Mapping the Circumgalactic Medium: Connecting Galaxies to their Environment

Matthew C. Wilde

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Jessica Werk, Chair

Matthew McQuinn

Kirill Tchernyshyov

Program Authorized to Offer Degree:
Astronomy

University of Washington

Abstract

Mapping the Circumgalactic Medium: Connecting Galaxies to their Environment

Matthew C. Wilde

Chair of the Supervisory Committee:
Professor Jessica Werk
Astronomy

The work presented in this dissertation provides the groundwork for understanding key unresolved questions in the study of galaxy evolution: the circumgalactic medium (CGM) and the role of environment. I present a new survey that has the power to drastically improve our understanding of the CGM and use it to estimate the size of the CGM as a function of galactic mass. In addition, a novel method to reconstruct the cosmic web using Physarum polycephalum slime mold is presented and applied to the SDSS surveys. This cosmic web reconstruction was released publicly with SDSS DR17 and will allow exciting connections between galaxies and their environments to be illuminated, to distances greater than other previous catalogs.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

Table Number Page

ACKNOWLEDGMENTS

The work presented in this dissertation represents the culmination of seven years of graduate study at the University of Washington that would not have been possible without the support of so many individuals. I would like to start by thanking my graduate advisor, Jessica Werk, for her academic, professional, and funding support throughout my graduate career. It has been an honor to learn from you and collaborate with you on so many fascinating projects. I also acknowledge the generous financial support of the NSF IGERT fellowship in conjunction with the eScience Department, which has provided me with countless opportunities for inspiring conversations and collaborations over the years. I owe a debt of gratitude to my dissertation committee members, Jessica Werk, Kirill Tchernyshyov, for guiding me through the finish line, as well as Matthew McQuinn for serving on my committee.

There are countless close collaborators that I have sincerely appreciated working with and whose expertise I am grateful for. I would like to thank Joeseph Burchett, J. X. Prochaska, and Todd Tripp for advising the creation of the CGM^2 survey.

The trials of graduate school would have been too much to bear without the support of my friends and family. I deeply appreciate having had the opportunity to develop strong friendships through competitive foosball games with fellow UW Astronomy graduate students. I am grateful for the friendships with David Fleming, Jacob Lustig-Yaeger and Hannah Bish whose support both in and out of the Astronomy Department enriched the entire experience. I would also like to thank Daniel D'Orazio, Lauren Miller, Jeff Andrews, Dan Forman-Mackey and Michael Hahn for being the best of friends. In addition, I must thank the Al's Tavern family that provided a sense of home during my time in Seattle (as well as a place to work on this PhD).

The perspectives, approaches, and results presented here all emanate from collaborative efforts. Throughout this dissertation, I will use the royal “we” to acknowledge these collaborations and to acknowledge the continued collaborative journey between me, the author, and you, the reader.

DEDICATION

To my parents, Dr. James Wilde and Deborah Newman. I could not have been blessed with more supporting and loving parents. Thank you for the inspiration and patience to let me find my way in life.

Chapter 1

INTRODUCTION

Modern cosmological models predict the universe's mass budget is dominated by dark matter, forming the skeletal structure of the universe we now call the "cosmic web", upon which galaxies form and evolve. The cosmic web, along with galaxies, were formed from small perturbations (Tolman, 1934) in the very early universe which were seeded during an inflationary phase. Denser than average regions accelerated more slowly than regions of average density. The density contrasts of these regions grew slowly at first and later hastened as gravity locally halted the expansion and caused those regions to collapse upon themselves, the matter falling to the center of the local potential wells. This collapse caused the matter to settle into gravitationally bound but random orbits into objects known as "halos" in a process known as "virialization". These virialzed halos have a gravitational potential energy that is roughly equivalent to twice the kinetic energy of the matter (Rubin et al., 1980). Non-baryonic dark matter cannot lose orbital energy by radiating photons, but a halo's baryonic matter is able to lose orbital energy through inelastic collisions that do radiate photons. Baryonic matter therefore settles toward the halo's center, where some of it may further cool and forms stars, producing an observable galaxy, surrounded by a gaseous atmosphere.

Many questions remain about the details of this broad-brush picture. Despite the relatively straightforward origins of galaxies from gravitational growth of density fluctuations in the early universe along with radiative cooling, the appearance of galaxies varies wildly. This simple picture presents no mass scale, yet galaxy properties differ greatly as a function of mass.

