. 2015).

Some recent studies find that about half of $\text{Ly}\alpha$ absorbers lie within galaxy halos, at impact parameters $\rho < 350 \text{ kpc}$ (Côté et al. 2005, Prochaska et al. 2006, Wakker & Savage 2009). In addition, Wakker & Savage (2009) find that an absorber lies within 400 kpc and 400 km s^{-1} for 90% of galaxies brighter than $0.1L_*$, and all galaxies have a $\text{Ly}\alpha$ absorber within 1.5 Mpc. Higher redshift studies, such as Rudie et al. (2012a) at $2 < z < 3$, find evidence for an elevated density of absorbers up to 2 Mpc from galaxies. Wakker & Savage (2009) also discovered a correlation between $\text{Ly}\alpha$ absorption linewidth and impact parameter ρ , observing that the broadest lines ($\text{FWHM} > 150 \text{ km s}^{-1}$) are only seen within 350 kpc of a galaxy, while at $\rho > 1 \text{ Mpc}$, only lines with $\text{FWHM} < 75 \text{ km s}^{-1}$ occur. This suggests that the temperature and/or turbulence of gas increases in the presence of galaxies, a hypothesis that has been further supported by the results of Wakker et al. (2015).

Studying the enrichment of galaxy halos is necessary for constraining outflow models and informing stellar feedback prescriptions. Directly measuring the velocities and column densities of absorbers as a function of impact parameter and orientation around galaxies would provide the clearest evidence of inflow or outflow ac-

tivity, but results are still uncertain. Kacprzak et al. (2011b) claim to find that Mg II equivalent widths correlate with galaxy inclination, but Mathes et al. (2014) find no such correlation for Ly α and O VI absorbers. Furthermore, we should expect outflowing gas to be more highly enriched and trace the metallicity of the associated galaxy, with inflowing gas instead appearing only in H I. Both Stocke et al. (2013) and Liang & Chen (2014) find an “edge” to heavy ion absorption at $\sim 0.5R_{\text{vir}}$, but find Ly α covering fractions of $\sim 0.75 - 1$ continuing out to R_{vir} . However, Mathes et al. (2014) measure O VI absorption out to $\sim 3R_{\text{vir}}$, and Savage et al. (2014) find that more than half of O VI absorption occurs beyond $1 R_{\text{vir}}$ from the nearest galaxy. Additionally, Borthakur et al. (2015) find that Ly α absorption EW correlates with galaxy H I gas fraction, but only weakly with SFR, suggesting that accretion flow from the CGM is slow and continuous.

Recent results from Kacprzak et al. (2011b & 2012a) suggest that absorbing systems have a preferred orientation with respect to the major and minor axes of the galaxies they are associated with. This could be evidence of inflows and outflows, or an effect of the global structure of galaxy halos, but the statistics are not yet good enough to provide consistent answers. A larger-scale study of inclination and azimuthal angles vs. absorber properties is needed in order to elucidate the distribution of absorbing systems around galaxies. This is most easily done for the largest galaxies in the nearby universe, where it is possible to obtain inclinations and unambiguous absorber associations.

Previous studies have suffered from small sample sizes (e.g., Mathes et al. 2014 use 14 galaxies, Stocke et al. 2013 use 11, Werk et al. 2014 use 44), incompleteness due to their higher mean redshifts (e.g., the Mathes et al. 2014 sample is $0.12 < z < 0.67$), and limited impact parameter reach (e.g., Werk et al. 2014 probe CGM gas only within $\rho < 160$ kpc of galaxies). To address these shortcomings, we are conducting a large survey of the properties of intergalactic gas in the nearby universe, where we have good and relatively complete information on both faint and bright galaxies, in order to reveal how the IGM and galaxies affect each other.

We are taking advantage of the over 500 archived QSO and Seyfert spectra taken by the COS and Space Telescope Imaging Spectrograph (STIS) on HST, combined with the wealth of information available for the $\sim 100,000$ galaxies with $cz < 10,000$ km s $^{-1}$ found in the NASA Extragalactic Database (NED) to probe the environment of absorbing gas systems in the nearby universe. In this paper we introduce a new, **analytic** method for associating absorption lines with nearby galaxies. This approach will allow for an objective understanding of the distribution of the gas around galaxies, which requires looking for both detections and non-detections of gas, both near and far from galaxies, with a robust and reproducible metric for matching galaxies with absorption.

This paper presents our likelihood-matching method with initial results from a pilot study of 33 sight lines, chosen for their proximity to large galaxies and high signal-to-noise spectra. It is organized as follows: in Section 2 we present the data and analysis techniques, in Section 3 we present the results, in Section 4 we discuss possible interpretations of our results, and Section 5 presents a summary.

2. DATA AND ANALYSIS

2.1. Galaxy Data

Achieving the goal of this study relies on knowing the locations and properties of all galaxies near detected Ly α absorption lines. To facilitate this, we have constructed a database of all $z \leq 0.033$ ($cz \leq 10,000$ km s $^{-1}$) galaxies with published data available through the NASA Extragalactic Database (NED). A full description of this catalog will be presented in French & Wakker 2017 (in prep). Here we summarize its important aspects.

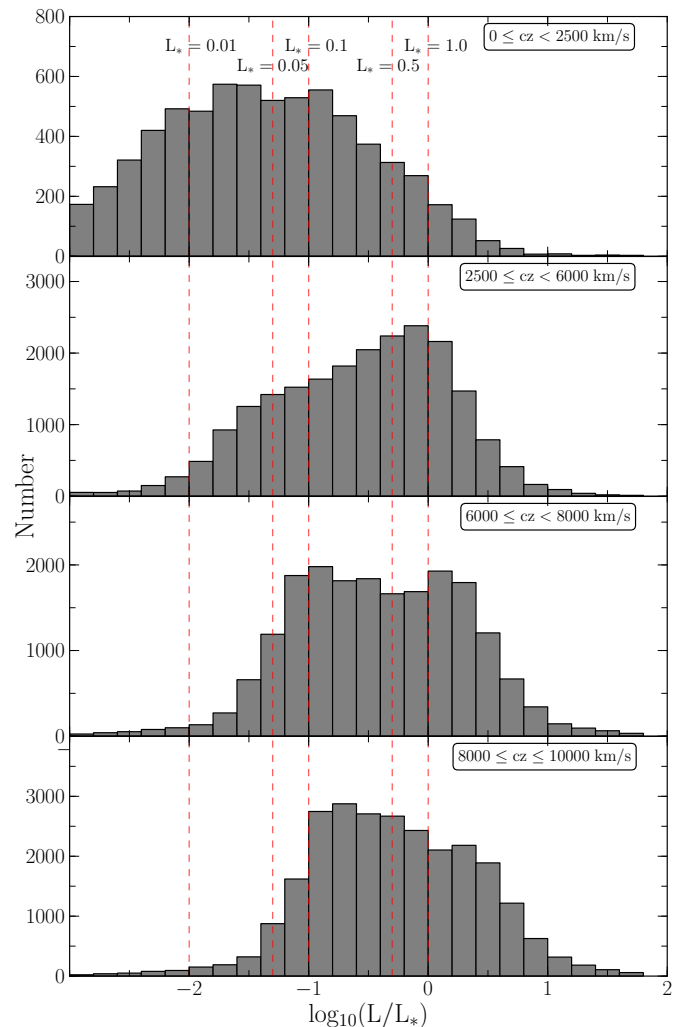


Figure 1. Distribution of L/L_* values for all galaxies in the dataset. Red dashed vertical lines highlight 1, 0.5, 0.1, 0.05 and 0.01 L_* .

