# Project AMIGA: The Circumgalactic Medium of Andromeda[*]

Nicolas Lehner[1] ⬥, Samantha C. Berek[1,2,20], J. Christopher Howk[1] ⬥, Bart P. Wakker[3] ⬥, Jason Tumlinson[4,5] ⬥,
Edward B. Jenkins[6] ⬥, J. Xavier Prochaska[7] ⬥, Ramona Augustin[4] ⬥, Suoqing Ji[8] ⬥, Claude-André Faucher-Giguère[9] ⬥,
Zachary Hafen[9] ⬥, Molly S. Peeples[4,5] ⬥, Kat A. Barger[10] ⬥, Michelle A. Berg[1] ⬥, Rongmon Bordoloi[11] ⬥, Thomas M. Brown[4] ⬥,
Andrew J. Fox[12] ⬥, Karoline M. Gilbert[4,5] ⬥, Puragra Guhathakurta[7] ⬥, Jason S. Kalirai[13], Felix J. Lockman[14] ⬥,
John M. O'Meara[15], D. J. Pisano[16,17,21], Joseph Ribaudo[18] ⬥, and Jessica K. Werk[19] ⬥

[1] Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA
[2] Department of Astronomy, Yale University, New Haven, CT 06511, USA
[3] Department of Astronomy, University of Wisconsin–Madison, WI 53706, USA
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[5] Department of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA
[6] Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA
[7] UCO/Lick Observatory, Department of Astronomy & Astrophysics, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[8] TAPIR, Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA
[9] CIERA and Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA
[10] Department of Physics & Astronomy, Texas Christian University, Fort Worth, TX 76129, USA
[11] North Carolina State University, Department of Physics, Raleigh, NC 27695-8202, USA
[12] AURA for ESA, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[13] Johns Hopkins Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA
[14] Green Bank Observatory, Green Bank, WV 24944, USA
[15] W. M. Keck Observatory 65-1120 Mamalahoa Highway Kamuela, HI 96743, USA
[16] Department of Physics & Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA
[17] Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA
[18] Department of Engineering and Physics, Providence College, Providence, RI 02918, USA
[19] Department of Astronomy, University of Washington, Seattle, WA 98195, USA

## Abstract

Project AMIGA (Absorption Maps In the Gas of Andromeda) is a survey of the circumgalactic medium (CGM) of Andromeda (M31, $R_{\rm vir} \simeq 300$ kpc) along 43 QSO sightlines at impact parameters $25 \leqslant R \leqslant 569$ kpc (25 at $R \lesssim R_{\rm vir}$). We use ultraviolet absorption measurements of Si II, Si III, Si IV, C II, and C IV from the Hubble Space Telescope/Cosmic Origins Spectrograph and O VI from the Far Ultraviolet Spectroscopic Explorer to provide an unparalleled look at how the physical conditions and metals are distributed in the CGM of M31. We find that Si III and O VI have a covering factor near unity for $R \lesssim 1.2\,R_{\rm vir}$ and $\lesssim 1.9\,R_{\rm vir}$, respectively, demonstrating that M31 has a very extended $\sim 10^4$–$10^{5.5}$ K ionized CGM. The metal and baryon masses of the $10^4$–$10^{5.5}$ K CGM gas within $R_{\rm vir}$ are $\gtrsim 10^8$ and $\gtrsim 4 \times 10^{10}\,(Z/0.3\,Z_\odot)^{-1}\,M_\odot$, respectively. There is not much azimuthal variation in the column densities or kinematics, but there is with $R$. The CGM gas at $R \lesssim 0.5\,R_{\rm vir}$ is more dynamic and has more complicated, multiphase structures than at larger radii, perhaps a result of more direct impact of galactic feedback in the inner regions of the CGM. Several absorbers are projected spatially and kinematically close to M31 dwarf satellites, but we show that those are unlikely to give rise to the observed absorption. Cosmological zoom simulations of $\sim L^*$ galaxies have O VI extending well beyond $R_{\rm vir}$ as observed for M31 but do not reproduce well the radial column density profiles of the lower ions. However, some similar trends are also observed, such as the lower ions showing a larger dispersion in column density and stronger dependence on $R$ than higher ions. Based on our findings, it is likely that the Milky Way has a $\sim 10^4$–$10^{5.5}$ K CGM as extended as for M31 and their CGM (especially the warm–hot gas probed by O VI) are overlapping.

*Unified Astronomy Thesaurus concepts:* Circumgalactic medium (1879); Andromeda Galaxy (39); Local Group (929); Quasar absorption line spectroscopy (1317)

*Supporting material:* figure sets, machine-readable tables

## 1. Introduction

Over the past 10 yr, in particular since the installation of the Cosmic Origins Spectrograph (COS) on the Hubble Space Telescope (HST), we have made significant leaps in empirically characterizing the circumgalactic medium (CGM) of galaxies at low redshift, where a wide range of galaxy masses can be studied (see recent review by Tumlinson et al. 2017). We appreciate now that the CGM of typical star-forming or quiescent galaxies has a large share of galactic baryons and metals in relatively cool gas phases ($10^4$–$10^{5.5}$ K; e.g., Stocke et al. 2013; Bordoloi et al. 2014; Liang & Chen 2014; Peeples et al. 2014; Werk et al. 2014; Johnson et al. 2015; Burchett et al. 2016; Prochaska et al. 2017b; Chen et al. 2019; Poisson et al. 2020). We have come to understand that the CGM of galaxies at $z \lesssim 1$ is not just filled with metal-enriched gas ejected by successive galaxy outflows, but has also a large amount of

metal-poor gas (<1%–2% solar) in which little net chemical enrichment has occurred over several billions of years (e.g., Ribaudo et al. 2011; Thom et al. 2011; Lehner et al. 2013, 2018, 2019; Wotta et al. 2016, 2019; Prochaska et al. 2017b; Kacprzak et al. 2019; Zahedy et al. 2019; Poisson et al. 2020). The photoionized gas around $z \lesssim 1$ galaxies is very chemically inhomogeneous, as shown by large metallicity ranges and the large metallicity variations among kinematically distinct components in a single halo (Lehner et al. 2019; Wotta et al. 2019; see also Crighton et al. 2013; Muzahid et al. 2015, 2016; Rosenwasser et al. 2018). Such a large metallicity variation is observed not only in the CGM of star-forming galaxies but also in the CGM of passive and massive galaxies, where there appears to be as much cold, bound H I gas as in their star-forming counterparts (e.g., Thom et al. 2012; Tumlinson et al. 2013; Berg et al. 2019; Zahedy et al. 2019).

These empirical results have revealed both expected and unexpected properties of the CGM of galaxies, and they all provide new means to understand the complex relationship between galaxies and their CGM. Prior to these empirical results, the theory of galaxy formation and evolution was mostly left constraining the CGM properties indirectly by their outcomes, such as galaxy stellar mass and interstellar medium (ISM) properties. Thus, the balance between outflows, inflows, recycling, and ambient gas—and the free parameters controlling them—was tuned to match the optical properties of galaxies rather than implemented directly as physically rigorous and self-consistent models. These indirect constraints suffer from problems of model uniqueness: it is possible to match stellar masses and metallicities with very different treatments of feedback physics (e.g., Hummels et al. 2013; Liang et al. 2016). Recent empirical and theoretical advances offer a way out of this model degeneracy. New high-resolution, zoom-in simulations employ explicit treatments of the multiple gas-phase nature and feedback from stellar population models (e.g., Hopkins et al. 2014, 2018). It is also becoming clear that high resolution not only inside the galaxies but also in their CGM is required to capture more accurately the complex processes in the cool CGM, such as metal mixing (Hummels et al. 2019; Peeples et al. 2019; Suresh et al. 2019; van de Voort et al. 2019; Corlies et al. 2020).

A significant limitation in interpreting the new empirical results in the context of new high-resolution zoom simulations is that only average properties of the CGM are robustly derived from traditional QSO absorption-line techniques for examining halo gas. In the rare cases where there is a UV-bright QSO behind a given galaxy, the CGM is typically probed along a single "core sample" through the halo of each galaxy. These measurements are then aggregated into a statistical map, where galaxies with different inclinations, sizes, and environments are blended together and the radial–azimuthal dependence of the CGM is essentially lost. All sorts of biases can result: phenomena that occur strongly in only a subset of galaxies can be misinterpreted as being weaker but more common, and genuine trends with mass or star formation rate (SFR) can be misinterpreted as simply scatter with no real physical meaning (see also Bowen et al. 2016). Simulations also suggest that time-variable winds, accretion flows, and satellite halos can induce strong halo-to-halo variability, further complicating interpretation (e.g., Hafen et al. 2017; Oppenheimer et al. 2018a). Observational studies of single-galaxy CGM with multiple sightlines are therefore required to gain spatial information on the properties of the CGM.

Multi-sightline information on the CGM of single galaxies has been obtained in a few cases from binary or multiple (two to four) grouped QSOs behind foreground galaxies (e.g., Bechtold et al. 1994; Martin et al. 2010; Keeney et al. 2013; Bowen et al. 2016), gravitationally lensed quasars (e.g., Smette et al. 1992; Rauch et al. 2001; Ellison et al. 2004; Lopez et al. 2005; Zahedy et al. 2016; Rubin et al. 2018; Kulkarni et al. 2019), giant gravitational arcs (e.g., Lopez et al. 2020), or extended bright background objects observed with integral field units (e.g., Péroux et al. 2018). These observations provide better constraints on the kinematic relationship between the CGM gas and the galaxy and on the size of CGM structures. However, they yield limited information on the gas-phase structure owing to a narrow range of ionization diagnostics or poor-quality spectral data. Thus, it remains unclear how tracers of different gas phases vary with projected distance $R$ or azimuth $\Phi$ around the galaxy.

The CGM that has been pierced the most is that of the Milky Way (MW), with several hundred QSO sightlines (Wakker et al. 2003; Shull et al. 2009; Lehner et al. 2012; Putman et al. 2012; Richter et al. 2017) through the Galactic halo. However, our position as observers within the MW disk severely limits the interpretation of these data (especially for the extended CGM; see Zheng et al. 2015, 2020) and makes it difficult to compare with observations of other galaxies.

With a virial radius that spans over $23°$ on the sky, M31 is the only $L^*$ galaxy where we can access more than five sightlines without awaiting the next generation of UV space-based telescopes (e.g., The LUVOIR Team 2019). With current UV capabilities, it is the only single galaxy where we can study the global distribution and properties of metals and baryons in some detail.

In our pilot study (Lehner et al. 2015, hereafter LHW15), we mined the HST/COS G130M/G160M archive available at the Barbara A. Mikulski Archive for Space Telescopes (MAST) for sightlines piercing the M31 halo within a projected distance of $\sim 2\,R_{\rm vir}$ (where $R_{\rm vir} = 300\,{\rm kpc}$ for M31; see below). There were 18 sightlines, but only 7 at projected distance $R \lesssim R_{\rm vir}$. Despite the small sample, the results of this study were quite revealing, demonstrating a high covering factor (6/7) of M31 CGM absorption by Si III (and other ions, including, e.g., C IV, Si II) within $1.1\,R_{\rm vir}$ and a covering factor near zero (1/11) between $1.1\,R_{\rm vir}$ and $2\,R_{\rm vir}$. We found also a drastic change in the ionization properties, as the gas is more highly ionized at $R \sim R_{\rm vir}$ than at $R < 0.2\,R_{\rm vir}$. The LHW15 results strongly suggest that the CGM of M31 as seen in absorption of low ions (C II, Si II) through intermediate (Si III, Si IV) and high ions (C IV, O VI) is very extended out to at least the virial radius. However, owing to the small sample within $R_{\rm vir}$, the variation of the column densities ($N$) and covering factors ($f_{\rm c}$) with projected distances and azimuthal angle remains poorly constrained.

Our Project Absorption Maps In the Gas of Andromeda (AMIGA) is a large HST program (PID: 14268; PI: Lehner) that aims to fill the CGM with 18 additional sightlines at various $R$ and $\Phi$ within $1.1\,R_{\rm vir}$ of M31 using high-quality COS G130M and G160M observations, yielding a sample of 25 background QSOs probing the CGM of M31. We have also searched MAST for additional QSOs beyond $1.1\,R_{\rm vir}$ up to $R = 569\,{\rm kpc}$ from M31 ($\sim 1.9\,R_{\rm vir}$) to characterize the

extended gas around M31 beyond its virial radius. This archival search yielded 18 suitable QSOs. Our total sample of 43 QSOs probing the CGM of a single galaxy from 25 to 569 kpc is the first to explore simultaneously the azimuthal and radial dependence of the kinematics, ionization level, surface densities, and mass of the CGM of a galaxy over its entire virial radius and beyond. With these observations, we can also test how the CGM properties derived from one galaxy using multiple sightlines compare with a sample of galaxies with single sightline information, and we can directly compare the results with cosmological zoom-in simulations.

With the COS G130M and G160M wavelength coverage, the key ions in our study are C II, C IV, Si II, Si III, and Si IV (other ions and atoms include Fe II, S II, O I, N I, and N V, but these are typically not detected, although the limit on O I constrains the level of ionization). These species span ionization potentials from <1 to ∼4 ryd and thus trace neutral to highly ionized gas at a wide range of temperatures and densities. We have also searched the Far Ultraviolet Spectroscopic Explorer (FUSE) to have coverage of O VI, which resulted in 11 QSOs in our sample having both COS and FUSE observations. The H I Lyα absorption can unfortunately not be used because the MW dominates the entire Lyα absorption. Instead, we have obtained deep H I 21 cm observations with the Robert C. Byrd Green Bank Telescope (GBT) toward all the targets in our sample and several additional ones (Howk et al. 2017, hereafter Paper I), showing no detection of any H I down to a level of $N_{\mathrm{H\,I}} \simeq 4 \times 10^{17}\,\mathrm{cm}^{-2}$ (5σ; averaged over an area that is 2 kpc at the distance of M31). Our nondetections place a limit on the covering factor of such optically thick H I gas around M31 to $f_{\mathrm{c}} < 0.051$ (at 90% confidence level) for $R \lesssim R_{\mathrm{vir}}$.

This paper is organized as follows. In Section 2, we provide more information about the criteria used to assemble our sample of QSOs and explain the various steps to derive the properties (velocities and column densities) of the absorption. In that section, we also present the line identification for each QSO spectrum, which resulted in the identification of 5642 lines. In Section 2.5, we explain in detail how we remove the foreground contamination from the Magellanic Stream (MS; e.g., Putman et al. 2003; Nidever et al. 2008; Fox et al. 2014), which extends to the M31 CGM region of the sky with radial velocities that overlap with those expected from the CGM of M31. For this work, we have developed a more systematic and automated methodology than in LHW15 to deal with this contamination. In Section 3, we present the sample of the M31 dwarf satellite galaxies to which we compare the halo gas measurements. In Section 4, we derive the empirical properties of the CGM of M31, including how the column densities and velocities vary with $R$ and $\Phi$, the covering factors of the ions and how they change with $R$, and the metal and baryon masses of the CGM of M31. In Section 5, we discuss the results derived in Section 4 and compare them to observations from the COS-Halos survey (Tumlinson et al. 2013; Werk et al. 2014) and to state-of-the-art cosmological zoom-ins from, in particular, the Feedback in Realistic Environments (FIRE; Hopkins et al. 2020) and the Figuring Out Gas and Galaxies In Enzo (FOGGIE; Peeples et al. 2019) simulation projects. In Section 6, we summarize our main conclusions.

To properly compare to other work and to simulations, we must estimate a characteristic radius for M31. We use the radius $R_{200}$ enclosing a mean overdensity of 200 times the critical density: $R_{200} = (3M_{200}/4\pi\,\Delta\,\rho_{\mathrm{crit}})^{1/3}$, where $\Delta = 200$ and $\rho_{\mathrm{crit}}$ is the critical density. For M31, we adopt $M_{200} = 1.26 \times 10^{12}\,M_{\odot}$ (e.g., Watkins et al. 2010; van der Marel et al. 2012), implying $R_{200} \simeq 230$ kpc. For the virial mass and radius ($M_{\mathrm{vir}}$ and $R_{\mathrm{vir}}$), we use the definition that follows from the top-hat model in an expanding universe with a cosmological constant where $M_{\mathrm{vir}} = 4\pi/3\,\rho_{\mathrm{vir}}\,R_{\mathrm{vir}}^{3}$, where the virial density $\rho_{\mathrm{vir}} = \Delta_{\mathrm{vir}}\Omega_{\mathrm{m}}\rho_{\mathrm{crit}}$ (Klypin et al. 2011; van der Marel et al. 2012). The average virial overdensity is $\Delta_{\mathrm{vir}} = 360$ assuming a cosmology with $h = 0.7$ and $\Omega_{\mathrm{m}} = 0.27$ (Klypin et al. 2011). Following, e.g., van der Marel et al. (2012), $M_{\mathrm{vir}} \simeq 1.2M_{200} \simeq 1.5 \times 10^{12}\,M_{\odot}$ and $R_{\mathrm{vir}} \simeq 1.3R_{200} \simeq 300$ kpc. The escape velocity at $R_{200}$ for M31 is then $v_{200} \simeq 212\,\mathrm{km\,s}^{-1}$. A distance of M31 of $d_{\mathrm{M31}} = 752 \pm 27$ kpc based on the measurements of Cepheid variables (Riess et al. 2012) is assumed throughout. We note that this distance is somewhat smaller than the other often-adopted distance of M31 of 783 kpc (e.g., Brown et al. 2004; McConnachie et al. 2005), but for consistency with our previous survey, as well as the original design of Project AMIGA, we have adopted $d_{\mathrm{M31}} = 752$ kpc. All the projected distances were computed using the 3D separation (coordinates of the target and distance of M31).

## 2. Data and Analysis

### 2.1. The Sample

The science goals of our HST large program require estimating the spatial distributions of the kinematics and metal column densities of the M31 CGM gas within about 1.1 $R_{\mathrm{vir}}$ as a function of azimuthal angle and impact parameter. The search radius was selected based on our pilot study, where we detected M31 CGM gas up to ∼1.1 $R_{\mathrm{vir}}$, but essentially not beyond (LHW15) (a finding that we revisit in this paper with a larger archival sample; see below). With our HST program, we observed 18 QSOs at $R \lesssim 1.1\,R_{\mathrm{vir}}$ that were selected to span the M31 projected major axis, minor axis, and intermediate orientations. The sightlines do not sample the impact parameter space or azimuthal distribution at random. Instead, the sightlines were selected to probe the azimuthal variations systematically. The sample was also limited by a general lack of identified UV-bright active galactic nuclei (AGNs) behind the northern half of M31's CGM owing to higher foreground MW dust extinction near the plane of the MW disk. Combined with seven archival QSOs, these sightlines probe the CGM of M31 in azimuthal sectors spanning the major and minor axes with a radial sample of 7–10 QSOs in each ∼100 kpc bin in $R$.

In addition to target locations, the 18 QSOs were optimized to be the brightest available QSOs (to minimize exposure time) and to have the lowest available redshifts (in order to minimize the contamination from unrelated absorption from high-redshift absorbers). For targets with no existing UV spectra prior to our observations, we also required that the Galaxy Evolution Explorer near-UV (NUV) and far-UV (FUV) flux magnitudes are about the same to minimize the likelihood of an intervening Lyman limit system (LLS) with optical depth at the Lyman limit $\tau_{\mathrm{LL}} > 2$. An intervening LLS could absorb more or all of the QSO flux we would need to measure foreground absorption in M31. This strategy successfully kept QSOs with intervening LLS out of the sample.
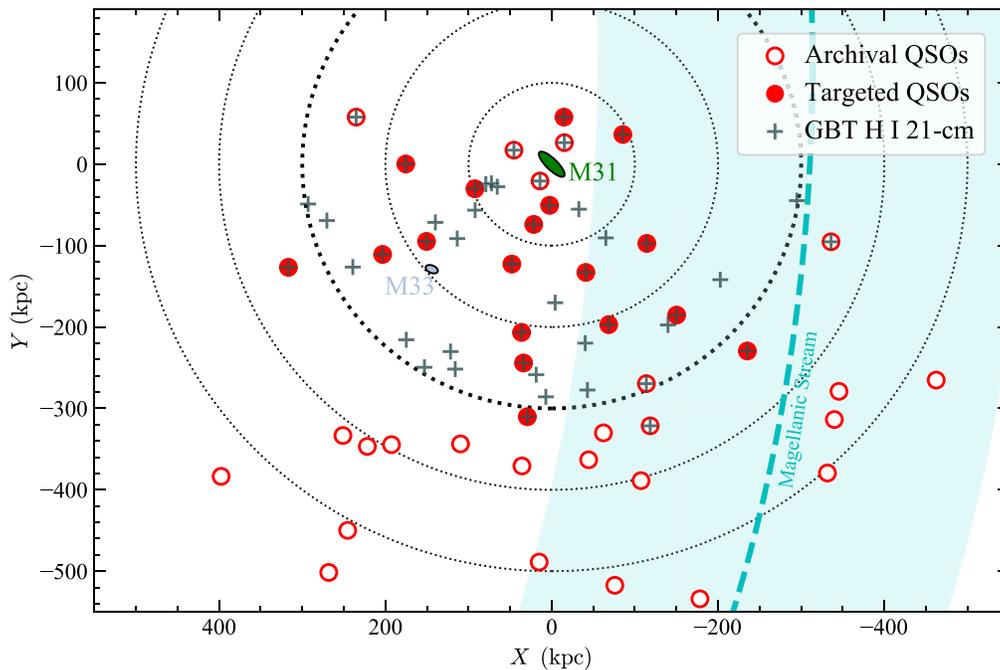
**Figure 1.** Locations of the Project AMIGA pointings relative to the M31–M33 system. The axes show the angular separations converted into physical coordinates relative to the center of M31. North is up and east to the left. The 18 sightlines from our large HST program are shown with red filled circles; the 25 archival COS targets are shown with open red circles. Plus signs show the GBT H I 21 cm observations described in Paper I. Dotted circles show impact parameters $R = 100$, 200, 300, 400, and 500 kpc. $R_{\rm vir} = 300$ kpc is marked with a thick dashed line. The sizes and orientations of the two galaxies are taken from RC3 (de Vaucouleurs et al. 1991) and correspond to the optical $R_{25}$ values. The light-blue dashed line shows the plane of the MS ($b_{\rm MS} = 0°$) as defined by Nidever et al. (2008). The shaded region within $b_{\rm MS} \pm 20°$ of the MS midplane is the approximate region where we identify most of the MS absorption components contaminating the M31 CGM absorption (see Section 2.5).

As we discuss below and as detailed by LHW15, the MS crosses through the M31 region of the sky at radial velocities that can overlap with those of M31 (see also Nidever et al. 2008; Fox et al. 2014). To understand the extent of MS contamination and the extended gas around M31 beyond the virial radius, we also searched for targets beyond 1.1 $R_{\rm vir}$ with COS G130M and/or G160M data. This search identified another 18 QSOs at $1.1 \lesssim R/R_{\rm vir} < 1.9$ that met the data quality criteria for inclusion in the sample.[22] Our final sample consists of 43 sightlines probing the CGM of M31 from 25 to 569 kpc; 25 of these probe the M31 CGM from 25 to 342 kpc, corresponding to 0.08–1.1 $R_{\rm vir}$. Figure 1 shows the locations of each QSO in the M31–M33 system (the filled circles being targets obtained as part of our HST program PID: 14268 and the open circles being QSOs with archival HST COS G130M/ G160M data), and Table 1 lists the properties of our sample QSOs ordered by increasing projected distances from M31. In this table, we list the redshift of the QSOs ($z_{\rm em}$), the J2000 right ascension (R.A.) and declination (decl.), the MS coordinate ($l_{\rm MS}$, $b_{\rm MS}$; see Nidever et al. 2008 for the definition of this coordinate system), the radially ($R$) and Cartesian ($X$, $Y$) projected distances, the program identification of the HST program (PID), the COS grating used for the observations of the targets, and the signal-to-noise ratio (S/N) per COS resolution element of the COS spectra near the Si III transition.

## 2.2. UV Spectroscopic Calibration

To search for M31 CGM absorption and to determine the properties of the CGM gas, we use ions and atoms that have their wavelengths in the UV (see Section 2.4). Any transitions with $\lambda > 1144$ Å are in the HST COS bandpass. All the targets in our sample were observed with HST using the COS G130M grating ($R_\lambda \approx 17{,}000$). All the targets observed as part of our new HST program were also observed with COS G160M, and all the targets but one within $R < 1.1$ $R_{\rm vir}$ have both G130M and G160M wavelength coverage.

We also searched for additional archival UV spectra in MAST, including the FUSE ($R_\lambda \approx 15{,}000$) archive to complement the gas-phase diagnostics from the COS spectra with information from the O VI absorption. We use the FUSE observations for 11 targets with adequate S/N near O VI (i.e., $\gtrsim 5$): RX J0048.3+3941, IRAS F00040+4325, Mrk 352, PG 0052+251, Mrk 335, UGC 12163, PG 0026+129, Mrk 1502, NGC 7469, Mrk 304, and PG 2349−014 (only the first six targets in this list are at $R \lesssim 1.1$ $R_{\rm vir}$). We did not consider FUSE data for quasars without COS observations because the available UV transitions in the FUV spectrum (O VI, C II, C III, Si II, Fe II) are either too weak or too contaminated to allow for a reliable identification of the individual velocity components in their absorption profiles.

There are also three targets (Mrk 335, UGC 12163, and NGC 7469) with HST STIS E140M ($R_\lambda \simeq 46{,}500$) observations that provide higher-resolution information.[23]

---

[22] This search found eight additional targets at $R > 1.6$ $R_{\rm vir}$ that are not included in our sample. SDSS J021348.53+125951.4, 4C 10.08, and LBQS 0052-0038 were excluded because of low S/N in the COS data. NGC 7714 has smeared absorption lines. LBQS 0107-0232/3/5 lie at $z_{\rm em} \simeq 0.7–1$ and have extremely complex spectra. HS 2154+2228 at $z_{\rm em} = 1.29$ has no G130M wavelength coverage, making the line identification highly uncertain.

[23] For two targets, we also use COS G225M (3C 454.3) and FOS NUV (3C 454.3, PG 0044+030) observations to help with the line identification (see Section 2.3). The data processing follows the same procedure as the other data.

**Table 1**
Sample Summary

| Target | $z_{em}$ | R.A. (deg) | Decl. (deg) | $l_{MS}$ (deg) | $b_{MS}$ (deg) | $R$ (kpc) | $X$ (kpc) | $Y$ (kpc) | PID | COS Grating | S/N |
|---|---|---|---|---|---|---|---|---|---|---|---|
| RX J0048.3+3941 | 0.134 | 12.079 | 39.687 | −125.0 | 24.8 | 25.0 | 14.1 | −20.7 | 11632 | G130M-G160M | 31.5 |
| HS 0033+4300 | 0.120 | 9.096 | 43.278 | −129.1 | 22.7 | 30.5 | −15.2 | 26.5 | 11632 | G130M-G160M | 5.9 |
| HS 0058+4213 | 0.190 | 15.380 | 42.493 | −128.1 | 27.3 | 48.6 | 45.4 | 17.3 | 11632 | G130M-G160M | 8.1 |
| RX J0043.6+3725 | 0.080 | 10.927 | 37.422 | −122.6 | 23.7 | 50.5 | 2.5 | −50.5 | 14268 | G130M-G160M | 17.9 |
| Zw 535.012 | 0.048 | 9.087 | 45.665 | −131.7 | 22.9 | 59.7 | −14.7 | 57.8 | 14268 | G130M-G160M | 17.8 |
| RX J0050.8+3536 | 0.058 | 12.711 | 35.612 | −120.5 | 25.0 | 77.1 | 21.6 | −74.0 | 14268 | G130M-G160M | 18.7 |
| IRAS F00040+4325 | 0.163 | 1.652 | 43.708 | −130.1 | 17.4 | 93.0 | −85.5 | 36.5 | 14268 | G130M-G160M | 24.9 |
| RXS J0118.8+3836 | 0.216 | 19.706 | 38.606 | −123.7 | 30.7 | 97.2 | 92.3 | −30.2 | 14268 | G130M-G160M | 14.4 |
| Mrk 352 | 0.015 | 14.972 | 31.827 | −116.1 | 26.5 | 131.7 | 48.0 | −122.7 | 14268 | G130M-G160M | 12.2 |
| RX J0028.1+3103 | 0.500 | 7.045 | 31.063 | −116.3 | 19.7 | 139.1 | −41.1 | −132.9 | 14268 | G130M-G160M | 13.6 |
| KAZ 238 | 0.043 | 0.242 | 33.344 | −119.7 | 14.5 | 150.2 | −114.4 | −97.3 | 14268 | G130M-G160M | 11.5 |
| FBS 0150+396 | 0.212 | 28.278 | 39.929 | −125.9 | 37.2 | 175.5 | 175.5 | 0.2 | 14268 | G130M-G160M | 10.3 |
| 3C 48.0 | 0.367 | 24.422 | 33.160 | −117.3 | 34.6 | 177.9 | 150.6 | −94.9 | 14268 | G130M-G160M | 16.6 |
| 4C 25.01 | 0.284 | 4.916 | 26.048 | −111.4 | 17.0 | 208.7 | −68.6 | −197.1 | 14268 | G130M-G160M | 18.6 |
| PG 0052+251 | 0.155 | 13.717 | 25.427 | −109.2 | 24.7 | 209.8 | 36.3 | −206.7 | 14268 | G130M-G160M | 28.6 |
| RXS J0155.6+3115 | 0.135 | 28.900 | 31.255 | −115.0 | 38.4 | 231.6 | 203.4 | −110.8 | 14268 | G130M-G160M | 17.2 |
| RBS 2055 | 0.038 | 357.970 | 26.326 | −113.2 | 11.0 | 238.6 | −150.3 | −185.4 | 14268 | G130M-G160M | 22.3 |
| 3C 66A | 0.444 | 35.665 | 43.035 | −131.3 | 41.9 | 242.2 | 235.2 | 57.9 | 12612 | G130M-G160M | 21.4 |
| RX J0053.7+2232 | 0.148 | 13.442 | 22.539 | −106.2 | 23.9 | 246.6 | 33.9 | −244.2 | 14268 | G130M-G160M | 14.2 |
| Mrk 335 | 0.026 | 1.581 | 20.203 | −106.3 | 12.6 | 292.6 | −113.8 | −269.6 | 11524 | G130M-G160M | 29.8 |
| Mrk 1148 | 0.064 | 12.978 | 17.433 | −100.8 | 22.5 | 311.6 | 29.3 | −310.3 | 14268 | G130M-G160M | 20.9 |
| RBS 2005 | 0.120 | 351.476 | 21.887 | −110.5 | 4.0 | 328.6 | −235.3 | −229.5 | 14268 | G130M-G160M | 16.2 |
| RX J0023.5+1547 | 0.412 | 5.877 | 15.796 | −100.8 | 15.5 | 335.9 | −62.2 | −330.1 | 14071 | G130M | 6.4 |
| Mrk 1179 | 0.038 | 38.343 | 27.937 | −111.0 | 46.7 | 341.0 | 316.6 | −126.6 | 14268 | G130M-G160M | 9.5 |
| PG 0003+158 | 0.451 | 1.497 | 16.164 | −102.3 | 11.5 | 342.5 | −118.4 | −321.4 | 12038 | G130M-G160M | 21.4 |
| UGC 12163 | 0.025 | 340.664 | 29.725 | −121.2 | −2.5 | 349.2 | −335.9 | −95.3 | 12212 | G130M-G160M | 9.9 |
| SDSS J011623.06+142940.6 | 0.394 | 19.096 | 14.495 | −96.2 | 27.6 | 360.7 | 109.7 | −343.6 | 13774 | G130M-G160M | 7.7 |
| PG 0026+129 | 0.142 | 7.307 | 13.268 | −97.9 | 16.2 | 365.8 | −44.5 | −363.1 | 12569 | G130M | 16.9 |
| Mrk 1502 | 0.061 | 13.396 | 12.693 | −95.7 | 21.8 | 372.4 | 35.8 | −370.7 | 12569 | G130M-G160M | 23.2 |
| SDSS J014143.20+134032.0 | 0.045 | 25.430 | 13.676 | −93.7 | 33.4 | 394.8 | 192.7 | −344.5 | 12275 | G130M | 4.6 |
| Mrk 1501 | 0.089 | 2.629 | 10.975 | −96.9 | 11.2 | 403.3 | −107.4 | −388.8 | 12569 | G130M | 3.4 |
| IRAS 01477+1254 | 0.147 | 27.618 | 13.150 | −92.6 | 35.4 | 411.6 | 221.8 | −346.7 | 11727 | G130M-G160M | 4.9 |
| SDSS J015952.95+134554.3 | 0.504 | 29.971 | 13.765 | −92.7 | 37.7 | 417.3 | 251.1 | −333.3 | 12603 | G130M | 11.2 |
| 3C 454.3 | 0.859 | 343.491 | 16.148 | −107.5 | −5.0 | 444.1 | −345.5 | −279.0 | 13398 | G130M-G160M | 6.6 |
| SDSS J225738.20+134045.0 | 0.595 | 344.409 | 13.679 | −104.9 | −5.0 | 462.5 | −339.9 | −313.7 | 11598 | G130M-G160M | 7.9 |
| PG 0044+030 | 0.623 | 11.775 | 3.332 | −86.6 | 17.7 | 489.0 | 15.1 | −488.8 | 12275 | G130M | 8.4 |
| NGC 7469 | 0.016 | 345.815 | 8.874 | −99.9 | −5.3 | 503.9 | −331.6 | −379.4 | 12212 | G130M-G160M | 32.7 |
| PHL 1226 | 0.404 | 28.617 | 4.805 | −82.5 | 34.1 | 512.4 | 245.4 | −449.8 | 12536 | G130M | 8.8 |
| UM 228 | 0.098 | 5.254 | 0.880 | −86.3 | 10.7 | 522.8 | −75.9 | −517.3 | 13017 | G130M-G160M | 7.3 |
| Mrk 304 | 0.066 | 334.301 | 14.239 | −109.0 | −13.9 | 533.1 | −462.4 | −265.2 | 12569 | G130M-G160M | 23.9 |
| Mrk 595 | 0.027 | 40.395 | 7.187 | −80.6 | 46.0 | 552.4 | 397.6 | −383.5 | 12275 | G130M | 11.1 |
| PG 2349−014 | 0.174 | 357.984 | −1.154 | −86.6 | 3.2 | 562.9 | −178.3 | −534.0 | 12569 | G130M | 27.1 |
| Mrk 1014 | 0.163 | 29.959 | 0.395 | −76.9 | 33.9 | 568.7 | 268.1 | −501.5 | 12569 | G130M | 24.3 |

**Note.** The 18 sightlines from our large HST program have the PID 14268; for the supplemental archival 25 other targets, the HST PID is listed. All the projected distances are computed using the 3D separation (coordinates of the target and distance of M31 assumed to be 752 kpc). The coordinates $l_{MS}$ and $b_{MS}$ are the MS longitudes and latitudes as defined by Nidever et al. (2008). The S/N is given per COS resolution element (assuming $R \sim 17{,}000$) and estimated in the continuum near Si III $\lambda1206$.

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

Information on the design and performance of COS, STIS, and FUSE can be found in Green et al. (2012), Woodgate et al. (1998), and Moos et al. (2000), respectively. For the HST data, we use the pipeline-calibrated final data products available in MAST. The HST STIS E140M data have an accurate wavelength calibration, and the various exposure and echelle orders are combined into a single spectrum by interpolating the photon counts and errors onto a common grid, adding the photon counts and converting back to a flux.

The processing of the FUSE data is described in detail by Wakker et al. (2003) and Wakker (2006). In short, the spectra

are calibrated using version 2.1 or version 2.4 of the FUSE calibration pipeline. The wavelength calibration of FUSE can suffer from stretches and misalignments. To correct for residual wavelength shifts, the central velocities of the MW interstellar lines are determined for each detector segment of each individual observation. The FUSE segments are then aligned with the interstellar velocities implied by the STIS E140M spectra or with the velocity of the strongest component seen in the 21 cm H I spectrum. Since the O VI absorption can be contaminated by $H_2$ absorption, we remove this contamination following the method described in Wakker (2006). This

contamination can be removed fairly accurately with an uncertainty of about $\pm 0.1$ dex on the O VI column density (Wakker et al. 2003).

For the COS G130M and G160M spectra, the spectral lines in separate observations of the same target are not always aligned, with misalignments of up to $\pm 40$ km s$^{-1}$ that vary as a function of wavelength. This is a known issue that has been reported previously (e.g., Savage et al. 2014; Wakker et al. 2015). While the COS team has improved the wavelength solution, we find that this problem can still be present sometimes. Since accurate alignment is critical for studying multiple gas phases probed by different ions, and since there is no way to determine a priori which targets are affected, we uniformly apply the Wakker et al. (2015) methodology to co-add the different exposures of the COS data to ensure proper alignment of the absorption lines. In short, we identify the various strong ISM and intergalactic medium (IGM) weak lines and record the component structures and identify possible contamination of the ISM lines by IGM lines. We cross-correlate each line in each exposure, using a $\sim 3$ Å wide region, and apply a shift as a function of wavelength to each spectrum. To determine the absolute wavelength calibration, we compare the velocity centroids of the Gaussian fits to the interstellar UV absorption lines (higher velocity absorption features being Gaussian fitted separately) and the H I emission observed from our 9' GBT H I survey (Paper I) or otherwise from 21 cm data from the Leiden/Argentine/Bonn (LAB) survey (Kalberla et al. 2005) or the Parkes Galactic All Sky Survey (GASS; Kalberla et al. 2010). The alignment is coupled with the line identification into an iterative process to simultaneously determine the most accurate alignment and line identification (see Section 2.3). To combine the aligned spectra, we add the total counts in each pixel and then convert back to flux, using the average flux/count ratio at each wavelength (see also Tripp et al. 2011; Tumlinson et al. 2011); the flux error is estimated from the Poisson noise implied by the total count rate.

### 2.3. Line Identification

We are interested in the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, where absorption from the M31 CGM may occur (see Section 2.4 for the motivation of this velocity range). It is straightforward to identify M31 absorption or its absence in this predefined velocity range, but we must ensure either that there is no contamination from higher-redshift absorbers or, if there is, that we can correct for it.