The ΛCDM cosmological model, consisting of gravitational clumping of cold dark matter (CDM) combined with Einstein's cosmological constant Λ which accelerates the halos away from one another, has been hugely successful in reproducing the large scale structure of the universe (Alam et al., 2017, e.g.). In this model, galaxies form in a hierarchical process where small halos form first and massive halos grow via mergers of smaller ones. This would imply that larger and more massive galaxies could appear younger than smaller ones, which is the opposite effect seen in observations. Massive elliptical galaxies formed most of their stars very early in the history of the universe, while lower-mass disk galaxies formed their stars later and are still making them (Behroozi et al., 2013). Clearly, additional physics beyond the physics of gravitational infall and cosmological shock heating governs the baryon cycle that fuels star formation.

Most of the outstanding problems in galaxy evolution can be solved by a better understanding of the circumgalactic medium (CGM), or the galactic atmosphere (See Tumlinson et al. (2017) for a review). This rarefied, gaseous medium can extend to $10 - 100$ times the typical size ascribed to galaxy disks, and have been found to contain more regular, baryonic matter than the galaxies themselves. Astronomers have found this gaseous reservoir to harbor complex physical processes due to the interaction with processes in the galaxy itself, such as the birth and death of stars and the feeding of the central black hole, known as active galactic nuclei (AGN). These phenomena, known collectively as "feedback" can expel photons and gas out of galaxies, where the photons and gas interact with the gas that resides well beyond the reach of galaxies, known as the intergalactic medium (IGM), as it falls into the gravitational potential of the galaxies. The interplay between these forces may limit a galaxy's ability to accrue more material to fuel star formation, or may heat up the gas to a regime where the gas cannot cool to form new stars, thus governing galaxy evolution itself. Although the CGM is a region of great importance in understanding galaxies holistically, there remain many aspects of the CGM which are still unknown such as the size and structure of the CGM and its dependence of galaxy properties, as well as the dependence on properties of the larger scale environment surrounding galaxies.

1.1 A Brief History of the CGM

As early as 1956, Lyman Spitzer predicted the existence of a hot "galactic corona" surrounding the Milky Way (MW) based on on observations taken by Guido Münch (Münch and Zirin, 1961) of Na and Ca in spectroscopic observations of distant stars away from the galactic plane. This prediction ushered in the era of using distant background objects to observe the absorption signatures of intervening gas. After the discovery of Quasars — extremely bright extragalactic sources — astronomers quickly began to use their spectra to study the intervening extragalactic gas. It was Bahcall and Spitzer (1969) who first proposed that these absorption lines at multiple redshifts are caused by gas in "extended halos of normal galaxies". It wasn't until the 1980s that some absorption lines in quasar spectra were associated with galaxies near the line of sight while others, the Lyman alpha forest, were attributed to the IGM (Sargent et al., 1980; Bergeron and Stasińska, 1986; Bergeron et al., 1992).

The advent of the Hubble Space Telescope (HST) and the spectroscopic instruments aboard led to great leaps in the 1990s, leading to the discovery that galaxy halos give rise to strong Lyα, CIV, and other metal lines (Lanzetta et al., 1995; Chen et al., 1998, e.g.) in a gaseous medium that is multi-phase and richly structured.

The Sloan Digital Sky Survey (SDSS) created a massive database of galaxy observations and led to a leap in the understanding of galaxy evolution, while also posing new questions, such as how galaxies come to quench, or cease star formation. SDSS also uncovered another mystery: Galaxies were missing a large fraction of the baryons they were predicted to have, known as the missing baryon problem (Bregman, 2007). These questions led theorists to implement new physical prescriptions for feedback and inflow to describe galaxy evolution. The computing revolution of this time allowed simulations to begin to describe the feedback and inflow of gas into galaxies. However, new observations and theories are still necessary to fully capture the complex processes at play and their subsequent effect on the evolution of galaxies.

1.2 Major Problems in Galaxy Evolution

As mentioned above, there remain many puzzles in galaxy evolution. Many of these challenges in our model of galaxy evolution involve the regulation of gas flows into and out of galaxies, passing through the CGM.

How do galaxies sustain their star formation? The ISM contained in star forming gas is only sufficient to produce stars for a small fraction of the life of the galaxy. This implies an external supply of gas to replenish the ISM and keep it in a semi-equilibrium. The depletion timescale, $\tau_{\rm dep}$ $M_{\rm gas}/\dot{M}$, does not vary greatly of a large range in galaxy masses. Lower mass galaxies generally have bursty star formation histories while massive galaxies tend towards a more continuous star formation history, suggesting the means to which these galaxies acquire gas must be different, or regulated in some way. Accounting for all of the gas mass in the CGM, which is fueling the star formation, requires knowledge of the size of the CGM, in addition to the physical means of gas regulation.