Target	R.A.	Dec.	z	Program	Grating	Obs ID	Obs Date	T_{exp}^* [ks]	S/N* [1238]
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1H0717+714	07 21 53.3	+71 20 36	0.5003	12025	G130M	LBG812	11-12-27	6.0	37
2dFGRS_S393Z082	02 45 00.8	-30 07 23	0.3392	12988	G130M	LC1040	13-05-27,28	17.7	10
						LC1045			
FBQSJ1431+2442	14 31 25.8	+24 42 20	0.4069	13342	G130M	LC8903	15-03-29	16.5	17
				12603		LBS314	13-03-08		
H1101-232	11 03 37.7	-23 29 31	0.1860	12025	G130M	LBG804	11-07-05	13.3	16
HE0241-3043	02 43 37.7	-30 30 48	0.6693	12988	G130M	LC1070	13-06-21	7.0	14
LBQS1230-0015	12 33 04.1	-00 31 34	0.4709	11598	G130M	LB5N15	10-08-01	10.3	13
				12486		LBP250	12-04-26		
MRC2251-178	22 54 05.9	-17 34 55	0.0661	12029	G130M	LBGB03	11-09-29	5.5	42
Mrk290	15 35 52.3	+57 54 09	0.0296	11524	G130M	LB4Q02	09-10-28	3.9	38
Mrk876	16 13 57.2	+65 43 10	0.1290	11524	G130M	LB4Q03	10-04-08,09	12.6	65
				11686		LB4F05			
Mrk1014	01 59 50.2	+00 23 41	0.1630	12569	G130M	LBP404	12-01-25	1.8	17
PG0832+251	08 35 35.9	+24 59 41	0.3310	12025	G130M	LBG808	12-04-19	6.1	14
PG0003+158	00 05 59.3	+16 09 49	0.4509	12038	G130M	LBGL17	11-10-22	10.4	25
PG1001+054	10 04 20.1	+05 13 01	0.1610	13347	G130M	LCCV02	14-06-19	5.2	14
				13423		LC9W02	14-04-04		
PG1302-102	13 05 33.0	-10 33 20	0.2784	12038	G130M	LBGL04	11-08-16	6.0	27
RBS1768	21 38 49.7	-38 28 40	0.1830	12936	G130M	LC1201	13-06-25	7.0	24
RX J0714.5+7408	07 14 36.2	+74 08 11	0.3710	12275	G130M	LBH402	11-03-18	8.3	18
RX J1017.5+4702	10 17 30.9	+47 02 25	0.3354	13314	G130M	LC9M04	14-01-29	8.7	12
RX J1117.6+5301	11 17 40.5	+53 01 50	0.1587	14240	G130M	LCWM05	16-04-13	4.9	11
RX J1236.0+2641	12 36 04.1	+26 41 36	0.2092	12248	G130M	LBH087	12-01-29	4.2	11
RX J1330.8+3119	13 30 53.2	+31 19 32	0.2423	12248	G130M	LBHO85	11-07-11	4.3	11
RX J1356.4+2515	13 56 25.6	+25 15 23	0.1640	12248	G130M	LBH057	12-02-03	2.3	10
RX J1503.2+6810	15 03 16.5	+68 10 06	0.1140	12276	G130M	LBI609	10-12-31	1.9	11
RX J1544.5+2827	15 44 30.5	+28 27 56	0.2314	13423	G130M	LC9W08	14-02-25	2.1	10
RX J2043.1+0324	20 43 06.2	+03 24 50	0.2710	13840	G130M	LCJW02	14-10-23	7.8	15
RX J2139.7+0246	21 39 44.2	+02 46 05	0.2600	13840	G130M	LCJW03	14-10-27	7.9	16
SBS0957+599	10 01 02.6	+59 44 15	0.7475	12248	G130M	LBHO65	11-03-18,19	3.3	12
SDSSJ021218.32-073719.8	02 12 18.3	-07 37 20	0.1739	12248	G130M	LBHO83	11-06-26	6.5	12
						LBHO92	11-08-21		
						LBHO92	11-08-21		
SDSSJ080838.80+051440.0	08 08 38.8	+05 14 40	0.3606	12603	G130M	LBS330	12-03-17	4.7	10
SDSSJ091728.60+271951.0	09 17 28.6	+27 19 51	0.0756	14071	G130M	LCX202	15-11-30	15.5	11
						LCX2Z2	16-02-06		
SDSSJ112224.10+031802.0	11 22 24.1	+03 18 02	0.4753	12603	G130M	LBS318	13-03-29	7.6	13
SDSSJ130524.30+035731.0	13 05 24.3	+03 57 31	0.5457	12603	G130M	LBS321	12-06-25,26	7.6	13
SDSSJ135726.27+043541.4	13 57 26.2	+04 35 41	1.2345	12264	G130M	LBJ005	11-06-22	14.1	21
						LBJ007	11-06-26		
SDSSJ140428.30+335342.0	14 04 28.3	+33 53 42	0.5500	12603	G130M	LBS320	13-03-03	7.7	10
TON1009	09 09 06.1	+32 36 31	0.8103	12603	G130M	LBS328	12-04-22	4.7	12

Table 1. COS targets in this sample. *Total exposure time and S/N ratio is given for multi-orbit exposures.

The galaxy dataset contains over 108,000 entries, and includes data from SDSS, 2MASS, 2dF, 6dF, RC3, and many other, smaller surveys. Our criteria for including a galaxy in this dataset is only an accurate, spectroscopic redshift which places the galaxy in the $cz \leq 10,000 \text{ km s}^{-1}$ velocity range. This restriction leads to a completeness limit of $B \lesssim 18.7 \text{ mag}$, or $\sim 0.2L_*$, at $cz = 10,000 \text{ km s}^{-1}$, and progressively better towards

lower velocities (see Figure 1). This limit will vary depending on which major surveys include a particular region of the sky. The major contributor is whether or not SDSS data is available, which begins around $cz = 5,000 \text{ km s}^{-1}$. Figure 1 is split into 4 velocity bins to illustrate this. Our data is complete down to $\sim 0.1L_*$ in the first bin, $0 \leq cz \leq 2,500 \text{ km s}^{-1}$. At slightly higher velocity, $2500 \leq cz \leq 6000 \text{ km s}^{-1}$, the completeness falls

to barely better than $\sim 1.0L_*$ as we move past the near and well studied galaxies, but have yet to reach the footprint of deep surveys. SDSS data becomes available in the last two bins, spanning $6000 \leq cz \leq 10,000 \text{ km s}^{-1}$, and correspondingly completeness remains high down to the SDSS limits of $B \lesssim 18.7 \text{ mag}$, or $\sim 0.2L_*$ at $cz = 10,000 \text{ km s}^{-1}$.

Additionally, we have homogenized the galaxy data beyond the steps taken by NED by normalizing all measurements of galaxy inclination, position angle, and diameter to 2MASS K_s -band values. Most galaxies in NED have measures of inclination, position angle and diameter available in several different bands, so in order to make meaningful comparisons it is necessary to automatically choose one band for all measurements. We chose 2MASS values for this because it was an all-sky survey, and represents the largest fraction of available galaxy data. Physical galaxy diameters are derived from 2MASS K_s “total” angular diameter measurements and galaxy distances. 2MASS K_s “total” diameter estimates are surface brightness extrapolation measurements and are derived as

$$r_{\text{tot}} = r' + a(\ln(148))^b, \quad (1)$$

where r_{tot} is defined as the point where the surface brightness extends to 5 disk scale lengths, r' is the starting point radius ($> 5'' - 10''$ beyond the nucleus, or core influence), and a and b are Sersic exponential function scale length parameters ($f = f_0 \exp(-r/a)^{(1/b)}$, see Jarret et al. 2003 for a full description). Approximately 50% of all the galaxies have this 2MASS K_s “total” diameter. Of the remainder, 20% have SDSS diameters, 27% have no published diameter, and 3% have diameters from other surveys. We convert values in these other bands to 2MASS K_s “total” diameters via a simple least squares linear fit when necessary.