For ions with multiple transitions, it is relatively simple to determine whether contamination is at play by comparing the column densities and the shapes of the velocity profiles of the available transitions. The profiles of atoms or ions with a single transition can be compared to other detected ions to check whether there is some obvious contamination in the single transition absorption. However, some contamination may still remain undetected if it directly coincides with the absorption under consideration. Furthermore, when only a single ion with a single transition is detected (Si III $\lambda 1206$ being the prime example), the only method that determines whether it is contaminated or not is to undertake a complete line identification of all absorption features in each QSO spectrum.

For the 18 targets in our large HST program, our instrument setup ensures that we have the complete wavelength coverage with no gap between 1140 and 1800 Å. As part of our target

selection, we also favor QSOs at low redshift (44% are at $z_{\rm em} \leqslant 0.1$, 89% at $z_{\rm em} \leqslant 0.3$). This assures that Ly$\alpha$ remains in the observed wavelength range out to the redshift of the QSO (Ly$\alpha$ redshifts out the long end of the COS band at $z = 0.48$) and greatly reduces the contamination from EUV transitions in the COS bandpass. The combination of wavelength coverage and low QSO redshift ensures the most accurate line identification. At $R < 351$ kpc (i.e., $\lesssim 1.2\ R_{\rm vir}$), 93% have Ly$\alpha$ coverage down to $z = z_{\rm em}$ that remains in the observed wavelength range (one target has only observation of G130M, and another QSO is at $z = 0.5$; see Table 1). On the other hand, for the targets at $R > 351$ kpc, the wavelength coverage is not as complete over 1140–1800 Å (55% of the QSOs have only one COS grating—all but one have G130M, and four QSOs have $z_{\rm em} \gtrsim 0.48$). We note that the QSOs of 6/10 G130M observations have $z_{\rm em} < 0.17$, setting all the Ly$\alpha$ transitions within the COS G130M bandpass.

The overall line identification process is as follows. First, we mark all the ISM absorption features (i.e., any absorption that could arise from the MW or M31) and the velocity components (which is done as part of the overall alignment of the spectra; see Section 2.2). Local (approximate) continua are fitted near the absorption lines to estimate the equivalent widths ($W_\lambda$), and their ratios for ions with several transitions are checked to determine whether any are potentially contaminated. Next, we search for any absorption features at $z = z_{\rm em}$, again identifying any velocity component structures in the absorption. We then identify possible Ly$\alpha$ absorption and any other associated lines (other H I transitions and metal transitions) from the redshift of QSOs down to $z = 0$. In each case, if there are simultaneous detections of Ly$\alpha$, Ly$\beta$, and/or Ly$\gamma$ (and weaker transitions), we check that the equivalent width ratios are consistent. If there are any transitions left unidentified, we check whether it could be O VI $\lambda\lambda 1031$, 1037, as this doublet can sometimes be detected without any accompanying H I (Tripp et al. 2008). Finally, we check that the alignment in each absorber with multiple detected absorption lines is correct, or whether it needs some additional adjustment.

In the region $R \lesssim 1.1\ R_{\rm vir}$ and for 84% of the sample at any $R$, we believe that the line identifications are reliable and accurate at the 98% confidence level. In the appendices, we provide some additional information regarding the line identification, in particular for the troublesome cases. We also make available in a machine-readable format the full line identification for all the targets listed in Table 1 (see Appendix A).

### 2.4. Determination of the Properties of the Absorption at $-700 \leqslant {\rm v}_{LSR} \leqslant -150$ km s$^{-1}$

Our systematic search window for absorption that may be associated with the CGM of M31 is $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$ (LHW15). The $-700$ km s$^{-1}$ cutoff corresponds to about $-100$ km s$^{-1}$ less than the most negative velocities from the rotation curve of M31 ($\sim -600$ km s$^{-1}$; see Chemin et al. 2009). The $-150$ km s$^{-1}$ cutoff is set by the MW lines that dominate the absorption in the velocity range $-150 \lesssim v_{\rm LSR} \lesssim +50$ km s$^{-1}$. At $-100 \lesssim v_{\rm LSR} \lesssim -50$ km s$^{-1}$, the absorption is dominated by low- and intermediate-velocity clouds that are observed in and near the MW disk. Absorption from Galactic high-velocity clouds (HVCs) is seen to velocities $v_{\rm LSR} \sim -150$ km s$^{-1}$ toward distant Galactic halo stars in the

general direction of M31 (Lehner & Howk 2011; Lehner et al. 2012, 2015). Since the M31 disk rotation velocities extend to about $-150\,\mathrm{km\,s^{-1}}$ in the northern tip of M31, there is a small window that is inaccessible for studying the CGM of M31 (see also Lehner et al. 2015 and Section 3.2).[24]

To search for M31 CGM gas and determine its properties, we use the following atomic and ionic transitions: O I $\lambda1302$, C II $\lambda\lambda1036$, 1334, C IV $\lambda\lambda1548$, 1550, Si II $\lambda\lambda1190$, 1193, 1260, 1304, 1526, Si III $\lambda1206$, Si IV $\lambda\lambda1393$, 1402, O VI $\lambda1031$, Fe II $\lambda\lambda1144$, 1608, and Al II $\lambda1670$. We also report results (mostly upper limits on column densities) for N V $\lambda\lambda1238$, 1242, N I $\lambda1199$ (N I $\lambda\lambda1200$, 1201 being typically blended in the velocity range of interest, $-700 \leqslant v_{\mathrm{LSR}} \leqslant -150\,\mathrm{km\,s^{-1}}$), P II $\lambda1301$, S III $\lambda1190$, and S II $\lambda\lambda1250$, 1253, 1259.

To determine the column densities and velocities of the absorption, we use the apparent optical depth (AOD) method (see Section 2.4.2), but in Appendix D we confront the AOD results with column densities estimated from Voigt profile fitting (PF; see also Section 2.4.3). As much as possible at COS resolution, we derive the properties of the absorption in individual components. Especially toward M31, this is important since along the same line of sight in the velocity window $-700 \leqslant v_{\mathrm{LSR}} \leqslant -150\,\mathrm{km\,s^{-1}}$, there can be multiple origins of the gas (including the CGM of M31 or MS; see Figure 1 and LHW15) as we detail in Section 2.5. However, the first step to any analysis of the absorption imprinted on the QSO spectra is to model the QSO's continuum.

### 2.4.1. Continuum Placement

To fit the continuum near the ions of interest, we generally use the automated continuum fitting method developed for the COS CGM Compendium (CCC; Lehner et al. 2018). Figure 3 in Lehner et al. (2018) shows an example of an automatic continuum fit. In short, the continuum is fitted near the absorption features using Legendre polynomials. A velocity region of about $\pm1000$–$2000\,\mathrm{km\,s^{-1}}$ around the relevant absorption transition is initially considered for the continuum fit, but this could be changed depending on the complexity of the continuum placement in this region. In all cases the interval for continuum fitting is never larger than $\pm2000\,\mathrm{km\,s^{-1}}$ or smaller than $\pm250\,\mathrm{km\,s^{-1}}$. Within this predefined region, the spectrum is broken into smaller subsections and then rebinned. The continuum is fitted to all pixels that did not deviate by more than $2\sigma$ from the median flux, masking pixels from the fitting process that may be associated with small-scale absorption or emission lines. Legendre polynomials of orders

---

between 1 and 5 are fitted to the unmasked pixels, with the goodness of the fit determining the adopted polynomial order. Typically the adopted polynomials are of orders between 1 and 3 owing to the relative simplicity of the QSO continua when examined over velocity regions of 500–4000 $\mathrm{km\,s^{-1}}$. The only systematic exception is Si III, where the polynomial order is always between 2–3 and 5 owing to this line being in the wing of the extended local Ly$\alpha$ absorption profile.

This procedure is applied to our predefined set of transitions, with the continuum defined locally for each. Each continuum model is visually inspected for quality control. In a few cases, the automatic continuum fitting fails owing to a complex continuum (e.g., near the peak of an emission line or where many absorption lines were present within the predefined continuum window). In these cases, we first try to adjust the velocity interval of the spectrum to provide better-constrained fits; if that still fails, we manually select the continuum region to be fitted.

### 2.4.2. Velocity Components and AOD Analysis

The next step of the analysis is to determine the velocity components and integrate them to determine the average central velocities and column densities for each absorption feature. In Figure 2, we show an example of the normalized velocity profiles. In the online figure set, we provide a similar figure for each QSO in our sample. Although we systematically search for absorption in the full velocity range $-700 \leqslant v_{\mathrm{LSR}} \leqslant -150\,\mathrm{km\,s^{-1}}$, the most negative velocity of detected absorption in our sample is $v_{\mathrm{LSR}} = -508\,\mathrm{km\,s^{-1}}$; that is, we do not detect any M31 absorption in the range $-700 \lesssim v_{\mathrm{LSR}} \lesssim -510\,\mathrm{km\,s^{-1}}$. In Figure 2, MW absorption at $-100 \lesssim v_{\mathrm{LSR}} \lesssim 100\,\mathrm{km\,s^{-1}}$ is clearly seen in all species but N V. Absorption observed in the $-510 \leqslant v_{\mathrm{LSR}} \leqslant -150\,\mathrm{km\,s^{-1}}$ that is not color-coded is produced by higher-redshift absorbers or other MW lines.

To estimate the column density in each observed component, we use the AOD method (Savage & Sembach 1991). In this method, the absorption profiles are converted into apparent optical depth per unit velocity, $\tau_a(v) = \ln[F_c(v)/F_{\mathrm{obs}}(v)]$, where $F_c(v)$ and $F_{\mathrm{obs}}(v)$ are the modeled continuum and observed fluxes as a function of velocity. The AOD, $\tau_a(v)$, is related to the apparent column density per unit velocity, $N_a(v)$, through the relation $N_a(v) = 3.768 \times 10^{14}\tau_a(v)/(f\lambda\,(\text{\AA}))\,\mathrm{cm^{-2}}\,(\mathrm{km\,s^{-1}})^{-1}$, where $f$ is the oscillator strength of the transition and $\lambda$ is the wavelength in Å. The total column density is obtained by integrating the profile over the predefined velocity interval, $N = \int_{v_1}^{v_2} N_a(v)dv$, where $[v_1, v_2]$ are the boundaries of the absorption. We estimate the line centroids with the first moment of the AOD $v_a = \int v\tau_a(v)dv \big/ \int \tau_a(v)dv\,\mathrm{km\,s^{-1}}$. As part of this process, we also estimate the equivalent widths, which we use mainly to determine whether the absorption is detected at the $\geqslant 2\sigma$ level. In cases where the line is not detected at $\geqslant 2\sigma$ significance, we quote a $2\sigma$ upper limit on the column density, which is defined as twice the $1\sigma$ error derived for the column density assuming that the absorption line lies on the linear part of the curve of growth.

For features that are detected above the $2\sigma$ level, the estimated column densities are stored for further analysis. Since we have undertaken a full identification of the absorption features in each spectrum (see Section 2.3, Appendix A), we can reliably assess whether a given transition is contaminated using in particular the conflict plots described in the appendix (see Appendix B). If there is evidence of some line
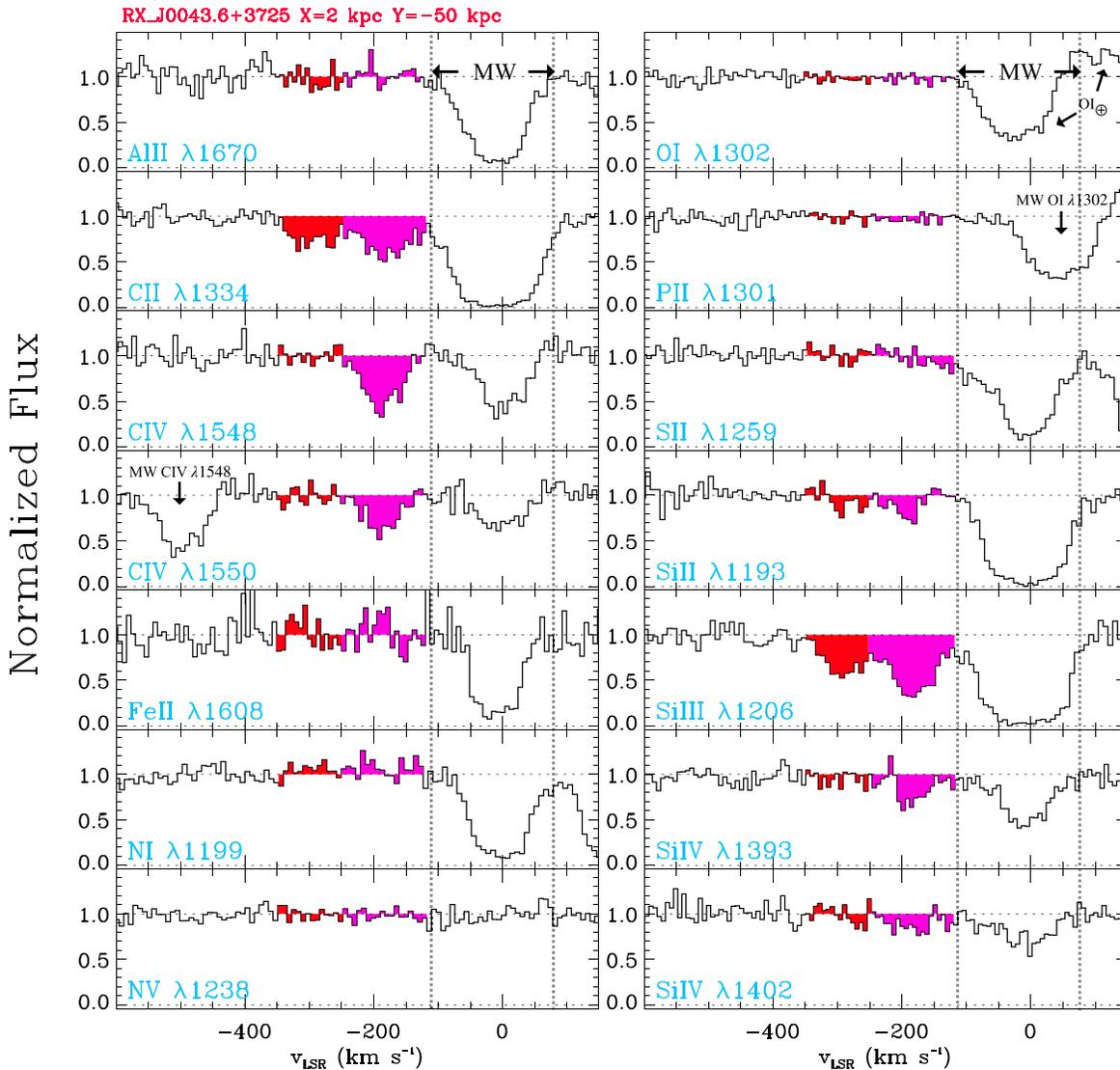
---

**Figure 2.** Example of normalized absorption lines as a function of the LSR velocity toward RX J0043.6+3725 showing the typical atoms and ions probed in our survey. High negative velocity components likely associated with M31 are shown in colors, and each color represents a different component identified at the COS G130M-G160M resolution. In this case, significant absorption is observed in the two identified components in C II, Si II, and Si III. Higher ions (Si IV, C IV) are observed in only one of the components, showing a change in the ionization properties with velocity. Some species are not detected, but their limits can still be useful in assessing the physical properties of the gas. The MW absorption is indicated between the two vertical dotted lines and is observed in all the species but N V. At $v_{LSR} \gtrsim -100 \text{ km s}^{-1}$, airglow emission lines can contaminate O I, and hence the MW absorption is contaminated, but typically that is not an issue for the surveyed velocity range $-700 \leqslant v_{LSR} \leqslant -150 \text{ km s}^{-1}$. The complete figure set (43 images) is available in the online journal.

(The complete figure set (43 images) is available.)

contamination and several transitions are available for this ion (e.g., Si II, Si IV, C IV), we exclude it from our list.

We find that contamination affects the Si III and C II in the velocity range $-700 \leqslant v_{LSR} \leqslant -150 \text{ km s}^{-1}$ in a few rare cases (six components of Si III and three components of C II $\lambda1334$).[25] For all but one of these contaminated Si III components, we can correct the contamination because the interfering line is a Lyman series line from a higher redshift and the other H I transitions constrain the equivalent width of the contamination. The one case we cannot correct this way is the $-340 \text{ km s}^{-1}$ component toward PHL 1226 (see also

Appendix A), which is associated with the MS. In the footnote of Table 2, we list the ions that are found to be contaminated at some level. For any column density that is corrected for contamination, the typical correction error is about 0.05–0.10 dex depending on the level of contamination, as well as the S/Ns of the spectrum in that region.

The last step is to check for any unresolved saturation. When the absorption is clearly saturated (i.e., the flux level reaches zero flux in the core of the absorption), the line is automatically marked as saturated and a lower limit is assigned to the column density. In Section 2.5, we will show how we separate the MS from the M31 CGM absorption, but we note that only the Si III components associated with the MS and the MW have their absorption reaching zero-flux level, not the components associated with the CGM of M31.

---

[25] Toward RX J0048.3+3941, C II $\lambda1334$ is contaminated in the third component, but C II $\lambda1036$ is available to correct for it in this case.

**Table 2**
Summary of the Results

| Target | Ion | $v_1$ (km s$^{-1}$) | $v_2$ (km s$^{-1}$) | $v$ (km s$^{-1}$) | $\sigma_v$ (km s$^{-1}$) | $\log N$ [cm$^{-2}$] | $\sigma^1_{\log N}$ | $\sigma^2_{\log N}$ | $f_N$ | $f_{MS}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| RX J0048.3+3941 | Al II | −480.0 | −320.0 | −397.2 | 19.6 | 11.92 | 0.15 | 0.23 | 0 | −1 |
| RX J0048.3+3941 | C II | −480.0 | −320.0 | −381.4 | 1.2 | 14.15 | 0.01 | 0.01 | −2 | −1 |
| RX J0048.3+3941 | C IV | −480.0 | −320.0 | −390.8 | 5.3 | 13.25 | 0.05 | 0.05 | 0 | −1 |
| RX J0048.3+3941 | Fe II | −480.0 | −320.0 | −432.9 | 25.2 | 13.40 | 0.15 | 0.23 | 0 | −1 |
| RX J0048.3+3941 | N I | −480.0 | −320.0 | ⋯ | ⋯ | 12.88 | 0.18 | 0.30 | −1 | −1 |
| RX J0048.3+3941 | N V | −480.0 | −320.0 | ⋯ | ⋯ | 12.77 | 0.18 | 0.30 | −1 | −1 |
| RX J0048.3+3941 | O I | −480.0 | −320.0 | ⋯ | ⋯ | 13.47 | 0.18 | 0.30 | −1 | −1 |
| RX J0048.3+3941 | O VI | −480.0 | −320.0 | −375.5 | 6.2 | 13.85 | 0.05 | 0.06 | 0 | −1 |
| RX J0048.3+3941 | S II | −480.0 | −320.0 | ⋯ | ⋯ | 13.74 | 0.18 | 0.30 | −1 | −1 |
| RX J0048.3+3941 | Si II | −480.0 | −320.0 | −374.0 | 2.4 | 12.93 | 0.02 | 0.02 | 0 | −1 |
| RX J0048.3+3941 | Si III | −480.0 | −320.0 | −376.0 | 1.6 | 13.06 | 0.02 | 0.02 | −2 | −1 |
| RX J0048.3+3941 | Si IV | −480.0 | −320.0 | ⋯ | ⋯ | 12.39 | 0.18 | 0.30 | −1 | −1 |

**Note.** The velocities $v_1$ and $v_2$ correspond to the integration of the absorption component. The $1\sigma$ errors $\sigma^1_{\log N}$ and $\sigma^2_{\log N}$ are the positive and negative errors, respectively, on the logarithm of the column density. The flag $f_N$ has the following definition: $0$ = detection (not saturated or contaminated); $-1$ = upper limit; $-2$ = lower limit (due to saturation of the line). The flag $f_{MS}$ has the following definition: $0$ = not contaminated by the MS; $-1$ = contaminated by the MS (see Section 2.5).

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

When the flux does not reach a zero-flux level, the procedure for checking saturation depends on the number of transitions for a given ion or atom. We first consider ions with several transitions (Si II, C IV, Si IV, sometimes C II) since they can provide information about the level of saturation for a given peak optical depth. For ions with several transitions, we compare the column densities with different $f\lambda$-values to determine whether there is a systematic decrease in the column density as $f\lambda$ increases. If there is not, we estimate the average column density using all the available measurements and propagate the errors using a weighted mean. For the Si II transitions, Si II $\lambda1526$ shows no evidence for saturation when detected based on the comparison with stronger transitions, while Si II $\lambda1260$ or $\lambda1193$ can be saturated if the peak optical $\tau_a \gtrsim 0.9$. For doublets (e.g., C IV, Si IV), we systematically check whether the column densities of each transition agree within $1\sigma$ error; if they do not and the weak transition gives a higher value (and there is no contamination in the weaker transition), we correct for saturation following the procedure discussed in Lehner et al. (2018) (and see also Savage & Sembach 1991). For C IV and Si IV, there is rarely any evidence for saturation (we only correct once for saturation of C IV in the third component observed in the Mrk 352 spectrum; in that component the peak optical $\tau_a \sim 0.9$). For single strong transitions (in particular Si III and often C II), if the peak optical depth is $\tau_a \geqslant 0.9$, we conservatively flag the component as saturated and adopt a lower limit for that component. We adopt $\tau_a \geqslant 0.9$ as the threshold for saturation based on other ions with multiple transitions (in particular Si II) where the absorption starts to show some saturation at this peak optical depth.

To estimate how the column density of silicon varies with $R$ (which has a direct consequence for the CGM mass estimates derived from silicon in Sections 4.5 and 4.8), it is useful to assess the level of saturation of Si III, which is the only silicon ion that cannot be directly corrected for saturation.[26] The lower limits of the Si III components associated with the CGM of

M31 are mostly observed at $R \lesssim 140$ kpc (only two are observed at $R > 140$ kpc), but they do not reach zero-flux level; these components are conservatively marked as saturated because their peak apparent optical depth is $\tau_a > 0.9$ (not because $\tau_a \gg 2$) and because the comparison between the different Si II transitions shows in some cases evidence for saturation (see above). Hence, the true values of the column densities of these saturated components are most likely higher than the adopted lower-limit values but are very unlikely to be overestimated by a factor $\gg$3–4. We can estimate how large the saturation correction for Si III might be using the strong Si II lines (e.g., Si II $\lambda1193$ or Si II $\lambda1260$) compared to the weaker ones (e.g., Si II $\lambda1526$). Going through the eight sightlines showing some saturation in the components of Si III associated with the CGM of M31 (see Table 2), for all the targets beyond 50 kpc, the saturation correction is likely to be small, $<$0.10–0.15 dex, based on the fact that many show no evidence of saturation in Si II $\lambda1260$ (when there is no contamination for this transition) or Si II $\lambda1193$. On the other hand, for the two innermost targets, the saturation correction is at least 0.3 dex and possibly as large as 0.6 dex based on the column density comparison between saturated Si II and weaker, unsaturated transitions. The latter would put $N_{Si} \simeq 14.5$ close to the maximum values derived with photoionization modeling in the COS-Halos sample (see Section 5.2). Therefore, for the components associated with the CGM of M31 at $R > 50$ kpc when we estimate the functional form of $N_{Si}$ with $R$, we adopt an increase of 0.1 dex of the lower limits. For the two inner targets at $R < 50$ kpc, we explore how an increase of 0.3 and 0.6 dex affects the estimation of $N_{Si}(R)$.

### 2.4.3. High-resolution Spectra and Profile Fitting Analysis

In Appendices C and D we explore the robustness of the AOD results by comparing high- and low-resolution spectra and by comparing to a Voigt PF analysis. There is good overall agreement in the column densities derived from the STIS and COS data, and our conservative choice of $\tau_a \sim 0.9$ as the threshold for saturation in the COS data is adequate (see Appendix C). For the PF analysis, we consider the most

---

[26] Some of the Si II transitions (especially Si II $\lambda\lambda1193, 1260$) have evidence for saturation, but weaker transitions are always available (e.g., Si II $\lambda1526$), and therefore we can determine a robust value of the column density of Si II.
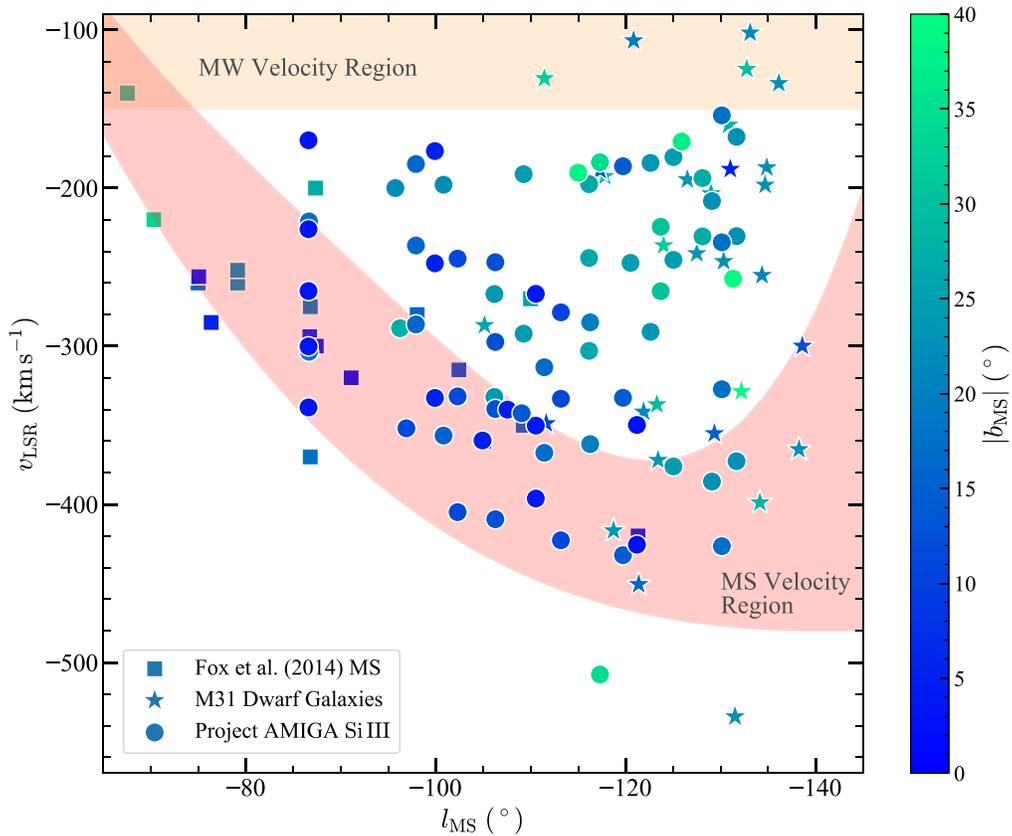
**Figure 3.** LSR velocity of the Si III components (circles) observed in our sample as a function of the MS longitude $l_{MS}$, color-coded according to the absolute MS latitude. Shaded regions show the velocities that can be contaminated by the MS and MW (by definition of our search velocity window, any absorption at $v_{LSR} > -150$ km s$^{-1}$ was excluded from our sample). We also show the data (squares) from the MS survey from Fox et al. (2014) and the radial velocities of the M31 dwarf galaxies (stars).

complicated blending of components in our sample and demonstrate that there are some small systematic differences between the AOD and PF-derived column densities (see Appendix D). However, these differences are small, and a majority of our sample is not affected by heavy blending. Hence, the AOD results are robust and are adopted for the remaining of the paper.

### 2.5. Correcting for Magellanic Stream Contamination

Prior to determining the properties of the gas associated with the CGM of M31, we need to identify that gas and distinguish it from the MW and the MS. We have already removed from our analysis any contamination from higher-redshift intervening absorbers and from the MW (defined as $-150 \lesssim v_{LSR} \lesssim 100$ km s$^{-1}$). However, as shown in Figure 1 and discussed in LHW15, the MS is another potentially large source of contamination: in the direction of M31, the velocities of the MS can overlap with those expected from the CGM of M31. The targets in our sample have MS longitudes and latitudes in the range $-132° \leqslant l_{MS} \leqslant -86°$ and $-14° \leqslant b_{MS} \leqslant +41°$. The H I 21 cm emission GBT survey by Nidever et al. (2010) finds that the MS extends to about $l_{MS} \simeq -140°$. Based on this and previous H I emission surveys, Nidever et al. (2008, 2010) found a relation between the observed LSR velocities of the MS and $l_{MS}$ that can be used to assess contamination in our targeted sightlines based on their MS coordinates. Using Figure 7 of Nidever et al. (2010), we estimate the upper and lower boundaries of the H I velocity range as a function of $l_{MS}$,

which we show in Figure 3 by the curved colored area. The MS velocity decreases with decreasing $l_{MS}$ up to $l_{MS} \simeq -120°$, where there is an inflection point where the MS LSR velocity increases. We note that the region beyond $l_{MS} \lesssim -135°$ is uncertain but cannot be larger than shown in Figure 3 (see also Nidever et al. 2010); however, this does not affect our survey since all our data are at $l_{MS} \gtrsim -132°$.

We take a systematic approach to removing the MS contamination that does not reject entire sightlines based on their MS coordinates since not all velocity components may be contaminated even on sightlines close to the MS. In Figure 3, we show LSR velocity of the Si III components as a function of the MS longitude. We choose Si III, as this ion is the most sensitive to detect both weak and strong absorption and is readily observed in the physical conditions of the MS and M31 CGM (Fox et al. 2014; Lehner et al. 2015). We consider the individual components, as for a given sightline, several components can be observed falling in or outside the boundary region associated with the MS as illustrated in Figure 3. We find that $28/74 \simeq 38\%$ of the detected Si III components are within the MS boundary region shown in Figure 3. We note that changing the upper boundary by $\pm 5$ km s$^{-1}$ would change this number by about $\pm 3\%$.

To our own sample, we also add data from two different surveys: the HST/COS MS survey by Fox et al. (2014) and the M31 dwarfs (McConnachie 2012; see also Section 3). For the MS survey, we restrict the sample to $-150° \leqslant l_{MS} \leqslant -20°$, i.e., overlapping with our sample but also including higher $l_{MS}$ values while still avoiding the Magellanic Clouds region,

where conditions may be different. The origin of the sample for the M31 dwarf galaxies is fully discussed in Section 3. The larger galaxy M33 is excluded here from that sample, as its large mass is not characteristic. The LSR velocities of the M31 dwarfs as a function of $l_{MS}$ are plotted with a star symbol in Figure 3. For the MS survey, we select the LSR velocities of Si III for the MS survey (note that these are average velocities that can include multiple components), which are shown with squares in Figure 3. Most (~90%) of the squares fall between the two curves in Figure 3, confirming the likelihood that these sightlines probe the MS (although we emphasize that this test was not initially used by Fox et al. 2014 to determine the association with the MS).

The M31 dwarf galaxies are of course not contaminated by the MS as the gas may be. However, if we assume that the dwarfs' companions have a similar kinematic distribution, the frequency with which the dwarfs fall within the velocity range where MS contamination is likely gives us guidance as to how frequently CGM absorption might be flagged as contaminated. For $l_{MS} \gtrsim -132°$ (the range probed by the background QSOs), only 9% (2/22) of the dwarfs are within the velocity region where MS contamination occurs. If the velocity distributions of the M31 dwarfs and M31 CGM gas are similar, this suggests that gas components flagged as MS material are highly likely to be MS gas. We note, however, that two additional dwarfs just miss being included in the MS velocity region; a small increase to our MS region would change the frequency of the dwarfs consistent with MS velocities to 18%.

Observations of H I 21 cm emission toward the QSOs observed with COS in the MS survey (Fox et al. 2014) and Project AMIGA (Howk et al. 2017) show only H I detections within $|b_{MS}| \lesssim 11°$. In the region defined by $-150° \leqslant l_{MS} \leqslant -20°$, the bulk of the H I 21 cm emission is observed within $|b_{MS}| \lesssim 5°$ (Nidever et al. 2010). We therefore expect the metal ionic column densities to have a strong absorption when $|b_{MS}| \lesssim 10°$ and a weaker absorption as $|b_{MS}|$ increases. In Figure 4, we show the total column densities of Si III for the velocity components from the Project AMIGA sample found within the MS boundary region shown in Figure 3, i.e., we added the column densities of the components that are likely associated with the MS. We also show in the same figure the results from the Fox et al. (2014) survey. Both data sets show the same behavior of the total Si III column densities with $|b_{MS}|$, an overall decrease in $N_{Si III}$ as $|b_{MS}|$ increases. Treating the limits as values, combining the two samples, and using the Spearman rank order, the test confirms the visual impression that there is a strong monotonic anticorrelation between $N_{Si III}$ and $|b_{MS}|$ with a correlation coefficient $r_S = -0.72$ and a p-value $\ll 0.1\%$.[27] There is a large scatter (about $\pm 0.4$ dex around the dotted line) at any $b_{MS}$, making it difficult to determine whether any data points may not be associated with the MS (e.g., the three very low $N_{Si III}$ at $12° < |b_{MS}| < 18°$ from our sample or the very high value at $|b_{MS}| \sim 27°$ from the Fox et al. 2014 sample).

In Figure 5, we show the individual column densities of Si III as a function of the impact parameter from M31 for the Project AMIGA sightlines, where we separate components associated with the MS from those that are not. Looking at Figures 1 and 4, we expect the strongest column densities
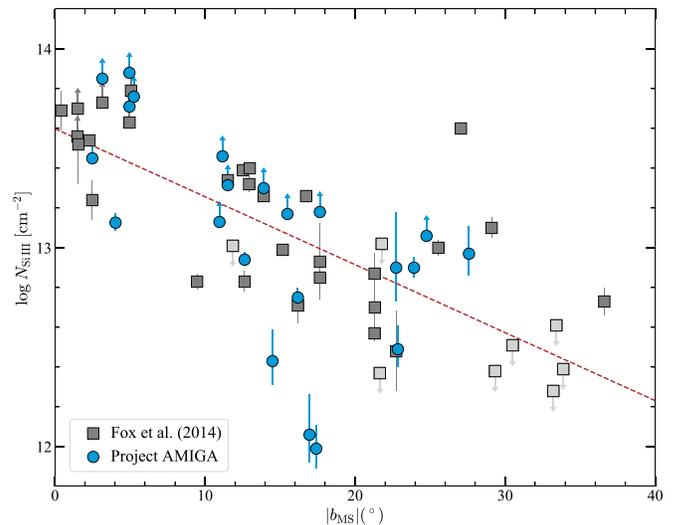
---

[27] We note that if we increase the lower limits by 0.15 dex or more and similarly decrease the upper limits, the significance of the anticorrelation would be similar.