One of the largest unsolved problems in galaxy evolution is quenching. Why do galaxies quench and how do they stay quenched? Cool gas has been observed surrounding massive quenched galaxies yet this gas is not being converted into star formation (Thom et al., 2012). Solutions to this involve controlling the gas supply, via shutting of IGM accretion or heating the CGM to the point that it cannot cool and enter the ISM. Low mass galaxies, which would otherwise continue to form stars, quench if they are a satellite or near a more massive galaxy (Geha et al., 2012; Putman et al., 2021). This form of quenching is thought to occur by ram pressure stripping of the satellite CGM via the CGM of the host galaxy. The boundary of the CGM is then clearly important to understanding when this stripping will occur. This appears to occur well outside the virial radius, at least in large clusters (Burchett et al., 2018, e.g.). In addition, galaxies in denser regions of the cosmic web have higher quenching fractions that those in low density regions such as near voids (Blanton and Moustakas, 2009; Peng et al., 2010; Prochaska et al., 2011, e.g.), implying the large scale environment is disturbing the CGM to some degree. The scales at which this takes place is of vital importance to understanding the mechanisms of environmental quenching.

Another puzzle presented by ΛCDM cosmology is the predictions that baryons follow the dark matter into halos, where the gas cools and falls into the ISM. Observed galaxies harbor only a small fraction of the halo's expected baryons in the ISM and stars. Werk et al. (2014) showed that these "missing" baryons are contained in the CGM (see also Bordoloi et al., 2014). However, calculating the total mass of baryons in the CGM required them to use a radius to make this claim. They sensibly integrated out to 150 kpc but the calculation goes as R^2 and thus the mass in of baryons will change if the CGM is indeed larger or smaller than this number and indeed there are hints that the size of the CGM is larger than the virial radius (Prochaska et al., 2011).

Galaxies form via the accretion of baryonic matter, gas, from their surroundings, the IGM, passing through the CGM. This gas is then gravitationally bound to the halo, and must subsequently cool, shedding the energy gained via assembly. However, the details in which the gas cools and collapses into the interstellar medium and eventually form stars remains a puzzle. I will review two of the outstanding problems in galaxy evolution that the CGM can help inform, namely: cold mode accretion and precipitation.

In the conventional picture of galaxy formation, the gas follows the track of shock heating to roughly the virial temperature of the galaxy potential well ($T \sim 10^6$ K for a Milky Way type galaxy) before cooling, condensing and forming stars. However, if this process proceeded unabated, it would lead to a continuous flow of cooling gas to the center of halos, which is not borne out in observations (Lea et al., 1973). One solution to this problem is the known as "chaotic cold accretion" However, galaxies seem to be able to self regulate by inducing feedback onto the accreting gas. Kereš et al. (2005), however, found in their simulations that roughly half of the gas radiates its acquired gravitational energy at much lower temperatures, typically $T < 10^5$ K, and there exists a bimodal temperature distribution. The "cold mode" of gas accretion dominates for low-mass galaxies (baryonic mass $M_\star \lesssim 10^{10} M_\odot$), while the conventional "hot mode" dominates the growth of high-mass systems. Cold accretion is often directed along filaments, allowing galaxies to efficiently draw gas from large distances, while hot accretion is quasi-spherical. This model, however, has some challenges with theory and observation: the prevalence of cool gas observed in halos is a challenge for the validity of subsonic cooling flow solutions, in which the halo gas is expected to be predominantly single-phased (Stern et al., 2019); furthermore, as the homogeneous inflow becomes increasingly prone to runaway thermal instability and formation of an inhomogeneous multiphase atmosphere in which stars can form. The precipitation hypothesis proposes that this transition to a multiphase state fuels feedback that places upper limits on the pressure, density, and cooling flow rate of the ambient medium. It plausibly explains how self regulating feedback can maintain the resulting limits on atmospheric density, pressure, and cooling time but does not explain how galaxy evolution drives massive galaxies toward such a self-regulating state (Donahue and Voit, 2022). Distinguishing between these models will require observations that better characterize the CGM. See Figure 1.1 for an overview of the CGM.

1.3 Atmospheric Boundary

The CGM in the context of this thesis is defined to include all of the gas gravitationally bound to a galaxy's halo outside of the galaxies interstellar medium (ISM). We can treat the CGM as an atmosphere because many of the concepts that apply to the atmospheres of planets and stars also apply to the gas in and around galaxies. The primary reason we consider all of the gas gravitationally bound to a galaxy is that there exists no consensus on where the CGM ends and where the IGM begins. Indeed, much of what was once considered to be the IGM is now considered part of the CGM. O VI observations show that the Warm-Hot Intergalactic Medium (WHIM) ($T \sim 10^5 - 10^6$K gas) cannot be distinguished from the CGM (McQuinn and Werk, 2018).