We used B -band magnitudes to estimate each galaxy’s luminosity in units of L_* as follows:

$$\frac{L}{L_*} = 10^{-0.4(M_B - M_{B*})}. \quad (2)$$

We adopt the CfA galaxy luminosity function by Marzke et al. (1994), which sets $B_* = -19.57$. Direct B band measurements are available for $\sim 30\%$ of galaxies, and most of the rest have SDSS g and r magnitudes, which can be converted to B via $B = g + 0.39(g - r) + 0.21$ (Jester et al. 2005). Finally, we also compute an estimate of the virial radius of each galaxy as $\log R_{\text{vir}} = 0.69 \log D + 1.24$. This follows the parametrization of Stocke et al. (2013) relating a galaxy’s luminosity to its virial radius, and the Wakker & Savage (2009) empirical relation between diameter and luminosity (see Wakker et al. 2015 and references therein for further details). Errors are propagated from the original published magnitude errors.

This homogeneous galaxy data table allows us to draw direct comparisons between the properties of the ab-

sorbers and the properties, separations, and environments of nearby galaxies with unprecedented completeness. The full dataset will be publicly released and discussed in further detail in a forthcoming paper (French & Wakker 2017, in prep).

2.2. Spectra

This initial pilot study contains 33 sightlines to bright QSOs observed with COS. We chose sightlines by first sorting the galaxy data table described above by galaxy diameter. This sorted list is then correlated with the full list of publicly available sightlines, and only those systems with impact parameter less than 500 Mpc and galaxy diameter, D , greater than 25 kpc are kept. Finally, we select 33 sightlines with $S/N \geq 10$ from this list.

All COS spectra for the target sightlines were obtained through the Barbara A. Mikulski Archive for Space Telescopes (MAST), and processed with CALCOS v3.0 or later. We combined individual exposures by the method of Wakker et al. (2015), which corrects the COS wavelength scale by cross-correlating all ISM and IGM lines in each exposure. This method addresses the up to $\pm 40 \text{ km s}^{-1}$ misalignments produced by CALCOS, and produces a corrected error array based on Poisson noise, which better matches the measured errors than the errors delivered in the x1d files. We then combine multiple exposures by aligning Galactic absorption lines with 21-cm spectra, and adding up the total counts in each pixel before converting to flux using the original, average flux-count ratio at each wavelength.

3. RESULTS

We have identified 48 Ly α absorption lines in the spectra of our initial 33 QSO sample, each of which has been associated with a single nearby galaxy of diameter $D \geq 25 \text{ kpc}$. Each absorption component is treated individually, resulting in several cases where multiple absorbers are associated with the same galaxy. In order to be considered for a pairing, a galaxy and absorption feature must appear within 400 km s^{-1} in velocity and 500 kpc in physical impact parameter from each other. When multiple galaxies pass these criteria for a particular line, we are left with two options. 1) one galaxy is obviously far larger and closer in physical and velocity space to the sightline, and may have several satellite galaxies, or 2) there are multiple galaxies near the absorber, making any association ambiguous; we do not include these cases in the further analysis.

To facilitate this decision, we calculate the likelihood, \mathcal{L} , of every possible galaxy-absorber pairing as follows:

$$\mathcal{L} = A e^{-\left(\frac{\rho}{R_{\text{eff}}}\right)^2} e^{-\left(\frac{\Delta v}{v_{\text{norm}}}\right)^2}. \quad (3)$$

Here ρ is the physical impact parameter, Δv the velocity difference between the absorber and the galaxy ($\Delta v = v_{\text{galaxy}} - v_{\text{absorber}}$), v_{norm} is the velocity normalization constant, and A is a factor included to increase the likelihood in the case that $\rho \leq R_{\text{eff}}$ (in

Target	Galaxy	R_{vir}	L/L_*	v_{galaxy}	Inc.	Az.	ρ	$v_{Ly\alpha}$	$W_{Ly\alpha}$	Δv	\mathcal{L}
(1)	(2)	[kpc]	(4)	[km s ⁻¹]	[deg]	[deg]	[kpc]	[km s ⁻¹]	[km s ⁻¹]	[km s ⁻¹]	(12)
1H0717+714	UGC03804	173	1.9	2887	55	7	207	2870	343±6	17	0.24
1H0717+714	UGC03804	173	1.9	2887	55	7	207	2956	39±4	-69	0.21
2dFGRS_S393Z082	NGC1097	304	6.1	1271	58	27	112	1239	570±21	32	1.9*
H1101-232	MCG-04-26-019	173	1.1	3623	68	26	179	3580	573±12	43	0.33
HE0241-3043	NGC1097	304	6.1	1271	58	77	219	1221	83±12	50	1.6*
HE0241-3043	NGC1097	304	6.1	1271	58	77	219	1310	184±15	-39	1.6*
LBQS1230-0015	NGC4517	208	0.5	1128	90	90	110	1127	473±16	1	1.6*
MRC2251-178	MCG-03-58-009	319	2.3	9030	61	39	320	9051	60±4	-21	1.4*
Mrk1014	NGC0768	231	3.0	7021	64	85	486	7080	117±11	-59	0.042*
Mrk290	NGC5987	322	3.0	3010	67	12	486	3105	511±5	-95	0.77*
Mrk290	NGC5987	322	3.0	3010	67	12	486	3207	319±4	-197	0.37*
Mrk876	UGC10294	165	0.1	3504	51	7	274	3478	280±3	26	0.063
PG0003+158	NGC7814	171	1.2	1050	68	47	197	833	131±15	217	0.081
PG0832+251	KUG0833+252	165	0.7	6964	62	55	294	6980	133±14	-16	0.041
PG0832+251	KUG0833+252	165	0.7	6964	62	55	294	7201	48±10	-237	0.01
PG1001+054	UGC05432	164	1.3	3995	36	78	217	4092	222±10	-97	0.14
PG1302-102	NGC4939	235	4.4	3110	48	61	265	3448	71±5	-338	0.05*
RBS1768	RFGC3781	253	1.0	9162	90	74	464	9360	364±4	-198	0.056*
RBS1768	RFGC3781	253	1.0	9162	90	74	464	9434	160±5	-272	0.024*
RX J0714.5+7408	UGC03717	202	1.2	4188	63	83	271	4074	58±7	114	0.13*
RX J0714.5+7408	UGC03717	202	1.2	4188	63	83	271	4264	410±9	-76	0.15*
RX J1017.5+4702	NGC3198	191	1.3	663	73	55	378	629	60±17	34	0.02
RX J1117.6+5301	NGC3631	187	1.9	1156	16	47	198	1131	356±20	25	0.32
RX J1117.6+5301	NGC3631	187	1.9	1156	16	47	198	1259	57±17	-103	0.25
RX J1236.0+2641	NGC4559	165	0.7	807	64	31	188	795	295±37	12	0.27
RX J1236.0+2641	NGC4565	292	1.7	1230	90	39	159	1012	337±32	218	0.54*
RX J1236.0+2641	NGC4565	292	1.7	1230	90	39	159	1188	288±24	42	1.7*
RX J1330.8+3119	UGC08492	204	2.0	7414	16	41	335	7401	330±15	13	0.081*
RX J1356.4+2515	CGCG132-055	206	1.3	8671	36	25	190	8475	126±18	196	0.35*
RX J1503.2+6810	CGCG318-012	250	2.1	9765	52	1	325	10122	44±14	-357	0.031*
RX J1544.5+2827	CGCG166-047	175	1.8	9646	43	61	326	9642	183±14	4	0.031
RX J1544.5+2827	CGCG166-047	175	1.8	9646	43	61	326	9759	169±12	-113	0.023
RX J2043.1+0324	NGC6954	166	1.4	4067	56	66	301	4080	82±10	-13	0.037
RX J2139.7+0246	UGC11785	203	0.7	4074	90	69	108	4083	490±7	-9	1.5
RX J2139.7+0246	UGC11785	203	0.7	4074	90	69	108	4181	529±7	-107	1.2*
SBS0957+599	MCG+10-14-058	261	1.1	9501	75	19	206	9469	78±12	32	1.4*
SDSSJ021218.32-073719.8	SDSSJ021315.79-073942.7	174	1.8	4800	52	10	268	4756	528±15	44	0.09
SDSSJ021218.32-073719.8	SDSSJ021315.79-073942.7	174	1.8	4800	52	10	268	4833	500±17	-33	0.092
SDSSJ080838.80+051440.0	UGC04239	279	2.1	8763	45	38	378	8740	883±24	23	0.87*
SDSSJ080838.80+051440.0	UGC04239	279	2.1	8763	45	38	378	8927	130±19	-164	0.45*
SDSSJ091728.60+271951.0	UGC04895	204	2.1	7073	61	32	408	7141	374±23	-68	0.022*
SDSSJ112224.10+031802.0	NGC3640	180	2.8	1251	38	22	139	1049	288±30	202	0.4
SDSSJ112224.10+031802.0	NGC3640	180	2.8	1251	38	22	139	1264	424±27	-13	1.1
SDSSJ130524.30+035731.0	UGC08186	268	1.1	7006	82	14	249	7039	480±14	-33	1.3*
SDSSJ135726.27+043541.4	NGC5364	211	2.4	1241	57	84	183	1124	85±11	117	0.74*
SDSSJ135726.27+043541.4	NGC5364	211	2.4	1241	57	84	183	1296	98±9	-55	0.97*
SDSSJ140428.30+335342.0	KUG1402+341	204	1.1	7919	72	63	118	7884	889±28	35	1.4
TON1009	NGC2770	204	1.9	1947	87	41	274	1961	350±21	-14	0.19*