**Figure 4.** Total column density of Si III associated with the MS as a function of the absolute MS latitude. We also show the MS survey by Fox et al. (2014) restricted to data with $-150° \leqslant l_{MS} \leqslant -20°$. The light-gray squares with downward-pointing arrows are nondetections in the Fox et al. sample. The dashed line is a linear fit to the data treating the limits as values. A Spearman ranking correlation test implies a strong anticorrelation with a correlation coefficient $r_S = -0.72$ and $p \ll 0.1\%$.
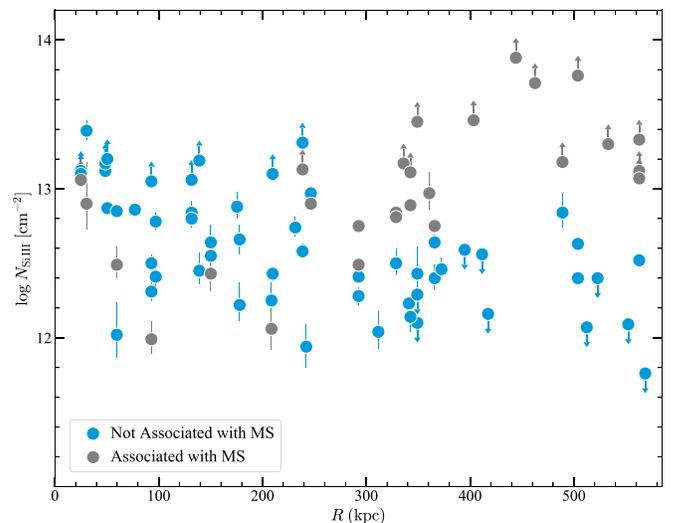


**Figure 5.** Logarithm of the column densities of the individual components for Si III as a function of the projected distances from M31 of the background QSOs, where the separation is made for the components associated or not with the MS.

associated with the MS to be at $|b_{MS}| \lesssim 10°$ and $R \gtrsim 300$ kpc, which is where they are located on Figure 5. We also expect a positive correlation between $N_{Si III}$ and $R$ for the MS contaminated components, while for uncontaminated components we expect the opposite (see LHW15). Treating again limits as values, the Spearman rank order test demonstrates a strong monotonic correlation between $N_{Si III}$ and $R$ ($r_S = 0.68$ with $p \ll 0.1\%$), while for uncontaminated components there is a strong monotonic anticorrelation ($r_S = -0.57$ with $p \ll 0.1\%$), in agreement with the expectations. Based on these results, it is therefore reasonable to consider any absorption components observed in the COS spectra within the MS boundary region defined in Figure 3 as most likely

**Table 3**
Summary of the M31 Dwarf Galaxies

| Name | Type | $l_{MS}$ (deg) | $b_{MS}$ (deg) | $v_{LSR}$ (km s$^{-1}$) | $D_{dwarf}$ (kpc) | $R$ (kpc) | $X$ (kpc) | $Y$ (kpc) | $M_*$ ($10^5 M_\odot$) | $M_{200}$ ($10^8 M_\odot$) | $R_{200}$ (kpc) | $v_{200}$ (km s$^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M32 | cE | −126.5 | 23.8 | −194.6 | 805.0 | 5.3 | −0.1 | −5.3 | 3200.0 | 852.2 | 92.8 | 88.9 |
| NGC 205 | dE/dSph | −127.5 | 23.4 | −241.3 | 824.0 | 8.0 | −5.8 | 5.5 | 3300.0 | 864.8 | 93.2 | 89.3 |
| And IX | dSph | −129.0 | 25.7 | −203.5 | 766.0 | 35.3 | 24.3 | 25.6 | 1.5 | 21.5 | 27.2 | 26.1 |
| And XVII | dSph | −130.3 | 22.9 | −246.3 | 794.0 | 42.3 | −13.2 | 40.2 | 2.6 | 28.0 | 29.7 | 28.5 |
| And I | dSph | −123.4 | 24.2 | −372.0 | 745.0 | 43.0 | 7.6 | −42.3 | 39.0 | 102.7 | 45.8 | 43.9 |
| And XXVII | dSph | −131.5 | 23.1 | −534.2 | 828.0 | 55.5 | −12.2 | 54.1 | 1.2 | 19.3 | 26.3 | 25.2 |
| And III | dSph | −121.9 | 22.0 | −341.5 | 748.0 | 65.2 | −18.9 | −62.4 | 8.3 | 48.9 | 35.8 | 34.3 |
| And X | dSph | −130.8 | 28.2 | −160.0 | 701.0 | 73.5 | 55.4 | 48.3 | 1.0 | 17.4 | 25.3 | 24.3 |
| And XXV | dSph | −133.1 | 21.9 | −101.8 | 813.0 | 79.0 | −28.3 | 73.8 | 6.8 | 44.4 | 34.7 | 33.2 |
| And XV | dSph | −123.3 | 29.8 | −336.9 | 631.0 | 89.7 | 81.4 | −37.7 | 4.9 | 38.0 | 32.9 | 31.5 |
| NGC 185 | dE/dSph | −134.7 | 23.4 | −198.0 | 617.0 | 93.1 | −8.2 | 92.8 | 680.0 | 405.1 | 72.4 | 69.4 |
| NGC 147 | dE/dSph | −134.8 | 22.4 | −187.0 | 676.0 | 97.5 | −20.8 | 95.3 | 620.0 | 387.6 | 71.4 | 68.4 |
| And XXVI | dSph | −134.3 | 20.8 | −255.1 | 762.0 | 97.8 | −41.8 | 88.4 | 0.6 | 13.9 | 23.5 | 22.5 |
| And XI | dSph | −118.7 | 23.9 | −416.6 | 759.0 | 98.4 | 9.8 | −97.9 | 0.5 | 12.6 | 22.8 | 21.8 |
| And XIX | dSph | −120.8 | 18.6 | −106.7 | 933.0 | 101.2 | −62.4 | −79.6 | 4.3 | 35.7 | 32.2 | 30.9 |
| And V | dSph | −134.1 | 28.6 | −398.9 | 773.0 | 105.3 | 60.9 | 85.9 | 3.9 | 34.0 | 31.7 | 30.4 |
| And XXIV | dSph | −132.7 | 30.2 | −124.7 | 600.0 | 107.7 | 80.8 | 71.2 | 0.9 | 17.1 | 25.2 | 24.2 |
| Cas II | dSph | −136.1 | 23.0 | −133.7 | 681.0 | 110.8 | −13.1 | 110.0 | 1.4 | 20.8 | 26.9 | 25.8 |
| And XIII | dSph | −117.7 | 24.9 | −192.5 | 912.0 | 110.9 | 25.1 | −108.1 | 0.4 | 11.5 | 22.1 | 21.2 |
| And XXI | dSph | −129.3 | 15.1 | −355.1 | 859.0 | 118.0 | −115.5 | 23.8 | 7.6 | 46.9 | 35.3 | 33.8 |
| And XX | dSph | −121.4 | 16.2 | −450.6 | 802.0 | 121.1 | −94.5 | −75.9 | 0.3 | 9.8 | 20.9 | 20.0 |
| And XXIII | dSph | −124.0 | 32.7 | −236.3 | 769.0 | 121.6 | 118.9 | −25.6 | 11.0 | 56.0 | 37.4 | 35.9 |
| And II | dSph | −117.8 | 30.1 | −192.5 | 652.0 | 135.2 | 92.5 | −98.6 | 76.0 | 141.5 | 51.0 | 48.9 |
| Cas III | dSph | −138.2 | 22.9 | −365.2 | 772.0 | 135.7 | −13.8 | 135.0 | 71.1 | 137.1 | 50.5 | 48.3 |
| And XXIX | dSph | −117.3 | 13.5 | −188.8 | 731.0 | 179.7 | −123.7 | −130.4 | 1.8 | 23.5 | 28.0 | 26.8 |
| And XXII | dSph | −111.4 | 32.2 | −130.6 | 794.0 | 210.1 | 130.5 | −164.7 | 0.3 | 10.5 | 21.5 | 20.6 |
| And VII | dSph | −138.6 | 12.1 | −299.8 | 762.0 | 211.5 | −157.1 | 141.5 | 95.0 | 157.5 | 52.9 | 50.6 |
| IC 10 | dIrr | −146.6 | 20.7 | −340.0 | 794.0 | 240.1 | −38.0 | 237.1 | 860.0 | 453.5 | 75.2 | 72.0 |
| Lac I | dSph | −131.0 | 4.6 | −188.0 | 756.0 | 255.4 | −252.4 | 38.9 | 40.9 | 105.1 | 46.2 | 44.3 |
| And VI | dSph | −111.6 | 10.5 | −348.8 | 783.0 | 258.2 | −153.1 | −207.9 | 28.0 | 87.6 | 43.5 | 41.6 |
| LGS 3 | dIrr/dSph | −105.1 | 26.2 | −286.8 | 769.0 | 259.8 | 65.4 | −251.4 | 9.6 | 52.4 | 36.6 | 35.1 |
| Per I | dSph | −132.2 | 49.4 | −328.6 | 785.0 | 337.5 | 331.3 | 64.4 | 11.3 | 56.7 | 37.6 | 36.0 |

**Note.** Galaxy parameters are from McConnachie (2012), Martin et al. (2014, 2016), and references therein.

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

associated with the MS. We therefore flag any of these components (28 out of 74 components for Si III) as contaminated by the MS, and they are not included further in our sample.

Finally, we noted above that only a small fraction of the dwarfs are found in the MS contaminated region. While that fraction is small (9%), this could still suggest that in the MS contaminated region some of the absorption could be a blend between both MS and M31 CGM components. However, considering the uncontaminated velocities along sightlines in (29 components) and outside (17 components) the contaminated regions, with a $p$-value of 0.74 the Kolmogorov–Smirnov (K-S) comparison of the two samples cannot reject the null hypothesis that the distributions are the same. This strongly suggests that the correction from the MS contamination does not bias much the velocity distribution associated with the CGM of M31 (assuming that there is no strong change of the velocity with the azimuth $\Phi$; as we explore this further in Sections 3.2 and 4.10, there is, however, no strong evidence of a velocity dependence with $\Phi$).

### 3. M31 Dwarf Galaxy Satellites

While Project AMIGA is dedicated to understanding the CGM of M31, our survey also provides a unique probe of the dwarf galaxies found in the halo of M31. In particular, we have the opportunity to assess whether the CGM of dwarf satellites plays an important role in the CGM of the host galaxy, as studied by cosmological and idealized simulations (e.g., Anglés-Alcázar et al. 2017; Bustard et al. 2018; Hafen et al. 2019, 2020). When considering the dwarf galaxies in our analysis, we have two main goals: (1) to determine whether the velocity distributions of the dwarfs and the absorbers are similar, and (2) to assess whether some of the absorption observed toward the QSOs could be associated directly with the dwarfs, either as gas that is gravitationally bound or as gas that has been recently stripped.

The sample for the M31 dwarf galaxies is mostly drawn from the McConnachie (2012) study of Local Group dwarfs, in which the properties of 29 M31 dwarf satellites were summarized. Four additional dwarfs (Cas II, Cas III, Lac I, Per I) are added from recent discoveries (Collins et al. 2013; Martin et al. 2014, 2016, 2017). M33 is excluded from that sample, as its large mass is not characteristic of satellites.[28] Table 3 summarizes our adopted sample of M31 dwarf galaxies

---

[28] In Appendix E, we further discuss and present some evidence that the CGM of M33 is unlikely to contribute much to the observed absorption in our sample.

(sorted by increasing projected distance from M31), listing some of their key properties. As listed in this table, most of the M31 satellite galaxies are dwarf spheroidal (dSph) galaxies, which have been shown to have been stripped of most of their gas, most likely via ram pressure stripping (Grebel et al. 2003), a caveat that we keep in mind as we associate these galaxies with absorbers.

### 3.1. Velocity Transformation

So far we have used LSR velocity to characterize MW and MS contamination of gas in the M31 halo. However, as we now consider relative motions over 30° on the sky, we cannot simply subtract M31's systemic radius velocity to place these relative motions in the correct reference frame. Over such large sky areas, tangential motion must be accounted for because the "systemic" projected radial velocity of the M31 system changes with sightline. To eliminate the effects of "perspective motion," we follow Gilbert et al. (2018) (and see also Veljanoski et al. 2014) by first transforming the heliocentric velocity ($v_\odot$) into the Galactocentric frame, $v_{\rm Gal}$, which removes any effects the solar motion could have on the kinematic analysis. We converted our measured radial velocities from the heliocentric to the Galactocentric frame using the relation from Courteau & van den Bergh (1999) with updated solar motions from McMillan (2011) and Schönrich et al. (2010):

$$v_{\rm Gal} = v_\odot + 251.24 \sin(l)\cos(b)$$
$$+ 11.1 \cos(l)\cos(b) + 7.25 \sin(b), \qquad (1)$$

where $(l, b)$ are the Galactic longitude and latitude of the object. To remove the bulk motion of M31 along the sightline to each object, we use the heliocentric systemic radial velocity for M31 of $-301\,{\rm km\,s^{-1}}$ (van der Marel & Guhathakurta 2008; Chemin et al. 2009), which is $v_{\rm M31,r} = -109\,{\rm km\,s^{-1}}$ in the Galactocentric velocity frame. The systemic transverse velocity of M31 is $v_{\rm M31,t} = -17\,{\rm km\,s^{-1}}$ in the direction on the sky given by the position angle $\theta_t = 287°$ (van der Marel et al. 2012). The removal of M31's motion from the sightline velocities resulting in peculiar line-of-sight velocities for each absorber or dwarf, $v_{\rm M31}$, is then given by (van der Marel & Guhathakurta 2008)

$$v_{\rm M31} = v_{\rm Gal} - v_{\rm M31,r} \cos(\rho)$$
$$+ v_{\rm M31,t} \sin(\rho)\cos(\phi - \theta_t), \qquad (2)$$

where $\rho$ is the angular separation between the center of M31 and the QSO or dwarf position and $\phi$ is the position angle of the QSO or dwarf with respect to M31's center. We note that the transverse term in Equation (2) is more uncertain (van der Marel & Guhathakurta 2008; Veljanoski et al. 2014), but its effect is also much smaller, and indeed including it or not would not quantitatively change the results; we opted to include that term in the velocity transformation. We apply these transformations to change the LSR velocities to heliocentric velocities to Galactocentric velocities to peculiar velocities for each component observed in absorption toward the QSOs and for each dwarf. With this transformation, an absorber or dwarf with no peculiar velocity relative to M31's bulk motion has $v_{\rm M31} = 0\,{\rm km\,s^{-1}}$, regardless of its position on the sky (Gilbert et al. 2018).
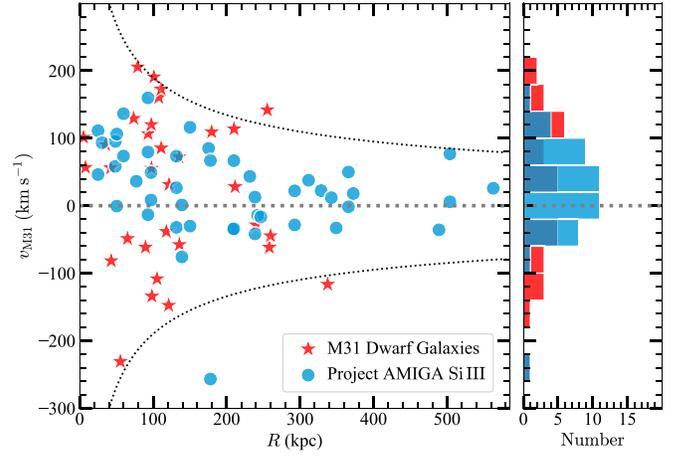


**Figure 6.** Left: M31 peculiar velocity (as defined by Equation (2)) against the projected distances for the observed absorption components associated with M31 (using Si III) and M31 dwarf galaxies. The dotted curves show the escape velocity divided by $\sqrt{3}$ to account for the unknown tangential motions of the absorbers and galaxies. Right: M31 velocity distributions with the same color-coding definition.

### 3.2. Velocity Distribution

In Figure 6, we compare the M31 peculiar velocities of the absorbers using Si III and dwarfs against the projected distance (see Section 3.1). In Figure 6, we also show the expected escape velocity, $v_{\rm esc}$, as a function of $R$ for a $1.3 \times 10^{12}\,M_\odot$ point mass. We conservatively divide $v_{\rm esc}$ by $\sqrt{3}$ in that figure to account for remaining unconstrained projection effects. Nearly all the CGM gas traced by Si III within $R_{\rm vir}$ is found at velocities consistent with being gravitationally bound, and this is true even at larger $R$ for most of the absorbers. This finding also holds for most of the dwarf galaxies, and, as demonstrated by McConnachie (2012), it holds when the galaxies' 3D distances are used (i.e., using the actual distance of the dwarf galaxies, instead of the projected distances used in this work). Therefore, both the CGM gas and galaxies probed in our sample at both small and large $R$ are consistent with being gravitationally bound to M31.

Figure 6 also informs us that the dwarf satellite and CGM gas velocities overlap to a high degree but do not follow identical distributions. The mean and standard deviation of the M31 velocities for the dwarfs are $+27.1 \pm 109.5$ and $+25.3 \pm 67.3\,{\rm km\,s^{-1}}$ for the CGM (Si III) gas. There is therefore a slight asymmetry favoring more positive peculiar motions. A simple two-sided K-S test of the two samples rejects the null hypothesis that the distributions are the same at 95% level confidence ($p = 0.04$). And indeed, while the two distributions overlap and the means are similar, the velocity dispersion of the dwarfs is larger than that of the QSO absorbers.[29] For the QSO absorbers, all the components but one have their M31 velocities in the interval $-80 \leqslant v_{\rm M31} \leqslant +160\,{\rm km\,s^{-1}}$, but 9/32 (28%) of the dwarfs are outside that range. Four of the dwarfs are in the range $+160 < v_{\rm M31} \leqslant +210\,{\rm km\,s^{-1}}$, a velocity interval that cannot be probed in absorption owing to foreground MW contamination. The other five dwarfs have $v_{\rm M31} < -80\,{\rm km\,s^{-1}}$, while only one out

---

[29] Considering the dwarf velocities outside the MS region, then we have $\langle v_{\rm M31} \rangle = 65.1 \pm 102.3\,{\rm km\,s^{-1}}$. In that case, the two-sided K-S test of the two samples rejects the null hypothesis that the distributions are the same at 99.0% level confidence.

**Table 4**
QSO Absorbers within $R_{200}$ of M31 Dwarf Galaxies

| Dwarf | QSO | $v_{\rm LSR,Si III}$ (km s$^{-1}$) | $\Delta_{\rm sep}$ | $v_{\rm esc}$ (km s$^{-1}$) | $\delta v$ (km s$^{-1}$) |
|-------|-----|------|------|------|------|
| M32 | RX J0048.3+3941 | −550.4 | 0.23 | 187.3 | 55.1 |
| M32 | RX J0048.3+3941 | −485.4 | 0.23 | 187.3 | 9.9 |
| M32 | HS 0033+4300 | −513.2 | 0.38 | 144.3 | 7.9 |
| M32 | HS 0058+4213 | −535.6 | 0.55 | 120.1 | 42.9 |
| M32 | HS 0058+4213 | −498.7 | 0.55 | 120.1 | 6.0 |
| M32 | RX J0043.6+3725 | −595.9 | 0.49 | 127.3 | 101.9 |
| M32 | RX J0043.6+3725 | −489.1 | 0.49 | 127.3 | 4.9 |
| M32 | Zw 535.012 | −535.4 | 0.70 | 106.4 | 27.8 |
| M32 | Zw 535.012 | −472.6 | 0.70 | 106.4 | 35.0 |
| M32 | RX J0050.8+3536 | −552.3 | 0.78 | 100.9 | 65.1 |

**Note.** The match between dwarfs and QSO absorbers was made so that the projected separation between the dwarf and QSO is within $R_{200}^{\rm dwarf}$ of the dwarf, i.e., $\Delta_{\rm sep} = r/R_{200}^{\rm dwarf} \leqslant 1$ (where $r$ is the projected distance between the QSO and dwarf). Here $v_{\rm esc}$ is the escape velocity of the dwarf at the projected distance of the QSO assuming a point-like mass with halo mass listed in Table 3. The velocity separation between the QSO absorber and the dwarf is $\delta v = |v_{\rm M31,Si III} - v_{\rm M31,dwarf}|$.

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

of 46 Si III components (2%) have $v_{\rm M31} < -80\,{\rm km\,s^{-1}}$. Both the small fraction of dwarfs at $v_{\rm M31} > +160\,{\rm km\,s^{-1}}$ and $v_{\rm M31} < -80\,{\rm km\,s^{-1}}$ and the even smaller fraction of absorbers at $v_{\rm M31} < -80\,{\rm km\,s^{-1}}$ suggest that there is no important population of absorbers at the inaccessible velocities $v_{\rm M31} > +160\,{\rm km\,s^{-1}}$ (see also Section 2.5).

### 3.3. The Association of Absorbers with Dwarf Satellites

Using the information from Table 3, we cross-match the sample of dwarf galaxies and QSOs to determine the QSO sightlines that are passing within a dwarf's $R_{200}$ radius. There are 11 QSOs (with 58 Si III components) within $R_{200}$ of 16 dwarfs. In Table 4, we summarize the results of this cross-match. Figure 7, we show the map of the QSOs and dwarf locations in our survey, where the M31 velocities of the Si III components and dwarfs are color-coded on the same scale and the circles around each dwarf represent their $R_{200}^{\rm dwarf}$ $R_{200}$ radius.

Table 4 and Figure 7 show that several absorbers can be found within $R_{200}$ of several dwarfs when Si III is used as the gas tracer. For example, the two components observed in Si III toward Zw 535.012 are found within $R_{200}$ of six dwarf galaxies. In Table 4, we also list the escape velocity ($v_{\rm esc}$) at the observed projected distance of the QSO relative to the dwarf, as well as the velocity separation between the QSO absorber and the dwarf ($\delta v \equiv |v_{\rm M31,Si III} - v_{\rm M31,dwarf}|$). So far we have not considered the velocity separation $\delta v$ between the dwarf and the absorber, but it is likely that if $\delta v \gg v_{\rm esc}$ then the observed gas traced by the absorber is unlikely to be bound to the dwarf galaxy even if $\Delta_{\rm sep} = R/R_{200} < 1$.

If we set $\delta v < v_{\rm esc}/\sqrt{3}$, then the sample of components would be reduced to 31 instead of 58. The sample is reduced still further down to 12 if the two most massive dwarfs (M32 and NGC 205) are removed from the sample, and down to 6 if the most massive dwarfs with $M_{\rm h} > 3.9 \times 10^{10}\,M_{\odot}$ are removed from the sample. Applying a cross-match where $\delta v < v_{\rm esc}$ and $\Delta_{\rm sep} < 1$ can reduce the degeneracy between different galaxies, especially if one excludes the four most massive galaxies. For example, RXS J0118.8+3836 is

located at $0.40R_{200}$ and $0.72R_{200}$ from Andromeda XV and Andromeda XXIII, but only in the latter case is $\delta v \ll v_{\rm esc}$ (and in the former case $\delta v > v_{\rm esc}$), making the two components observed toward RXS J0118.8+3836 more likely associated with Andromeda XXIII.

Several sightlines therefore pass within $\Delta_{\rm sep} < 1$ of a dwarf galaxy and show a velocity absorption within the escape velocity. This gas could be gravitationally bound to the dwarf. However, there are also five absorbers where $\delta v < v_{\rm esc}/\sqrt{3}$, but the QSO is at $1 < \Delta_{\rm sep} \leqslant 2$ from the dwarf, i.e., the velocity separation is small, but the spatial projected separation makes it unlikely to be bound to the dwarf. Here the velocity match may be a coincidence or a result of the relative proximity of the dwarfs and QSOs in the CGM of M31 assuming that the gas and dwarfs both follow the same global velocity motion of the M31 CGM. As illustrated in Figure 7, there are, however, some dwarfs with $\Delta_{\rm sep} \lesssim 2$ with a radial velocity very different from that observed in absorption toward the QSO or vice versa, implying that not all the dwarfs and gas velocities are tightly connected.

In summary, it is plausible that absorbers with $\Delta_{\rm sep} < 1$ and $\delta v \ll v_{\rm esc}$ trace gas associated with the CGM of a dwarf, but we cannot confirm unambiguously this association. We inspected a number of gas properties (e.g., column densities, ionization levels, kinematics) but did not find any that can differentiate clearly a dwarf CGM origin from an M31 CGM origin. Nothing in the properties of the components found within $\Delta_{\rm sep} < 1$ of a dwarf and having $\delta v \ll v_{\rm esc}$ makes them outliers. This is certainly not surprising since any association assumes that the dwarf galaxies have a rich gas CGM. Yet all the satellites listed in the cross-matched Table 4 are dSph galaxies, which are known to be neutral gas poor (Grebel et al. 2003). The dSph galaxies are also likely ionized gas deficient since the favored mechanism to strip their gas is ram pressure, a stripping mechanism efficient on both the neutral and ionized gas (Grebel et al. 2003; Mayer et al. 2006). Therefore, these galaxies are unlikely to have gas-rich CGM, and based on our observations, we do not find any persuasive evidence that gas associated with M31 satellites causes the absorption we see in the M31 CGM.

## 4. Properties of the M31 CGM

We now focus on determining the properties of the CGM of M31 using only the velocity components that are not contaminated by the MS (see 2.5). We use the following atoms and ions to characterize the M31 CGM: O I, Si II, Si III, Si IV, C II, C IV, O VI, and Fe II. O I and Fe II are not commonly detected, but even so they are useful in assessing the ionization and depletion levels of the CGM gas. Note that we use the terminology "low ions" for singly ionized species, "intermediate ions" for Si III and Si IV, and "high ions" for C IV and O VI. Also note that we adopt here the solar relative abundances from Asplund et al. (2009).

### 4.1. Metallicity of the CGM

Radio observations have not detected any H I 21 cm emission toward any of the QSO targets in Project AMIGA down to a $5\sigma$ level of $\log N_{\rm H I} \gtrsim 17.6$ (Paper I); many sightlines could therefore have $\log N_{\rm H I} \ll 17.6$. As a consequence of this, we cannot directly estimate the metallicities of the CGM in our sample. However, we have some weak detections of O I in four
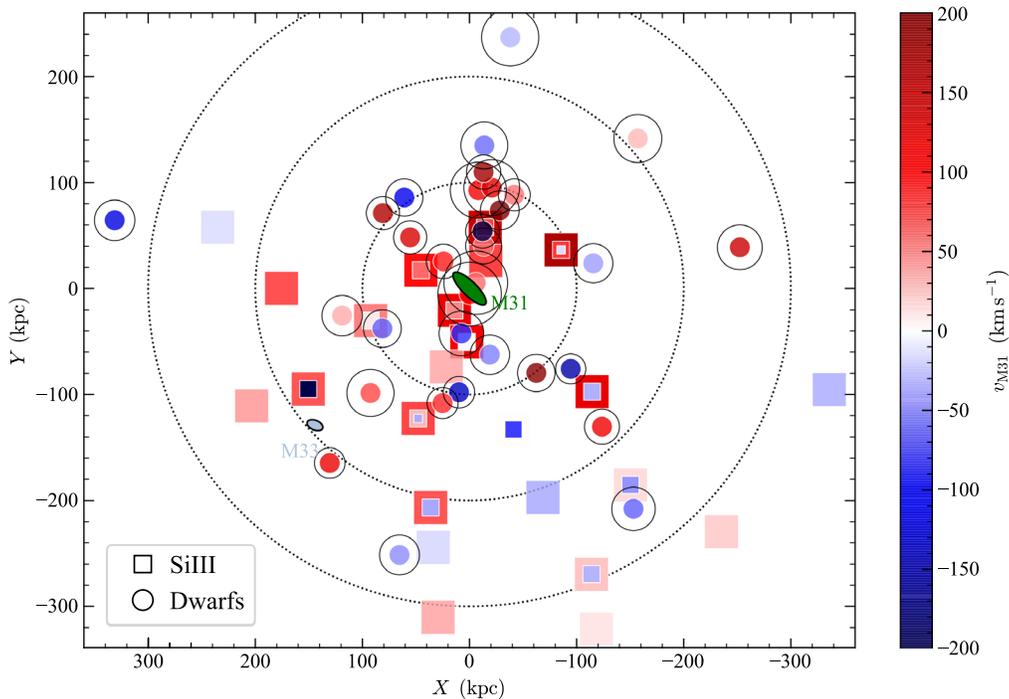
**Figure 7.** Locations of the QSOs (squares) and dwarfs (circles) relative to M31 (see Figure 1). The data are color-coded according to the relative velocities of the detected Si III (multiple colors in a symbol indicate multiple detected components) or the dwarfs. The black circles centered on the dwarfs indicate their individual $R_{200}$. The components associated with the MS have been removed.

components at better than the $3\sigma$ level. Since O I and H I have nearly identical ionization potentials and are strongly coupled through charge exchange reactions (Field 1971), O I is an excellent proxy for H I, requiring no or very small ionization correction as long as the photoionization spectrum is not too hard (e.g., Lehner et al. 2003). Therefore, O I can be compared to the limit of H I to put a lower limit on the metallicity. The O I logarithmic column densities are in the range of $13.3-13.7$ dex (see Table 2), with a mean of 13.5 dex. This implies [O I/ H I] $= \log(N_{\rm O\,I}/N_{\rm H\,I}) - \log({\rm O/H})_\odot \gtrsim -0.7$ or a metallicity $Z \gtrsim 0.2 Z_\odot$. This lower limit, however, assumes that there is no beam dilution effect, i.e., we assume that the limit on the H I column density in the 2 kpc beam (at the distance of M31) would be the same as in a pencil beam observed in absorption. Any beam dilution would increase the limit on H I, and therefore the metallicity limit could be less stringent. We therefore caution the reader not to take this limit as a hard lower limit.

### 4.2. Relative Abundances

While the metallicity remains quite uncertain, from the relative abundances of detected ions we can assess the level of ionization, dust depletion, and nucleosynthetic history. For assessing depletions, we can compare refractory elements like Fe to less refractory elements like Si (e.g., Savage & Sembach 1996; Jenkins 2009). Fe II and Si II have similar ionization energies (8–16 eV), and their observed ratio should be minimally affected by differential ionization. Hence, the ratio [Fe II/Si II] $= \log(N_{\rm Fe\,II}/N_{\rm Si\,II}) - \log({\rm Fe/Si})_\odot$ traces dust depletion levels. Unfortunately (but perhaps not surprisingly), Fe II is only detected in the sightline closest to M31, and in that sightline we derive [Fe II/Si II] $= -0.13 \pm 0.16$. In 10 other sightlines, we place upper limits on that ratio where the two smallest upper limits imply [Fe II/Si II] $\lesssim 0$, while all the

others are above 0 dex. While the information is minimal, this still demonstrates that there is no evidence for significant dust depletion in the CGM of M31. As depletions get stronger in denser gas, it is perhaps not surprising that we find little evidence for it in a sample where the sightlines all have $\log N_{\rm H\,I} \lesssim 17.6$ and low ions are not commonly detected. While we assume that dust would be the major factor to deplete Fe relative to Si, the lack of evidence for depletion of Fe also points to a negligible nucleosynthesis effect on that ratio that would produce a nonsolar $\alpha$-particle (e.g., Si) enhancement relative to Fe (e.g., Welty et al. 1997).

Using ratios of elements with different nucleosynthetic origins, we can assess the chemical enrichment history of the M31 halo gas by measuring departures from a solar relative abundance ratio in elements of different nucleosynthetic origin. For instance, the [C/$\alpha$] ratio should be sensitive to nucleosynthesis effects since there is a time lag between the production of $\alpha$-elements and carbon (see, e.g., Cescutti et al. 2009; Mattsson 2010). This analysis would be complicated by large depletions, but as we have shown above, the Fe/Si ratios show little if any evidence of large depletions. As Fe is typically the most depleted element in these conditions (Savage & Sembach 1996; Welty et al. 1999; Jenkins 2009), we can reliably assume that [C/$\alpha$] does not suffer large depletions and can therefore be used as a nucleosynthetic indicator. We use C II/Si II as a proxy for [C/$\alpha$]. However, we must also consider ionization effects since differential ionization can affect the C II/Si II ratio (C II has a higher ionization energy range [12–25 eV] than Si II [8–16 eV]). To assess this, we use the nine absorbers with detections of both Si II and C II to estimate [C II/Si II] $= \log(N_{\rm C\,II}/N_{\rm Si\,II}) - \log({\rm C/Si})_\odot$. Since this subsample includes both detections and lower limits owing to saturation of C II, we use a survival analysis where the four censored lower limits are included (Feigelson & Nelson 1985;

Isobe et al. 1986). We find that the mean [C II/Si II] = 0.07 ± 0.09 (where the error is the error on the mean from the Kaplan–Meier estimator) and the $1\sigma$ dispersion is 0.19 dex. This ratio is consistent with a solar value, i.e., nonsolar nucleosynthesis effects are negligible. If nondetections of Si II are included, the mean rises to [C II/Si II] = 0.52 ± 0.11, strongly indicating that ionization affects this ratio owing to photons ionizing Si II into Si III.

Therefore, based on the relative abundances of Fe and C to Si, there is no evidence for strong dust depletion or nonsolar nucleosynthesis effects in the CGM of M31. We emphasize that this does not mean that there is no dust in the CGM of M31, and indeed several studies have shown that the CGM of galaxies can have a substantial mass of dust (e.g., Ménard et al. 2010; Peek et al. 2015). However, its effect on elemental abundances must be smaller than in the dense regions of galaxies. The lack of nucleosynthesis effects on the abundance of Fe or C relative to Si strongly suggests that the overall metallicity of the gas is not extremely low, as enhancements of $\alpha$-elements are seen in low-metallicity MW halo stars and in low-metallicity gas in CGM absorbers over a range of redshift. For a sample of H I-selected absorbers with $15 \lesssim \log N_{\mathrm{H\,I}} \lesssim 18$ at $0.2 \lesssim z \lesssim 1$, Lehner et al. (2019) found little correlation between $[\mathrm{C}/\alpha]$ and the metallicity. However, in stars and H II regions in the local universe, there is evidence of a trend between $[\mathrm{C}/\alpha]$ and the metallicity where $[\mathrm{C}/\alpha] \simeq -0.6$ at $-2 \lesssim \log Z/Z_\odot \lesssim -0.5$ and $[\mathrm{C}/\alpha] \simeq 0$ near solar metallicities (e.g., Akerman et al. 2004; Fabbian et al. 2010). Therefore, the metallicity of the M31 CGM could still be subsolar, but it is unlikely to be much below $1/3\,Z_\odot$. This is consistent with the rough metallicity estimate set in Section 4.1.

### 4.3. Ionization Fractions

The ionization fraction of the CGM gas can be estimated directly by comparing the column densities of O I to those of Si II, Si III, and Si IV (e.g., Lehner et al. 2001; Zech et al. 2008). O I is an excellent proxy for neutral gas (see Section 4.1). Si II is found in both neutral and ionized gas, and Si III and Si IV arise only in ionized gas. O and Si are both $\alpha$-elements with similar nucleosynthetic origins and have similar levels of dust depletion in the diffuse gas (Savage & Sembach 1996; Jenkins 2009). Therefore, if the ratio $[\mathrm{O\,I/Si}] = \log(N_{\mathrm{O\,I}}/N_{\mathrm{Si}}) - \log(\mathrm{O/Si})_\odot$ is subsolar, ionization is important in the M31 CGM.

To obtain the total Si column density, we use the individual ion columns listed in Table 2. In the case of nondetections of the Si ions or O I, we conservatively add the upper limits to the column densities. When there are lower limits present, we add the column densities using the lower limit values. When both detections and nondetections are present, we consider the two extreme possibilities, where we either set the column density of the nondetection to the upper limit value (i.e., the absorption is nearly detected—case 1) or neglect the upper limit (i.e., it is a true nondetection—case 2). For 28 targets, we can estimate the [O I/Si] ratio. Considering case 1, we find that the mean and dispersion are [O I/Si] < −0.95 ± 0.38 and the full range is [<−1.78, <−0.34], i.e., on average the gas is ionized at the >89% level. In case 2, the mean and dispersion are [O I/Si] < −0.74 ± 0.51, so that the ionization fraction is still >81% on average. These are upper limits because typically O I is not detected. However, even in the five cases where O I is detected, four of five are upper limits too because Si III is saturated and hence only a lower limit on the column density of

Si can be derived. In that case, [O I/Si] ranges from <−1.78 to −0.43 (or to <−0.85 if we remove the absorber where the O I absorption is just detected at the $2\sigma$ level), i.e., even when O I is detected to more than $3\sigma$, the gas is still ionized at levels >86%–98%.

The combination of Si II, Si III, and Si IV allows us to probe gas within the ionization energies 8–45 eV, i.e., the bulk of the photoionized CGM of M31. The high ions, C IV and O VI, have ionization energies 48–85 eV and 114–138 eV, respectively, and are not included in the above calculation. The column density of H can be directly estimated in the ionization energy 8–45 eV range from the observations via $\log N_{\mathrm{H}} = \log N_{\mathrm{Si}} - \log Z/Z_\odot$. As we show below, Si varies strongly with $R$ with values $\log N_{\mathrm{Si}} \gtrsim 13.7$ at $R \lesssim 100$ kpc and $\log N_{\mathrm{Si}} \lesssim 13.3$ at $R \gtrsim 100$ kpc, which implies $N_{\mathrm{H}} \gtrsim 1.5 \times 10^{18}\ (Z/Z_\odot)^{-1}\ \mathrm{cm}^{-2}$ and $\lesssim 0.6 \times 10^{18}\ (Z/Z_\odot)^{-1}\ \mathrm{cm}^{-2}$, respectively. For the high ions, a ionization correction needs to be added, and, e.g., for O VI, $\log N_{\mathrm{H}} = \log N_{\mathrm{O\,VI}} - \log Z/Z_\odot - \log f^i_{\mathrm{O\,VI}}$, where $f^i_{\mathrm{O\,VI}} \lesssim 0.2$ is the ionization fraction of O VI that peaks around 20% for any ionizing models (e.g., Gnat & Sternberg 2007; Oppenheimer & Schaye 2013a; Lehner et al. 2014). As discussed below, there is little variation of $N_{\mathrm{O\,VI}}$ with $R$, and it is always such that $\log N_{\mathrm{O\,VI}} \gtrsim 14.4$–14.9 within 300 kpc from M31, which implies $N_{\mathrm{H}} \gtrsim (2.5$–$8.1) \times 10^{18}(Z/Z_\odot)^{-1}\ \mathrm{cm}^{-2}$. Therefore, the CGM of M31 not only is mostly ionized (often at levels close to 100%) but also contains a substantial fraction of highly ionized gas, with even higher column densities than the weakly photoionized gas at $R \gtrsim 100$ kpc.

### 4.4. Ion Column Densities versus R

In Figure 8, we show the logarithmic (left) and linear (right) values of the total column densities of the components associated with M31 for C II, Si II, Si III, Si IV, C IV, and O VI as a function of the projected distances from M31. Gray data points are upper limits, while blue data with upward-pointing arrows are lower limits owing to saturated absorption. Overall, the column densities decrease at higher impact parameter. As the ionization potentials of the ions increase, the decrease in the column densities becomes shallower; O VI is almost flat. These conclusions were already noted in LHW15, but now that the region from 50 to 350 kpc is filled with data, these trends are even more striking. However, our new sample shows also an additional feature: there is a remarkable change in the quantity and quality of absorption around $R_{200} \simeq 230$ kpc. This change is especially visible for the low and intermediate ions, whereby high column densities of C II, Si II, Si III, and Si IV are observed solely at $R \lesssim R_{200}$. Low column densities C II, Si II, Si III, and Si IV are observed at all $R$, but strong absorption is observed only at $R \lesssim R_{200}$. The frequency of strong absorption is also larger at $R \lesssim 0.6 R_{200}$ than at larger $R$ for all ions. While the higher ionization gas traced by C IV and O VI is also weaker beyond $R_{200}$, the changes are less extreme. For C II, Si II, Si III, and Si IV, the difference between low and high column densities is a factor $\gtrsim 5$–10, while it drops to a factor of 2–4 for C IV, and possibly even less for O VI.