The most commonly used scale to define the boundary of the CGM is the virial radius, itself not well defined. We will define the virial radius in terms the mean matter density contained in a halo as a fraction of the of cosmological critical overdensity

$$\rho_c = \frac{3H(z)^2}{8\pi G} \tag{1.1}$$

where G is the gravitational constant and $H(z)$ is the Hubble expansion parameter at a redshift z. The radius encompassing a matter density Δ times the critical density $\Delta\rho_c$, is therefore

$$r_\Delta = \left(\frac{3M_\Delta}{4\pi\Delta\rho_c}\right)^{1/3} = \left(\frac{2GM_\Delta}{\Delta H(z)^2}\right)^{1/3} \tag{1.2}$$

where M_Δ is the mass contained within r_Δ. Using the relation $v_c(r_\Delta)^2 = GM_\Delta/r_\Delta^2$ we arrive at

$$r_\Delta = \left(\frac{2}{\Delta}\right)^{1/2} \frac{v_c(r_\Delta)}{H(z)}. \tag{1.3}$$

Simulations show that orbits of dark matter particles in the outer parts of a cosmological halo gradually shift from mostly infalling to mostly isotropic in the neighborhood of $\Delta \sim 200$ (Cole and Lacey, 1996; Bryan and Norman, 1998) so we will adopt $R_{\rm vir} = r_{200c}$ as the virial radius. The virial radius is $\sim$ 250 kpc for a MW mass halo and so dwarfs the size of the disk by an order of magnitude. However, this boundary is rather arbitrary in the context of the CGM. At least some of the gas outside of r_{200c} was previously within that boundary and may have been gently pushed beyond it by a thermally-driven expansion of the CGM. The virial radius, based on collapse of dark matter overdensity into the halo is however, unobservable and thus untestable. Recently, an alternative approach to define a halo boundary was put forward by several authors. This approach, known as the “splashback” radius, specifies the radius at which matter that is bound to the halo can orbit to upon the first collapse (Diemer and Kravtsov, 2014; More et al., 2015, e.g.). This radius my also correspond to the radius at which gas can shock heat at halo assembly, and at which infalling substructures my be stripped by their host halos. The splashback radius, however, has not been widely adopted (yet) so I will continue to refer to the virial radius as a comparison of the atmospheric boundary. By observing the gas around galaxies, the CGM, we can estimate the actual extent of galaxies. This is necessary for if one wished to account for all of the baryons that belong to galaxies, for example. In this thesis, I will show the the boundary of the CGM can be estimated from observations, at least for the cool gas ($\lesssim 10^4$K).

1.4 Observations of the CGM

One of the most successful means of observing the CGM is to use the spectra of bright background objects, measuring the absorption of intervening gas from bright extragalactic sources. Since its installation in the 2009 HST servicing mission, the Cosmic Origins Spectrograph (COS) aboard the Hubble Space Telescope (HST) has made significant strides in