Table 2. All associated systems. The largest \mathcal{L} value is given, with a (*) indicating that this corresponds to $\mathcal{L}_{D^{1.5}}$, otherwise the quoted \mathcal{L} was computed with R_{vir} .

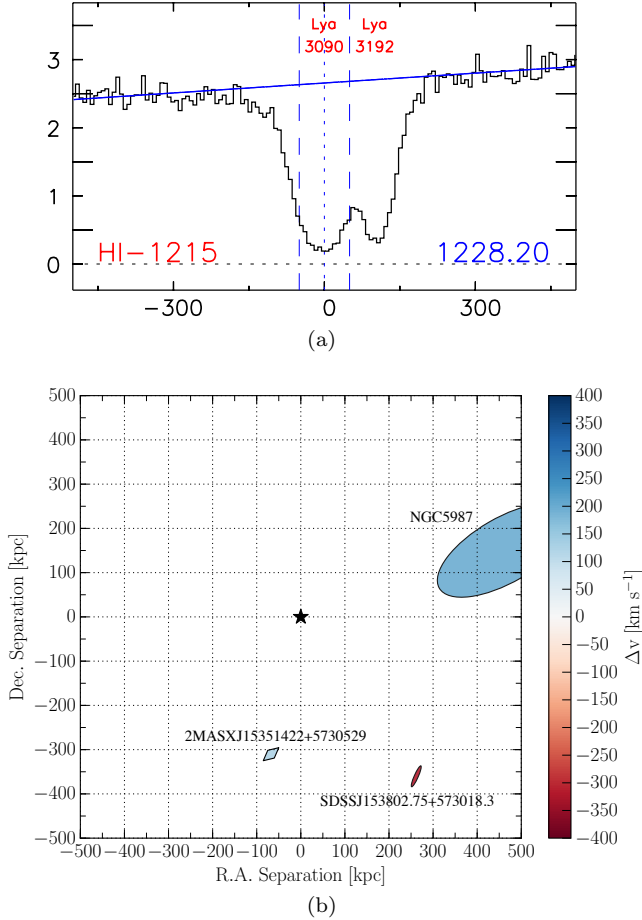


Figure 2. a) An example of 2 Ly α lines found in the Mrk290 sightline at 3090 and 3192 . b) A map of *all* galaxies within a 500 kpc impact parameter of target Mrk290 sightline and with velocity (cz) within 400 km s⁻¹ of absorption detected at 3192 km s⁻¹ (central black star). The galaxy NGC5987 ($v = 3010$ km s⁻¹, inclination = 65°) has been paired with the Ly α absorption features at $v = 3090, 3192$ km s⁻¹ because it is the largest and closest galaxy in both physical and velocity space to the absorption feature.

which case $A = 2$, otherwise $A = 1$). Many similar studies and simulations (e.g., Wakker & Savage 2009, Liang & Chen 2014, Mathes et al. 2014) suggest that Ly α absorbers lie within 400 km s⁻¹ of their associated galaxies, so throughout this paper we adopt a halfway point of $v_{\text{norm}} = 200$ km s⁻¹. Future work will explore the result of varying this normalization parameter and making refinements such as, e.g., relating v_{norm} to the galaxy’s rotation velocity.

We compute \mathcal{L} for two different values of R_{eff} : R_{vir} , the virial radius of the galaxy, and $D^{1.5}$, the major diameter of the galaxy to the power of 1.5. \mathcal{L} computed with R_{vir} is liable to select satellite galaxies instead of the larger hosts, so including a version with $D^{1.5}$ serves as a two-tiered selection system. An absorber separated by 200 km s⁻¹ in velocity and $1R_{\text{vir}}$ in impact parameter from a $D = 30$ kpc galaxy would have $\mathcal{L}_{R_{\text{vir}}} = 0.27$ and $\mathcal{L}_{D^{1.5}} = 0.11$. In order for an absorber to be marked

as “associated” with a particular galaxy, we require that its \mathcal{L} must be a factor of 5 larger than the next best possible association, and $\mathcal{L} \geq 0.01$ for at least one of $\mathcal{L}_{R_{\text{vir}}}$ or $\mathcal{L}_{D^{1.5}}$. We visually inspect systems with only one \mathcal{L} meeting these criteria, and decide to reject or include it based on the complexity of the nearby galaxy environment. In Table 2 we quote only the largest value of \mathcal{L} , and use an asterisk to denote when this corresponds to $\mathcal{L}_{D^{1.5}}$.

Figures 2(a) and 2(b) show an example of a Ly α absorption line with a map of its galaxy environment, showing an unambiguous pairing between the absorption features at 3090, 3192 km s⁻¹ toward Mrk290 and galaxy NGC5987 ($\mathcal{L}^* = 0.37$). All analysis that follows concerns similarly “associated” systems.

Additionally, we split the absorber-galaxy catalog based on the velocity difference of the two, Δv . With this scheme, we refer to an absorber with a *lower* velocity than the associated galaxy as *blueshifted*, while an absorber with a *higher* velocity is referred to as *redshifted*. The rest of the results will be analyzed based upon this splitting. In all figures blue and red-shifted absorbers are represented as blue diamonds and red circles, respectively, and red diamonds correspond to systems where *both* a red and blue-shifted absorber is detected. We use open symbols for systems with $\rho \leq R_{\text{vir}}$.