In Figure 9, we show the logarithmic values of the column densities derived from the individual components for C II, Si II, Si III, Si IV, C IV, and O VI as a function of the projected distances from M31. Similar trends are observed in Figure 8, but Figure 9 additionally shows that (1) more complex velocity structures (i.e., multiple velocity components) are predominantly
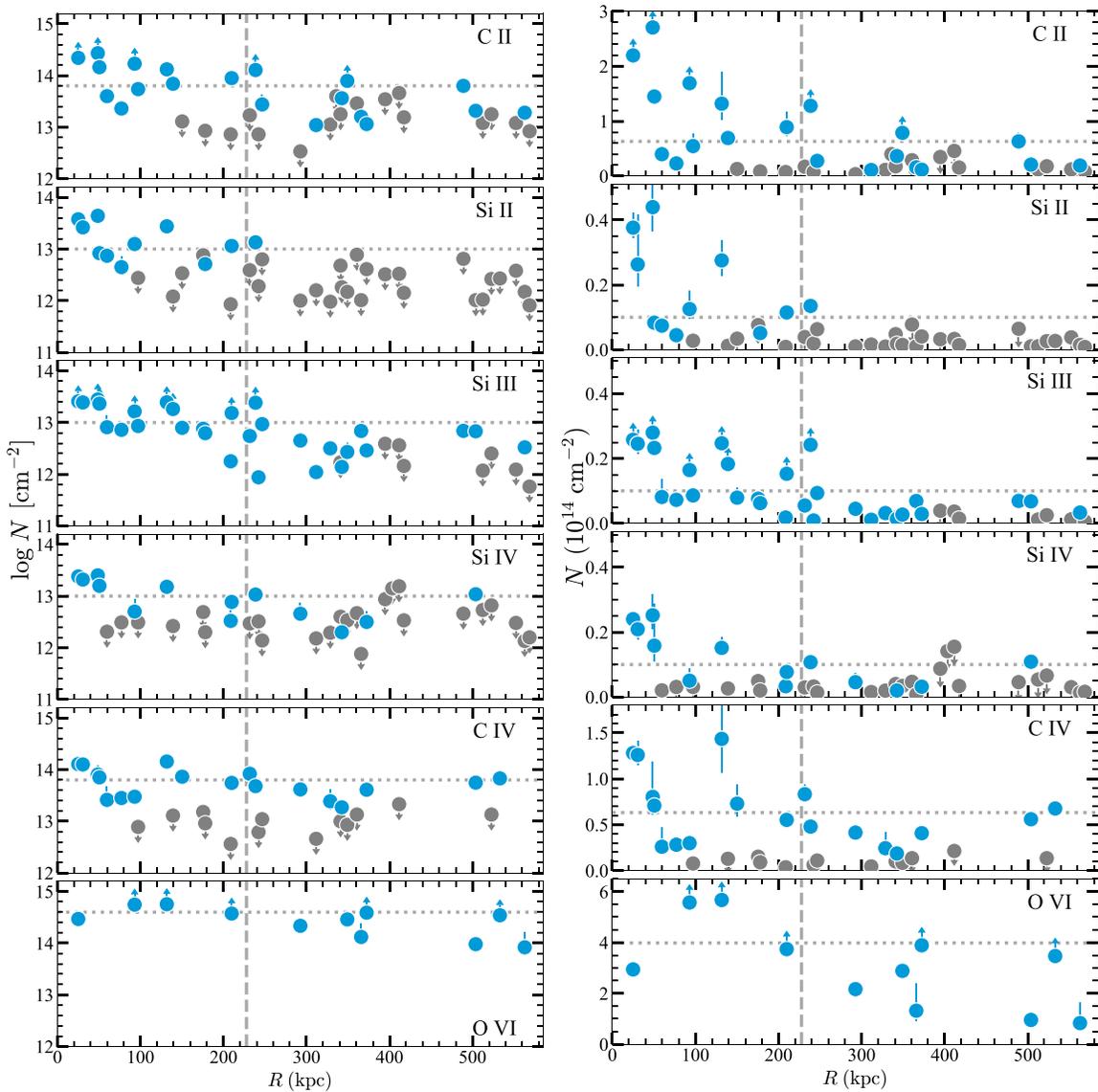
**Figure 8.** Total column densities of the ions as a function of $R$ with overall ionization potential increasing from top to bottom panels. The column densities are shown as logarithmic values (with the same relative vertical scale of about 3 dex in each panel) on the left and in linear units on the right. Blue circles are detections, while gray circles with downward-pointing arrows are nondetections. A blue circle with an upward-pointing arrow denotes that the absorption is saturated, resulting in a lower limit. The components associated with the MS have been removed. The dashed vertical line marks $R_{200}$. The dotted horizontal line in each panel corresponds to the threshold column density ($N_{th}$) for each ion that separates strong absorbers (mostly observed at $R \lesssim R_{200}$) from weak ones (observed at any $R$; Section 4.6 for more detail). Note how Si III and O VI are detected at high frequency well beyond $R_{200}$ and even $R_{vir}$.

observed at $R \lesssim R_{200}$ and (2) factor $\gtrsim$2–10 changes in the column densities are observed across multiple velocity components along a given sightline.

### 4.5. Silicon Column Densities versus R

With Si II, Si III, and Si IV, we can estimate the total column density of Si within the ionization energy range 8–45 eV without any ionization modeling. Gas in this range should constitute the bulk of the cool photoionized CGM of M31 (see Section 4.3). In Figure 10, we show the total column density of Si (estimated following Section 4.3) against the projected distance $R$ from M31. The vertical ticked bars in Figure 10 indicate data with some upper limits, and the length of the vertical bar represents the range of $N_{Si}$ values allowed between cases 1 and 2 (see Section 4.3).

Figure 10 reinforces the conclusions observed from the individual low ions in Figures 8 and 9. Overall there is a decrease of the column density of Si at larger $R$. This decrease has a much stronger gradient in the inner region of the M31 CGM between $R \lesssim 25$ kpc and about $R \sim 100$–150 kpc than at $R \gtrsim 150$ kpc. $N_{Si}$ changes by a factor $>$5–10 between about 25 and 150 kpc, while it changes by a factor $\lesssim$2 between 150 and 300 kpc. The scatter in $N_{Si}$ is also larger in the inner regions of the CGM than beyond $\gtrsim$120–150 kpc.

To model this overall trend (which is also useful to determine the baryon and metal content of the CGM; see Section 4.8), we consider three models, a hyperbolic (H) model, a single power-law (SPL) model, and a Gaussian process (GP) model. We refer the reader to Appendix F, where we fully explain the modeling process and how lower and upper limits are accounted for in the modeling. Figure 10
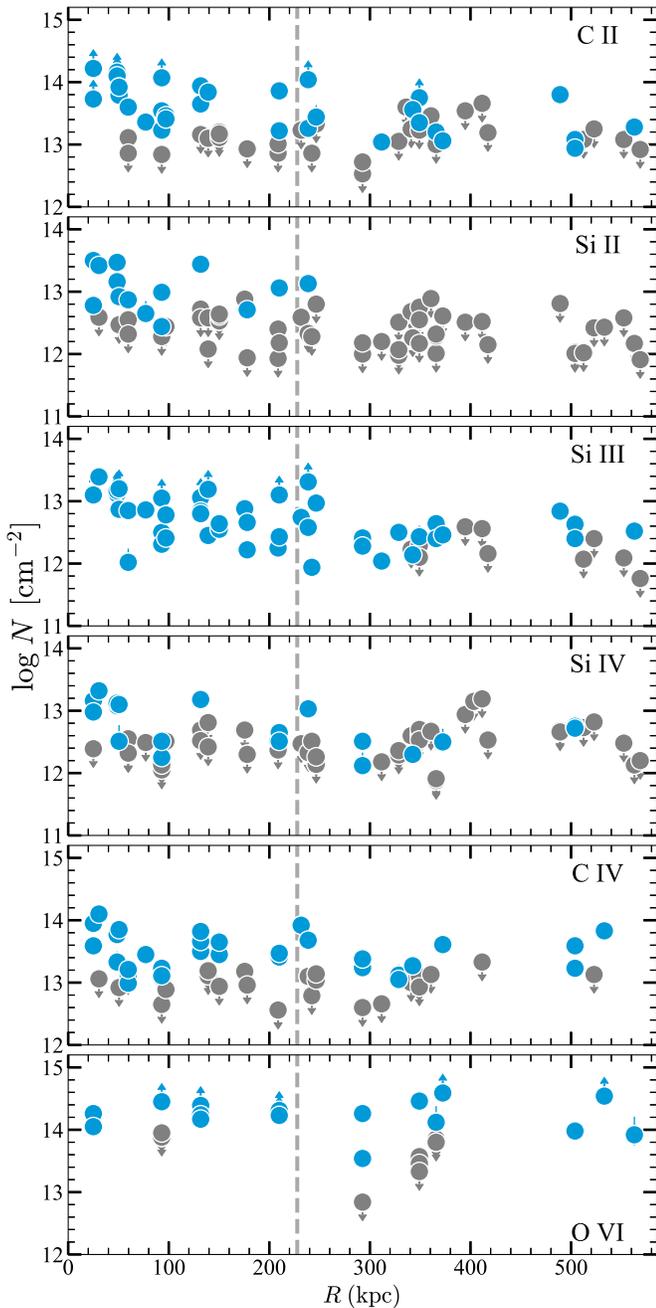
17

**Figure 9.** Logarithm of the column densities for the individual components of various ions (low to high ions from top to bottom) as a function of the projected distances from M31 of the background QSOs. Blue circles are detections, while gray circles with downward-pointing arrows are nondetections. A blue circle with an upward-pointing arrow denotes that the absorption is saturated, resulting in a lower limit. The components associated with the MS have been removed. The dashed vertical lines show the $R_{200}$ location. The same relative vertical scale of about 3 dex is used in each panel for comparison between the different ions.

shows that these three models greatly overlap. The nonparametric GP model overlaps more with the SPL model than with the H model in the range 250 kpc $\lesssim R \lesssim$ 400 kpc and at $R < 90$ kpc (especially for the high H model; see Figure 10). While there are some differences between these models (and we will explore in Section 4.8 how these affect the mass estimates of the CGM), they all further confirm the strong evolution of the column density of Si with $R$ between $\lesssim 25$ and
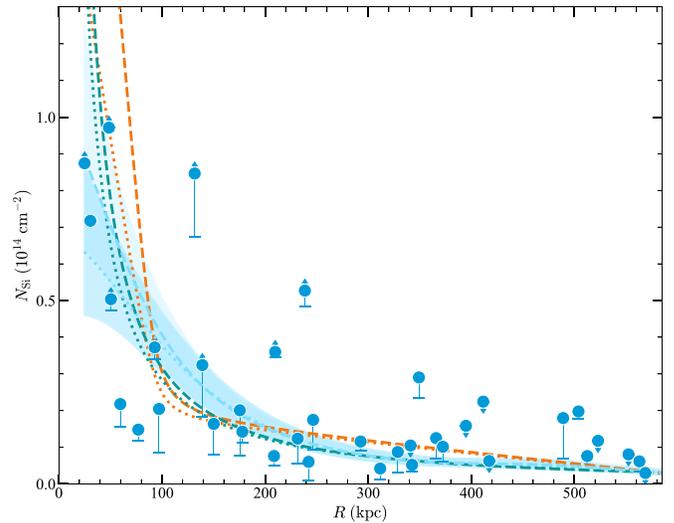


**Figure 10.** Total column densities of Si (i.e., $N_{\rm Si} = N_{\rm Si\,II} + N_{\rm Si\,III} + N_{\rm Si\,IV}$) as a function of the projected distances from M31 of the background QSOs. The vertical ticked bars show the range of values allowed if the upper limit of a given Si ion is negligible or not. The lower limits have upward-pointing arrows, and the upper limits are flagged using downward-pointing arrows. The orange, green, and blue curves are the H-, SPL-, and GP-derived models to the data, respectively (see text for details regarding how censoring is treated in each model). The dotted and dashed curves correspond to models where the lower limits at $R < 50$ kpc are increased by 0.3 or 0.6 dex. The blue areas correspond to the dispersion derived from the GP models (see Appendix F for more details).

90–150 kpc and a much shallower evolution with $R$ beyond 200 kpc. In Section 4.8, we use these models to constrain the metal and baryon masses of the cool CGM gas probed by Si II, Si III, and Si IV.

### 4.6. Covering Factors

As noted in Section 4.4, the diagnostic ions behave differently with $R$ in a way that reflects the underlying physical conditions. For example, Si II has a high detection rate within $R < 100$ kpc, a sharp drop beyond $R > 100$ kpc, and a total absence at $R \gtrsim 240$ kpc (see Figures 8 and 9). On the other hand, Si III and O VI are detected at all $R$ traced by our survey. The column densities of Si III, however, fall significantly with $R$, while O VI remains relatively flat. In this section, we further quantify the detection rates, or the covering factors, for each ion.

To calculate the covering factors of the low and high ions, we follow the methodology described in Paper I for H I by assuming a binomial distribution. We assess the likelihood function for values of the covering factor given the number of detections against the total sample, i.e., the number of targets within a given impact parameter range (see Cameron 2011). As demonstrated by Cameron (2011), the normalized likelihood function for calculating the Bayesian confidence intervals on a binomial distribution with a noninformative (uniform) prior follows a $\beta$-distribution.

In Figure 11, we show the *cumulative* covering factors ($f_c$) for the various ions, where each point represents the covering factor for all impact parameters less than the given value of $R$. The vertical bars are 68% confidence intervals. As discussed in Sections 4.4 and 4.5, the highest column densities are found exclusively at $R \leqslant R_{200}$ (and for some ions at $R \lesssim$ 100–150 kpc) for all ions except O VI. For the covering factors, we therefore consider (1) the entire sample (most of the upper
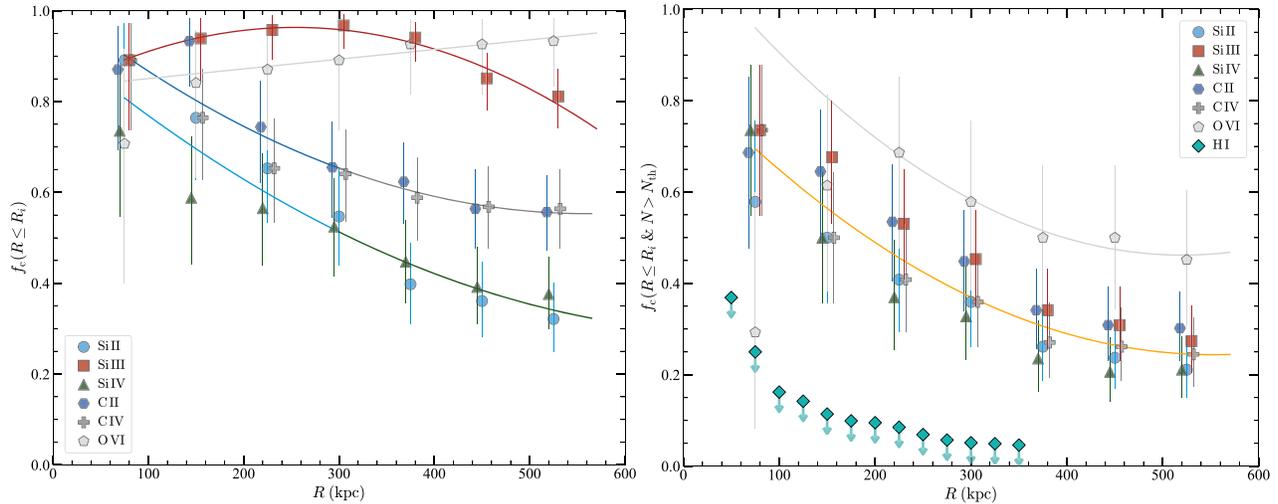
**Figure 11.** Cumulative covering factors for impact parameters less than $R$ without (left) and with (right) some threshold cut on the column densities (for Si ions, log $N_{th} = 13$; for C ions, log $N_{th} = 13.8$; for O VI, log $N_{th} = 14.6$; and for H I, log $N_{th} = 17.6$; see text for more details, and Paper I for H I). A slight artificial offsetting has been applied in the data points in order to avoid overlapping in the error bars. Confidence intervals (vertical bars) are at the 68% level, and data points are the median values. On the left panel, the solid lines are polynomial fits to median values of $f_c$ for Si III, O VI, C II–C IV, and Si II–Si IV (i.e., taking the mean value of $f_c$ between these two ions at a given $R$). On the right panel, the orange line is a polynomial fit to the mean values of $f_c$ for C II, C IV, Si II, Si III, and Si IV, while the gray line is a polynomial fit to the median values of $f_c$ for O VI.

limits—nondetections—are at the level of lowest column densities of a detected absorption, so it is adequate to do that) and (2) the sample where we set a threshold column density ($N_{th}$) to be included in the sample. In the left panel of Figure 11 we show the first case, while in the right panel we focus on the strong absorbers only. For the Si ions, we use log $N_{th} = 13$; for the C ions, log $N_{th} = 13.8$; for O VI, log $N_{th} = 14.6$. These threshold column densities are chosen to separate strong absorbers that are mostly observed within $R \lesssim R_{200}$ from the weaker absorbers that are observed at any $R$ (see Figure 8). We also show in the right panel of Figure 11 the results for the H I emission from Paper I.

These results must be interpreted in light of the fact that the intrinsic strength of the diagnostic lines varies by ion. The oscillator strengths, $f\lambda$, of these transitions are listed in Table 5, along with the solar abundances of these elements. The optical depth scales as $\tau \propto f\lambda N$ (see Section 2.4.2), and $f\lambda$ is a good representation of the strength of a given transition and $N$ of the abundance. For the Si ions, Si III has the strongest transition, a factor of 2.7–5.5 stronger than Si IV and a factor of 1.3–5.7 stronger than Si II (the weaker Si II $\lambda$1526 is sometimes used to better constrain the column density of Si II if the absorption is strong). Si II and Si IV have more comparable strength, which is also the case between C II and C IV. Comparing between different species, $(f\lambda)_{\text{Si III}} \simeq 14.4 (f\lambda)_{\text{O VI}}$, but this is counterbalanced by oxygen being 15 times more abundant than silicon (and a similar conclusion applies comparing Si III with C II or C IV).

With that in mind, we first consider the left panel of Figure 11, showing the cumulative covering factors for detections at any column density. We fitted four low-degree polynomials to the data: Si III, O VI, and treating in pairs C II+C IV and Si II+Si IV, which follow each other well. For C II+C IV and Si II+Si IV, we fit the mean covering factors of each ionic pair. For O VI, we only fitted data beyond 200 kpc owing to the smaller size sample (there are only three data points within 200 kpc and 11 in total; see Figure 8). It is striking how the cumulative covering factors of Si III and O VI

**Table 5**
Strength and Abundance of Key Ions

| Ion | $f\lambda$ | $[X/H]_\odot$ |
| --- | --- | --- |
| C II $\lambda$1334 | 171 | $-3.57$ |
| C IV $\lambda$1550 | 147 | $-3.31$ |
| C IV $\lambda$1548 | 294 | $-3.57$ |
| Si II $\lambda$1526 | 203 | $-4.49$ |
| Si II $\lambda$1190 | 348 | $-4.49$ |
| Si II $\lambda$1193 | 695 | $-4.49$ |
| Si II $\lambda$1260 | 1487 | $-4.49$ |
| Si III $\lambda$1206 | 1967 | $-4.49$ |
| Si IV $\lambda$1402 | 358 | $-4.49$ |
| Si IV $\lambda$1393 | 715 | $-4.49$ |
| O VI $\lambda$1031 | 137 | $-3.31$ |

**Note.** $[X/H]_\odot$ is the solar abundance of the element X, i.e., C, O, or Si (solar abundances from Asplund et al. 2009).

vary with $R$ quite differently from each other and from the other ions. The cumulative covering factor of Si III appears to increase with $R$, but overall it is essentially consistent with about a unity covering out to 360 kpc, where it starts to decreases. At any $R$, the covering factor of Si III remains much higher than $f_c$ of C II+C IV or Si II+Si IV. The cumulative covering factor of O VI monotonically increases with $R$ up to $R \sim 569$ kpc. In contrast, while the cumulative covering factors of C II+C IV or Si II+Si IV are offset from each other, they both monotonically decrease with $R$.

Turning to the right panel of Figure 11, where we show $f_c$ with column density thresholds that change with species (see above), the relation between $f_c$ and $R$ is quite different. For all the ions, the cumulative covering factors monotonically decrease with increasing $R$. For C II, C IV, Si II, Si III, and Si IV, the covering factors are essentially the same within $1\sigma$, and the orange line in Figure 11 shows a second-degree polynomial fit to the mean values of $f_c$ between these different ions. Ignoring data at $R < 200$ kpc owing to the small sample

size, O VI has a similar evolution of $f_c$ with $R$, but overall $f_c$ is tentatively $\sim 1.5$ times larger than for the other ions at any $R$.

The contrast between the two panels of Figure 11 and insight from Figures 8 and 9 strongly suggest that the CGM of M31 has three main populations of absorbers: (1) the strong absorbers that are found mostly at $R \lesssim 100$–150 kpc (0.3–0.5 $R_{\rm vir}$) probing the denser regions and multiple gas phase (singly to highly ionized gas) of the CGM, (2) weak absorbers probing the diffuse CGM traced principally by Si III (but also observed in higher ions and more rarely in C II) that are found at any surveyed $R$ but more frequent at $R \lesssim R_{200}$, and (3) hotter, more diffuse CGM probed by O VI, O VI having the unique property compared to the ions that its column density remains largely invariant with the radius of the M31 CGM.

### 4.7. Ion Ratios and Their Relation with R

In Section 4.3, we show that the ratio of O I to Si ions provides a direct estimate of the ionization fraction of the CGM gas of M31. Using ratios of the main ions studied here (C II, C IV, Si II, Si III, Si IV, O VI), we can further constrain the ionization and physical conditions in the CGM of M31 and how they may change with $R$. To estimate the ionic ratios, we consider the component analysis of the absorption profiles, i.e., we compare the column densities estimated over the same velocity range, noting, however, that coincident velocities do not necessarily mean that they probe the same gas, especially if their ionization potentials are quite different (such as for C II and C IV). In Figure 12, we show the results for several ion ratios as a function of $R$.

#### 4.7.1. The Si II/Si III and Si IV/Si III Ratios

The Si II/Si III and Si IV/Si III ratios are particularly useful because they trace different ionization levels independently of relative elemental abundances. The ionization potentials for these ions are 8.1–16.3 eV for Si II, 16.3–33.5 eV for Si III, and 33.5–45.1 eV for Si IV. The top two panels of Figure 12 show the ratios Si II/Si III and Si IV/Si III as a function of $R$. In both panels, there are many upper limits, and any trends are made difficult to decipher by the presence of so many limits. However, one trend is clear: all the Si II detections are confined to $R \lesssim R_{200}$ (see also Figure 9).

With so many upper limits, we use the Kaplan–Meier estimator (see Section 4.2) to estimate the mean of these ratios: $\langle \log N_{\rm Si\,II}/N_{\rm Si\,III} \rangle = (-0.50 \pm 0.04) \pm 0.23$ (mean, error on the mean from the Kaplan–Meier estimator, and standard deviation) based on 44 data points with 38 upper limits. Similarly, we find $\langle \log N_{\rm Si\,IV}/N_{\rm Si\,III} \rangle = (-0.49 \pm 0.07) \pm 0.20$ (43 data points with 32 upper limits). There are only 4/44 components with $\log N_{\rm Si\,II}/N_{\rm Si\,III} \gtrsim 0$ and 8/43 with $\log N_{\rm Si\,IV}/N_{\rm Si\,III} \gtrsim 0$. In the latter cases, Si IV could trace lower-density photoionized gas or hotter, collisionally ionized material than the lower ionization states. Among the three Si ions in our survey, Si III is the dominant ion at any $R$ from M31 in the ionizing energy range 8.1–45.1 eV. Ions (of any element) with ionizing energies in the range of 16.3–33.5 eV are therefore expected to be dominant ions at least for processes that are dominated by photoionization.

The Si II/Si III ratio has previously been used to constrain the properties of the photoionized gas. According to photoionization modeling produced by Oppenheimer et al. (2018a), an ionic ratio of $\langle \log N_{\rm Si\,II}/N_{\rm Si\,III} \rangle = (-0.50 \pm 0.04) \pm 0.23$
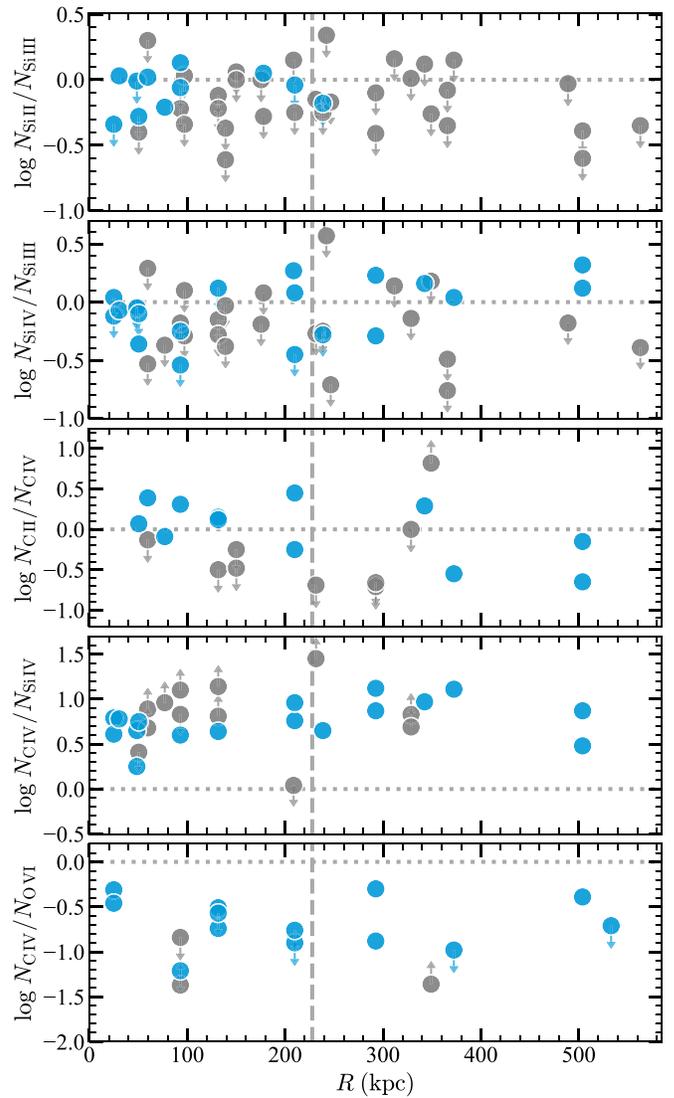


**Figure 12.** Logarithmic column density ratios of different ions as a function of the projected distances from M31 of the background QSOs. The column densities in individual components are compared to estimate the ionic ratios. Blue symbols indicate that both ions in the ratio are detected. Blue downward- or upward-pointing arrows indicate that the absorption is saturated for the ion in the denominator or numerator of the ratio, respectively. Gray symbols indicate that one of the ions in the ratio is not detected at $>2\sigma$. The components associated with the MS have again been removed. The dashed vertical lines mark $R_{200}$.

would imply gas density in the range of $-3 \lesssim \log n_{\rm H} \lesssim -2.5$ and a temperature of the gas around $10^4$ K (see Figure 16 in Oppenheimer et al. 2018a).

#### 4.7.2. The C II/C IV Ratio

For the C II/C IV ratio the ionizing energy ranges are well separated with 11.3–24.3 eV for C II and 47.9–64.4 eV for C IV. In fact, with an ionization potential above the He II ionization edge at 54.4 eV, C IV can also be produced not just by photoionization but also by collisional ionization. Therefore, C II and C IV are unlikely to probe the same ionization mechanisms or be in a gas with the same density. We note that C II has ionization energies that overlap with Si III and larger than those of Si II, which partially explains the presence of C II

beyond $R_{200}$, where Si II is systematically not detected (see also Figure 9).

The third panel of Figure 12 shows the C II/C IV ratios. There is an apparent overall decrease of C II/C IV with increasing $R$, but with a large scatter: $\log N_{\rm C\,II}/N_{\rm C\,IV} \gtrsim 0$ is more frequently observed at $R < R_{200}$ (6/12) than at $R > R_{200}$ (2/9), consistent with the observation made in Section 4.4 that the gas becomes more highly ionized as $R$ increases. With the survival analysis (considering the only lower limit as a detection), we find $\langle \log N_{\rm C\,II}/N_{\rm C\,IV} \rangle = (-0.21 \pm 0.11) \pm 0.40$ (21 data points with eight upper limits). Considering data at $R < R_{200}$, we have $\langle \log N_{\rm C\,II}/N_{\rm C\,IV} \rangle = (-0.07 \pm 0.10) \pm 0.28$ (12 data points with four upper limits), while at $R \geqslant R_{200}$ we find $\langle \log N_{\rm C\,II}/N_{\rm C\,IV} \rangle = (-0.33 \pm 0.18) \pm 0.22$ (nine data points with four upper limits), confirming again that the gas is more ionized and also more highly ionized at $R > R_{200}$.

### 4.7.3. The C IV/Si IV Ratio

For the C IV/Si IV ratio, ions from different elements are compared, but as we discuss in Section 4.2, the relative abundances of C and Si are consistent with the solar ratio owing to little evidence of any strong dust depletion or nucleosynthesis effects, i.e., these effects should not impact the observed ratio of C IV/Si IV. Si IV and C IV have near-adjacent ionization energies of 33.5–45.1 eV and 47.9–64.4 eV, respectively. Both photoionization and collisional ionization processes can be important at these ionizing energies. However, for $\log N_{\rm C\,IV}/N_{\rm Si\,IV} > 0$, ionization from hot stars is unimportant (see Figure 13 in Lehner et al. 2011), which is nearly always the case, as illustrated in Figure 12. A harder photoionizing spectrum or collisional ionization must be at play to explain the origin of these ions.

Figure 12 suggests a moderate correlation between $\log N_{\rm C\,IV}/N_{\rm Si\,IV}$ and $R$. If the two data points beyond 400 kpc are removed (and treating the limits as actual values), a Spearman rank order implies a monotonic correlation between $\log N_{\rm C\,IV}/N_{\rm Si\,IV}$ and $R$ with a correlation coefficient $r_{\rm S} = +0.45$ and $p = 0.019$ for the gas at $R < 1.2\,R_{\rm vir}$. Considering the entire sample, the Spearman rank test yields $r_{\rm S} = 0.34$ and $p = 0.07$. This is again consistent with our earlier conclusion that the gas becomes more highly ionized as $R$ increases. With the survival analysis (considering the three upper limits as detections),[30] we find $\langle \log N_{\rm C\,IV}/N_{\rm Si\,IV} \rangle = (+0.87 \pm 0.07) \pm 0.24$ (29 data points with nine lower limits). This is about a factor of 1.9 larger than the mean derived for the broad C IV and Si IV components in the MW disk and low halo (Lehner et al. 2011), which is about one standard deviation larger.

### 4.7.4. The C IV/O VI Ratio

Finally, in the last panel of Figure 12, we show the C IV/O VI ratio as a function of $R$. As for the C IV/Si IV ratio, ions from different elements are compared, and for the same reasons, the relative dust depletion or nucleosynthesis effects should be negligible. With 113.9–138.1 eV ionizing energies needed to produce O VI, this is the most highly ionized species in our survey, and, as we demonstrated in the previous section, the O VI properties (covering factor and column density as a function of $R$) are quite unique. Figure 12 does not reveal any relation between $\log N_{\rm C\,IV}/N_{\rm O\,VI}$ and $R$.

If we treat the two lower limits as detections, then the survival analysis yields $\langle \log N_{\rm C\,IV}/N_{\rm O\,VI} \rangle = (-0.93 \pm 0.11) \pm 0.32$ (16 data points with six upper limits). The mean and range of $\log N_{\rm C\,IV}/N_{\rm O\,VI}$ are smaller than observed in the MW disk and low halo, where the full range varies from $-1$ to $+1$ dex (see, e.g., Figure 14 of Lehner et al. 2011). This demonstrates that the highly ionized gas in the 113.9–138.1 eV range is much more important than in the 47.9–64.4 eV range at any $R$ of the M31 CGM.

### 4.8. Metal and Baryon Mass of the M31 CGM

With a better understanding of the column density variation with $R$, we can estimate with more confidence the metal and baryon masses of the M31 CGM than in LHW15, which had very little information between 50 and 300 kpc. The metal mass can be directly estimated from the column densities of the metal ions. With the silicon ions, we have information on its three dominant ionization stages in the $T < 7 \times 10^4$ K ionized gas (ionizing energies in the range of 8–45 eV; see Section 4.5), so we can obtain a direct measured metal mass without any major ionization corrections. Following LHW15 (see also Peeples et al. 2014), the metal mass of the cool photoionized CGM is

$$M_Z^{\rm cool} = 2\pi\, \mu_{\rm Si}^{-1}\, m_{\rm Si} \int R\, N_{\rm Si}(R)\, dR,$$

where $\mu_{\rm Si} = 0.064$ is the solar mass fraction of metals in silicon (i.e., $12 + \log({\rm Si/H})_\odot = 7.51$ and $Z_\odot = 0.0142$ from Asplund et al. 2009), $m_{\rm Si} = 28 m_{\rm p}$, and for $N_{\rm Si}(R)$ we use the H (Equation (F1)), SPL (Equation (F2)), and GP models that we determine in Section 4.5 and Appendix F (see Figure 10).

A direct method to estimate the total mass is to convert the total observed column density of Si to total hydrogen column density via $N_{\rm H} = N_{\rm H\,I} + N_{\rm H\,II} = N_{\rm Si}\,({\rm Si/H})_\odot^{-1}\,(Z/Z_\odot)^{-1}$. The baryonic mass of the CGM of M31 is then

$$M_{\rm g}^{\rm cool} = 2\pi\, m_{\rm H}\, \mu\, f_{\rm c} \left(\frac{\rm Si}{\rm H}\right)_\odot^{-1} \left(\frac{Z}{Z_\odot}\right)^{-1} \int R\, N_{\rm Si}(R)\, dR,$$

where $\mu \simeq 1.4$ (to correct for the presence of He), $m_{\rm H} = 1.67 \times 10^{-24}$ g is the hydrogen mass, $f_{\rm c}$ is the covering fraction (which is 1 over the considered radii), and $\log({\rm Si/H})_\odot = -4.49$ is the solar abundance of Si. Inserting the values for each parameter, $M_{\rm g}^{\rm cool}$ can be simply written in terms of $M_Z^{\rm cool}$: $M_{\rm g}^{\rm cool} \simeq 10^2 (Z/Z_\odot)^{-1} M_Z^{\rm cool}$.

In Table 6, we summarize the estimated metal mass over different regions of the CGM for the three models of $N_{\rm Si}(R)$, within $R_{200}$ (first entry), within $R_{\rm vir}$ (second entry), within $1/2\,R_{\rm vir}$ (third entry), between $1/2\,R_{\rm vir}$ and $R_{\rm vir}$ (fourth entry), and within 360 kpc (fifth entry), which corresponds to the radius where at least one of the Si ions is always detected (beyond that, the number of detections drastically plummets). A key difference between the H/SPL models and the GP model is that the range of values for the H/SPL models is derived using the low (dotted) and high (dashed) curves in Figure 10, while for the GP models we actually use the confidence intervals from the low and high models (i.e., the top and bottom of the shaded blue curve in Figure 10). Hence, it is not surprising that the mass ranges for the GP model are larger. Nevertheless, there is a large overlap between the three models. As the GP results overlap with the other models and provide empirical confidence

---

[30] If these three upper limits are included or excluded from the sample, the means are essentially the same.

**Table 6**
Metal Mass of the Cool CGM of M31

| Range | $M_Z$ H Model | $M_Z$ SPL Model | $M_Z$ GP Model |
|---|---|---|---|
| (kpc) | ($10^7 M_\odot$) | ($10^7 M_\odot$) | ($10^7 M_\odot$) |
| 5–230 | 1.5–2.1 | 1.4–1.7 | 1.2–2.1 |
| 5–300 | 2.0–2.6 | 1.8–2.0 | 1.5–2.5 |
| 5–150 | 1.0–1.5 | 1.0–1.2 | 0.73–1.4 |
| 150–300 | 1.0–1.1 | 0.77–0.78 | 0.72–1.1 |
| 5–360 | 2.5–3.0 | 2.1–2.3 | 1.7–2.9 |

**Note.** $M_Z$ corresponds to the metal mass traced by Si II, Si III, and Si IV. The H model corresponds to the low and high hyperbola models, the SPL model represents the high and low single power-law models, and the GP model encompasses the range allowed by the errors from the low and high model (see Section 4.5 and Appendix F).

intervals, we adopt them for the remaining of the paper. Within $R_{\rm vir}$, the metal and cool gas masses are therefore $(2.0 \pm 0.5) \times 10^7$ and $2 \times 10^9 (Z/Z_\odot)^{-1} M_\odot$, respectively. Owing to the new functional form of $N_{\rm Si}(R)$ and how the lower limits are treated, this explains the factor of 1.4 increase in the metal mass compared to that derived in LHW15.

These masses do not include the more highly ionized gas traced by O VI, C IV, or even higher (unobserved) ionization states. Even though the sample with O VI is smaller than C IV, we use O VI to probe the higher-ionization gas phase because, as we show above, the properties of O VI (column density and covering fraction as a function of $R$) are quite different from those of all the lower ions, including C IV, which behaves more like the other, lower ions. Furthermore, Lehner et al. (2011), using 1.5–3 km s$^{-1}$ resolution UV spectra, show that C IV can probe cool and hotter gas, while the profiles of N V and O VI are typically broad and more consistent with hotter gas. Since O VI is always detected and there is little evidence for variation with $R$ (see Figure 8), we can simply use the mean column density $\log N_{\rm O VI} = 14.46 \pm 0.10$ (error on the mean using the survival method for censoring) to estimate the baryon mass assuming a spherical distribution:

$$M_{\rm g}^{\rm warm} = \pi r^2 \, m_{\rm H} \, \mu \, f_{\rm c} \, \frac{N_{\rm O VI}}{f_{\rm O VI}^i} \left(\frac{\rm O}{\rm H}\right)_\odot^{-1} \left(\frac{Z}{Z_\odot}\right)^{-1},$$

where the O VI ionization fraction is $f_{\rm O VI}^i \lesssim 0.2$ (see Section 4.3) and $f_{\rm c} = 1$ for O VI at any $R$ (see Figure 8). Within $R_{\rm vir}$, we find $M_{\rm g}^{\rm warm} \gtrsim 9.3 \times 10^9 (Z/Z_\odot)^{-1} M_\odot$ or $M_{\rm g}^{\rm warm} \gtrsim 4.4 M_{\rm g}^{\rm cool}$ (assuming that the metallicity is about similar in the cool and warm gas phases). At $R_{200}$, we find $M_{\rm g}^{\rm warm} \gtrsim 5.5 \times 10^9 (Z/Z_\odot)^{-1} M_\odot$. These are lower limits because the fraction of O VI could be much smaller than 20% and the metallicity of the cool or warm ionized gas is also likely to be less than solar (see below). In terms of metal mass in the highly ionized gas phase, this gives $M_{\rm g}^{\rm warm} \simeq 10^2 (Z/Z_\odot)^{-1} M_Z^{\rm warm}$ and $M_Z^{\rm warm} \gtrsim 4.4 M_Z^{\rm cool}$. Since O VI is detected out to the maximum surveyed radius of 569 kpc, within that radius (i.e., 1.9 $R_{\rm vir}$) we find $M_{\rm g}^{\rm warm} \gtrsim 34 \times 10^9 (Z/Z_\odot)^{-1} M_\odot$.

By combining both the cool and hot gas-phase masses, we can find the baryon mass for gas in the temperature range of $\sim 10^4$–$10^{5.5}$ K within $R_{\rm vir}$:

$$M_{\rm g} = M_{\rm g}^{\rm cool} + M_{\rm g}^{\rm warm}$$
$$\gtrsim 1.1 \times 10^{10} \left(\frac{Z}{Z_\odot}\right)^{-1} M_\odot.$$

Within $R_{200}$, the total mass $M_{\rm g} \gtrsim 7.2 \times 10^9 M_\odot$. As the stellar mass of M31 is about $10^{11} M_\odot$ (e.g., Geehan et al. 2006; Tamm et al. 2012), the mass of the diffuse weakly and highly ionized CGM of M31 within 1 $R_{\rm vir}$ is therefore at least 10% of the stellar mass of M31 and could be significantly larger than 10%.