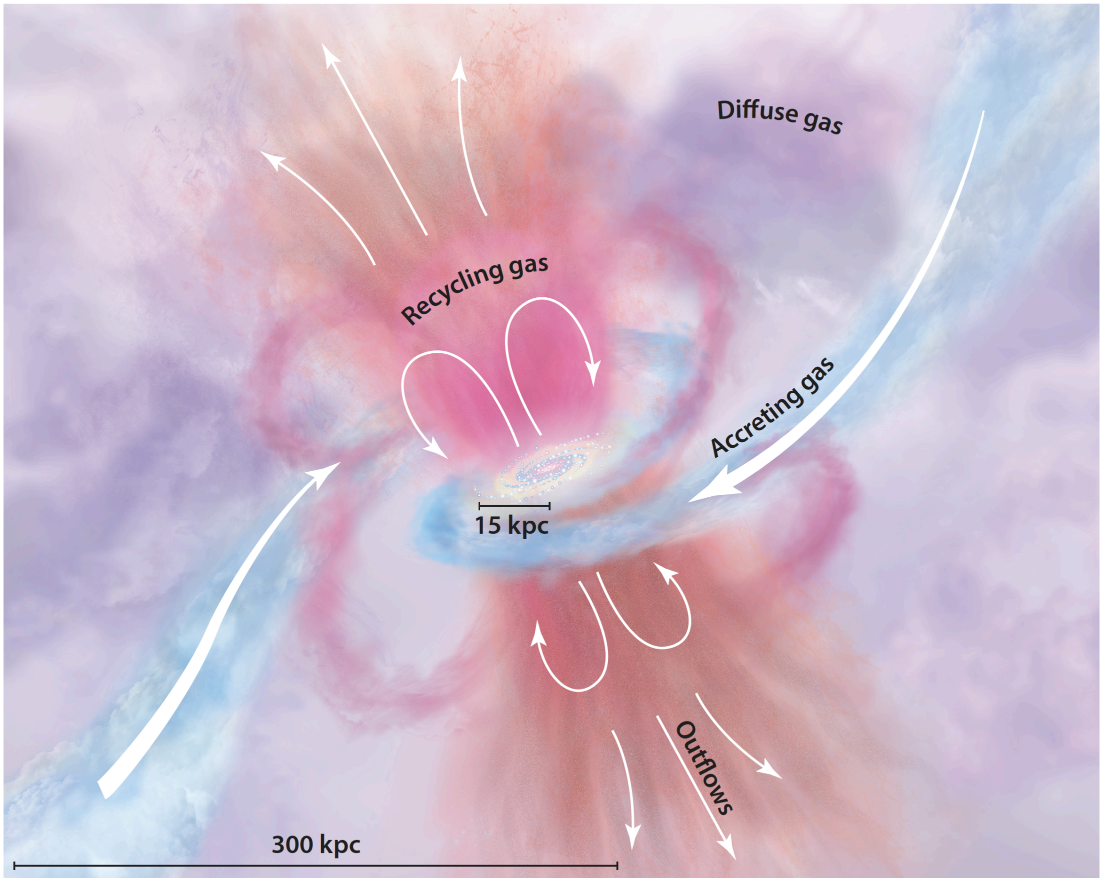


Figure 1.1: An artists depiction of the CGM. The galaxy's disk are fed by filamentary, or cold flow accretion from the IGM (blue). Outflows are driven from the disk in pink and orange, whereas gas that was previously ejected is recycling, known as the galactic fountain. The diffuse gas halo in varying tones of purple includes gas that is likely contributed by all these sources and mixed together over time. The scale of this region is given in the lower left of the diagram and is larger than the a MW mass virial radius ($R_{\rm vir} \sim 250$ kpc). Diagram from Tumlinson et al. (2017).

characterizing the low-redshift CGM over a wide range of physical phases (T $\approx 10^4$ K - 10^6 K). Because COS is ~10 times more sensitive than previous-generation UV spectrographs (e.g. STIS), it has thus only recently become possible to assemble statistically significant samples of galaxy-QSO pairs. Commonly observed lines such as HI (via the Lyman series), SiII, SiIII, CIV, OVI, and NeVIII possess a wide range of ionization potentials (from 0.8 − 15 Rydberg) and allow us to assess the density, temperature, ionization mechanism, and kinematics of the absorbing material in conjunction with photoionization modeling (e.g., Werk et al., 2014; ?).

Absorption-based work over the last decade has shown that: the CGM of L^* galaxies contains a major share of galactic baryons ($M_{\rm CGM} \geq M_*$; Werk et al. (2014)), serves as a massive reservoir of galactic metals (at least as much as in the ISM; (Peeples et al. (2014)), and is shaped by both enriched galactic outflows ($0.5\ Z_\odot$; Prochaska et al. (2017)) and a metal-poor component that resembles "cold accretion" from the IGM ($0.03\ Z_\odot$?) While actively star-forming galaxies exhibit a highly-ionized component to their CGM characterized by strong OVI absorption out to at least 150 kpc, non-star-forming, elliptical galaxies show only weak detections of OVI if OVI is detected at all (Tumlinson et al. (2011)). However, these same passive galaxies exhibit a high incidence of strong HI absorption in their CGM, as much cold, bound gas as their star-forming counterparts (Thom et al., 2012). These results emphasize that the processes that transform galaxies from star-forming disks to passive ellipticals do affect the physical state of the CGM, but do not completely evacuate it of cool, 10^4 K gas.

1.5 Simulating the CGM

Theorists have begun to directly trace the origin and lifecycle of the CGM gas using both cosmological zoom-in simulations of individual halos and cosmological hydrodynamic simulations with statistical samples of galaxies (Hummels et al., 2013; Ford et al., 2013; Anglés-Alcázar et al., 2017, e.g.,). From these simulations, we have learned that in order to reproduce the observed gas column densities out to ~ 200 kpc, bursty star-formation and feedback able to generate significant outflows are required (e.g., Liang et al. 2015). Cosmological hydrodynamical simulations that incorporate strong galactic winds indicate that at any given time, $65 - 80\%$ of the total baryons in a galactic halo reside outside of the galaxy stellar disk, and $> 50\%$ of these CGM baryons recycle through the ISM of the galaxy, typically on one-

Gyr timescales (?Christensen et al., 2016; Anglés-Alcázar et al., 2017). Nonetheless, crucial questions remain regarding the physical timescales and the spatial extent of gas outflow, inflow, and recycling, and the separate contributions of black hole and supernova feedback at different galaxy masses and redshifts. One major discrepancy between observations and simulations is that, thus far, SPH, AMH, and moving mesh cosmological hydrodynamical simulations universally under-predict the observed OVI by a factor of 3 −5. Resolving this discrepancy may involve incorporating additional gas physics (e.g. cosmic rays; Butsky et al. (2020)) or feedback mechanisms (e.g. 'flickering AGN'; Oppenheimer et al. (2018)). Furthermore, simulations have yet to examine the influence of large-scale environment and halo assembly histories on the nature of the CGM in a self-consistent manner. In addition, most simulations that focus on the CGM fail to integrate past the virial radius, thus limiting their ability to connect the CGM to the IGM, despite hints (Burchett et al., 2018, e.g.) that the CGM extends beyond this rather arbitrary distance. Thus determining an atmospheric boundary or, extent of the CGM, is necessary for simulations to compare to observations.