3.1. EW - ρ Anti-correlation

Numerous previous studies have suggested that Ly α equivalent width (EW) is anti-correlated with impact parameter (ρ) to the nearest galaxy. We find a weak anti-correlation, as shown in Figure 3(a). However, as Churchill et al. (2013a) also found with Mg II absorption, we find a stronger anti-correlation when we normalize ρ by R_{vir} . Figure 3(b) shows this expected anti-correlation when plotting EW vs ρ/R_{vir} . A possible explanation for this trend is that larger galaxies host larger, more physically extended CGM halos. We would thus expect the absorber EW to also correlate positively with R_{vir} . Figure 4(a) shows EW as a function of R_{vir} , with the blue-dashed and red-dotted lines show the average EW in bins of 50 kpc of R_{vir} , showing little evidence of a correlation. However, by similarly plotting ρ as a function of R_{vir} , we instead find some evidence that absorbers around larger galaxies tend to be found at higher impact parameters. While we expect the upper-left quadrant of this figure to be sparsely populated (our likelihood-based method would tend not to choose small galaxies at large impact parameters), it is unclear to us why the lower-right quadrant (large galaxies with absorbers at low impact parameter) is also sparsely populated. The full-sized sample at the completion of our study should provide a clearer picture.

3.2. Inclination

In this section we examine the inclinations of the associated galaxies compared to the distributions of absorbers. We correct for the finite thickness of galaxies,

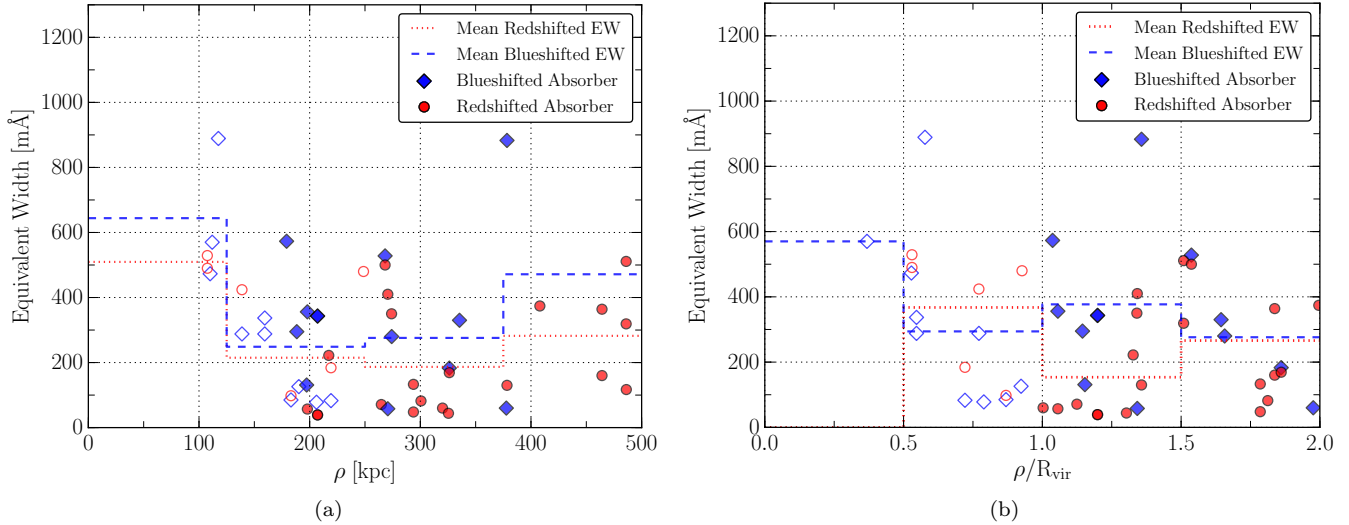


Figure 3. a) Equivalent width of each absorber as a function of impact parameter, ρ . b) Equivalent width as a function of ρ/R_{vir} . The anti-correlation is strongest when scaling ρ by the galaxy virial radius. Absorbers are separated into red and blue-shifted samples based on Δv . Bins of mean EW are overlaid in red-dotted, and blue-dashed lines for their respective samples. **Open symbols correspond to systems with $\rho \leq R_{vir}$.**

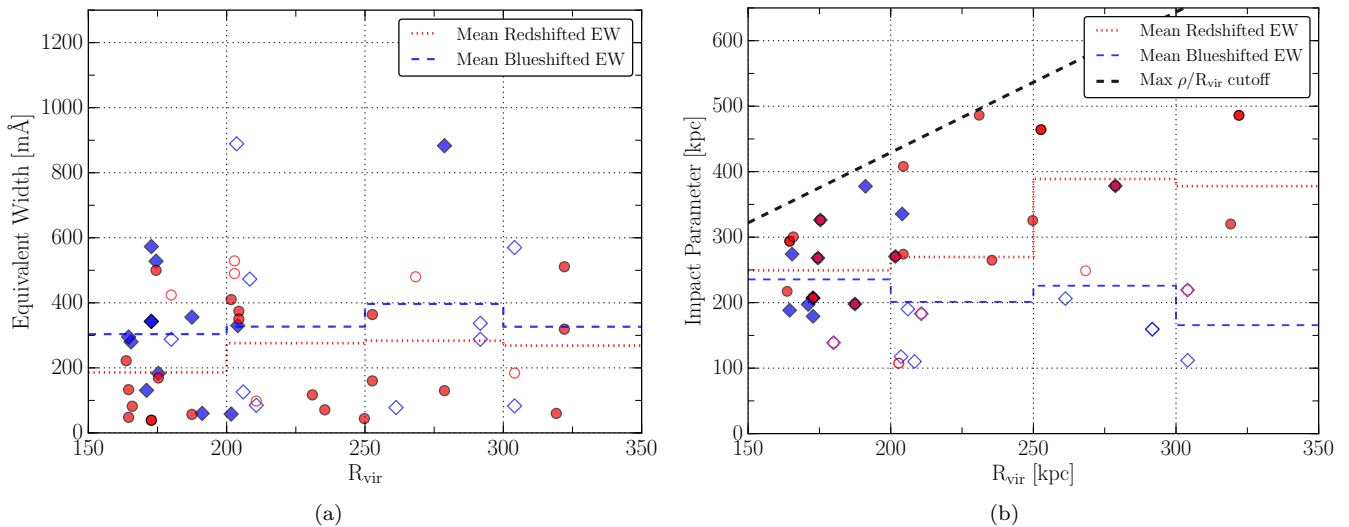


Figure 4. a) Equivalent width of each absorber as a function of the virial radius of the associated galaxy. b) Impact parameter to each absorber as a function of the virial radius of the associated galaxy. The black dashed line indicates the cutoff at $\rho/R_{vir} = 2.14$ imposed by our \mathcal{L} limit. In each, the blue-dashed and red-dotted lines shows the average EW in 50 kpc bins of impact parameter for the blueshifted and redshifted absorbers, respectively. **Open symbols correspond to systems with $\rho \leq R_{vir}$.**

which causes b/a to deviate from $\cos i$ at high inclinations, by computing galaxy inclinations with the following formula from Heidmann et al. (1972a):

$$\cos(i) = \sqrt{\frac{q^2 - q_0^2}{1 - q_0^2}}, \quad (4)$$

where $q = b/a$, the ratio of the minor to major axis, and q_0 is the intrinsic axis ratio, set to $q_0 = 0.2$ for all galaxies (e.g., Jones, Davies, and Trewhella 1996). 6 of the 48 total absorbers are associated with E or S0

type galaxies, but we have chosen to keep the value of $q_0 = 0.2$ uniform throughout. The calculated values of $\cos(i)$ we use for these galaxies are thus conservative under-estimates of their true inclinations.