This estimate does not take into account the hot ($T \gtrsim 10^6$ K) coronal gas. The diffuse X-ray emission is observed to extend to about 30–70 kpc around a handful of massive, nonstarbursting galaxies (Anderson & Bregman 2011; Bregman et al. 2018) or in stacked images of galaxies (Anderson et al. 2013; Bregman et al. 2018), but beyond 50 kpc, the CGM is too diffuse to be traced with X-ray imaging, even though a large mass could be present. Using the results summarized recently by Bregman et al. (2018), the hot gas mass of spiral galaxy halos is in the range $M_{\rm g}^{\rm hot} \simeq (1$–$10) \times 10^9 M_\odot$ within 50 kpc (assuming or estimating metallicities in the range 0.13–$0.5 Z_\odot$). For M31, $M_{\rm g} = M_{\rm g}^{\rm cool} + M_{\rm g}^{\rm warm} \gtrsim 0.4 \times 10^9 M_\odot$ within 50 kpc. Extrapolating the X-ray results to $R_{\rm vir}$, Bregman et al. (2018) find masses of the hot X-ray gas similar to the stellar masses of these galaxies in the range $M_{\rm g}^{\rm hot} \simeq (1$–$10) \times 10^{11} M_\odot$. For the MW hot halo within 1 $R_{\rm vir}$, Gupta et al. (2017; but see also Gupta et al. 2012, 2014; Wang & Yao 2012; Henley et al. 2014) derive $(3$–$10) \times 10^{10} M_\odot$, i.e., on the low side of the mass range listed in Bregman et al. (2018). The hot gas could therefore dominate the mass of the CGM of M31. There are, however, two caveats to that latter conclusion. First, if $f_{\rm O VI} \ll 0.2$, then $M_{\rm g}^{\rm warm}$ could become much larger than the lower limits we give. Second, the metallicity of the hot X-ray gas ranges from 0.1 to $0.5 Z_\odot$ with a mean metallicity of $0.3 Z_\odot$ (Bregman et al. 2018; Gupta et al. 2017), while for the cooler gas we have conservatively adopted a solar abundance. If instead we adopt a $0.3 Z_\odot$ metallicity (consistent with the rough limits set in Sections 4.1, 4.2), then $M_{\rm g} \simeq 3.7 \times 10^{10} M_\odot$ within $R_{\rm vir}$, which is now comparable to the hot halo mass of the MW. If we adopt the average metallicity derived for the X-ray gas, then $M_{\rm g}^{\rm cool} + M_{\rm g}^{\rm warm}$ would be comparable to the hot gas mass if $M_{\rm g}^{\rm hot} \sim 5 \times 10^{10} M_\odot$ within $R_{\rm vir}$ for M31. Depending on the true metallicities and the actual state of ionization, the cool and warm gas in the M31 halo could therefore contribute to a substantial enhancement of the total baryonic mass compared to our conservative assumptions.

### 4.9. Mapping the Metal Surface Densities in the CGM of M31

Thus far, we have ignored the distribution of the targets in azimuthal angle ($\Phi$) relative to the projected minor and major axes of M31, where different physical processes may occur. In Figure 13, we show the distribution of the column densities of each ion in the $X$-$Y$ plane near M31, where the circles represent detections and downward-pointing triangles are nondetections. Multiple colors in a given circle indicate multiple components along that sightline for that ion. We also show the projected minor and major axes of M31 (dashed lines). The overall trends that are readily apparent from Figure 13 are the ones already described in the previous sections: (1) overall the column density decreases with increasing $R$, (2) the decrease in $N$ is much stronger for low ions than high ions, and (3) Si III and
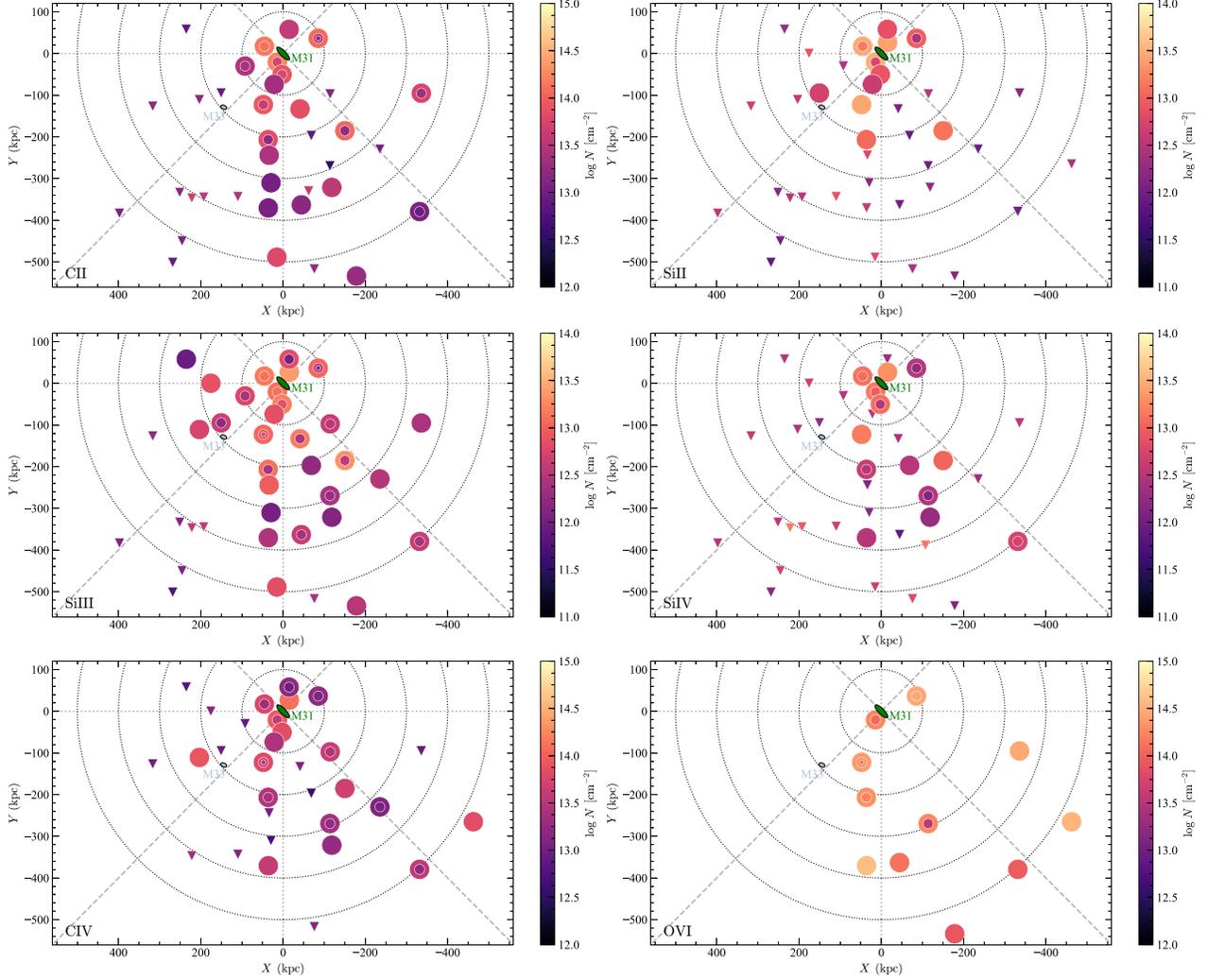
**Figure 13.** Positions of the Project AMIGA targets relative to the M31, where the axes show the physical impact parameter from the center of M31 (north is up, east to the left). Dotted circles are centered on M31 to mark 100 kpc intervals. The dashed lines represent the projected minor and major axes of M31, and the thin dotted lines are ±45° from the major/minor axes (which by definition of the coordinate systems also correspond to the vertical and horizontal zero axis). Each panel corresponds to a different ion. In each panel, the column densities of each velocity component are shown and color-coded according the vertical color bar. Circles represent detections, while triangles are nondetections. Circles with several colors indicate that the observed absorption along the sightlines has more than one component.

O VI are observed at all $R$ probed, while singly ionized species tend to be more frequently observed at small impact parameters. This figure (and Figure 9) also reveals that absorption with two or more components is observed more frequently at $R < 200$ kpc: using Si III, 64%–86% of the sightlines have at least two velocity components at $R < 200$ kpc, while this drops to 14%–31% at $R > 200$ kpc (68% confidence intervals using the Wilson score interval); similar results are found using the other ions. However, the complexity of the velocity profiles does not change with $\Phi$.

Considering various radius ranges (25–50 kpc, 50–100 kpc, etc.) up to 1 $R_{\rm vir}$, there is no indication that the column densities depend strongly on $\Phi$. Considering Si III first, it is equally detected along the projected major and minor axes and in between (wherever there is a sightline), and overall the strength of the absorption mostly depends on $R$, not $\Phi$. Considering the other ions, they all show a mixture of detections and nondetections, and the nondetections (which are mostly beyond 50 kpc) are not preferentially observed along a certain axis or

one of the regions shown in Figure 13. We therefore find no strong evidence of an azimuthal dependence in the column densities.

Beyond $\gtrsim 1.1 \, R_{\rm vir}$, the situation is different, with all but one detection (in C IV and O VI only) being near the southern projected major axis and about 52° west off near the $X = 0$ kpc axis. There is detection in this region of Si III, C IV, Si IV, O VI, and also C II. That is the main region where C II is detected beyond 200 kpc. In contrast, between the $X = 0$ kpc axis and southern projected minor axis, the only region where there are several QSOs beyond $R_{\rm vir}$, there is no detection in any of the ions (excluding O VI because there are no FUSE observations in these directions). Although that direction is suspiciously in the direction of the MS, it is very unlikely to be additional contamination from the MS because (1) the velocities are inconsistent with the MS expectations in these directions (see Figure 3 and also Section 4.10), and (2) we see a decrease in the column densities as $|b_{\rm MS}|$ decreases (closer to the core of the MS), the opposite of expectations if the material is associated
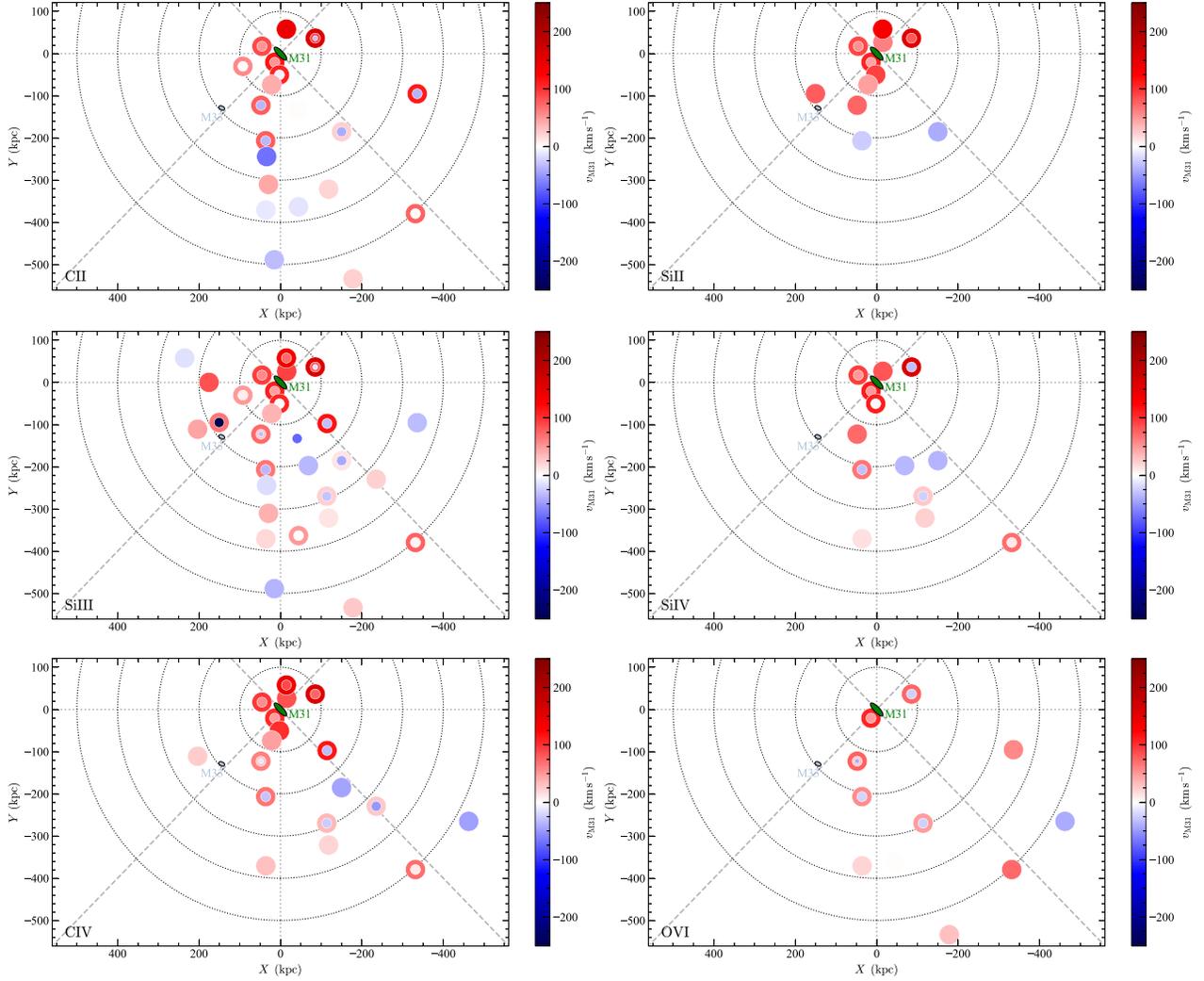
**Figure 14.** Similar to Figure 13, but we now show the distribution of the M31 velocities for each component observed for each ion. Circles with several colors indicate that the observed absorptions along the sightlines have more than one component. By definition, in the M31 velocity frame, an absorber with no peculiar velocity relative to M31's bulk motion has $v_{M31} = 0$ km s$^{-1}$.

with the MS. Therefore, while at $R < R_{vir}$ there is no apparent trend between $N$ and $\Phi$ for any ions (although we keep in mind that the azimuthal information for O VI is minimal), most of the detections at $R > R_{vir}$ are near the southern projected major axis and 52° west off of that axis.

The fact that the gas is observed mainly in a specific region of the CGM beyond $R_{vir}$ suggests an IGM filament feeding the CGM of M31, as is observed in some cosmological simulations. In particular, Nuza et al. (2014) study the gas distribution in simulated recreations of MW and M31 using a constrained cosmological simulation of the Local Group from the Constrained Local UniversE Simulations (CLUES) project. In their Figures 3 and 6, they show different velocity and density projection maps where the central galaxy (M31 or MW) is edge-on. They find that some of the gas in the CGM can flow in a filament-like structure, coming from outside the virial radius all the way down to the galactic disk.

### 4.10. Mapping the Velocities in the CGM of M31

How the velocity field of the gas is distributed in $R$ and $\Phi$ beyond 25–50 kpc is a key diagnostic of accretion and

feedback. However, a statistical survey using one sightline per galaxy (such as COS-Halos) cannot address this problem because it observes many galaxies in an essential random mix of orientations and inclinations, which necessarily washes out any coherent velocity structures. An experiment like Project AMIGA is needed to access information about large-scale flows in a sizable sample of lines of sight for a single galaxy. The velocity information remains limited because we have only the (projected) radial velocity along pencil beams piercing the CGM at various $R$ and $\Phi$. Nevertheless, as we show below, some trends are apparent thanks to the large size of the sample. We use here the $v_{M31}$ peculiar velocities as defined by Equation (2). By definition, in the M31 velocity frame, an absorber with no peculiar velocity relative to M31's bulk motion has $v_{M31} = 0$ km s$^{-1}$. In Section 3.2, we show that the M31 peculiar velocities of the absorbers seen toward the QSOs and the velocities of the M31 dwarf satellites largely overlap. We now review how the velocities of the absorbers are distributed in the CGM of M31 over the entire surveyed range of $R$.

In Figures 14 and 15, we show the distribution of the M31 peculiar velocities of the individual components identified for
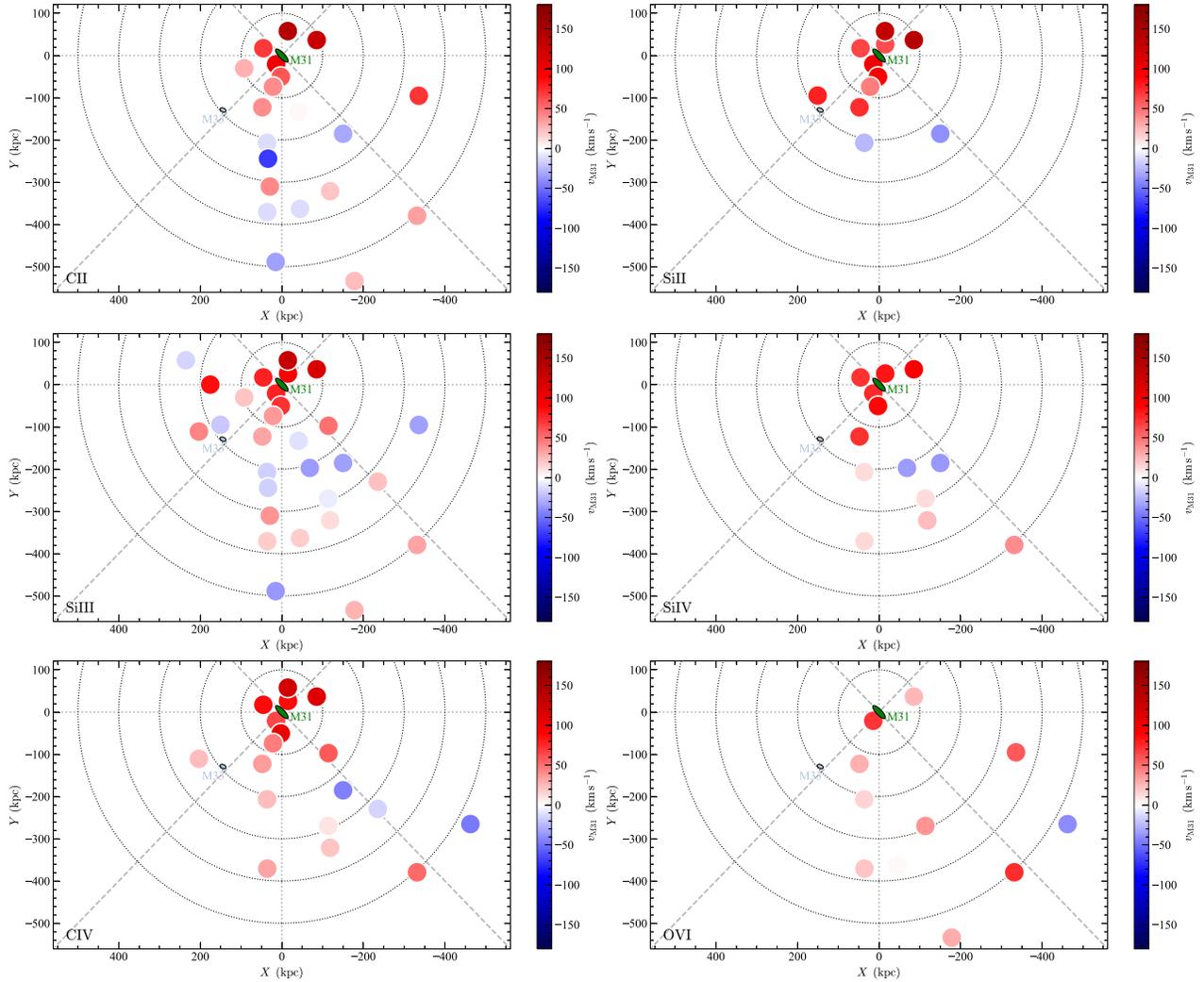
**Figure 15.** Same as Figure 14, but for the average velocities.

each ion and the column-density-weighted average velocities of each ion, respectively. Circles with several colors again indicate that the observed absorption appears in more than one component. Both Figures 14 and 15 demonstrate that in many cases there is some overlap in the velocities between lower ions (Si II, C II, Si III) and higher ions (Si IV, C IV, O VI). This strongly implies that the CGM of M31 has multiple gas phases with overlapping kinematics when they are observed in projection (a property also readily observed from the normalized profiles shown in Figure 2 and in the figure set in the online Journal). There are also some rarer cases where there is no velocity correspondence in the velocities between Si III and higher ions (see, e.g., near $X \simeq -335$, $Y \simeq -95$ kpc), indicating that the observed absorption in each ion is dominated by a single phase–that is, the components are likely to be distinct single-phase objects.

The full range of velocities associated with the CGM of M31 is $-249 \lesssim v_{\rm M31} \lesssim +175$ km s$^{-1}$ for Si III, but for all the other ions it is $-53 \lesssim v_{\rm M31} \lesssim +175$ km s$^{-1}$. Furthermore, there is only one absorber/component of Si III that has $v_{\rm M31} = -249$ km s$^{-1}$. We emphasize that the rarity of velocity $v_{\rm M31} < -249$ km s$^{-1}$ (corresponding to $v_{\rm LSR} < -510$ km s$^{-1}$ in the direction of this sightline) is not an artifact since

velocities below these values are not contaminated by any foreground gaseous features.

We show in Section 3.2 that the M31 dwarf satellites have a velocity dispersion that is larger (110 km s$^{-1}$ for the dwarfs vs. 68 km s$^{-1}$ for the Si III absorbers) and the M31 dwarfs have some velocities in the velocity range contaminated by the MW and MS. While the CGM gas velocity field distribution may not follow that of the dwarf satellites, it remains plausible that some of the absorption from the extended region of the M31 CGM could be lost owing to contamination from the MW or MS, i.e., we may not be fully probing the entire velocity distribution of the M31 CGM. However, as discussed in Section 2.5, there is no evidence that the velocity distributions of the Si III component in and outside the MS contamination zone are different (see also Figures 1 and 14), and hence it is quite possible that at least the MS contamination does not affect much the velocity distribution of the M31 CGM. With these caveats, we now proceed to describe the apparent trends of the velocity distribution in the CGM of M31.

From Figure 14, the first apparent property was already noted in the previous section: the velocity complexity (and hence full width) of the absorption profiles increases with decreasing $R$ (see Section 4.9). Within $R \lesssim 100$ kpc or

**Table 7**
Summary of the Velocities of the CGM of M31

| Ion | Region | $\langle v_{M31}\rangle$ (km s$^{-1}$) | IQR (km s$^{-1}$) |
|---|---|---|---|
| | | Individual Components | |
| C II | All $R$ | $31.3 \pm 52.3$ | $[-6.5, 68.0]$ |
| C II | $R \leqslant 100$ kpc | $80.6 \pm 39.9$ | $[37.9, 101.1]$ |
| C II | $R > 100$ kpc | $4.7 \pm 36.6$ | $[-21.2, 24.9]$ |
| Si II | All $R$ | $65.7 \pm 53.1$ | $[53.8, 95.2]$ |
| Si II | $R \leqslant 100$ kpc | $90.6 \pm 32.3$ | $[67.2, 117.8]$ |
| Si II | $R > 100$ kpc | $22.3 \pm 54.5$ | $[-27.0, 76.8]$ |
| Si III | All $R$ | $26.8 \pm 45.5$ | $[-10.1, 54.7]$ |
| Si III | $R \leqslant 100$ kpc | $77.9 \pm 34.1$ | $[59.9, 95.9]$ |
| Si III | $R > 100$ kpc | $7.4 \pm 32.3$ | $[-12.4, 36.1]$ |
| Si IV | All $R$ | $39.5 \pm 43.1$ | $[8.3, 71.0]$ |
| Si IV | $R \leqslant 100$ kpc | $82.2 \pm 7.5$ | $[76.3, 88.4]$ |
| Si IV | $R > 100$ kpc | $12.8 \pm 33.7$ | $[-0.6, 26.9]$ |
| C IV | All $R$ | $42.6 \pm 47.3$ | $[10.8, 70.2]$ |
| C IV | $R \leqslant 100$ kpc | $88.3 \pm 23.8$ | $[71.1, 102.1]$ |
| C IV | $R > 100$ kpc | $13.5 \pm 33.4$ | $[4.9, 40.4]$ |
| O VI | All $R$ | $28.9 \pm 30.8$ | $[13.1, 43.0]$ |
| | | Averaged Components | |
| C II | All $R$ | $34.2 \pm 58.9$ | $[-23.0, 68.1]$ |
| C II | $R \leqslant 100$ kpc | $67.0 \pm 55.3$ | $[22.7, 92.4]$ |
| C II | $R > 100$ kpc | $10.5 \pm 49.3$ | $[-33.2, 37.2]$ |
| Si II | All $R$ | $68.1 \pm 50.6$ | $[54.2, 98.7]$ |
| Si II | $R \leqslant 100$ kpc | $86.4 \pm 35.1$ | $[61.5, 111.2]$ |
| Si II | $R > 100$ kpc | $22.3 \pm 54.5$ | $[-27.0, 76.8]$ |
| Si III | All $R$ | $25.3 \pm 67.3$ | $[-18.9, 66.9]$ |
| Si III | $R \leqslant 100$ kpc | $69.3 \pm 48.1$ | $[43.8, 103.3]$ |
| Si III | $R > 100$ kpc | $4.1 \pm 64.9$ | $[-24.0, 47.7]$ |
| Si IV | All $R$ | $39.6 \pm 55.4$ | $[-2.5, 78.8]$ |
| Si IV | $R \leqslant 100$ kpc | $69.9 \pm 56.2$ | $[52.3, 112.4]$ |
| Si IV | $R > 100$ kpc | $14.9 \pm 40.4$ | $[-22.9, 50.3]$ |
| C IV | All $R$ | $43.3 \pm 55.7$ | $[8.4, 84.4]$ |
| C IV | $R \leqslant 100$ kpc | $91.3 \pm 33.7$ | $[69.9, 107.4]$ |
| C IV | $R > 100$ kpc | $14.0 \pm 45.2$ | $[-8.0, 52.8]$ |
| O VI | All $R$ | $27.9 \pm 44.0$ | $[-9.8, 68.5]$ |

**Note.** Average, standard deviation, and IQR velocity range are listed for the individual components and averaged M31 velocity components.

$\lesssim 200$ kpc, about 75% of the Si III absorbers have at least two components (at the COS G130M−G160M resolution). This drops to about 33% at $200 < R \lesssim 569$ kpc.

The second property evident from either Figure 14 or Figure 15 is that the M31 peculiar velocities are larger at $R \lesssim 100$ kpc than at higher $R$. Table 7 lists the average M31 velocities, their standard deviations, and their interquartile ranges (IQRs) for the individual components and averaged components in three samples: the full AMIGA set, the subset with $R \leqslant 100$ kpc, and the subset with $R > 100$ kpc. From this table and for all the ions besides O VI, $\langle v_{M31}\rangle = +90$ km s$^{-1}$ at $R \leqslant 100$ kpc, while at $R > 100$ kpc, $\langle v_{M31}\rangle = +20$ km s$^{-1}$, a factor of 4.5 smaller. There are only two data points for O VI, at $R \leqslant 100$ kpc, but the average at $R > 100$ kpc is also $\langle v_{M31}\rangle = +22$ km s$^{-1}$, following a similar pattern to that observed for the other ions. For all the ions but C II, the velocity dispersions or IQRs are smaller at $R \leqslant 100$ kpc than at $R > 100$ kpc.

The third property observed in Figure 14 or Figure 15 is that at $R \leqslant 100$ kpc there is no evidence for negative M31 velocities, while at $R > 100$ kpc about 40% of the Si III sample

has blueshifted $v_{M31}$ velocities. This partially explains the previous result, but even if we consider the absolute velocities, $\langle |v_{M31}|\rangle = +40$ km s$^{-1}$ at $R > 100$ kpc, implying $\langle |v_{M31}(R > 100)|\rangle = 0.44\langle |v_{M31}|(R \leqslant 100)\rangle$, i.e., in absolute terms or not, $v_{M31}$ is smaller at $R > 100$ kpc than at $R \leqslant 100$ kpc. Therefore, at $R > 100$ kpc, not only are the peculiar velocities of the CGM gas less extreme, but they are also more uniformly distributed around the bulk motion of M31. At $R < 100$ kpc, the peculiar velocities of the CGM gas are more extreme and systematically redshifted relative to the bulk motion of M31.

The fourth property appears in Figure 7, where we compare $v_{M31}$ velocities of the M31 dwarfs and Si III absorbers, which shows that overall the velocities of the satellites and the CGM gas do not follow each other. As noted in Section 3.3 (see also Table 4), some velocity components seen in absorption toward the QSOs are found with $\Delta_{sep} < 1$ and have $\delta v < v_{esc,dwarf}$. However, the last two trends described above for the CGM gas are not observed for the dwarfs. First, contrasting with the M31 CGM gas, Figure 7 shows that both blue- and redshifted $v_{M31}$ velocities for the dwarfs are observed at any $R$ and azimuth. Second, at $R > 100$ kpc and $R \leqslant 100$ kpc, the average $v_{M31}$ velocities for the dwarfs are also similar, while for the CGM gas they are quite different (see above and Table 7). These differences strongly suggest that the velocity fields of the dwarfs and CGM gas are decoupled. We infer from this decoupling that (1) gas bound to satellites does not make a significant contribution to CGM gas observed in this way, and (2) the velocities of gas removed from satellites via tidal or ram pressure interactions, if it is present, become decoupled from the dwarf that brought it in (as one might expect from its definition as unbound to the satellites).

The fifth property is more readily apparent considering the average velocities shown in Figure 15, where, considering the CGM gas in different annuli, there is an apparent change in the sign of the average $v_{M31}$ velocities with on average a positive velocity at $R < 200$ kpc ($\langle v_{M31}\rangle = +58.5 \pm 43.5$ km s$^{-1}$ for Si III), negative velocity at $200 < R < 300$ kpc ($\langle v_{M31}\rangle = -11.3 \pm 24.2$ km s$^{-1}$ for Si III), and again positive velocity at $300 < R < 400$ kpc ($\langle v_{M31}\rangle = +12.3 \pm 21.8$ km s$^{-1}$ for Si III). This is more evident with Si III, where the sample of absorbers is larger, but taking the average velocities in the different annuli, the same pattern is observed for C II, Si III, Si IV, and C IV.

We emphasize again that the MS contamination does not really alter these properties and neither is the source of these properties. As shown in Figure 1 (see also Section 2.5), the MS contamination predominantly occurs in the region $X < 0$. There is no evidence that these properties change with $\Phi$ and in particular between the halves of the map $X < 0$ and $X > 0$ (see Section 2.5). Absorption occurring in the velocity range $-50 \lesssim v_{M31} \lesssim +150$ km s$^{-1}$ is also not contaminated by the MS.

## 5. Discussion

The major goal of Project AMIGA is to determine the global distribution of the gas phases and metals through the entire CGM of a representative galaxy. With a large sample of QSOs accumulated over many surveys, and newly observed by HST/COS, we are able to probe multiple sightlines that pierce M31 at different radii and azimuthal angles. Undertaking this study in the UV has been critical since only in this wavelength band are there diagnostics and spectral resolution to constrain the

physical properties of the multiple gas phases existing in the CGM over $10^{4-5.5}$ K (for $z = 0$, the hottest phase can only be probed with X-ray observations). With 25 sightlines within about 1.1 $R_{\rm vir}$ and 43 within 569 kpc ($\lesssim 1.9\,R_{\rm vir}$) of M31, the size of the sample and the information as a function of $R$ and $\Phi$ are unparalleled. We will now consider the broad patterns and conclusions we can draw from this unique data set.

### 5.1. Pervasive Metals in the CGM of M31

A key finding of Project AMIGA is the ubiquitous presence of metals throughout the CGM of M31. While the search for H I with $\log N_{\rm H\,I} \gtrsim 17.5$ in the CGM of M31 toward pointed radio observations has been unsuccessful in the current sample (Paper I and see Figure 1), the covering factor of Si III (29 sightlines) is essentially 100% out to 1.2 $R_{\rm vir}$, while O VI associated with M31 is detected toward all 11 sightlines with FUSE data, all the way out to 1.9 $R_{\rm vir}$, the maximum radius of our survey (see Sections 4.4, 4.6). From the ionization range probed by Project AMIGA, we further show that Si III and O VI are key probes of the diffuse gas (see Section 4.7). With information from Si II, Si III, and Si IV, we demonstrate that Si III is the dominant ion in the ionizing energy range of 8–45 eV (see Section 4.7.1). The high covering factors of Si III and O VI imply that these ions could be pervasively distributed, but it is quite possible also that small structures could lead to these large covering factors in projection.

The finding of ubiquitous metals in the CGM of M31 is a strong indication of ongoing and past gas outflows that ejected metals well beyond their formation site. Based on a specific SFR of SFR/$M_\star = (5 \pm 1) \times 10^{-12}$ yr$^{-1}$ (using the stellar mass $M_\star$ and SFR from Geehan et al. 2006; Kang et al. 2009), M31 is not currently in an active star-forming episode. In fact, Williams et al. (2017) show that the bulk of star formation occurred in the first $\sim$6 billion years and the last strong episode happened over $\sim$2 billion years ago (see also Figure 6 in Telford et al. 2019 for a metal production model of M31). For a typical outflow with gas moving at 100 km s$^{-1}$, the gas would have traveled about 200 kpc in 2 Gyr, i.e., this is the maximum distance at which the last burst of star formation could have impacted the CGM of M31. Hence, many of the metals seen in the CGM of M31 must have also been ejected by previous star-forming episodes and/or stripped from its dwarfs and more massive companions. However, the most distant metals, those beyond $R_{\rm vir}$ (and especially given that they are predominantly found along a preferred direction from M31; see Section 4.9), may be supplied from elsewhere. For example, they could have originated in a Local Group medium bearing metals lost from other galaxies (e.g., the MW or dwarfs) or in a pre-polluted IGM filament.

In Section 4.8, we estimate that the mass of metals $M_Z^{\rm cool} = (2.0 \pm 0.5) \times 10^7\,M_\odot$ within $R_{\rm vir}$ for the predominantly photoionized gas probed by Si II, Si III, and Si IV. For the gas probed by O VI, we find that $M_Z^{\rm warm} > 4.4 M_Z^{\rm cool} \gtrsim 9 \times 10^7\,M_\odot$ within $R_{\rm vir}$ (this is a lower limit because the ionization fraction of O VI is an upper limit; see Section 4.8). The sum of these two phases yields a lower limit to the CGM metal mass because the hotter phase probed by X-rays and metals bound in dust is not included. If the hot baryon mass of M31 is not too different from that estimated for the MW (see Section 4.8), then we expect $M_Z^{\rm hot} \approx M_Z^{\rm warm}$. The CLUES simulation of the Local Group estimates that the mass of the hot ($>10^5$ K) gas is a factor of 3 larger than the cooler ($<10^5$ K) gas (Nuza et al. 2014). The CGM dust mass remains quite uncertain but could be at the level of $5 \times 10^7\,M_\odot$ according to estimates around 0.1–1$L^*$ galaxies (Ménard et al. 2010; Peeples et al. 2014; Peek et al. 2015). Hence, a plausible lower limit on the total metal mass of the CGM of M31 out to $R_{\rm vir}$ is $M_Z^{\rm CGM} \gtrsim 2.5 \times 10^8\,M_\odot$, a factor of $\sim$2 times larger than the cool+warm gas-phase metal mass.

The stellar mass of M31 is $(1.5 \pm 0.2) \times 10^{11}\,M_\odot$ (e.g., Williams et al. 2017). Using this result, Telford et al. (2019) estimated that the current metal mass in stars is $3.9 \times 10^8\,M_\odot$, i.e., about the same amount that is found in the entire CGM of M31 up to $R_{\rm vir}$. Telford et al. (2019) also estimated the metal mass of the gas in the disk of M31 to be around $(0.8–3.2) \times 10^7\,M_\odot$, while Draine et al. (2014) estimated the dust mass in the disk to be around $5.4 \times 10^7\,M_\odot$, yielding a total metal mass in the disk of M31 of about $M_Z^{\rm disk} \simeq 5 \times 10^8\,M_\odot$. Therefore, M31 has in its CGM within $R_{\rm vir}$ at least 50% of the present-day metal mass in its disk. As we show in Sections 4.4 and 4.8 and discuss above, metals are also found beyond $R_{\rm vir}$, especially in the more highly ionized phase traced by O VI (and even higher ions). These metals may have come from M31 or from other galaxies in the Local Group such as the MW or dwarf galaxies.

### 5.2. Comparison with COS-halos Galaxies

The Project AMIGA experiment is quite different from most of the surveys of the CGM of galaxies done so far. Outside the local universe, surveys of the CGM of galaxies involve assembling samples of CGM gas in aggregate by using one sightline per galaxy (see Section 1), in some nearby cases up to three to four sightlines (e.g., Bowen et al. 2016; Keeney et al. 2017). By assembling a sizable sample of absorbers associated with galaxies in a particular subpopulation (e.g., $L^*$, sub-$L^*$, passive or star-forming galaxies), one can then assess how the column densities change with radii around that kind of galaxy, and from this estimate average surface densities, mass budgets, etc., can then be evaluated. By contrast, Project AMIGA has assembled almost as many sightlines surrounding M31 as COS-Halos had for its full sample of 44 galaxies. We can now make a direct comparison between these two types of experiments. For this comparison, we use the COS-Halos survey of $0.3 < L/L^* < 2$ galaxies at $z \simeq 0.2$, which selected galaxies within about 160 kpc from the sightline (Tumlinson et al. 2011, 2013; Werk et al. 2013, 2014). The full mass range of the COS-Halos galaxies is quite large, $11.5 \lesssim \log M_{200} \lesssim 13.7$, but most of the star-forming galaxies are in the range $11.5 \lesssim \log M_{200} \lesssim 12.5$, and most of the passive quiescent galaxies have $13.0 \lesssim \log M_{200} \lesssim 13.7$. As a reminder, M31 has $\log M_{200} = 12.1$ (see Section 1).