1.6 The Intergalactic Medium and Large Scale Structure

Of paramount concern in galaxy evolution science is the impact of a galaxy's environment on its morphology and star formation activity. Correlations between environmental metrics and galaxy properties, such as morphology (e.g., Dressler (1980)), color (Abell, 1965, e.g.), and star formation (Peng et al. (2010)) have been known about for many decades, but the physical mechanisms and their relative importance remain heavily pursued problems. Galaxy-environment analyses typically fall along one of two paths: local environment-centric or large-scale environment-centric. In the former, one employs an environmental density metric, such as a nearest-neighbor distance or density within some aperture (Kauffmann et al., 2004; Peng et al., 2010), or galaxies are associated with a local group or cluster environment (Yang et al., 2007; Berlind et al., 2016) and galaxy properties are studied with respect to the properties of the group or cluster (Carollo et al., 2013; Catinella et al., 2013).

Studies of galaxy-cosmic web dependencies report somewhat mixed results. Kuutma et al. (2017) find higher elliptical-to-spiral ratio and decreasing Star Formation Rate (SFR) towards filament spines. Similarly, Crone Odekon et al. (2018) report that, at fixed stellar mass, galaxies closer to filaments or in higher density environments are more deficient in HI. These large-scale effects on galaxies have also been investigated in the cosmological simulation domain. Codis et al. (2018) measure the spin-filament alignment in IllustrisTNG (Vogelsberger et al., 2014), find a strong dependence on spin alignment with galaxy mass. Pasha et al. (2022) find that the collapse of large scale structure into sheets at higher redshifts ($z \sim 3$) can create shocks that explain quenching in dwarf galaxies similar to the effect seen in nearby dwarf galaxies that exist in the presence of clusters and groups.

Aside from the galaxies themselves, one may also study the intergalactic medium in context with the cosmic web environment. Wakker et al. (2015) measured the Lyα absorption in quasar spectra probing a foreground visually identified filament, finding increasing absorber equivalent width and linewidth with decreasing projected distance to the center of the filament. With a larger archival sample of QSOs and filaments, Bouma et al. (2021) find similar results, with Lyα absorbers showing both greater incidence and column density at small projected distance and velocity offsets from filaments first identified by Courtois et al. (2013). In the first application of a novel application of an algorithm based on the Physarum polycephalum slime mold used to reconstruct the cosmic web, Burchett et al. (2020a) analyze the Lyα optical depth as a function of cosmic web density probed by QSO sightlines and find three distinct regimes: a void regime at low matter overdensity with no detected absorption, an onset of absorption in the outer skins of filaments with monotonically increasing optical depth, and the highest-density regime where the absorption no longer increases with local density but rather turns over and declines at the highest densities. This has important implications for the interaction of the galaxies, and their CGM, through which the IGM accretes to form stars. Combining the environmental density measurements with cosmic web reconstruction methods with CGM measurements will illucidate the connection to effects on the CGM of galaxies, and thus galaxy evolution.

1.7 Structure of this Document

It is clear that the CGM plays a vital role in galaxy evolution as it connects the galaxies to their surroundings, which provide fuel for star formation and growth. These surroundings depend of the location of galaxies in the context of the cosmic web. In this manuscript, we determine a data driven means of determining the extent of the CGM and new tools that will allow these observations to be compared with cosmic density.

The first part of this manuscript is dedicated to determining the extent of the CGM. In Chapter 2, I present initial results from the COS and Gemini Mapping the Circumgalactic Medium (CGMCGM $\equiv$ CGM^2) survey. The CGM^2 survey consists of 1689 galaxies, all with high-quality Gemini GMOS spectra, within 1 Mpc of twenty-two $z \lesssim 1$ quasars. I use the HI absorption measurements to study the covering fraction as a function of galaxy mass and impact parameter. I also present a model to determine the extent of the HI CGM.

In Chapter 3, I combine the CGM^2 survey with another survey, The COS Absorption Survey of Baryon Harbors (CASBaH) in order to model the IGM/CGM interface and present a more sophisticated model of the HI CGM. We use this model to provide another estimate of the boundary of the CGM.