Figure 5 shows red and blueshifted absorbers' EW plotted against the inclinations of their associated galaxies. We note that there is a clear excess of absorbers near galaxies of high inclination, with 77% of redshifted and 73% of blueshifted absorbers being associated with galaxies of $i \geq 50$ deg, and only 3 absorbers associated with a galaxy of $i < 35$. The solid-black and dashed-

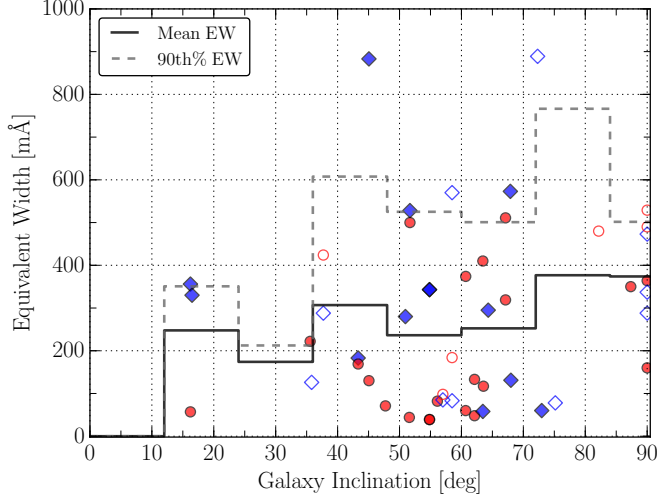


Figure 5. Equivalent width of each absorber as a function of the inclination angle of the associated galaxy. The black and dashed-grey lines show the mean and 90th percentile EW of all absorbers in bins of 15° . **Open symbols correspond to systems with $\rho \leq R_{vir}$.**

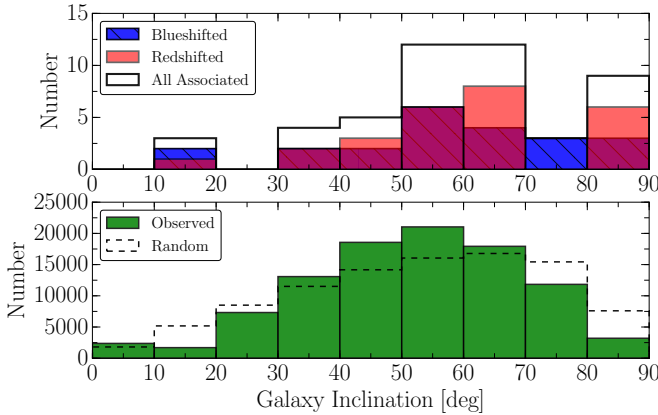


Figure 6. Top: Distribution of inclinations for all associated galaxies, split into red and blue shifted sets. **Bottom:** Distribution of inclinations of all observed galaxies in the $cz \leq 10,000 \text{ km s}^{-1}$ redshift range. The dashed line shows the inclination distribution for a truly random sample (i.e., no observational biases).

grey lines show mean and 90th percentile histograms, respectively, in bins of 12° . There does not appear to be much evolution of EW across galaxy inclination, although a slight increase of mean EW is possibly present towards higher inclinations.

In total 75% of absorbers are associated with highly inclined galaxies ($i \geq 50^\circ$). Only 56% of all galaxies in the survey volume are highly inclined, indicating a preference for detecting absorption around inclined galaxies. Figure 6 shows the distribution of galaxy inclinations for both the red and blue-shifted associated galaxies and all galaxies within the survey volume. We tested the difference between the full distribution of inclination angles for all galaxies in our survey volume and the distribution for all associated galaxies (red

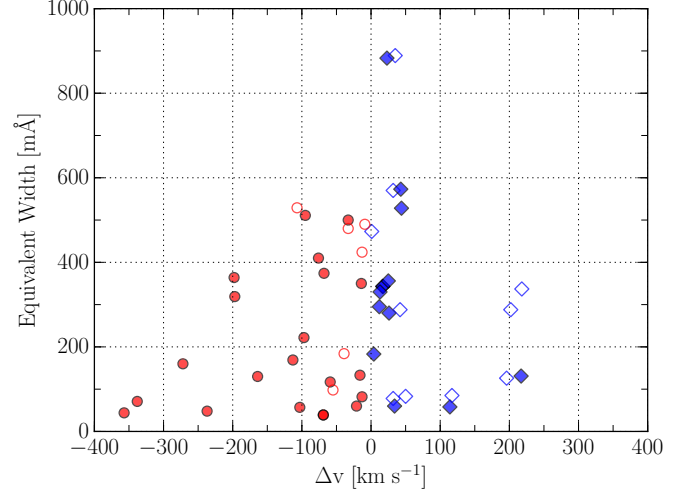


Figure 7. Equivalent width as a function of the velocity separation between the galaxy and absorption line. **Open symbols correspond to systems with $\rho \leq R_{vir}$.**

+ blue-shifted absorbers) using the Anderson-Darling (AD) statistical distribution test, yielding a p-value of $AD_p = 0.00037$. Thus at a 99.96% confidence level ($\sim 3.6\sigma$ for a normal distribution) the inclinations of our associated galaxies are not sampled from the average distribution of observed inclinations. Hence, we take this to mean that the shape of the CGM of these galaxies is not perfectly spherical.

It is worth noting here that the observed distribution of galaxy inclinations is *not* flat, as one might naively expect. The dashed line in Figure 6 shows the distribution of observable inclinations for a random, uniform sample (i.e., a uniform distribution of $q = b/a$ values between $0.2 - 1.0$). There could be a number of effects contributing to the difference between this expected distribution and the observed (shown in green). If our sample is magnitude limited and we assume galaxies are mostly optically thin but with an very-thin optically thick component (e.g., a dust lane), then mostly face-on and mostly edge-on galaxies would be underrepresented due to surface-brightness and dust obscuration effects (e.g., see Jones, Davies and Trewheila 1996). It is possible that a similar effect is also responsible for the over-abundance of $\text{Ly}\alpha$ detections around highly inclined galaxies. If we assume a disk or oblate spheroid halo shape and a covering fraction below unity for the CGM, the probability of encountering a cloud near an inclined galaxy would increase due to the increased path-length through the halo. We will produce a model to test this and other possible explanations in Paper II, when we have the much larger dataset available.

3.3. Velocity Difference (Δv)

We find evidence for an anti-correlation between absorber EW and the velocity difference between the galaxy and the associated absorption, Δv . The mean and maximal EW of absorption increases with decreas-

ing Δv (see Figure 7). In total, 32/48 (67%) of absorbers are found within $\pm 100 \text{ km s}^{-1}$. This $\pm 100 \text{ km s}^{-1}$ threshold also applies to absorber EW , with only 1 absorber of $EW \geq 400$ found with $\Delta v > 100 \text{ km s}^{-1}$. Blueshifted absorbers are on average closer to their associated galaxy, with $\overline{\Delta v}_{blue} = 68 \pm 16 \text{ km s}^{-1}$, compared to $\overline{\Delta v}_{red} = -108 \pm 20 \text{ km s}^{-1}$ for the redshifted sample, and correspondingly have higher average equivalent width, $\overline{EW}_{blue} = 329 \pm 52 \text{ mÅ}$ compared to $\overline{EW}_{red} = 245 \pm 34 \text{ mÅ}$.