#### 5.2.1. Column Densities of Si versus R

In Figure 16, we show the total column densities of Si as a function of $R/R_{200}$ for the COS-Halos galaxies and M31. For the COS-Halos survey, each data point corresponds to an absorber at some impact parameter from a galaxy, while for M31, each data point is an absorber probing the CGM at a different impact parameter from the same galaxy. For COS-Halos, we consider two cases: (1) the column densities of Si estimated in a similar fashion as those for M31, and (2) the
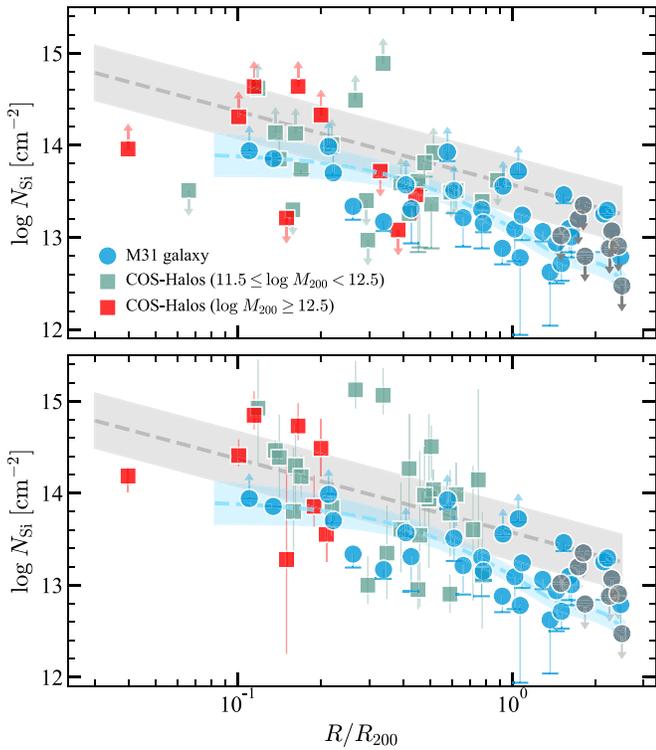
**Figure 16.** Comparison of the total column densities of Si from M31 and the COS-Halos galaxies as a function of $R/R_{200}$. Top panel: $N_{Si}$ are directly constrained by the estimates on $N_{Si\,II}$, $N_{Si\,III}$, and $N_{Si\,IV}$ from the observations for both Project AMIGA and COS-Halos (Werk et al. 2013). The error bars are less than the size of the symbols, and the vertical bars include the range of possible values if the nondetections are either near their $2\sigma$ upper limits or so low as to be negligible. Bottom panel: same as the top panel, but $N_{Si}$ for the COS-Halos data is derived from CLOUDY photoionization models (Werk et al. 2014). The gray dashed line and shaded region represent the best fit between $N_{Si}$ and $R/R_{200}$ and its dispersion using COS-Halos modeled data, respectively. The blue area shows the full range of the GP model of the Project AMIGA data (see Section 4.5).

column densities of Si estimated from photoionization modeling. For case 2 we use the results from Werk et al. (2014; see also Prochaska et al. 2017b), which were used to determine the metal mass of the CGM of the COS-Halos galaxies in Peeples et al. (2014). For case 1, we use the results from Werk et al. (2013) and follow the procedure in Section 4.5 to estimate $N_{Si}$ from the column densities of Si II, Si III, and Si IV. We require that all three Si ions are available, except in the cases where there are only lower limits for two ions (typically Si II and Si III) since in that case the resulting column density is a lower limit that encompasses any missing column density from the remaining Si ion (typically Si IV). The sample size in case 1 is 35, while in case 2 it is 33, with some overlap between the two subsamples. In Figure 16, we also show the modeled column density of Si as a function of $R$, in gray for COS-Halos and blue for M31 (we show the adopted GP model; see Section 4.5).

A striking difference between the COS-Halos and M31 data immediately apparent from Figure 16 is that a large number of very high Si column densities are observed in COS-Halos at $R/R_{200} \lesssim 0.3$ that are absent in Project AMIGA. The reason that the COS-Halos Si column densities have higher lower limits than those of M31 is because the weak transitions of Si II are saturated in COS-Halos, a situation not observed in M31 toward any of the sightlines—the lower limits of Si arise only because Si III is saturated. These high Si column densities also

correspond to the very strong $N_{H\,I}$ ($\log N_{H\,I} \gtrsim 18$) absorbers observed in COS-Halos, but again not in M31 (see Figure 5 and Section 4.1 in Paper I). Beam dilution could have affected somewhat the comparison between H I column densities from COS-Halos (H I absorption) and M31 (H I emission); for the metal ions, this is not an issue. Therefore, the higher frequency of saturated Si II transitions in COS-Halos compared to M31 is a real effect, not an artifact.

Besides this difference, the estimated Si column densities from the observations in the COS-Halos and Project AMIGA surveys are distributed with a similar scatter at larger impact parameters ($R/R_{200} \gtrsim 0.4$), where the gas is more ionized (see top panel in Figure 16). The photoionization-modeled COS-Halos Si displayed in the bottom panel has some higher values than observed in the top panel, but in the impact parameter region $0.4 \lesssim R/R_{200} \lesssim 0.8$ where they are observed, there are also several lower limits. Beyond $R > 0.9R_{200}$, there are no COS-Halos observations (owing to the design of the survey). The extrapolated model to the COS-Halos observations shown in gray in Figure 16 is a factor of 2–4 higher than the models of the Project AMIGA data shown in blue depending on $R/R_{200}$.

A likely explanation for the higher column density absorbers is that some of these COS-Halos absorbers could be fully or partly associated with a galaxy near to the initially targeted COS-Halos galaxies. In this case the gas would be more likely to show high columns of neutral and weakly ionized gas. Indeed, while the COS-Halos galaxies were selected to have no bright companions, that selection did not preclude fainter nearby companions such as dwarf satellites (see Tumlinson et al. 2013). Galaxy observation follow-up by Werk et al. (2012) found several $L > 0.1L^*$ galaxies within 160 kpc of the targeted COS-Halos galaxy. Comparing the results from other surveys of galaxies/absorbers at low redshift (Bowen et al. 2002; Stocke et al. 2013), Bregman et al. (2018) also noted a higher preponderance of high H I column density absorbers in the COS-Halos survey. However, the higher COS-Halos column densities at large radii could also be an effect of evolution in the typical CGM, as COS-Halos probed a slightly higher cosmological redshift. It is also possible that the M31 CGM is less rich in neutral gas at these radii than the typical $L^*$ galaxy at $z \sim 0.2$, because of its star formation history or environment.

### 5.2.2. CGM Mass Comparison

A key physical parameter of the CGM is its mass, which is obtained from the column density distribution of the gas and assuming a certain geometry of the gas. For M31, we cannot derive the baryonic mass of CGM gas without assuming a metallicity since the H I column density remains largely unknown and very uncertain toward all the targets in our sample (but see Sections 4.1 and 4.8). However, the metal mass of the cool gas probed by Si II, Si III, and Si IV can be straightforwardly estimated directly from the observations without any ionization modeling (see top panel of Figure 16).

Even though both Peeples et al. (2014) and Werk et al. (2014) use the Si column densities derived from photoionization models, as illustrated in Figure 16, this would not change the outcome that the metal mass of the cool CGM gas derived from the COS-Halos survey is about a factor of 2–3 higher than the metal mass derived in Project AMIGA. This is because there are seven COS-Halos Si column densities at $R/R_{200} < 0.3$ that are much higher than seen about M31,

driving the overall model of $N_{Si}(R)$ substantially higher. The fact that these high $N_{Si}$ values are not found in the CGM of M31 or lower-redshift galaxies at similar impact parameters (e.g., Bowen et al. 2002; Stocke et al. 2013) suggests a source of high column density H I and Si II absorbers in the COS-Halos sample that could be recent outflows, strong accretion/recycling, or gas associated with closer satellites to the sightline. With only five targets within $R/R_{200} < 0.3$ and none below $R/R_{200} < 0.1$ for M31, it would be quite useful to target more QSOs in the inner region of the CGM of M31 to better determine how $N_{Si}(R)$ varies with $R$ at small impact parameters.

For the warm–hot gas probed by O VI, the COS-Halos star-forming galaxies have $\langle N_{O\,VI} \rangle = 10^{14.5}$ cm$^{-2}$, a detection rate close to 100%, and no large variation of $N_{O\,VI}$ with $R$ (Tumlinson et al. 2011). For M31, we have a similar average O VI column density, similar hit rate, and little evidence for large variation of $N_{O\,VI}$ with $R$ (see Section 4.8). This implies that the masses of the warm–hot CGM of M31 and COS-Halos star-forming galaxies are similar. M31 has a specific SFR that is a factor $\gtrsim 10$ lower than the COS-Halos star-forming galaxies, but its halo mass is on the higher side of the COS-Halos star-forming galaxies (but lower than the COS-Halos quiescent galaxies). As discussed in Section 5.3.3 in more detail, M31 and the COS-Halos star-forming galaxies have halo masses in the range $M_{200} \simeq 10^{11.7}$–$10^{12.3}\,M_\odot$, corresponding to a virial temperature range that overlaps with the temperature at which the ionization fraction of O VI peaks, which may naturally explain some of the properties of the O VI in the CGM of "$L^*$" galaxies (Oppenheimer et al. 2018b). It is also possible that some O VI arises in photoionized gas or combinations of different phases (see Section 5.3.3).

Based on the comparison above, we find that the O VI is less subject to the uncertainty in the association of the absorber to the correct galaxy owing to its column density being less dependent on $R$ (see also Section 5.3). Therefore, this leads to similar metal masses of the CGM of the $z \sim 0.2$ COS-Halos galaxies and M31 for the O VI gas phase. For the lower ions, their column densities are more dependent on $R$ (see also Section 5.3). Therefore, the association of the absorber to the correct galaxy is more critical to derive an accurate column density profile with $R$ and hence derive an accurate CGM metal mass. However, we note that despite these uncertainties, the metal mass of the cool CGM of the COS-Halos galaxies is only a factor of 2–3 higher than that derived for M31.

### 5.3. A Changing CGM with Radius

A key discovery from Project AMIGA is that the properties of the CGM of M31 change with $R$. This is reminiscent of our earlier survey (LHW15), but the increase in the size sample has transformed some of the tentative results of our earlier survey into robust findings. In particular, the radius around $R \sim 100$–150 kpc appears critical in view of several properties changing near this threshold radius:

1. For any ions, the frequency of strong absorption is larger at $R \lesssim 100$–150 kpc than at larger $R$.
2. The column densities of Si and C ions change by a factor $>5$–10 between about 25 and 150 kpc, while they change only by a factor $\lesssim 2$ between 150 and 300 kpc.
3. The detection rate of singly ionized species (C II, Si II) is close to 100% at $R < 150$ kpc but sharply decreases

beyond (see Figure 9), and therefore the gas has a more complex gas-phase structure at $R < 150$ kpc.
4. The peculiar velocities of the CGM gas are more extreme and systematically redshifted relative to the bulk motion of M31 at $R \lesssim 100$ kpc, while at $R \gtrsim 100$ kpc the peculiar velocities of the CGM gas are less extreme and more uniformly distributed around the bulk motion of M31.

There are also two other significant regions: (1) beyond $R_{200} \simeq 230$ kpc the gas is becoming more ionized and more highly ionized than at lower $R$ (e.g., there is a near total absence of Si II absorption beyond $R_{200}$ [see Figure 9], or a higher C II/C IV ratio on average at $R \gtrsim R_{200}$ than at lower $R$ [see Section 4.7.2]), and (2) beyond 1.1 $R_{vir}$ the gas is not detected in all the directions away from M31, as it is at smaller radii, but only in a cone near the southern projected major axis and about 52° west of the $X = 0$ kpc axis (see Section 4.9).

The overall picture that can be drawn out from these properties is that the inner regions of the CGM of M31 are more dynamic and complex, while the more diffuse regions at $R \gtrsim 0.5\ R_{vir}$ are more static and "simpler." Zoom-in cosmological simulations capture in more detail and more accurately the structures of the CGM than large-scale cosmological simulations thanks to their higher mass and spatial resolution. Below we use several results from zoom simulations to gain some insights on these observed changes with $R$. However, the results laid out in Section 4 also now provide a new test bed for zoom simulations, so that not only qualitative but also quantitative comparison can be undertaken. We note that most of the zoom simulations discussed here have only a single massive halo. However, according to the ELVIS simulations of Local Group analogs (Garrison-Kimmel et al. 2014), there should be no major difference at least within about $R_{vir}$ for the distribution of the gas between isolated and paired galaxies.

### 5.3.1. Visualization and Origins of the CGM Variation

To help visualize the properties described above and gain some insights into the possible origins of these trends, we begin by qualitatively examining two zoom simulations. First, we consider the Local Group zoom simulations from the CLUES project (Nuza et al. 2014), where the gas distribution around MW and M31-like galaxies is studied. Their work does not show the distribution of the individual ions but examines the two main gas phases above and below $10^5$ K in an environment that is a constrained analog to the Local Group. Interestingly, considering Figure 3 (simulated M31) or Figure 6 (simulated MW) in Nuza et al., the region within 100–150 kpc appears more complex, with a large covering factor for both cool and hot gas phases and higher velocities than at larger radii. In these simulations, this is a result of the combined effects of cooling and supernova heating affecting the closer regions of the CGM of M31. This simulation also provides an explanation for the gas observed beyond 1.1 $R_{vir}$ that is preferentially observed in a limited region of the CGM of M31 (see Figure 13 and see middle right panel of their Figure 3) whereby the $\lesssim 10^5$ K gas might be accreting onto the CGM of M31. We also note that Nuza et al. (2014) find a mass for the $\lesssim 10^5$ K CGM gas of $1.7 \times 10^{10}\,M_\odot$, broadly consistent with our findings (see Section 4.8). More quantitative comparisons between the CLUES (or Local Group analog simulations like ELVIS-FIRE simulations; Garrison-Kimmel et al. 2014, 2019) and Project

AMIGA results are beyond the scope of this paper, but they would be valuable to undertake in the future.

Second, we consider the zoom Eris2 simulation of a massive, star-forming galaxy at $z = 2.8$ presented in Shen et al. (2013). The Eris2 galaxy being $z = 2.8$ and with an SFR of 20 $M_\odot$ yr$^{-1}$ is nothing like M31, but this paper shows the distribution of the gas around the central galaxy using some of the same ions that are studied in Project AMIGA, specifically Si II, Si IV, C II, C IV, and O VI (see their Figures 3(a) and 4(a), (b)). Because Eris2 is so different from M31, we would naively expect their CGM properties to be different, and yet (1) Eris2 is surrounded by a large diffuse O VI halo with a near-unity covering factor all the way out to about 3 $R_{\rm vir}$; (2) the covering factor of absorbing material in the CGM of Eris2 declines less rapidly with impact parameter for C IV or O VI compared to C II, Si II, or Si IV; and (3) beyond $R_{\rm vir}$, the covering factor of Si II drops more sharply than C II. There are also key differences, like the strongest absorption in any of these ions being observed in the bipolar outflows perpendicular to the plane of the disk, which is unsurprisingly not observed in M31 since it currently has a low SFR (e.g., Williams et al. 2017). However, the broad pictures of the CGM of M31 and the simulated Eris2 galaxy are remarkably similar. This implies that some of the properties of the CGM may depend more on the microphysics producing the various gas phases than the large-scale physical processes (outflow, accretion) that vary substantially over time. In fact, the Eris2 simulation shows that inflows and outflows coexist and are both traced by diffuse O VI; in Eris2, a high covering factor of strong O VI absorbers seems to be the least unambiguous tracer of large-scale outflows.

### 5.3.2. Quantitative Comparison in the CGM Variation between Observations and Simulations

Two simulations of M31-like galaxies in different environments at widely separated epochs show some similarity with some of the observed trends in the CGM of M31. We now take one step further by quantitatively comparing the column density variation of the different ions as a function of $R$ in three different zoom-in cosmological simulations, two being led by members of the Project AMIGA team (FIRE and FOGGIE collaborations), and a zoom-in simulation from the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulation project (Crain et al. 2015; Schaye et al. 2015; Oppenheimer et al. 2018a).

(1) *Comparison with FIRE-2 Zoom Simulations*

We first compare our observations with column densities modeled using cosmological zoom-in simulations from the FIRE project.[31] Details of the simulation setup and CGM modeling methods are presented in Ji et al. (2020). Briefly, the outputs analyzed here are FIRE-2 simulations evolved with the GIZMO code using the meshless finite mass (MFM) solver (Hopkins 2015). The simulations include a detailed model for stellar feedback including core-collapse and Type Ia supernovae, stellar winds from OB and AGB stars, photoionization, and radiation pressure (for details, see Hopkins et al. 2018). We focus on the "m12i" FIRE halo, which has a mass $M_{\rm vir} \approx 1.2 M_{200} \approx 1.2 \times 10^{12}\, M_\odot$ at $z = 0$, which is comparable to the halo mass of M31. However, neither the SFR history nor the present-day SFR is similar. The "m12i" FIRE halo has a factor

of 10–12 higher SFR (see Figure 3 in Hopkins et al. 2020) than the present-day SFR of M31 of 0.5 $M_\odot$ yr$^{-1}$ (e.g., Kang et al. 2009). We compare Project AMIGA to FIRE-2 simulations with two different sets of physical ingredients. The "MHD" run includes magnetic fields, anisotropic thermal conduction, and viscosity, and the "CR" run includes all these processes plus the "full physics" treatment of stellar cosmic rays (CRs). The CR simulation assumes a diffusion coefficient $\kappa_{||} = 3 \times 10^{29}$ cm$^{-2}$ s$^{-1}$, which was calibrated to be consistent with observational constraints from $\gamma$−ray emission of the MW and some other nearby galaxies (Chan et al. 2019; Hopkins et al. 2020). Ji et al. (2020) showed that CRs can potentially provide a large or even dominant nonthermal fraction of the total pressure support in the CGM of low-redshift $\sim L^*$ galaxies. As a result, in the fiducial CR run analyzed here, the volume-filling CGM is much cooler ($\sim 10^4$–$10^5$ K) and is thus photoionized in regions where in the run without CRs it is filled with hot gas that is more collisionally ionized.

The column densities are generated as discussed in Ji et al. (2020) using the Trident code (Hummels et al. 2017) to model the ion populations and density projections. For the photo-ionization modeling, a hybrid treatment combining the FG09 (Faucher-Giguère et al. 2009) and HM12 (Haardt & Madau 2012) UV background models is used.[32]

In Figure 17, we compare the ion column densities from FIRE-2 simulations with observationally derived total column densities around M31 as a function of $R/R_{200}$. The green and orange curves show the median simulated column densities for the MHD and CR runs, respectively, while the shaded regions show the full range of columns for all sightlines at a given impact parameter (the lowest values are truncated to match the scales that are adequate for the observations; see Ji et al. 2020 for the full range of values). The CR run produces higher column densities and better agreement with observations than the MHD run for all ions presented. The much higher column densities of low/intermediate ions (C II, Si II, Si III, and Si IV) in the CR run are due to the more volume-filling and uniform cool phase, which produces higher median values of ion column densities and smaller variations across different sightlines. In contrast, in the MHD run the cool phase is pressure confined by the hot phase to compact and dense regions, leading to smaller median columns but larger scatter for the low and intermediate ions. We note, however, that even in the CR runs the predicted column densities are lower than observations at the larger impact parameters $R \gtrsim 0.5 R_{200}$. This might be due to insufficient resolution to resolve fine-scale structure in outer halos, or it may indicate that feedback effects are more important at large radii than in the present simulations. This difference is quite notable owing to the fact that the star formation of the "m12i" galaxy has been continuous with an SFR in the range of 5–20 $M_\odot$ yr$^{-1}$ (Hopkins et al. 2020) over the past $\sim 8$ billion years, while M31 had only a continuous SFR around 6–8 $M_\odot$ yr$^{-1}$ over its first 5 billion years, while over the past 8 billion years it had only two short bursts of star formation about 4 and 2 billion years ago

---

[31] FIRE project website: http://fire.northwestern.edu.

[32] We use this mixture because, based on the recent UV background analysis of Faucher-Giguère (2020), the FG09 model is in better agreement with the most up-to-date low-redshift empirical constraints at energies relevant for low and intermediate ions (C II, Si II, Si III, Si IV, and C IV). However, the HM12 model is likely more accurate for high ions such as O VI because the FG09 model used a crude AGN spectral model that underpredicted the higher-energy part of the UV/X-ray background. Ji et al. (2020) show how some ion columns depend on the assumed UV background model.
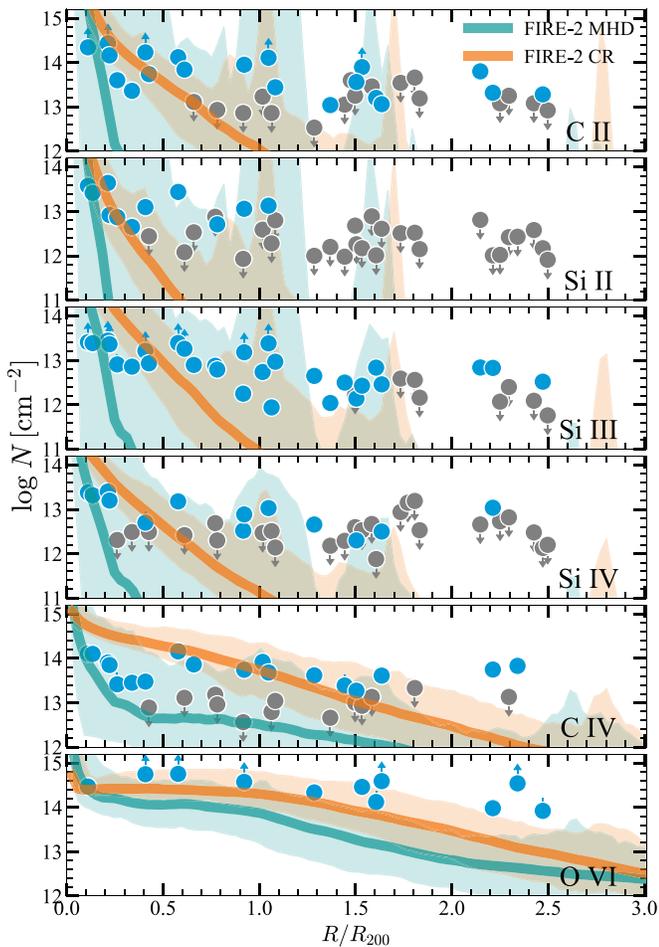
**Figure 17.** Comparison of ion column density profiles between Project AMIGA (total column densities) and FIRE-2 simulations, with "MHD" and "CR" runs. Thick curves show median values of an ensemble of sightlines produced from simulations, and shaded regions show the full range across all model sightlines.

(Williams et al. 2017). While there are some clear discrepancies, the simulations also follow some similar trends to those seen in the empirical results detailed in Section 4: (1) the simulated column densities of the low ions decrease more rapidly with $R$ than the high ions; (2) O VI is observed beyond $1.7R_{200}$, where there is no substantial amount of low/intermediate ions; and (3) a larger scatter is observed in the column densities of the low and intermediate ions than in O VI.

In the FIRE-2 simulations, both collisional ionization and photoionization can contribute significantly to the simulated O VI columns, typically with an increasing contribution from photoionization with increasing impact parameter, driven by decreasing gas densities. In the MHD run, most of the O VI in the inner halo ($R \lesssim 0.5R_{200}$) is produced by collisional ionization, but photoionization can dominate at larger impact parameters. In the CR run, collisional ionization and photo-ionization contribute comparably to the O VI mass at radii $50 < R < 200$ kpc (Ji et al. 2020). The actual origins of the CGM in terms of gas flows in FIRE-2 simulations without magnetic fields or CRs were analyzed in Hafen et al. (2020), although the results are expected to be similar for simulations with MHD only. In these simulations, O VI exists as part of a well-mixed hot halo, with contributions from all the primary channels of CGM mass growth: IGM accretion, wind, and

contributions from satellite halos (reminiscent of the Eris2 simulations; see above and Shen et al. 2013). The metals responsible for O VI absorption originate primarily in winds, but IGM accretion may contribute a large fraction of total gas mass traced by O VI since the halo is well mixed and IGM accretion contributes $\gtrsim 60\%$ of the total CGM mass. In the simulations, the hot halo gas persists in the CGM for billions of years, and the gas that leaves the CGM does so primarily by accreting onto the central galaxy (Hafen et al. 2019).

(2) *Comparison with FOGGIE Simulations*

We also compare the observed total column densities to the MW-mass "Tempest" ($M_{200} \approx 4.2 \times 10^{11} M_\odot$) halo from the FOGGIE simulations,[33] which has a halo mass of $M_{200} \approx 4.2 \times 10^{11} M_\odot$ (Peeples et al. 2019). Again we employ the software Trident (Hummels et al. 2017) to generate the ion populations and density projections, limiting potential sources of discrepancy. We use the $z = 0$ output (see Zheng et al. 2020 for simulation details), but because of the size difference between M31 and the Tempest galaxy, we again scale all distances by $R_{200}$ ($R_{200} = 159$ kpc for the simulated halo compared to 230 kpc for M31). The only "feedback" included in this FOGGIE run is thermal explosion-driven supernova outflows. While this feedback is limited in scope compared to FIRE, FOGGIE achieves higher mass resolution than FIRE-2 by using a "forced refinement" scheme that applies a fixed computational cell size of $\sim 381\ h^{-1}$ pc within a moving cube centered on the galaxy that is $\sim 200\ h^{-1}$ ckpc on a side. This refinement scheme enforces constant *spatial resolution* on the CGM, resulting in a variable and very small mass resolution in the low-density gas, with typical cell masses of ($\lesssim 1$–$100\ M_\odot$). The individual small-scale structures that contribute to the observed absorption profiles can therefore be resolved. These small-scale structures that become only apparent in high-resolution simulations are hosts to a significant amount of cool gas, enhancing the column densities in especially the low ionization state of the gas (Peeples et al. 2019; Corlies et al. 2020; see also Hummels et al. 2019; Rhodin et al. 2019; van de Voort et al. 2019).

As for the FIRE-2 simulations, we compare the Project AMIGA total column densities to FOGGIE because in the simulation we do not (yet) separate individual components, but look at the projected column densities through the halo. We note that the CGM is not necessarily self-similar, so some differences between the simulation predictions and M31 observations at rescaled impact parameter could be due to the halo mass difference. This is especially so since the halo mass range $M_h \approx 3 \times 10^{11}$–$10^{12} M_\odot$ corresponds to the expected transition between cold and hot accretion (e.g., Birnboim & Dekel 2003; Keres et al. 2003; Faucher-Giguère et al. 2011; Stern et al. 2020).

In Figure 18, we compare the simulated and observed column densities for each ion probed by our survey. The pink and green shaded areas are the data points from the simulation (with and without satellite contribution, respectively) and show the total column density in projection through the halo. The scatter in the simulated data points comes from variation in the structures along the mock sightline, and most of the scatter is in fact below $10^{11}$ cm$^{-2}$. The peaks in the column densities are due to small satellites in the halo, which enhance primarily the low-ion column densities. We show the green points to

---

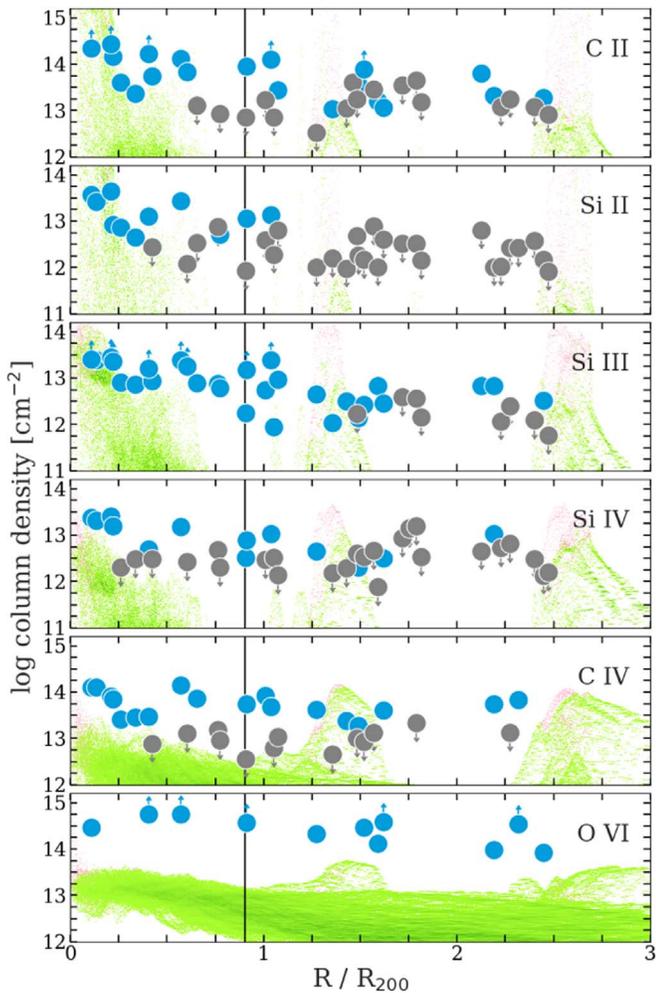[33] FOGGIE project website: http://foggie.science.

**Figure 18.** Comparison of ion column density profiles between Project AMIGA (total column densities) and the "Tempest" halo from the FOGGIE simulations. The pink and green shaded areas are projected total column densities from the simulated halo with and without galaxy/satellite contributions, respectively, while the rest of the figure is analogous to Figure 8. The vertical line shows the extent of the forced resolution cube in the FOGGIE simulation.

highlight the difference between the mock column densities with and without satellites. For the high-ionization lines the difference is negligible, while the difference in the low ions is significant.

Overall, the metal-line column densities are systematically lower than in the observations at any $R$. Only at $R \lesssim 0.3R_{200}$ is there some overlap for the singly ionized species between the FOGGIE simulation and observations. However, the discrepancy is particularly striking for Si III and the high ions. This can be understood by the current feedback implementation in FOGGIE, which does not expel enough metals from the stellar disk into the CGM (Hamilton-Campos et al. 2020) to be consistent with known galactic metal budgets (Peeples et al. 2014). This effect is expected to be stronger for the high ions than the low ions, due to the additional heating and ionization of the CGM that would be expected from stronger feedback, and indeed the discrepancy between the simulation and observations is larger for the high ions (and Si III) than for the singly ionized species. However, while the absolute scale of the column densities is off, there are also some similarities between the simulation and observations in the behavior of the

relative scale of the column density profiles with $R$: (1) the column densities of the low ions drop more rapidly with $R$ than the high ions; (2) despite the inadequate feedback in the current simulations, the O VI-bearing gas (and C IV to a lesser extent) is observed well beyond $R_{200}$; and (3) a larger scatter is observed in the column densities of the low and intermediate ions than in O VI. It is striking that the overall slope of the O VI profile resembles the observations but at significantly lower absolute column density. In the FOGGIE simulation, the low ions tracing mainly dense, cool gas are preferentially found in the disk or satellites, while the hotter gas traced by the higher ions is more homogeneously distributed in the halo.

(3) *Comparison with EAGLE Simulations*

Finally, we compare our results with the EAGLE zoom-in simulations (EAGLE *Recal-L025N0752* high-resolution volume) discussed in length in Oppenheimer et al. (2018a). The EAGLE simulations have successfully reproduced a variety of galaxy observables (e.g., Crain et al. 2015; Schaye et al. 2015) and achieved "broad but imperfect" agreement with some of the extant CGM observations (e.g., Turner et al. 2016; Oppenheimer et al. 2018a; Rahmati & Oppenheimer 2018; Lehner et al. 2019; Wotta et al. 2019).

Oppenheimer et al. (2018a) aimed to directly study the multiphase CGM traced by low metal ions and to compare with the COS-Halos survey (see Section 5.2). As such, they explored the circumgalactic metal content traced by the same ions explored in Project AMIGA in the CGM galaxies with masses that comprise that of M31. Overall Oppenheimer et al. find agreement between the simulated and COS-Halos samples for Si II, Si III, Si IV, and C II within a factor of two or so and larger disagreement with O VI, where the column density is systematically lower. With Project AMIGA, we can directly compare the results with one of the EAGLE galaxies that has a mass very close to M31 and also compare the column densities beyond 160 kpc, the maximum radius of the COS-Halos survey (Tumlinson et al. 2013; Werk et al. 2013). We refer the reader to Oppenheimer et al. (2016), Rahmati & Oppenheimer (2018), and Oppenheimer et al. (2018a) for more details on the EAGLE zoom-in simulations, which employ the SpecExBin code (Oppenheimer & Davé 2006) for ion modeling and column density projections. We also refer the reader to Figure 1 in Oppenheimer et al. (2018a), where in the middle column they show the column density map for galaxy halo mass of $\log M_{200} = 12.2$ at $z \simeq 0.2$, which qualitatively shows similar trends described in Section 5.3.1.

In Figure 19, we compare the EAGLE and observed column densities as a function of the impact parameter out to $R_{vir}$. As in the previous two figures, the blue and gray circles are detections and nondetections in the halo of M31, respectively. The green curve in each panel represents the mean column density for each ion as a function of the impact parameter for the EAGLE galaxy with $\log M_{200} = 12.1$ at $z = 0$. In contrast to FIRE-2 or FOGGIE simulations, the EAGLE simulations appear to produce a better agreement between $N$ and $R$ for low and intermediate ions (Si II, Si III, Si IV), and C IV out to larger impact parameters. However, as already noted in Oppenheimer et al. (2018a), this agreement is offset by producing too much column density for the low and intermediate ions at small impact parameters (see, e.g., Si II, which is not affected by lower limits and is clearly overproduced at $R \lesssim 80$ kpc). The flat profile of O VI, with very little dependence on $R$, is similar to the observations and other models, but overall the EAGLE
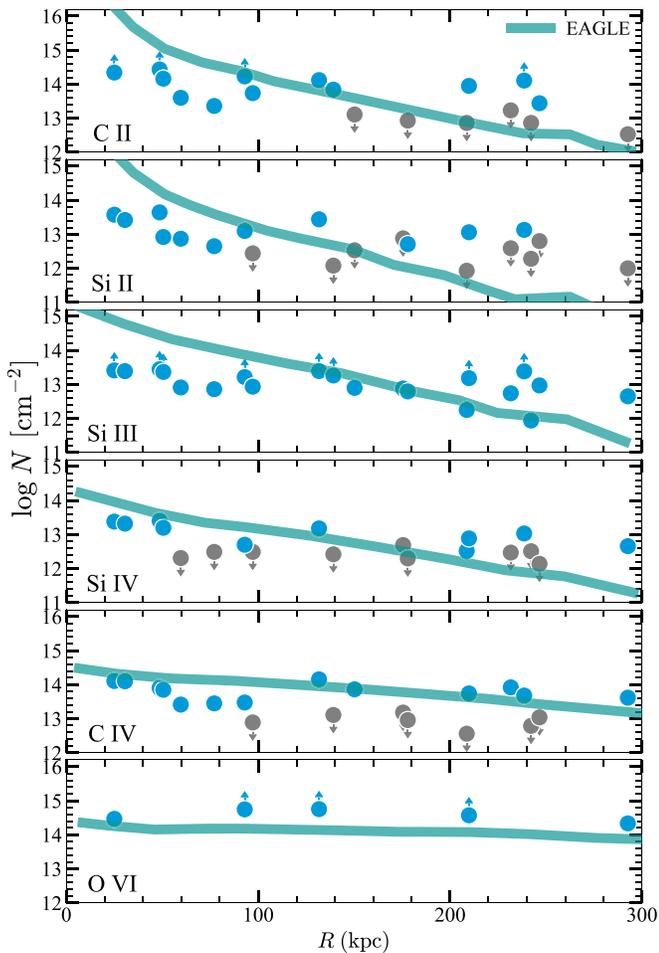
**Figure 19.** Comparison of ion column density profiles between Project AMIGA and EAGLE zoom-in simulations of a galaxy with $\log M_{200} \simeq 12.1$ at $z = 0$ (from the models presented in Oppenheimer et al. 2018a). For the EAGLE simulation, the mean column densities are shown. Note that here we only plot the column density profiles out to about $R_{\rm vir}$.

O VI column densities are a factor of 0.2–0.6 dex smaller than observed. Oppenheimer et al. (2018a) (and also Oppenheimer et al. 2016) already noted that issue from their comparison with the COS-Halos galaxies (see also Section 5.2), requiring additional source(s) of ionization for the O VI such as AGN flickering (Oppenheimer & Schaye 2013b; Oppenheimer et al. 2018b) or possibly CRs as shown for the FIRE-2 simulations (see Ji et al. 2020 and above). While the results are shown only to $R_{\rm vir}$, as in the other simulations and M31, O VI is also observed well beyond $R_{\rm vir}$ in the EAGLE simulations (see Figure 1 in Oppenheimer et al. 2018a).

### 5.3.3. Insights from the Observation/Simulation Comparison

The comparison with the simulations shows that the CGM is changing in zoom-in simulations on length scales roughly similar to those observed in M31. The low ions and high ions follow substantially different profiles with radius, in both data and simulations. In the zoom-in simulations described above, the inner regions of the CGM of galaxies are more directly affected by large-scale feedback and recycling processes between the disk and CGM of galaxies. Therefore, it is not surprising that the M31 CGM within 100–150 kpc shows a large variation in column density profiles with $R$, a more

complex gas-phase structure, and larger peculiar velocities even though the current SFR is low. While both accretion and large-scale outflow coexist in the CGM and are responsible for the gas flow properties, stellar feedback is required to produce a substantial amount of metals in the CGM at large impact parameters (see Figures 17–19). M31 has currently a low SFR, but it had several episodes of bursting star formation in the past (e.g., Williams et al. 2017), likely ejecting a large portion of its metals in the CGM during these episodes.

Various models simulating different galaxy masses at different epochs, with distinct SFRs or feedback processes, can reproduce at some level the diffuse O VI observed beyond $R_{\rm vir}$. All the simulations we have reviewed produce O VI profiles that are flatter than the low ions and extend to beyond $R_{\rm vir}$ with significant column density. While the galaxy halo masses are different, they are all roughly in the range of about $10^{11.5}$–$10^{12.3}\,M_\odot$, which is a mass range where their virial temperatures overlap with the temperature at which the ionization fraction of O VI peaks (Oppenheimer et al. 2016). Using the EAGLE simulations, Oppenheimer et al. (2016) show that the virial temperature of the galaxy halos can explain the presence of strong O VI in the CGM of star-forming galaxies with $M_{200} \simeq 10^{11.5}$–$10^{12.3}\,M_\odot$ and the absence of strong O VI in the CGM of quiescent galaxies that have overall higher halo masses ($M_{200} \simeq 10^{12.5}$–$10^{13.5}\,M_\odot$) and hence higher virial temperatures, i.e., halo mass, not SFR, largely drives the presence of strong O VI in the CGM of galaxies according to these simulations (see Section 5.2). Production of the O VI in volume-filling virialized gas could explain why O VI is widely spread in the CGM of simulated galaxies and the real M31. Additional ionization mechanisms from CRs (Ji et al. 2020 and see Figure 17) or fluctuating AGNs (Oppenheimer & Schaye 2013b; Oppenheimer et al. 2018b) can further boost the O VI production, but halo masses with their virial temperatures close to the temperature at which the ionization fraction of O VI peaks appear to provide a natural source for the diffuse, extended O VI in the CGM of $L^*$ galaxies. Conversely, a number of studies have shown that significant O VI can arise in active outflows, with the outflow column densities varying strongly with the degree of feedback (Hummels et al. 2013; Hafen et al. 2020). Right now, no clear observational test can distinguish O VI in warm virialized gas and direct outflows. However, any model that attempts to distinguish them will be constrained by the flat profile and low scatter seen by Project AMIGA.