In Chapter 4, I present a novel data product: Using the Physarum polycephalum slime mold algorithm, MCPM, to reconstruct the cosmic web. We apply the MCPM to much deeper SDSS spectroscopic galaxy surveys than has been previously done. I present the methods used to create this catalog, which was publicly released as part of SDSS DR17 (Abdurro'uf et al., 2022).

Chapter 5 provides a summary and discussion of the work in this manuscript. I also discuss future avenues of study and provide some concluding remarks.

Chapter 2

CGM^2 I: THE EXTENT OF THE CIRCUMGALACTIC MEDIUM TRACED BY NEUTRAL HYDROGEN

Portions of this chapter were originally published in collaboration with Jessica Werk in the May 2021 edition of the Astrophysical Journal (Wilde et al., 2021, ApJ, Vol. 912, Issue 1, id.9; 2021 © American Astronomical Society, DOI: 10.3847/1538-4357/abea14), and are reproduced below with the permission of the American Astronomical Society.

Summary

We present initial results from the COS and Gemini Mapping the Circumgalactic Medium (CGMCGM $\equiv$ CGM^2) survey. The CGM^2 survey consists of 1689 galaxies, all with high-quality Gemini GMOS spectra, within 1 Mpc of twenty-two $z \lesssim 1$ quasars, all with S/N$\sim$10 HST/COS G130M+G160M spectra. For 572 of these galaxies having stellar masses $10^7 M_\odot < M_\star < 10^{11} M_\odot$ and $z \lesssim 0.5$, we show that the H I covering fraction above a threshold of $N_{\rm HI} > 10^{14}{\rm cm}^{-2}$ is $\gtrsim 0.5$ within 1.5 virial radii ($R_{\rm vir} \sim R_{200m}$). We examine the H I kinematics and find that the majority of absorption lies within $\pm$ 250 km s^{-1} of the galaxy systemic velocity. We examine H I covering fractions over a range of impact parameters to infer a characteristic size of the CGM, $R^{14}_{\rm CGM}$, as a function of galaxy mass. $R^{14}_{\rm CGM}$ is the impact parameter at which the probability of observing an absorber with $N_{\rm HI} > 10^{14}$ cm^{-2} is $> 50\%$. In this framework, the radial extent of the CGM of $M_\star > 10^{9.9} M_\odot$ galaxies is $R^{14}_{\rm CGM} = 346^{+57}_{-53}$ kpc or $R^{14}_{\rm CGM} \simeq 1.2 R_{\rm vir}$. Intermediate-mass galaxies with $10^{9.2} < M_\star/M_\odot < 10^{9.9}$ have an extent of $R^{14}_{\rm CGM} = 353^{+64}_{-50}$ kpc or $R^{14}_{\rm CGM} \simeq 2.4 R_{\rm vir}$. Low-mass galaxies, $M_\star < 10^{9.2} M_\odot$, show a smaller physical scale $R^{14}_{\rm CGM} = 177^{+70}_{-65}$ kpc and extend to $R^{14}_{\rm CGM} \simeq 1.6 R_{\rm vir}$. Our analysis suggests that using $R_{\rm vir}$ as a proxy for the characteristic radius of the CGM likely underestimates its extent.

2.1 Introduction

The Circumgalactic Medium (CGM) represents the complex interface between stellar evolution, feedback from super-massive massive black holes, and the cosmic web, dictated by the large scale cosmology of the universe (e.g., ΛCDM) Tumlinson et al. (2017). These phenomena span many orders of magnitude in relevant spatial scales and involve several different sub-disciplines of astrophysics. In the canonical cosmological picture of galaxy evolution of White and Rees (1978), gas accretes onto galactic halos from the intergalactic medium (IGM), cools and condenses to form the interstellar medium, and then eventually stars. At the same time, mass and energy is returned to the CGM via supernovae and AGN feedback, with some fraction potentially deposited back into the IGM.

In spite of its role as the harbor for a significant fraction of galactic baryons (e.g. Davé et al., 2010; Shull et al., 2012; Werk et al., 2014), the CGM is predicted to have low gas densities (n $\ll$ 0.01 cm^{-3}) due to its vast volume. Such low-density gas is most efficiently observed using absorption lines found in the spectra of bright background objects, usually quasars (QSOs). Generally, QSO absorption spectroscopy provides very high sensitivity to extremely low column densities in various ions, down to $N \sim 10^{12}$ cm^{-2}, and access to a broad range of ionized metal transitions that trace gas at both high and low densities in the CGM.