Additionally, of the 48 associated absorbers, 29 are matched with the same galaxy as another absorber (for a total of 14 unique galaxies in this subset). All but one of these cases involve two absorbers in the same sightline yet separated in velocity around a galaxy. 23/29 of these are oriented such that the higher EW absorber has the smaller Δv , and the 6 others are close in either velocity or EW . The one galaxy with 3 associated absorbers, NGC1097, shows this trend across two sightlines as well, with absorbers at $\Delta v = 32 \text{ km s}^{-1}$ and $EW = 570 \text{ mÅ}$ towards 2dFGRS.S393Z082, and $\Delta v = -39 \text{ km s}^{-1}$ and $EW = 184 \text{ mÅ}$ and $\Delta v = 50 \text{ km s}^{-1}$ and $EW = 83 \text{ mÅ}$ towards HE0241-3043.

This result is opposite what we might expect selection effects associated with our likelihood method to produce. Because \mathcal{L} is small for both high Δv and high ρ/R_{vir} , there should be mostly low ρ/R_{vir} systems at high Δv . Low ρ/R_{vir} systems should also have higher EW on average, as evidenced in the $EW - \rho$ anti-correlation discussed above. Figure 7 shows the opposite however, with only low EW systems at high Δv . It must therefore be the case that EW tends to anti-correlate with both Δv and ρ/R_{vir} .

3.4. Azimuth

In this section we examine properties of absorbers as a function of their azimuthal angle with respect to their associated galaxy. Azimuth is defined as the angle between the major axis of a galaxy and the vector connecting the absorption feature and the midpoint of the galaxy plane. Figure 8 illustrates this.

The mean azimuth angle for blueshifted absorbers is $43 \pm 5^\circ$, and $49 \pm 5^\circ$ for redshifted absorbers. Figure 9 shows the distribution of azimuth angles for both red and blue-shifted absorbers. Unlike the findings of Kacprzak et al. (2011b, 2012a), who find a bimodal distribution of Mg II absorbers around galaxies, our distributions of Ly α absorbers are generally consistent with a flat, or random distribution. There is possibly a slight overabundance of redshifted absorbers around 0° (minor-axis) and blueshifted absorbers between $20 - 50^\circ$ (just off major-axis), but we cannot assign this observation much significance yet given the small sample size. We additionally find no significant correlation between azimuth angle and EW or Δv . See Figure 10 for a map of the locations of absorbers relative to their associated galaxies, split between red and blue-shifted absorbers and into three bins of inclination.

These contrasting results may indicate a genuine difference between the properties of Ly α and metal lines, since the distribution of the latter is thought to be influenced by outflows, which may be focused along the minor axis.

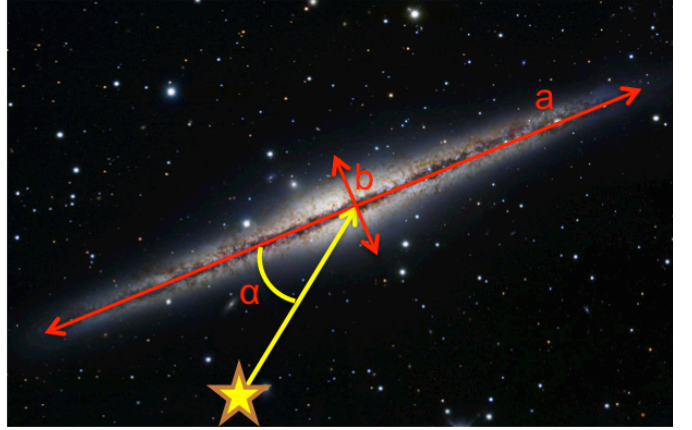


Figure 8. Azimuth is the angle, α , between the major axis of the galaxy, a , and a vector extending from the QSO target to the galaxy center. Image of NGC891 credit: Composite Image Data - Subaru Telescope (NAOJ), Hubble Legacy Archive, Michael Joner, David Laney (West Mountain Observatory, BYU); Processing - Robert Gendler.

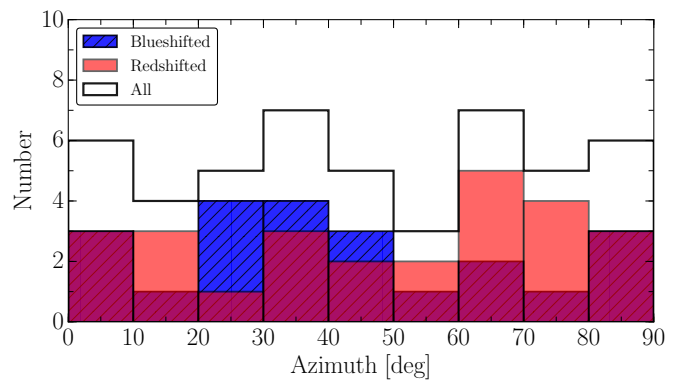


Figure 9. The distributions of azimuth angles for red and blue-shifted samples, with the combined sample plotted in black. Azimuth = 0 corresponds to absorption detected along the projected major axis of the galaxy, and Azimuth = 90 is along the minor axis.

4. DISCUSSION

In this paper we have chosen to separate absorption systems into their individual components, whereas many comparable CGM studies have instead chosen to group together absorption components within some velocity window. While making the assumption that absorbers within some velocity interval are physically linked certainly has merit, it is also possible that absorption components close in velocity may in fact be physically distinct (see, e.g., Churchill et al. 2015).