On the other hand, the cooler, diffuse ionized gas probed predominantly by Si III, as well as low ions (C II, Si II) at smaller impact parameters, is not well reproduced in the simulations. In the FIRE-2 and FOGGIE simulations, the column densities of Si III and low ions within $\lesssim 0.3 R_{200}$ are reasonably matched, but their covering fractions drop sharply and much more rapidly than observed for M31 when $R > 0.3 R_{200}$. Only near satellite galaxies within $0.3 R_{200}$ do the column densities of these ions increase. This is, however, not a fair comparison, as M31 lacks gas-rich satellites within this radius. Furthermore, the near-unity covering factor of Si III out to $1.65 R_{200}$ in the CGM of M31 could not be explained by dwarf satellites anyway. For the EAGLE simulation, this problem is not as extreme as in the other simulations, but EAGLE does overproduce low and intermediate ions in the inner regions ($\lesssim 0.3 R_{200}$) of the CGM. Possibly maintaining a high resolution out to $R_{\rm vir}$ would be needed to accurately model

the small-scale structures of the cool gas content and preserve it over longer periods of times (Hummels et al. 2019; Peeples et al. 2019; van de Voort et al. 2019).

While the observations of M31 and simulations discussed above show some discrepancy, there is an overall trend that is universally observed: when the ionization energies increase from the singly ionized species (Si II, C II) to intermediate ions (Si III, Si IV) to C IV to O VI, the column density dispersions and dependence on $R$ decrease. While the larger scatter in the low and intermediate ion column densities compared to O VI was observed previously (e.g., Werk et al. 2013; Liang et al. 2016), that trend with $R$ was not as obvious owing to a larger scatter at any $R$, in part caused by neighboring galaxies or different galaxy masses (Oppenheimer et al. 2018a). This general trend is the primary point of agreement between the observations and simulations, especially considering that the simulations were not tuned to match the CGM properties. This trend most likely arises from the physical conditions of the gas: in the inner regions of the CGM the gas takes on a density that favors the production of the low and intermediate ions. At these densities O VI would need to be collisionally ionized or distributed in pockets of low-density photoionized gas. In the outer regions of the CGM, the overall gas must have a much lower density where O VI and weak Si III and nearly no singly ionized species can be produced predominantly by photo-ionization processes. This basic structure of the CGM appears in broad agreement between Project AMIGA, statistical sampling of many galaxies like COS-Halos, and three different suites of simulations.

### 5.4. Implications for the MW CGM

Based on our findings, it is likely that the MW has not only an extended hot CGM (Gupta et al. 2014, 2017; see also Henley & Shelton 2010; Miller & Bregman 2013) but also an extended CGM of cool (Si II, Si III, Si IV) and warm–hot (C IV, O VI) gas that extends all the way to about 300 kpc ($R_{vir}$), and even farther away for the O VI. In fact, the MW and M31 O VI CGMs most likely already overlap as can be seen, e.g., in the CLUES simulations of the Local Group (Nuza et al. 2014) since the distance between M31 and MW is only 752 kpc.

The multiple gas-phase MW halo has largely been studied using HVCs because the velocities of these clouds are high enough to separate them from the disk absorption (e.g., Wakker & van Woerden 1997; Putman et al. 2012; Richter et al. 2017). However, we emphasize that the large majority of HVCs, including the predominantly ionized HVCs, are not at hundreds of kiloparsecs from the MW, but most of them are within 15–20 kpc from the Sun (e.g., Wakker 2001; Thom et al. 2008; Wakker et al. 2008; Lehner & Howk 2011; Lehner et al. 2012), i.e., in a radius not even explored by Project AMIGA and many other surveys of the galaxy CGM at higher redshifts (e.g., Werk et al. 2013; Liang & Chen 2014; Borthakur et al. 2016; Burchett et al. 2016). Only the MS allows us to probe the interaction between the MW and the Magellanic Clouds in the CGM of the MW out to about 50–100 kpc (e.g., D'Onghia & Fox 2016). The results from Project AMIGA strongly suggest that the CGM of the MW is hidden in the low-velocity absorption arising from its disk (see also Zheng et al. 2015). To complicate the matter, the column densities of the low ions, intermediate ions, and C IV drop substantially beyond 100–150 kpc (see, e.g., Figures 8 and 10). Owing to its strength and little dependence on $R$, O VI is among the best

ultraviolet diagnostics of the extended CGM (see also the recent FOGGIE simulation results by Zheng et al. 2020).

### 6. Summary

With Project AMIGA, we have surveyed the CGM of a single galaxy (M31) with an unprecedented number of background targets (43) piercing it at various azimuths and impact parameters, 25 from 0.08 $R_{vir}$ to about 1.1 $R_{vir}$ and the additional 18 between $1.1 < R/R_{vir} \lesssim 1.9$. The 43 QSOs were all observed with COS G130M/G160M or G130M (providing in particular O I, C II, C IV, Si II, Si III, and Si IV), and 11 were also observed with FUSE (providing O VI). The resolution of the COS G130M/G160M and the S/Ns have been key for the success of this program. All the data were uniformly reduced and analyzed. For the 43 QSOs in our sample, we have identified all the absorption features in their spectra to determine whether any transitions used to probe the CGM of M31 could be contaminated. We provide the entire line identification in Appendix A. While we survey only a single galaxy, M31, the uniqueness of our experiment has allowed us to gain a wealth of new information that can be summarized as follows:

1. Ionized gas traced by Si III and O VI has near-unity covering factor all the way out to 1.2 $R_{vir}$ and 1.9 $R_{vir}$, respectively. All the other ions have their covering factors monotonically decreasing as $R$ increases.

2. We do not find that the properties of the CGM of M31 strongly depend on the azimuth with respect to the major and minor axes, but several properties of the CGM depend on the projected distance.

3. The gas has a more complex gas-phase structure at $R \lesssim 0.5 R_{vir}$ with high covering factors of all the ions. At larger $R$, the gas becomes more highly ionized, with a paucity of singly ionized species. Stronger absorbers are also observed closer to M31, with the column densities of all the ions but O VI decreasing sharply as $R$ increases up to $R \lesssim 0.5 R_{vir}$; beyond $R \gtrsim 0.5 R_{vir}$ the column densities decrease much more mildly with increasing $R$.

4. The velocity structure of the absorption profiles is more complex with $R \lesssim 0.5 R_{vir}$, where frequently more than one velocity component is observed, while at larger $R$ the absorption profiles predominantly show only one velocity component (at the COS resolution). The peculiar velocities of the CGM gas are also more extreme and systematically redshifted by about +90 km s$^{-1}$ relative to the bulk motion of M31 at $R \lesssim 0.5 R_{vir}$. On the other hand, at $R \gtrsim 0.5 R_{vir}$ the peculiar velocities are both blue- and redshifted relative to the bulk motion of M31 and only by 10–20 km s$^{-1}$ on average.

5. Cosmological zoom-in simulations of $\sim L^*$ galaxies (individual galaxies or galaxies in Local Group analogs) show that O VI does extend well beyond $R_{vir}$ as observed for M31. On the other hand, cosmological zoom-in simulations do not reproduce well the column density profiles of the low ions (Si II, C II) or intermediate ions (Si III, Si IV). All the zoom-in simulations explored in this work show some common traits with the observations of the CGM of M31: (i) the column densities of O VI do not vary much with $R$, while those of the lower ions have a strong dependence with $R$; (ii) the scatter in the column densities at $R$ is smaller in O VI than any lower ions; and

(iii) O VI is observed at $R \gg R_{\rm vir}$. In other words, the dispersion and the dependence of the column densities on the impact parameter decline going from singly through doubly to highly ionized species.

6. We estimate that the mass of the cool metal mass probed by Si II, Si III, and Si IV of the CGM within $R_{\rm vir}$ is $2 \times 10^7 \, M_\odot$ and by O VI is $>9 \times 10^7 \, M_\odot$. The total metal mass could be as large as $\gtrsim 2.5 \times 10^8 \, M_\odot$ if the dust and hot X-ray gas are accounted for. Since the total metal mass in the disk of M31 is about $M_Z^{\rm disk} \simeq 5 \times 10^8 \, M_\odot$, the CGM of M31 could have as much as half of the present-day metal mass of its disk and possibly much more.

7. We estimate that the baryon mass of the $\sim 10^4$–$10^{5.5}$ K gas is $\gtrsim 3.7 \times 10^{10} (Z/0.3\,Z_\odot)^{-1} \, M_\odot$ within $R_{\rm vir}$. The dependence on the largely unknown metallicity of the CGM makes the baryon mass estimate uncertain, but it is broadly comparable to other recent observational results or estimates in zoom-in simulations.

8. We study whether any of the M31 dwarf satellites could give rise to some of the observed absorption associated with the CGM of M31. We find that several absorbers are within close spatial and velocity proximity of the dwarfs and hence could be associated with the CGM of dwarfs if they have a gaseous CGM. However, these are dSph galaxies, which have had their gas stripped via ram pressure and unlikely to have much gas left in their CGM. And, indeed, none of the properties of the absorbers in close proximity to these dwarf galaxies show any peculiarity that would associate them with the CGM of these satellites rather than the CGM of M31.

9. Based on the Project AMIGA results, it is likely that the MW has similarly a cool and warm–hot ionized CGM, which may extend to similar radii for Si III ($R \sim 320$ kpc) and O VI ($R \sim 570$ kpc). The signature of the MW CGM gas must be hidden in and blended with the absorption from the MW disk since a large fraction of the high-velocity absorption (HVC) has been observed in the halo of MW at $d < 10$–20 kpc (excluding the MS). This implies that the MW and M31 CGM probed by O VI must most likely already overlap and interact with each other.

PyIGM (Prochaska et al. 2017a), yt (Turk et al. 2011), Trident (Hummels et al. 2017).

## Appendix A
## Line Identification

In Table A1, we provide the line identification for each absorption feature detected at about the $2\sigma$ level (the line list is complete at this level but can also include less significant absorption). The table is ordered by alphabetical order of the QSO name and for each QSO in order of increasing observed wavelength (second column). In this table, we define the various types of absorption features as follows (third column): "ISMLG" is any ISM/CGM/IGM absorption from the Local Group environment (mostly the MW and M31); "IGMABS" is any intervening IGM/CGM absorber at $\Delta v > 3000 \, {\rm km \, s^{-1}}$ from the QSO redshift; "PROXIMATE" is a proximate/ associated absorber at $500 < \Delta v < 3000 \, {\rm km \, s^{-1}}$ from the QSO redshift; "INTRINSIC" is an absorber at $\Delta v < 500 \, {\rm km \, s^{-1}}$ from the QSO redshift. Any "UNIDENTIFIED" feature at the $>2\sigma$ level is marked with that denomination. Finally, "OTHER" includes the known fixed-pattern noise feature ("FPN"), the special case of fixed-pattern noise feature occurring near the edge of the COS detector ("EDGE"), or the 1043 Å detector flaw in the FUSE data that causes a fake line ("FLAW"). FPN, EDGE, and FLAW all appear in the fourth column, which is otherwise used to list the atom or ion detected. The fifth column gives the rest wavelength of the atom/ion. The sixth column provides the information regarding the frame into which the velocity (sixth column) and redshift (seventh column) are defined ("L": LSR frame—any absorption at $|v_{\rm LSR}| \leqslant 700 \, {\rm km \, s^{-1}}$; otherwise, "H": heliocentric frame). Finally, the last two columns give the approximate equivalent widths ($W_\lambda$) and errors that are only provided as guidelines, i.e., these should not be used for quantitative scientific purposes since the continuum placement is only approximate. We finally note that the $H_2$ lines are not individually measured but are based on a model of the $H_2$ absorption (see Wakker 2006), which is the reason for not providing an error on $W_\lambda$.

The process for identifying the absorption lines in the COS spectra is reviewed in Section 2.3. As discussed in this section, some of the QSOs do not have the full FUV wavelength coverage, or their redshifts put Ly$\alpha$ beyond the observed wavelength. There are seven such cases that are reviewed below:

1. 3C 454.3: With $z_{\rm em} = 0.859$, the highest-redshift Ly$\alpha$ absorber would be at 2259 Å. However, there are also COS G225M and FOS G190H/G270H data that help disentangle any possible Ly$\beta$ from Ly$\alpha$. In the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150 \, {\rm km \, s^{-1}}$, there is detection of absorption in several ions, with all the velocity profiles being consistent, suggesting no contamination in the surveyed velocity range $-700 \, {\rm km \, s^{-1}} \leqslant v_{\rm LSR} \leqslant -150 \, {\rm km \, s^{-1}}$.

2. PG 0044+030: With $z_{\rm em} = 0.623$, the highest-redshift Ly$\alpha$ absorber would be at 1973 Å. The FOS G190H/ G270H data help securely identify Ly$\beta$ above 1347 Å in the G130M spectrum. Many lines between 1215 and 1347 Å are clearly identified as higher-redshift Lyman series and metal lines, leaving just five absorption features identified as Ly$\alpha$, which is about the expected number of absorbers given the S/N in the COS spectrum

**Table A1**
Line List Identification

| Target | $\lambda_{\rm obs}$ (Å) | Type | Ion | $\lambda_{\rm rest}$ (Å) | Frame | $cz$ (km s$^{-1}$) | $z$ | $W_\lambda$ (mÅ) | $\sigma_{W_\lambda}$ (mÅ) |
|---|---|---|---|---|---|---|---|---|---|
| 3C 454.3 | 1144.95 | ISMLG | Fe II | 1144.9 | L | 2.0 | −0.00002 | 309.4 | 96.6 |
| 3C 454.3 | 1146.29 | IGMABS | H I | 1025.7 | H | 35240.0 | 0.11758 | 260.5 | 71.2 |
| 3C 454.3 | 1152.82 | ISMLG | P II | 1152.8 | L | −1.0 | −0.00002 | 166.3 | 38.3 |
| 3C 454.3 | 1153.25 | IGMABS | O VI | 1031.9 | H | 35246.0 | 0.11758 | 155.6 | 30.3 |
| 3C 454.3 | 1158.39 | IGMABS | O III | 832.9 | H | 117142.0 | 0.39074 | 142.0 | 28.6 |
| 3C 454.3 | 1159.58 | IGMABS | O VI | 1037.6 | H | 35239.0 | 0.11758 | 93.0 | 44.3 |
| 3C 454.3 | 1182.72 | IGMABS | H I | 1025.7 | H | 45885.0 | 0.15307 | 63.0 | 27.3 |
| 3C 454.3 | 1183.43 | IGMABS | H I | 1025.7 | H | 46094.0 | 0.15376 | 548.3 | 24.0 |
| 3C 454.3 | 1183.77 | IGMABS | H I | 1025.7 | H | 46192.0 | 0.15408 | 296.1 | 22.1 |
| 3C 454.3 | 1188.85 | ISMLG | Si II | 1190.4 | L | −393.0 | −0.00127 | 121.2 | 22.0 |
| 3C 454.3 | 1189.16 | ISMLG | Si II | 1190.4 | L | −317.0 | −0.00106 | 140.3 | 26.0 |
| 3C 454.3 | 1190.41 | ISMLG | Si II | 1190.4 | L | −3.0 | −0.00002 | 412.4 | 22.7 |

**Note.** Definition of the identified types of absorption features: ISMLG: ISM/CGM/IGM absorption from the Local Group environment (mostly the MW and M31); IGMABS: intervening IGM/CGM absorber at $\Delta v > 3000$ km s$^{-1}$ from the QSO redshift; PROXIMATE: proximate/associated absorber at $500 < \Delta v < 3000$ km s$^{-1}$ from the QSO redshift; INTRINSIC: intrinsic absorber at $\Delta v < 500$ km s$^{-1}$ from the QSO redshift; UNIDENTIFIED: unknown origin of the absorption; OTHER: FPN: fixed-pattern noise feature; OTHER: EDGE: special case of fixed-pattern noise feature occurring near the edge of the COS detector; OTHER: FLAW: refers to the 1043 Å detector flaw in the FUSE data that causes a fake line. In the frame column, "L" stands for LSR frame (any absorption at $|v_{\rm LSR}| \leqslant 500$ km s$^{-1}$), "H" for heliocentric frame (any absorption at $|v_{\rm LSR}| > 500$ km s$^{-1}$). Note that the equivalent widths ($W_\lambda$) and errors are provided for guidelines and should not be used for quantitative scientific purposes. The H$_2$ lines are not individually measured but are based on a model of the H$_2$ absorption, which is the reason for not providing an error on $W_\lambda$ (i.e., in the table it is set to "0.0").

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

of this target. In the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, there is detection in absorption of several ions, with all the velocity profiles and central velocities being consistent with each other, suggesting no contamination in the surveyed velocity range.

3. PHL1226: This target has only COS G130M observations, and since $z_{\rm em} = 0.404$, the highest-redshift Ly$\alpha$ absorber would be at 1705 Å. There are 11 lines that are listed as Ly$\alpha$ in Table A1, which might be Ly$\beta$. However, some are very unlikely Ly$\beta$, as they are so strong that Ly$\gamma$ and/or metal lines would be expected to be detected and are not. In all the ions but Si III there is no detection in the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$. The absorption feature at $-332$ km s$^{-1}$ from Si III $\lambda 1206$ is clearly identified as O I $\lambda 1039$ associated with the super Lyman limit system at $z = 0.15974$, and therefore there is no detection of M31 CGM gas toward this sightline.

4. RX J0023.5+1547: This target has also only COS G130M observations, and since $z_{\rm em} = 0.412$, the highest-redshift Ly$\alpha$ absorber would be at 1716 Å. There are seven lines that are listed as Ly$\alpha$ in Table A1 which could be Ly$\beta$. Given the S/N of the COS spectrum, the expected number of Ly$\alpha$ lines between 1215 and 1460 Å is 15, while we identify 13. On the other hand, the expected number of Ly$\beta$ is 4, and only one is identified. This is well within the possible cosmic variance, but it is possible that two to three lines that are identified as Ly$\alpha$ could actually be Ly$\beta$. In the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, only Si III is detected; there is no other likely origin for this absorption feature.

5. RX J0028.1+3103: This target has also COS G130M and G160M observations, but since $z_{\rm em} = 0.500$, the highest-redshift Ly$\alpha$ absorber would be at 1823 Å. Only one

possible Ly$\beta$ identified as Ly$\alpha$ is above the spectral limit of 1792 Å. In the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, two components are detected, with one observed in C II, Si III, and Si IV, and a weaker one only in Si III. Based on our line identification, there is no other likely origin for the weaker Si III-identified absorption feature.

6. SDSS J015952.95+134554.3: This target has also only COS G130M observations, and since $z_{\rm em} = 0.504$, the highest-redshift Ly$\alpha$ absorber would be at 1828 Å. Given the S/N of the COS spectrum, we expect about 21 Ly$\alpha$ and 7 Ly$\beta$, while 19 and 2 are listed in Table A1, respectively. So it is possible that two to three lines identified as Ly$\alpha$ could actually be Ly$\beta$. There is no detection of Si III (or other ions) in the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, and hence no contamination issue.

7. SDSS J225738.20+134045.0: This target has COS G130M and G160M observations, but since $z_{\rm em} = 0.595$, the highest-redshift Ly$\alpha$ absorber would be at 1938 Å. There are three Ly$\alpha$ with no corresponding Ly$\beta$ that possibly could be Ly$\beta$, but it is very unlikely to be the case for the three identified Ly$\alpha$. In the velocity range $-700 \leqslant v_{\rm LSR} \leqslant -150$ km s$^{-1}$, there is detection of absorption in several ions, with all the velocity profiles and central velocities being consistent with each other, suggesting no contamination in the surveyed velocity range.

## Appendix B
## Visually Identifying Contamination

In order to visually assess possible line contamination, as well as conflicts in independent line identifications, we have designed the "conflict plot" that is shown in Figure B1. In this

**Figure B1.** Example of conflict plot used to determine potential contamination in the velocity range $-700 \lesssim v_{LSR} \lesssim -150$ km s$^{-1}$. Locations on the x-axis are scaled according to log $\lambda_{obs}$, but the markers show the wavelength. The y-axis scale is proportional to log $(1 + z)$, but the markers show z. The diagonal lines represent loci of different transitions and how their observed wavelengths change with z. The gray lines correspond to metal transitions, while the green ones indicate H$_2$ transitions. Lyman series lines of H I are shown in blue. Some of the transitions are identified near the top and on the sides of the plot. Absorption features identified in the UV spectrum are shown with gray filled circles, with varying sizes proportional to $\sqrt{W_\lambda}$. The red horizontal lines represent redshift systems with any line (usually Ly$\alpha$) that has $W_\lambda \geqslant 100$ mÅ; systems with Ly$\alpha$ having strengths below this threshold do not have lines that stretch across the whole plot, but they do have short line segments that cover just Ly$\beta$ and Ly$\gamma$. Most of the lines of interest appear along the yellow and orange lines, which represent a blueshift of $-500$ and $-150$ km s$^{-1}$, respectively. Potential conflicts arise at the intersection of the horizontal red line, a gray/green/blue sloping line, and the vertical orange/yellow line.

plot, the locations on the x-axis are scaled according to log $\lambda_{obs}$ and the y-axis scale is proportional to log$(1 + z)$. The diagonal lines represent loci of different transitions and how their observed wavelengths change with z. The diagonal lines are color-coded depending on whether they trace metal transitions (gray), H$_2$ (green), or H I (blue). We identify some of the transitions at the top and on the sides of the plot. Any absorption feature that is identified in the COS spectrum is identified by a gray filled circle. Although not shown here, an independent line identification would be represented with a different symbol, immediately identifying the similarities and differences between two independent line identifications. The red horizontal lines represent redshift systems with any line (usually Ly$\alpha$) that has $W_\lambda \geqslant 100$ mÅ; systems with Ly$\alpha$ having strengths below this threshold do not have lines that stretch across the whole plot, but they can have short line segments that cover just Ly$\beta$ and Ly$\gamma$. Some of the common interstellar lines appear along the yellow and orange lines that represent a blueshift of $-500$ and $-150$ km s$^{-1}$, respectively (i.e., within the velocity range where we observe M31 CGM gas). For Project AMIGA, we mostly concentrate on transitions at $\lambda > 1145$ Å that are available in the COS G130M and G160M wavelength bandpass. However, if there are FUSE data, which provide wavelength coverage $\lambda < 1145$ Å, we also use those.

In this figure, potential conflicts are readily detected at the intersection of the horizontal red line, a gray/green/blue sloping line, and the vertical orange/yellow line. If, for one of the lines of interest, there is a gray circle at $z > 0.01$ at the one

of the intersection, there is a potential contamination in the velocity range $-700 \leqslant v_{LSR} \leqslant -150$ km s$^{-1}$ that needs to be checked. For each sightline in our sample, we have produced these conflict plots to easily identify any potential contamination of the absorption in the velocity range $-700 \leqslant v_{LSR} \leqslant -150$ km s$^{-1}$.

For the specific example shown in Figure B1, there are two potential conflicts for the components identified in the velocity range $-700 \leqslant v_{LSR} \leqslant -150$ km s$^{-1}$. One is near the N I $\lambda1134.1$ line, where an O VI absorber at $z = 0.09753$ is present at 1132.58 Å (shifted by $-400$ km s$^{-1}$ relative to N I; see Table A1). However, there is no contamination near N I $\lambda1199.1$ in the velocity range $-700 \leqslant v_{LSR} \leqslant -150$ km s$^{-1}$, so we know that there is no N I absorption. The second conflict is for C II $\lambda1334$, where Ly$\alpha$ at $z = 0.09723$ contaminates the $-140$ km s$^{-1}$ C II component. Although there are several weaker Lyman series transitions that could have been used to correct for that contamination (which we did in some other cases), in this case we have also FUSE observations that provide C II $\lambda1036$ where the absorption is not contaminated in this component.

**Appendix C**
**Comparison between COS G130M/G160M and STIS E140M Spectra**

For three targets (Mrk 335, UGC 12163, and NGC 7469) in our sample we have higher-resolution STIS E140M spectra. Using the same integration velocity ranges that we used for

**Table C1**
Comparison between COS and STIS

| Ion/Instrument | $[v_1, v_2]$ (km s$^{-1}$) | $v$ (km s$^{-1}$) | $\log N$ [cm$^{-2}$] |
|---|---|---|---|
| | Mrk 335 | | |
| C II λ1334 COS | −450, −372 | −406.9 ± 3.7 | 13.22 ± 0.08 |
| C II λ1334 STIS | −450, −372 | −403.2 ± 4.7 | 13.34 ± 0.11 |
| C II λ1334 COS | −372, −310 | −334.6 ± 2.1 | 13.34 ± 0.06 |
| C II λ1334 STIS | −372, −310 | −338.8 ± 2.0 | 13.52 ± 0.06 |
| Si III λ1206 COS | −450, −372 | −409.7 ± 1.8 | 12.49 ± 0.04 |
| Si III λ1206 STIS | −450, −372 | −406.5 ± 5.3 | 12.51 ± 0.13 |
| Si III λ1206 COS | −372, −310 | −339.9 ± 0.8 | 12.75 ± 0.02 |
| Si III λ1206 STIS | −372, −310 | −337.8 ± 2.2 | 12.82 ± 0.07 |
| Si III λ1206 COS | −310, −273 | −296.6 ± 0.8 | 12.40 ± 0.04 |
| Si III λ1206 STIS | −310, −273 | −301.8 ± 2.2 | 12.64 ± 0.17 |
| Si III λ1206 COS | −273, −190 | −246.9 ± 3.7 | 12.28 ± 0.06 |
| Si III λ1206 STIS | −273, −190 | ⋯ | <12.2 |
| | UGC 12163 | | |
| C II λ1334 COS | −475, −375 | −423.3 ± 3.6 | >13.94 |
| C II λ1334 STIS | −475, −375 | −425.2 ± 5.0 | 14.05 ± 0.13 |
| C II λ1334 COS | −375, −310 | −353.0 ± 7.3 | 13.35 ± 0.20 |
| C II λ1334 STIS | −375, −310 | ⋯ | <13.28 |
| C II λ1334 COS | −220, −180 | −196.6 ± 3.1 | 13.26 ± 0.18 |
| C II λ1334 STIS | −220, −180 | −200.8 ± 3.6 | $13.37^{+0.16}_{-0.26}$ |
| Si III λ1206 COS | −475, −375 | −426.0 ± 2.3 | >13.30 |
| Si III λ1206 STIS | −475, −375 | −425.5 ± 21.4 | >13.45 |
| Si III λ1206 COS | −375, −310 | −349.8 ± 6.2 | 12.43 ± 0.18 |
| Si III λ1206 STIS | −375, −310 | ⋯ | <12.57 |
| | NGC 7469 | | |
| C II λ1334 COS | −400, −268 | −335.3 ± 0.7 | >14.27 |
| C II λ1334 STIS | −400, −268 | −337.6 ± 0.8 | >14.44 |
| C II λ1334 COS | −268, −210 | −251.4 ± 3.8 | 13.06 ± 0.09 |
| C II λ1334 STIS | −268, −210 | −251.1 ± 3.3 | 13.10 ± 0.08 |
| C II λ1334 COS | −202, −150 | −176.2 ± 3.4 | 12.94 ± 0.11 |
| C II λ1334 STIS | −202, −150 | −184.9 ± 6.8 | $12.66^{+0.14}_{-0.20}$ |
| Si III λ1206 COS | −400, −268 | −325.2 ± 0.5 | >13.57 |
| Si III λ1206 STIS | −400, −268 | −332.7 ± 5.9 | >13.76 |
| Si III λ1206 COS | −268, −210 | −246.9 ± 0.9 | 12.63 ± 0.02 |
| Si III λ1206 STIS | −268, −210 | −250.1 ± 1.7 | 12.76 ± 0.16 |
| Si III λ1206 COS | −202, −150 | −178.8 ± 1.3 | 12.32 ± 0.04 |
| Si III λ1206 STIS | −202, −150 | −174.8 ± 1.2 | 12.58 ± 0.04 |

**Note.** COS stands here for COS G130M and STIS for STIS E140M. The S/Ns (per resolution element) near C II and Si III are, respectively, Mrk 335: 36.6, 32.2 (COS), 9.5, 4.8 (STIS); UGC 12163: 10.7, 10.8 (COS), 5.1, 2.3 (STIS); NGC 7469: 37.5, 35.4 (COS), 16.5, 9.0 (STIS).

COS (which is justified since the spectra were initially all aligned), we estimate the velocities and column densities in the STIS E140M spectra. The results are summarized in Table C1. In the footnote of this table we also list the S/Ns in the continuum near C II and Si III since not only resolution but also S/Ns can explain some of the differences. The STIS data have systematically lower S/Ns than the COS spectra. Fox et al. (2005) show that in low-S/N STIS E140M spectra (four to nine per resolution element), the AOD method can over-estimate the apparent column densities by a factor of 0.1–0.4 dex, especially when the absorption is weak. As for the COS spectra, we use the original binning sampling of the data to estimate the column densities. Binning by 2 or 3 pixels the STIS spectra did not change the results in contrast to the study of Fox et al. (2005), but in this study the simulated spectra were affected only by Poisson noise, while the STIS E140M are affected by both Poisson and fixed-pattern noises.

As illustrated in Figure C1 (see also normalized profiles in the figure set in the online Journal) with the spectra of Mrk 335, where we show for C II and Si III the COS and STIS spectra, more components can be revealed in the higher-resolution spectrum and the components appear sharper and deeper in the higher-resolution spectrum. However, the STIS spectrum is also much noisier. For Mrk 335, the absorption in all the components is weak, with a peak absorption depth at the 20% level in C II and 30%–40% in Si III. Within about the $1\sigma$–$2\sigma$ errors the column densities between COS and STIS are in agreement even though additional components are revealed in the STIS spectrum. The column densities derived from the STIS spectrum are systematically higher, but this effect is consistent with the lower S/Ns in the STIS data that were observed by Fox et al. (2005). For UGC 12163, with a peak absorption depth at 60% in C II and 90% in Si III in the component at $-425$ km s$^{-1}$ in the COS spectrum (see the online figure set), the absorption in these two transitions is marked as saturated. The STIS Si III reaches zero-flux level, confirming complete saturation in Si III. The STIS apparent column density of C II is 0.09 dex higher than the estimated lower limit from the COS spectrum, but in agreement within the $1\sigma$ error, implying that the adopted peak optical depth of $\tau_a > 0.9$ is about right for saturation in the COS spectra (see Section 2.4.2). The other components toward UGC 12163 are weak, and the STIS upper limits are in agreement with the COS detection. Toward NGC 7469, the most negative absorption is again the strongest component, and both COS C II and Si III were correctly identified as saturated (even though again these do not reach zero-flux levels in the COS spectrum, while they do in the STIS spectrum). The other two components in the spectra of NGC 7469 are very weak. For Si III, the S/N effect is observed with the STIS spectrum having 0.1–0.2 dex higher than the column densities derived in the very high S/N COS spectrum. On the other hand, the S/N near C II is higher in the STIS spectrum, and for the component at $-251$ km s$^{-1}$, the column densities derived from the COS and STIS spectra are in excellent agreement (see Table C1).
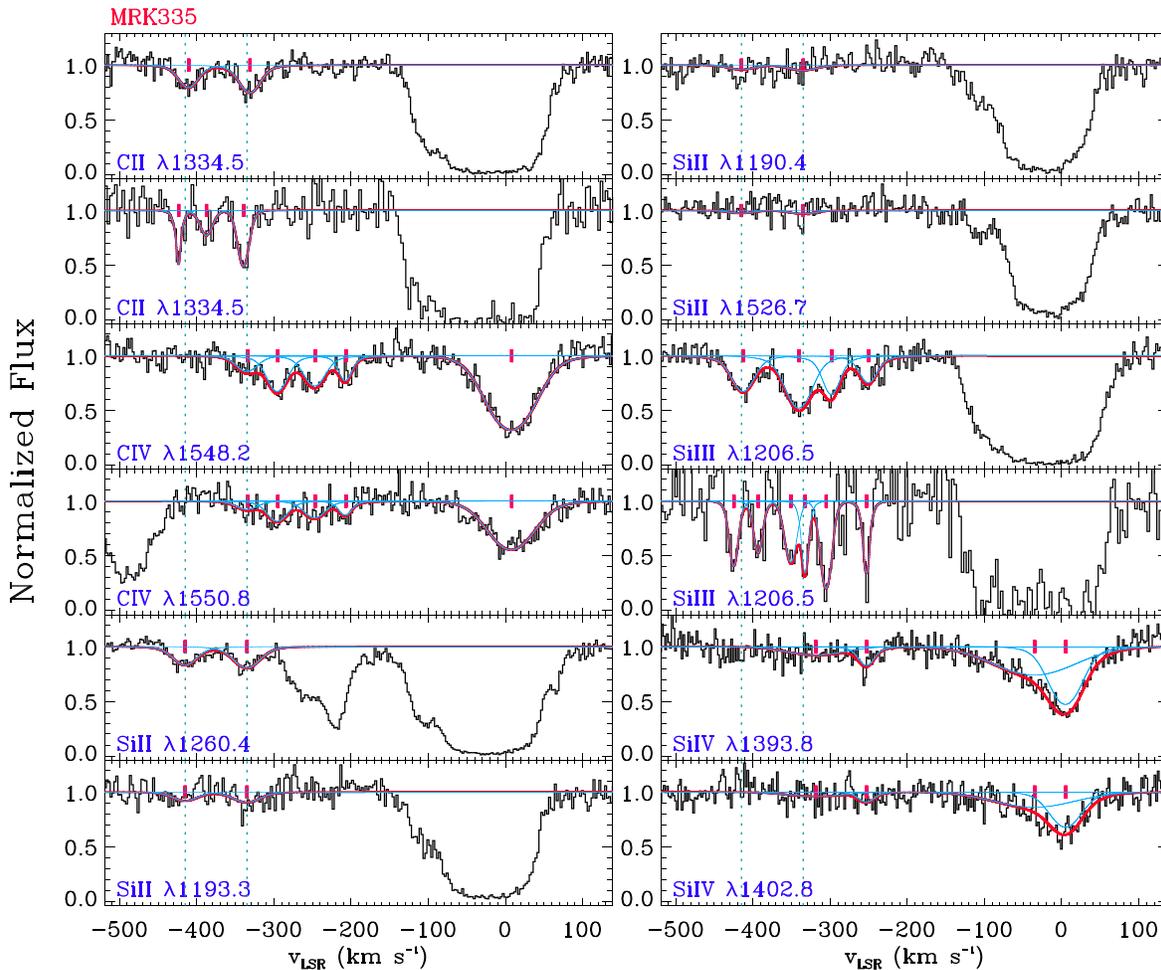
**Figure C1.** Example of normalized absorption lines as a function of the LSR velocity toward Mrk 335 with a Voigt PF model to the data. The red tick marks show the velocity centroids. In each panel, the red line shows the resulting PF, while the blue line shows the individual components. The green vertical dotted lines show the velocity centroids of Si II. When the same ions appear twice (here C II and Si III), the top and bottom panels show the COS and STIS data, respectively. Note that the higher resolution of STIS E140M shows additional components in C II and Si III, but the low S/N near Si III makes the results very uncertain. The complete figure set (6 images) is available in the online journal.

(The complete figure set (6 images) is available.)

While the sample with both STIS and COS spectra is small, the comparison shows that (1) there is overall a good agreement in the column densities derived from the STIS and COS data and (2) our conservative choice of $\tau_a \sim 0.9$ as the threshold for saturation in the COS data is adequate.

## Appendix D
## Confronting the AOD Results with a Line Profile Fitting Analysis

For the most blended profiles (six targets in our sample), we also use a Voigt PF analysis to separate the absorption profiles into individual components with the goal to assess differences in column density estimates between the PF and AOD methods. With the PF method, we model the absorption profile as individual components using a modified version of the software described in Fitzpatrick & Spitzer (1997; see also Lehner et al. 2011 for the updates). The best-fit values describing the gas are determined by comparing the model profiles convolved with the COS G130M or G160M (and STIS E140M when available) instrumental line-spread function (LSF) of the data. The COS and STIS LSFs are not purely

Gaussian, and we adopt the tabulated COS LSFs from the COS and STIS instrument handbooks (Fischer et al. 2019; Riley et al. 2019). As the COS LSFs vary with the FUV lifetime positions, we use the COS LSFs at the appropriate lifetime positions.[34] Three parameters $N_i$, $b_i$, and $v_i$ for each component, $i$, are input as initial guesses and were subsequently varied to minimize $\chi^2$. The fitting process enables us to find the best fit of the component structure using the data from one or more transitions of the same ionic species simultaneously. However, all the ions are fitted independently (i.e., we did not assume a common component structure for all the ions a priori). When STIS E140M data are available, we also fit the COS and STIS independently to assess how different these are (see also Appendix C). We apply this method to C II, C IV, Si II, Si III, Si IV, and O VI if it is available.

We always start each fit with the smallest number of components that reasonably model the profiles and add more

---

[34] For targets obtained at different lifetime positions, we adopt the one with the longest exposure time, but we note that the results would not change quantitatively if we adopted another lifetime position (less than 0.02–0.04 dex on the column densities).

**Table D1**
Summary of the Profile Fit Results

| Target | Ion | Comp. | $v$ (km s$^{-1}$) | $\sigma_v$ (km s$^{-1}$) | $b$ (km s$^{-1}$) | $\sigma_b$ (km s$^{-1}$) | $\log N$ [cm$^{-2}$] | $\sigma_{\log N}$ |
|---|---|---|---|---|---|---|---|---|
| RX J0048.3+3941 | C II | 1 | −471.4 | 4.3 | 13.8 | 7.7 | 12.90 | 0.13 |
| RX J0048.3+3941 | C II | 2 | −418.9 | 4.5 | 17.5 | 7.6 | 13.21 | 0.13 |
| RX J0048.3+3941 | C II | 3 | −373.8 | 0.9 | 17.3 | 1.7 | 14.07 | 0.02 |
| RX J0048.3+3941 | C II | 4 | −329.1 | 3.2 | 14.6 | 5.1 | 13.21 | 0.10 |
| RX J0048.3+3941 | C II | 5 | −244.3 | 0.8 | 17.3 | 1.3 | 13.85 | 0.02 |
| RX J0048.3+3941 | C II | 6 | −177.3 | 1.0 | 20.0 | 1.3 | 14.25 | 0.03 |
| RX J0048.3+3941 | C IV | 1 | −386.7 | 4.2 | 39.2 | 6.3 | 13.25 | 0.05 |
| RX J0048.3+3941 | C IV | 2 | −239.0 | 0.6 | 21.7 | 0.9 | 14.07 | 0.01 |
| RX J0048.3+3941 | C IV | 3 | −183.6 | 0.8 | 11.7 | 1.3 | 13.59 | 0.02 |
| RX J0048.3+3941 | C IV | 4 | −30.5 | 4.6 | 40.0 | 3.7 | 13.62 | 0.07 |
| RX J0048.3+3941 | C IV | 5 | −8.8 | 2.6 | 11.7 | 6.0 | 13.09 | 0.20 |
| RX J0048.3+3941 | Si II | 1 | −374.7 | 1.3 | 15.4 | 2.0 | 12.83 | 0.04 |
| RX J0048.3+3941 | Si II | 2 | −323.1 | 5.2 | 28.5 | 10.3 | 12.49 | 0.10 |
| RX J0048.3+3941 | Si II | 3 | −251.8 | 1.8 | 13.1 | 3.2 | 12.79 | 0.04 |

**Note.** Errors $v$, $b$, and $N$ are $1\sigma$ errors.

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

components as needed. We do not fix any of the input parameters, i.e., each input parameter is allowed to freely vary. This procedure is repeated for each profile until the best fit is achieved. We finally bear in mind that even though the $\chi^2$ goodness of fit may be good, the PF method may still not assess correctly the saturation level, especially since the COS resolution is relatively crude for PF. In a low-S/N or complicated profile, a broad component may also be fitted principally to reduce the $\chi^2$, while several narrower components could be more adequate and physically more appropriate (see also Lehner et al. 2011). This is a limitation of the PF, especially when the spectral resolution is only ∼17 km s$^{-1}$ and/or the S/N is low. For C IV and Si IV we fit all the components, including the MW low-velocity components since those are not saturated, while for the other ions we only fit the high-velocity components since the low-velocity components are saturated.

The results from component fitting are provided in Table D1. In Figure C1, we show an example of PF where both COS and STIS observations are available (in the online figure set, we show the PFs for the six targets). Considering first the COS data only, although C II and Si II are fitted independently, the velocity centroids of C II and Si II match each other. The two components seen in C II and Si II are also observed in Si III at the same velocities. However, only the component at −330 km s$^{-1}$ is observed in C IV (and possibly Si IV). Two additional components are observed in Si III but not in C II or Si II. These additional components seen in Si III are also observed in C IV, but shifted in velocity; an additional component is also observed at −227 km s$^{-1}$ in only C IV. This demonstrates a clear change in the gas properties with velocity. The STIS E140M observations show additional components, but the results are far more uncertain (especially in the narrow components) owing to the lower S/N of the STIS data (for STIS column densities with less than 0.15 dex, those are quite consistent with the COS-derived column densities). Much of these conclusions regarding the velocity structures and how the ionization properties change in the different components can be, however, drawn from the AOD analysis by comparing the velocity profiles and derived averaged velocities in each determined component.

The six targets that we model with Voigt PFs have the most complex velocity structure in our sample, and therefore a comparison with the AOD results provides a way to assess how similar or dissimilar are the results between the AOD and PF analyses in the worst-case scenarios. For C II, C IV, Si II, Si III, and Si IV, we use the information in Tables 2 and D1 to match the component. For 78% (69/88) of the components, we can match them directly. For another eight, we co-add two components in the PF or AOD to match the AOD or PF results, and therefore in total we have 88% matched AOD and PF components. For 3% (3/88) and 9% (8/88) of the components, the PF fits yield extremely narrow ($b < 4$ km s$^{-1}$) or broad ($b \gtrsim 45$ km s$^{-1}$) components, respectively; in these cases, the results are deemed uncertain because they appear in low-S/N spectra and/or complicated profiles (see above). In particular, the majority of broad components (7/8) appear in Si III and Si IV, with no counterpart broad components in C IV.

For the matched components, the PF and AOD velocity centroids are in good agreement, with a difference on average of 0.4–2.8 km s$^{-1}$ (and dispersion around the mean of 3–6 km s$^{-1}$) depending on the ions. In Figure D1, we show the comparison of the column densities for the individual components derived from the PF (y-axis) and AOD (x-axis) methods. Within $1\sigma$ error, the majority of the data are within the 1:1 relationship. There is a slight systematic since $\langle \log N_{\rm PF} - \log N_{\rm AOD} \rangle \simeq +0.06 \pm 0.10$ for C II, Si II, C IV, and S IV; for Si III that difference is somewhat larger with $+0.09 \pm 0.10$ dex, but within $1\sigma$ dispersion in agreement with the 1:1 relationship. This systematic can be understood as follows: (1) in blended absorption some extra absorption can be present in the wings of the profiles that is taken into account in the PF but not necessarily in the AOD method (this effect is more important for weak absorption features); (2) the width in the AOD integration is fixed, while in the PF it is a free parameter, which can increase the column if the width is larger than used in the AOD. Another systematic observed from Figure D1 is that the errors in the PF method are on average about $+0.07 \pm 0.15$ dex larger than those of the AOD. Several effects can explain this systematic: (1) broad shallow components can arise in PF but not in AOD (e.g., for Si IV,
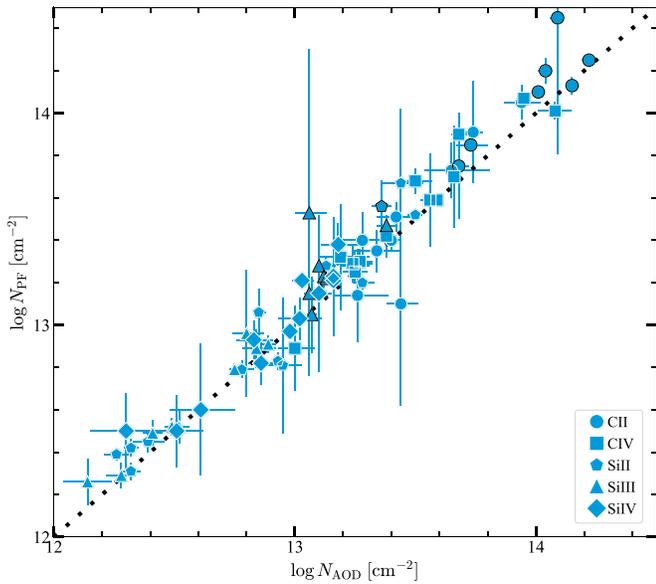
**Figure D1.** Comparison of the column densities derived using the AOD and PF methods for the matched components. Black outlined symbols indicate that the absorption is flagged as saturated using the AOD method.

removing any components with $b > 40 \, \mathrm{km \, s^{-1}}$ from the sample would change the error difference from 0.08 to 0.04 dex); (2) PF of saturated components (removing Si III saturated components from the sample would change the systematic from 0.08 to 0.04 dex—as noted above, with the AOD we, however, consider saturated components as lower limits only, which is likely to be the case also for the PF results); (3) in the PF, all the components influence the error in each component (typically relatively well separated and not too shallow absorbing components lead to similar errors as, e.g., C II, Si II, and Si III shown in Figure C1; see Table D1).

In conclusion, while there are some systematic differences between the AOD- and PF-derived column densities, those are on average small (less than 15%). Furthermore, and importantly, the targets considered in this section have the most complicated blending of components in our sample. Since a great part of this small systematic arises owing to the profiles being heavily blended, a majority of our sample is not affected by those.

## Appendix E
## M33 in the CGM of M31

M33 is separated from M31 by about 190 kpc (see Figure 1) and is the third most massive galaxy in the Local Group, but still has a mass about 20 times lower than M31 (Corbelli 2003). It is considered a dwarf spiral galaxy, but its stellar mass of $(3–6) \times 10^9 \, M_\odot$ (Corbelli 2003) is at least 10 times larger than the next two most massive satellites (M32 and NGC 205) of M31 (see Table 3), making M33 quite unique. Kam et al. (2017) show that the halo mass could be as large as $5.2 \times 10^{11} \, M_\odot$ within a virial radius of 168 kpc, but this would imply a very low baryonic mass fraction, suggesting a more plausible M33 virial radius (and hence halo mass) that is much smaller.

While M33 appears quite unique as a dwarf spiral galaxy, there are two main reasons that the CGM of M33 is unlikely to affect much the observed absorption observed toward the
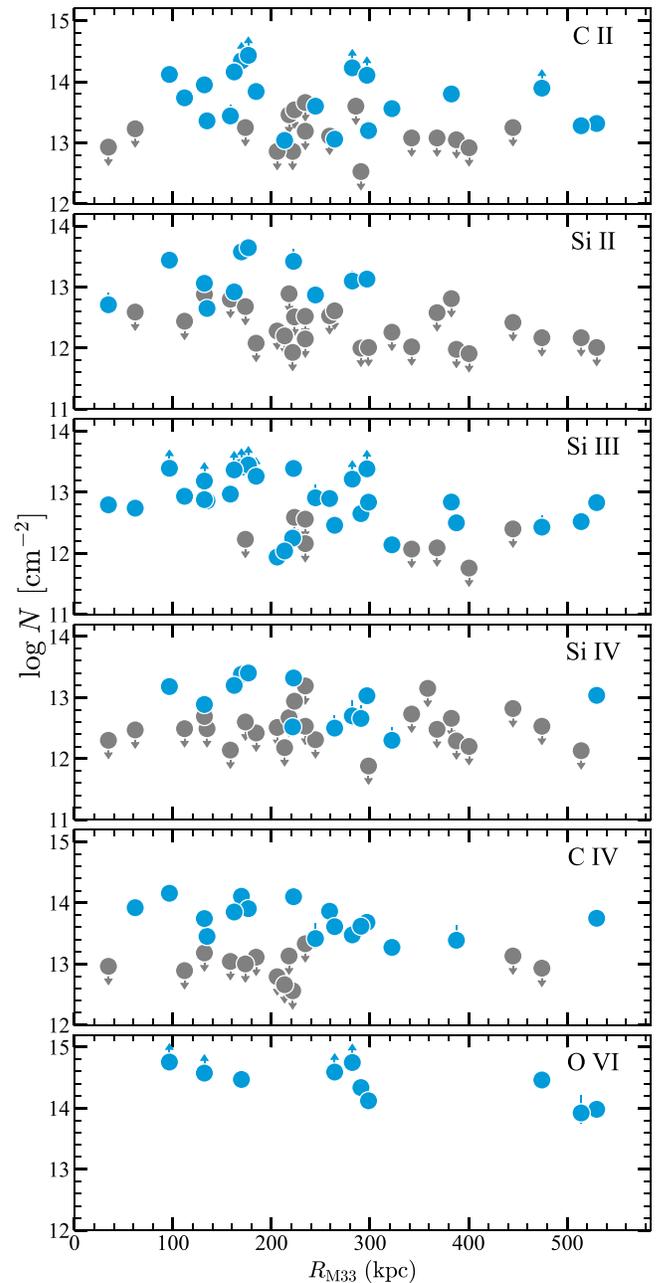


**Figure E1.** Logarithm of the total column densities of the detected ions as a function of $R_{M33}$ (the projected distance between the QSO sightline and M33). Blue circles are detections, while gray circles with downward-pointing arrows are nondetections. A blue circle with an upward-pointing arrow denotes that the absorption is saturated, resulting in a lower limit.

QSOs in our sample. First, the systemic velocity of M33 is $-180 \, \mathrm{km \, s^{-1}}$ and the rotation velocity range is from $-300$ to $-75 \, \mathrm{km \, s^{-1}}$. Therefore, a large fraction of the M33 CGM absorption may actually be lost in the MW HVC and disk absorption ($v_{LSR} > -150 \, \mathrm{km \, s^{-1}}$). Second, checking the column density maps shown in Figure 13, there is no apparent trend between $N$ and the projected distance from M33, and, furthermore the two closest sightlines to M33 show a lack of strong absorption from singly ionized species. To show explicitly this lack of trend, we plot in Figure E1 the column densities of the various ions in our sample as a function of the

projected distance from M33 ($R_{M33}$). Contrary to Figure 8, there is no trend between $N$ and $R_{M33}$. Furthermore, nondetections and detections are found at any projected distances from M33. All these strongly suggest that the CGM of M33 does not contribute significantly to the observed absorption associated with the CGM of M31.

## Appendix F
## Modeling the Column Density of Si as a Function of $R$

To model the functional form of $N_{Si}$ with $R$ (see Section 4.5), we consider three models: a hyperbolic (H) model, a single power-law (SPL) model, and a Gaussian Process (GP) model. Since our sample contains both lower and upper limits, we first need to determine how to treat censored points in the fit. Beyond 360 kpc, 8/10 data are upper limits, but there is only 1 upper limit at $R < 360$ kpc. Upper limits correspond to nondetection of any Si ions, and hence the value of $N_{Si}$ can only decrease. We have assessed the effect of these upper limits on the fit by refitting with these values decreased by a factor of 2, 4, and 10; the overall effect on the fits is small. We adopt a factor of 10 decrease for all the models. The lower limits are mostly observed at $R \lesssim 140$ kpc (only two are observed at $R > 140$ kpc; see Figure 10). As discussed in Section 2.4.2, none of the absorption of Si III associated with the CGM of M31 reaches the zero-flux level in the COS spectra. Based on our discussion in Section 2.4.2, we estimate that the lower limits at $R > 50$ kpc should be increased by 0.1 dex, while for the two inner lower limits the increase is larger, and likely as large as 0.3–0.6 dex; we consider these two extrema in the fit. Finally, for data with vertical ticked bars in Figure 10, we take the mean value between the high and low values.

For the H model, we first use the smaller correction on the inner region lower limits and consider the two populations of data below and above $R_{th} = 130$ kpc and fit each population with a linear fit. We note that $R_{th}$ can vary between 90 and 130 kpc without varying the fit results. We find $N_{Si}/(10^{14}\,\text{cm}^{-2}) = -1.76 \times 10^{-2} R + 1.80$ at $R_{th} < 130$ kpc and $N_{Si}/(10^{14}\,\text{cm}^{-2}) = -3.25 \times 10^{-4} R + 0.204$ at $R_{th} \geqslant 130$ kpc. The two lines intersect at $R_0 = 92$ kpc and $N_{Si}^0 = 2.0 \times 10^{13}$ cm$^{-2}$. If stronger correction on the lower limits is applied, then we find $N_{Si}/(10^{14}\,\text{cm}^{-2}) = -3.43 \times 10^{-2} R + 3.26$ at $R_{th} < 130$ kpc and $N_{Si}/(10^{14}\,\text{cm}^{-2}) = -3.24 \times 10^{-4} R + 0.204$ at $R_{th} \geqslant 130$ kpc. The two lines intersect at $R_0 = 90$ kpc and $N_{Si}^0 = 1.7 \times 10^{13}$ cm$^{-2}$. As expected, the effect of the different corrections on the lower limits is much stronger on the fit at $R < 130$ kpc. The two regimes can then be modeled with a single hyperbola (e.g., Watts & Bacon 1974):

$$N_{Si}/(10^{14}\,\text{cm}^{-2}) = N_{Si}^0/(10^{14}\,\text{cm}^{-2}) + \beta_1(R - R_0) + \beta_2[(R - R_0)^2 + \delta^2/4]^{0.5}, \quad \text{(F1)}$$

where $\beta_1 = (\alpha_1 + \alpha_2)/2 \simeq -0.0172$ ($\alpha_{1,2}$ being the slopes of each straight-line), $\beta_2 = (\alpha_2 - \alpha_1)/2 \simeq +0.0169$, and with the radius of curvature at $R = R_0$ being proportional to $\delta$ (an adjustable parameter that allows one to exactly follow the lines right to the intersection point, $\delta = 0$, or smoothly merge the two asymptotes, $\delta > 0$). We adopt $\delta = 30$. The dotted and dashed orange curves in Figure 10 show the models where the lower limits at $R < 50$ kpc are increased by a factor of 0.3 or 0.6 dex, respectively.

The second model is the SPL, which can be written as

$$N_{Si} = N_{Si}^0\, R^\beta \ (\text{cm}^{-2}), \quad \text{(F2)}$$

where $N_{Si}^0 = 10^{15.91}$ cm$^{-2}$ and $\beta = -1.23$ and $N_{Si}^0 = 10^{16.16}$ cm$^{-2}$ and $\beta = -1.33$ corresponding to lower limits at $R < 50$ kpc being increased by a factor of 0.3 and 0.6 dex, respectively. These two fits are shown in Figure 10 with the dotted and dashed green lines, respectively.

Finally, we use the GP model, which is a generic supervised learning method designed to solve a regression, here between log $N$ and $R$. The major advantages of this method are that the prediction interpolates the observations in a nonparametric way and is probabilistic so that empirical confidence intervals can be computed. We use the Python SCIKIT-LEARN GAUSSIAN PROCESS REGRESSION (Pedregosa et al. 2011; Buitinck et al. 2013) to model the data with a squared-exponential kernel (with a length scale of 0.1 and lower and upper bounds on length scale from 0.01 to 130). Changing the length scale has little effect on the model, but changing the bounds changes the smoothness of the model (a small length scale value means that function values can change quickly while large values characterize functions that change only slowly). The use of a more complex kernel like a Matern kernel would yield similar results when using similar bounds. We treat all the data with the same weight using an error of 0.3 dex on the column density. This is larger than the typical measurements (0.05–0.15 dex) in our sample, except for some of the data with vertical ticked bars in Figure 10. This error can be understood as a prior factor to smooth out the scatter of the data; empirically, 0.3 dex is the minimal value to make the model converge.[35] The effect of an increase on the error on each data point would flatten the relationship with a somewhat larger deviation. The same weight is justified in order to not favor detections versus upper or lower limits. With these assumptions, we model the data and show the mean values of the predictive distribution from the GP models in Figure 10, with the dotted and dashed blue curves corresponding again to the cases where the lower limits at $R < 50$ kpc are increased by a factor of 0.3 and 0.6 dex, respectively. The blue area around each curve show the standard deviation determined by the GP model.

## ORCID iDs

Nicolas Lehner ⓘ https://orcid.org/0000-0001-9158-0829
J. Christopher Howk ⓘ https://orcid.org/0000-0002-2591-3792
Bart P. Wakker ⓘ https://orcid.org/0000-0002-0507-7096
Jason Tumlinson ⓘ https://orcid.org/0000-0002-7982-412X
Edward B. Jenkins ⓘ https://orcid.org/0000-0003-1892-4423
J. Xavier Prochaska ⓘ https://orcid.org/0000-0002-7738-6875
Ramona Augustin ⓘ https://orcid.org/0000-0001-7472-3824
Suoqing Ji ⓘ https://orcid.org/0000-0001-9658-0588
Claude-André Faucher-Giguère ⓘ https://orcid.org/0000-0002-4900-6628
Zachary Hafen ⓘ https://orcid.org/0000-0001-7326-1736
Molly S. Peeples ⓘ https://orcid.org/0000-0003-1455-8788
Kat A. Barger ⓘ https://orcid.org/0000-0001-5817-0932

---

[35] Including another term in the kernel (e.g., adding another squared-exponential kernel to the original kernel) would produce a similar model with a larger standard deviation; in that case the error on each data point would have actually no effect on the model.

Michelle A. Berg ⓘ https://orcid.org/0000-0002-8518-6638
Rongmon Bordoloi ⓘ https://orcid.org/0000-0002-3120-7173
Thomas M. Brown ⓘ https://orcid.org/0000-0002-1793-9968
Andrew J. Fox ⓘ https://orcid.org/0000-0003-0724-4115
Karoline M. Gilbert ⓘ https://orcid.org/0000-0003-0394-8377
Puragra Guhathakurta ⓘ https://orcid.org/0000-0001-8867-4234
Jason S. Kalirai ⓘ https://orcid.org/0000-0001-9690-4159
Felix J. Lockman ⓘ https://orcid.org/0000-0002-6050-2008
John M. O'Meara ⓘ https://orcid.org/0000-0002-7893-1054
Joseph Ribaudo ⓘ https://orcid.org/0000-0003-3381-9795
Jessica K. Werk ⓘ https://orcid.org/0000-0002-0355-0134

## References

Akerman, C. J., Carigi, L., Nissen, P. E., Pettini, M., & Asplund, M. 2004, A&A, 414, 931
Anderson, M. E., & Bregman, J. N. 2011, ApJ, 737, 22
Anderson, M. E., Bregman, J. N., & Dai, X. 2013, ApJ, 762, 106
Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, MNRAS, 470, 4698
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Bechtold, J., Crotts, A. P. S., Duncan, R. C., & Fang, Y. 1994, ApJL, 437, L83
Berg, M. A., Howk, J. C., Lehner, N., et al. 2019, ApJ, 883, 5
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Bordoloi, R., Tumlinson, J., Werk, J. K., et al. 2014, ApJ, 796, 136
Borthakur, S., Heckman, T., Tumlinson, J., et al. 2016, ApJ, 833, 259
Bowen, D. V., Chelouche, D., Jenkins, E. B., et al. 2016, ApJ, 826, 50
Bowen, D. V., Pettini, M., & Blades, J. C. 2002, ApJ, 580, 169
Bregman, J. N., Anderson, M. E., Miller, M. J., et al. 2018, ApJ, 862, 3
Brown, T. M., Ferguson, H. C., Smith, E., et al. 2004, AJ, 127, 2738
Buitinck, L., Louppe, G., Blondel, M., et al. 2013, arXiv:1309.0328
Burchett, J. N., Tripp, T. M., Bordoloi, R., et al. 2016, ApJ, 832, 124
Bustard, C., Pardy, S. A., D'Onghia, E., Zweibel, E. G., & Gallagher, J. S. I. 2018, ApJ, 863, 49
Cameron, E. 2011, PASA, 28, 128
Cescutti, G., Matteucci, F., McWilliam, A., & Chiappini, C. 2009, A&A, 505, 605
Chan, T., Kereš, D., Hopkins, P., et al. 2019, MNRAS, 488, 3716
Chemin, L., Carignan, C., & Foster, T. 2009, ApJ, 705, 1395
Chen, H.-W., Johnson, S. D., Straka, L. A., et al. 2019, MNRAS, 484, 431
Collins, M. L. M., Chapman, S. C., Rich, R. M., et al. 2013, ApJ, 768, 172
Corbelli, E. 2003, MNRAS, 342, 199
Corlies, L., Peeples, M. S., Tumlinson, J., et al. 2020, ApJ, 896, 125
Courteau, S., & van den Bergh, S. 1999, AJ, 118, 337
Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, MNRAS, 450, 1937
Crighton, N. H. M., Hennawi, J. F., & Prochaska, J. X. 2013, ApJL, 776, L18
D'Onghia, E., & Fox, A. J. 2016, ARA&A, 54, 363
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Jr., et al. 1991, Third Reference Catalogue of Bright Galaxies (New York: Springer)
Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172
Ellison, S. L., Ibata, R., Pettini, M., et al. 2004, A&A, 414, 79
Fabbian, D., Khomenko, E., Moreno-Insertis, F., & Nordlund, Å. 2010, ApJ, 724, 1536
Faucher-Giguère, C.-A. 2020, MNRAS, 493, 1614
Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Feigelson, E. D., & Nelson, P. I. 1985, ApJ, 293, 192
Field, G. B. 1971, ApJ, 165, 29
Fischer, W. J., et al. 2019, Cosmic Origins Spectrograph Instrument Handbook, v.18.0 (Baltimore, MD: STScI)
Fitzpatrick, E. L., & Spitzer, L., Jr. 1997, ApJ, 475, 623
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Fox, A. J., Savage, B. D., & Wakker, B. P. 2005, AJ, 130, 2418
Fox, A. J., Wakker, B. P., Barger, K. A., et al. 2014, ApJ, 787, 147
Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Lee, K. 2014, MNRAS, 438, 2578
Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2019, MNRAS, 487, 1380
Geehan, J. J., Fardal, M. A., Babul, A., & Guhathakurta, P. 2006, MNRAS, 366, 996
Gilbert, K. M., Tollerud, E., Beaton, R. L., et al. 2018, ApJ, 852, 128
Gnat, O., & Sternberg, A. 2007, ApJS, 168, 213
Grebel, E. K., Gallagher, J. S. I., & Harbeck, D. 2003, AJ, 125, 1926
Green, J. C., Froning, C. S., Osterman, S., et al. 2012, ApJ, 744, 60
Gupta, A., Mathur, S., Galeazzi, M., & Krongold, Y. 2014, Ap&SS, 352, 775
Gupta, A., Mathur, S., & Krongold, Y. 2017, ApJ, 836, 243
Gupta, A., Mathur, S., Krongold, Y., Nicastro, F., & Galeazzi, M. 2012, ApJL, 756, L8
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2017, MNRAS, 469, 2292
Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2019, MNRAS, 488, 1248
Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2020, MNRAS, 494, 3581
Hamilton-Campos, K. A., Simons, R., Peeples, M., & Tumlinson, J. 2020, AAS Meeting, 52, 207.07
Henley, D. B., & Shelton, R. L. 2010, ApJS, 187, 388
Henley, D. B., Shelton, R. L., & Kwak, K. 2014, ApJ, 791, 41
Hopkins, P. F. 2015, MNRAS, 450, 53
Hopkins, P. F., Chan, T., Garrison-Kimmel, S., et al. 2020, MNRAS, 492, 3465
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800
Howk, J. C., Wotta, C. B., Berg, M. A., et al. 2017, ApJ, 846, 141
Hummels, C. B., Bryan, G. L., Smith, B. D., & Turk, M. J. 2013, MNRAS, 430, 1548
Hummels, C. B., Smith, B. D., Hopkins, P. F., et al. 2019, ApJ, 882, 156
Hummels, C. B., Smith, B. D., & Silvia, D. W. 2017, ApJ, 847, 59
Hunter, J. D. 2007, CSE, 9, 90
Isobe, T., Feigelson, E. D., & Nelson, P. I. 1986, ApJ, 306, 490
Jenkins, E. B. 2009, ApJ, 700, 1299
Ji, S., Chan, T., Hummels, C. B., et al. 2020, MNRAS, 496, 4221
Johnson, S. D., Chen, H.-W., & Mulchaey, J. S. 2015, MNRAS, 449, 3263
Kacprzak, G. G., Pointon, S. K., Nielsen, N. M., et al. 2019, ApJ, 886, 91
Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, A&A, 440, 775
Kalberla, P. M. W., McClure-Griffiths, N. M., Pisano, D. J., et al. 2010, A&A, 521, A17
Kam, S. Z., Carignan, C., Chemin, L., et al. 2017, AJ, 154, 41
Kang, Y., Bianchi, L., & Rey, S.-C. 2009, ApJ, 703, 614
Keeney, B. A., Stocke, J. T., Danforth, C. W., et al. 2017, ApJS, 230, 6
Keeney, B. A., Stocke, J. T., & Rosenberg, J. L., et al. 2013, ApJ, 765, 27
Keres, D., Yun, M. S., & Young, J. S. 2003, ApJ, 582, 659
Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102
Kulkarni, V. P., Cashman, F. H., Lopez, S., et al. 2019, ApJ, 886, 83
Lehner, N., & Howk, J. C. 2011, Sci, 334, 955
Lehner, N., Howk, J. C., Thom, C., et al. 2012, MNRAS, 424, 2896
Lehner, N., Howk, J. C., Tripp, T. M., et al. 2013, ApJ, 770, 138
Lehner, N., Howk, J. C., & Wakker, B. P. 2015, ApJ, 804, 79
Lehner, N., Jenkins, E. B., Gry, C., et al. 2003, ApJ, 595, 858
Lehner, N., Keenan, F. P., & Sembach, K. R. 2001, MNRAS, 323, 904
Lehner, N., O'Meara, J. M., Fox, A. J., et al. 2014, ApJ, 788, 119
Lehner, N., Wotta, C. B., Howk, J. C., et al. 2018, ApJ, 866, 33
Lehner, N., Wotta, C. B., Howk, J. C., et al. 2019, ApJ, 887, 5
Lehner, N., Zech, W. F., Howk, J. C., & Savage, B. D. 2011, ApJ, 727, 46
Liang, C. J., & Chen, H.-W. 2014, MNRAS, 445, 2061
Liang, C. J., Kravtsov, A. V., & Agertz, O. 2016, MNRAS, 458, 1164
Lopez, S., Reimers, D., Gregg, M. D., et al. 2005, ApJ, 626, 767
Lopez, S., Tejos, N., Barrientos, L. F., et al. 2020, MNRAS, 491, 4442
Martin, C. L., Scannapieco, E., Ellison, S. L., et al. 2010, ApJ, 721, 174
Martin, N. F., Chambers, K. C., Collins, M. L. M., et al. 2014, ApJL, 793, L14
Martin, N. F., Ibata, R. A., Lewis, G. F., et al. 2016, ApJ, 833, 167
Martin, N. F., Weisz, D. R., Albers, S. M., et al. 2017, ApJ, 850, 16
Mattsson, L. 2010, A&A, 515, A68
Mayer, L., Mastropietro, C., Wadsley, J., Stadel, J., & Moore, B. 2006, MNRAS, 369, 1021
McConnachie, A. W. 2012, AJ, 144, 4
McConnachie, A. W., Irwin, M. J., Ferguson, A. M. N., et al. 2005, MNRAS, 356, 979
McMillan, P. J. 2011, MNRAS, 414, 2446
Ménard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, MNRAS, 405, 1025
Miller, M. J., & Bregman, J. N. 2013, ApJ, 770, 118
Moos, H. W., Cash, W. C., Cowie, L. L., et al. 2000, ApJL, 538, L1

Muzahid, S., Kacprzak, G. G., Charlton, J. C., & Churchill, C. W. 2016, ApJ, 823, 66

Muzahid, S., Kacprzak, G. G., Churchill, C. W., et al. 2015, ApJ, 811, 132

Nidever, D. L., Majewski, S. R., & Butler Burton, W. 2008, ApJ, 679, 432

Nidever, D. L., Majewski, S. R., Butler Burton, W., & Nigra, L. 2010, ApJ, 723, 1618

Nuza, S. E., Parisi, F., Scannapieco, C., et al. 2014, MNRAS, 441, 2593

Oppenheimer, B. D., Crain, R. A., Schaye, J., et al. 2016, MNRAS, 460, 2157

Oppenheimer, B. D., & Davé, R. 2006, MNRAS, 373, 1265

Oppenheimer, B. D., & Schaye, J. 2013a, MNRAS, 434, 1043

Oppenheimer, B. D., & Schaye, J. 2013b, MNRAS, 434, 1063

Oppenheimer, B. D., Schaye, J., Crain, R. A., Werk, J. K., & Richings, A. J. 2018a, MNRAS, 481, 835

Oppenheimer, B. D., Segers, M., Schaye, J., Richings, A. J., & Crain, R. A. 2018b, MNRAS, 474, 4740

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Peek, J. E. G., Ménard, B., & Corrales, L. 2015, ApJ, 813, 7

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., Rahmani, H., Arrigoni Battaia, F., & Augustin, R. 2018, MNRAS, 479, L50

Poisson, M., Bustos, C., López Fuentes, M., Mandrini, C. H., & Cristiani, G. D. 2020, AdSpR, 65, 1629

Prochaska, J. X., Tejos, N., Wotta, C. B., et al. 2017a, pyigm/pyigm: Initial Release for Publications, UCSC, Zenodo, doi:10.5281/zenodo.1045480

Prochaska, J. X., Werk, J. K., Worseck, G., et al. 2017b, ApJ, 837, 169

Putman, M. E., Peek, J. E. G., & Joung, M. R. 2012, ARA&A, 50, 491

Putman, M. E., Staveley-Smith, L., Freeman, K. C., Gibson, B. K., & Barnes, D. G. 2003, ApJ, 586, 170

Rahmati, A., & Oppenheimer, B. D. 2018, MNRAS, 476, 4865

Rauch, M., Sargent, W. L. W., Barlow, T. A., & Carswell, R. F. 2001, ApJ, 562, 76

Rhodin, N. H. P., Agertz, O., Christensen, L., Renaud, F., & Fynbo, J. P. U. 2019, MNRAS, 488, 3634

Ribaudo, J., Lehner, N., Howk, J. C., et al. 2011, ApJ, 743, 207

Richter, P., Nuza, S. E., Fox, A. J., et al. 2017, A&A, 607, A48

Riess, A. G., Fliri, J., & Valls-Gabaud, D. 2012, ApJ, 745, 156

Riley, A., et al. 2019, STIS Instrument Handbook, v.11.0 (Baltimore, MD: STScI)

Rosenwasser, B., Muzahid, S., Charlton, J. C., et al. 2018, MNRAS, 476, 2258

Rubin, K. H. R., O'Meara, J. M., Cooksey, K. L., et al. 2018, ApJ, 859, 146

Savage, B. D., Kim, T.-S., Wakker, B. P., et al. 2014, ApJS, 212, 8

Savage, B. D., & Sembach, K. R. 1991, ApJ, 379, 245

Savage, B. D., & Sembach, K. R. 1996, ARA&A, 34, 279

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829

Shen, S., Madau, P., Guedes, J., et al. 2013, ApJ, 765, 89

Shull, J. M., Jones, J. R., Danforth, C. W., & Collins, J. A. 2009, ApJ, 699, 754

Smette, A., Surdej, J., Shaver, P. A., et al. 1992, ApJ, 389, 39

Stern, J., Fielding, D., Faucher-Giguère, C.-A., & Quataert, E. 2020, MNRAS, 492, 6042

Stocke, J. T., Keeney, B. A., Danforth, C. W., et al. 2013, ApJ, 763, 148

Suresh, J., Nelson, D., Genel, S., Rubin, K. H. R., & Hernquist, L. 2019, MNRAS, 483, 4040

Tamm, A., Tempel, E., Tenjes, P., Tihhonova, O., & Tuvikene, T. 2012, A&A, 546, A4

Telford, O. G., Werk, J. K., Dalcanton, J. J., & Williams, B. F. 2019, ApJ, 877, 120

The LUVOIR Team 2019, ApJ, 877, 120

Thom, C., Peek, J. E. G., Putman, M. E., et al. 2008, ApJ, 684, 364

Thom, C., Tumlinson, J., Werk, J. K., et al. 2012, ApJL, 758, L41

Thom, C., Werk, J. K., Tumlinson, J., et al. 2011, ApJ, 736, 1

Tripp, T. M., Meiring, J. D., Prochaska, J. X., et al. 2011, Sci, 334, 952

Tripp, T. M., Sembach, K. R., Bowen, D. V., et al. 2008, ApJS, 177, 39

Tumlinson, J., Peeples, M. S., & Werk, J. K. 2017, ARA&A, 55, 389

Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, Sci, 334, 948

Tumlinson, J., Thom, C., Werk, J. K., et al. 2013, ApJ, 777, 59

Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, ApJS, 192, 9

Turner, M. L., Schaye, J., Crain, R. A., Theuns, T., & Wendt, M. 2016, MNRAS, 462, 2440

van de Voort, F., Springel, V., Mandelker, N., van den Bosch, F. C., & Pakmor, R. 2019, MNRAS, 482, L85

van der Marel, R. P., Fardal, M., Besla, G., et al. 2012, ApJ, 753, 8

van der Marel, R. P., & Guhathakurta, P. 2008, ApJ, 678, 187

Veljanoski, J., Mackey, A. D., Ferguson, A. M. N., et al. 2014, MNRAS, 442, 2929

Wakker, B. P. 2001, ApJS, 136, 463

Wakker, B. P. 2006, ApJS, 163, 282

Wakker, B. P., Hernandez, A. K., French, D. M., et al. 2015, ApJ, 814, 40

Wakker, B. P., Savage, B. D., Sembach, K. R., et al. 2003, ApJS, 146, 1

Wakker, B. P., & van Woerden, H. 1997, ARA&A, 35, 217

Wakker, B. P., York, D. G., Wilhelm, R., et al. 2008, ApJ, 672, 298

Wang, Q. D., & Yao, Y. 2012, arXiv:1211.4834

Watkins, L. L., Evans, N. W., & An, J. H. 2010, MNRAS, 406, 264

Watts, D. G., & Bacon, D. W. 1974, Technometrics, 16, 369

Welty, D. E., Hobbs, L. M., Lauroesch, J. T., et al. 1999, ApJS, 124, 465

Welty, D. E., Lauroesch, J. T., Blades, J. C., Hobbs, L. M., & York, D. G. 1997, ApJ, 489, 672

Werk, J. K., Prochaska, J. X., Thom, C., et al. 2012, ApJS, 198, 3

Werk, J. K., Prochaska, J. X., Thom, C., et al. 2013, ApJS, 204, 17

Werk, J. K., Prochaska, J. X., Tumlinson, J., et al. 2014, ApJ, 792, 8

Williams, B. F., Dolphin, A. E., Dalcanton, J. J., et al. 2017, ApJ, 846, 145

Woodgate, B. E., Kimble, R. A., Bowers, C. W., et al. 1998, PASP, 110, 1183

Wotta, C. B., Lehner, N., Howk, J. C., et al. 2019, ApJ, 872, 81

Wotta, C. B., Lehner, N., Howk, J. C., O'Meara, J. M., & Prochaska, J. X. 2016, ApJ, 831, 95

Zahedy, F. S., Chen, H.-W., Johnson, S. D., et al. 2019, MNRAS, 484, 2257

Zahedy, F. S., Chen, H.-W., Rauch, M., Wilson, M. L., & Zabludoff, A. 2016, MNRAS, 458, 2423

Zech, W. F., Lehner, N., Howk, J. C., Dixon, W. V. D., & Brown, T. M. 2008, ApJ, 679, 460

Zheng, Y., Peeples, M. S., O'Shea, B. W., et al. 2020, ApJ, 896, 143

Zheng, Y., Putman, M. E., Peek, J. E. G., & Joung, M. R. 2015, ApJ, 807, 103