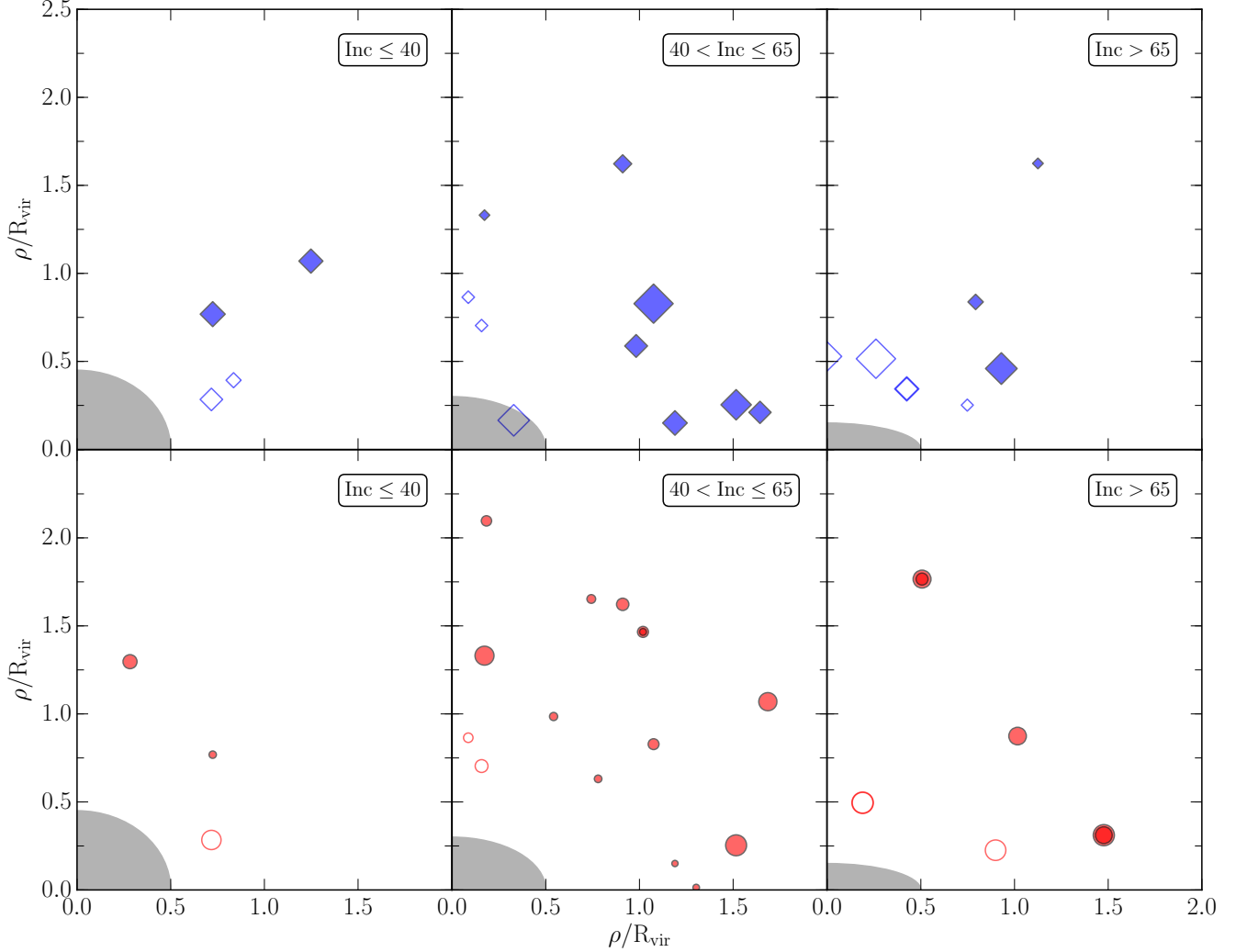


Figure 10. A map of where each absorber was detected with respect to the associated galaxies, separated into three bins of inclination (illustrated by the gray ellipse in the bottom left corner of each plot). **LEFT:** Absorbers associated with galaxies of inclination $0 \leq Inc \leq 40$, **CENTER:** $40 < Inc \leq 65$, and **RIGHT:** $65 \leq Inc$. Blueshifted and redshifted absorbers are separated into the top (diamonds) and bottom (circles) panels, and the marker size is scaled with EW . The sightlines containing multiple absorbers associated with the same galaxy can be identified by their darker color. **Open symbols correspond to systems with $\rho \leq R_{vir}$.**

Indeed, in Section 3.3 we identified several systems where the EW of components of a possible Ly α system individually anti-correlate with Δv . We will thus explore both methods in Paper II, where the much larger sample size will allow for a meaningful comparison.

Restricting ourselves to low redshift systems has several benefits and consequences. We are able to extend our search for associated galaxies to larger impact parameters than many related, higher-redshift CGM studies because of the availability of galaxy data. However, due to the observed anti-correlation between EW and impact parameter, this results in a larger fraction of low- EW and low-column density absorbers in our sample. Hence, we are likely tracing a region of the CGM not entirely analogous to that traced by, e.g., Kacprzak et al. (2011b), Mathes et al. (2014), and Borthakur et

al. (2015).

Studies focusing on metal-lines (e.g., Mg II and O VI) are generally associated with high-column density Lyman Limit Systems (LLS) which again tend to originate closer to their host galaxies. Most of the Ly α absorbers in this work are low-column density (generally $\log N(\text{H I}) \leq 14$), and originate near or beyond $1 R_{vir}$. At these distances we may actually be probing the interface between CGM associated with individual galaxies and the larger-scale network of intergalactic gas filaments, thus the lack of any correlation with azimuth angle is not wholly unexpected.

We do, however, detect an inclination effect on the density, and possibly EW , of Ly α absorbers. The combination of no azimuthal dependence and increased absorber density with inclination leads us to conclude that these galaxies have disc-like, oblate-spheroidal halos. A

perfectly spheroidal halo would show no correlation for either, and an extremely flattened halo would show up as enhanced number density along the major axis. These results are consistent with a picture where HI covering fraction steadily decreases from \sim unity very near to galaxy discs out to at least 1 Mpc, where gas associated with galaxies merges with the general IGM. Our larger upcoming dataset will provide the statistics necessary to probe this in finer detail, as well as give clues to the exact shape of this falloff and the level of clumpiness or filamentary structure in galaxies HI halos.

5. SUMMARY

We have introduced a novel, unbiased likelihood-method for associating absorption systems with nearby galaxies, and explored its implementation with a small subsample of 33 COS sightlines. Associating CGM absorbers with individual galaxies remains a difficult and ambiguous affair, but with this new metric we can at least do so in a reproducible manner.

In this pilot sample we have measured 48 Ly α absorption lines in the spectra of 33 COS targets and matched each to a single, large ($D \geq 25$ kpc) galaxy. Table 3 presents a breakdown of our results when separating absorber-galaxy pairs into red and blue-shifted samples. The following summarizes our findings:

- We introduce a likelihood parameter, \mathcal{L} , based on Gaussian profiles centered around ρ/R_{eff} and $\Delta v/v_{norm}$ to automate the matching of absorbers with associated galaxies. The response of \mathcal{L} can be tailored by choosing different values for R_{eff} and v_{norm} (we used $R_{eff} = [R_{vir}, D^{1.5}]$ and $v_{norm} = 200 \text{ km s}^{-1}$ in this work, and will explore other parameterizations in a future paper).
- Equivalent width (EW) anti-correlates most strongly with ρ when normalized by R_{vir} . It follows that EW weakly correlates and anti-correlates with R_{vir} and ρ , respectively.
- The mean and maximal EW of absorbers increases with decreasing Δv . The strongest absorbers are nearly all found within $\Delta v = \pm 100 \text{ km s}^{-1}$ of their associated galaxies.
- Ly α absorbers are most commonly associated with inclined galaxies. 73% of blueshifted and 77% of redshifted absorbers are associated with galaxies with $i \geq$

50 deg, whereas 56% of all galaxies in the survey volume have similarly high inclinations. The distributions of associated vs all galaxy inclinations differ at a greater than 99% confidence, or $\sim 3.6\sigma$, level according to the Anderson-Darling distribution test.

- We find no strong evidence for azimuth preference for absorption - Ly α absorbers appear to be distributed uniformly around galaxy major and minor axes.

In a future paper we will apply this method to a sample of hundreds of COS sightlines in an effort to produce the most statistically robust CGM study to date.

Statistic	Blueshifted Absorbers	Redshifted Absorbers
Number	22	26
Mean EW [mÅ]	329 ± 52	245 ± 34
Median EW [mÅ]	292 ± 16	177 ± 10
Mean R_{vir} [kpc]	215 ± 10	224 ± 10
Mean ρ [kpc]	218 ± 17	298 ± 23
Mean Δv [km s $^{-1}$]	68 ± 16	-108 ± 20
Mean Inc. [deg]	58 ± 4	61 ± 4
Mean Az. [deg]	43 ± 5	49 ± 5

Table 3. Average properties of the associated galaxy sample split into red and blue-shifted bins based on Δv . **All reported errors are standard errors in the mean.**

We would like to thank the referee for his/her valuable comments and suggestions. This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Based on observations with the NASA/ESA *Hubble Space Telescope*, obtained at the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. Spectra were retrieved from the Barbara A. Mikulski Archive for Space Telescopes (MAST) at STScI. Over the course of this study, D.M.F. and B.P.W. were supported by grant AST-1108913, awarded by the US National Science Foundation, and by NASA grants *HST*-AR-12842.01-A, *HST*-AR-13893.01-A, and *HST*-GO-14240 (STScI).

Facility: HST (COS)