Bahcall and Spitzer (1969) first posited that galaxies are responsible for the high-column density absorption lines in QSO spectra. In the 40 years since the prediction of Bahcall and Spitzer (1969), there has been much effort to link QSO absorption lines to both the IGM and CGM. In particular, the IGM traces the large scale structure of the universe, and is observed mainly through the Lyα forest (see Rauch, 1998 and McQuinn, 2016 for reviews). Previous works have found that the H I Lyα forest is associated with two distinct populations based on the strength of absorption systems; a low-density population that traces the cosmic-web, and a high-column density population that lies within dark matter halos (e.g., Morris et al., 1993; Tripp et al., 1998; Chen et al., 2005; Tejos et al., 2014; Burchett et al., 2020a). Linking sets of absorption lines arising at the same redshift (absorber systems) with redshifts of foreground galaxies in close proximity to the QSO sightline has essentially defined the last 20 years of CGM work (e.g. Chen and Tinker, 2008; Prochaska et al., 2011; Tumlinson et al., 2013; Rudie et al., 2019). Boksenberg and Sargent (1978) and Bergeron and Stasińska (1986) reported some of the first identifications of intervening galaxies responsible for intermediate redshift absorption line systems in quasar spectra. Subsequently, Lanzetta et al. (1995) found that all the luminous galaxies in their sample exhibit extended gaseous envelopes (not yet called CGM), out to at least ∼160 kpc.

Most of the early studies either searched for galaxies associated with known absorbers or they conducted blind surveys that started with no previously known information about either the absorbers or the galaxies at the outset of the survey. However, Bowen et al. (1991, 1995, 2002) used an inverted approach in which they searched for specific absorption features (e.g., Ca II Mg II , or H I) affiliated with galaxies with redshifts and properties that were known prior to the absorption survey. The advantage offered by this approach is that the CGM of galaxies with specific properties of interest could be accumulated more efficiently than in blind or absorption-selected surveys. Unfortunately, while many QSOs can be found behind low-z galaxies at interesting impact parameters, the first generation of HST spectrographs could not go deep enough to access most of those QSOs with a practical allocation of telescope time.

In 2009, the installation of the Cosmic Origins Spectrograph (COS) on HST, 20−30 times more sensitive than its UV spectrograph predecessors, allowed for targeted CGM studies of statistical samples of galaxies nearby in projection to background QSOs. The studies carried out with HST/COS over the last decade have revealed complex connections between the bulk properties of galaxies and their CGM. For example, the presence of O VI in the CGM, 10−150 kpc from the star-forming host galaxy's disk, is somehow linked to the present-day star-forming properties of the disk (Tumlinson et al., 2011). The H I content of the CGM scales with that of the host galaxies' H I disks, perhaps implying that the disk is being fed by accretion from the IGM and CGM (Borthakur et al., 2015). In addition, large quiescent galaxies that have long since ceased star formation were observed to contain rich reservoirs of cool (∼10^4 K) gas (e.g. Thom et al., 2012; Chen et al., 2018) which should provide fuel for new star formation.

Thus far, the empirical picture of the CGM remains somewhat piecemeal, assembled from a variety of targeted observations of small-moderate sized samples carried out by different teams. Generally, various surveys of the CGM at low-z have revealed a multi-phase medium extending as far as $0.5 - 1$ $R_{\rm vir}$ (R_{200m}), which exhibits absorption from H I and a range of metal ions with ionization potential energies ranging from 10 - 239 eV (e.g. Lanzetta et al., 1995; Stocke et al., 2013; Tripp et al., 2011; Muzahid et al., 2011; Tumlinson et al., 2013; Werk et al., 2013; Burchett et al., 2019). Absorption signatures of low-ionization state metals like Mg II and Si II are rarely detected greater than $\sim$ 100 kpc away from their host galaxies (Bordoloi et al., 2011; Nielsen et al., 2013; Werk et al., 2013; Bordoloi et al., 2014). By comparison, highly ionized metal lines such as O VI are detected out to $\sim$ 200 kpc (Tumlinson et al., 2011; Johnson et al., 2015, 2017; Prochaska et al., 2019) and in M31, out to $\sim$ 500 kpc (Lehner et al., 2020). This progress at low-z has been mirrored at z $\sim$2$-$3 with the Keck Baryonic Structure Survey (KBSS; Rudie et al. (2012)), which found typical scales of the CGM to be $\sim$ 300 kpc for L^* galaxies.

Observational progress enabled by HST/COS has presented several challenges to our current models of galaxy evolution. Photoionization modeling of the 10^4 K cool CGM gas around L* galaxies has revealed it to be an order of magnitude under-pressurized with respect to the envisioned 10^6 K virialized, ambient hot halo (e.g. Werk et al., 2014; Prochaska et al., 2017; Haislmaier et al., 2020), and thus likely not in thermal or pressure equilibrium as expected from simulations or simple theoretical arguments. Complex, non-equilibrium physical processes may include precipitation (e.g. Voit, 2019), shocks (e.g. McQuinn and Werk, 2018), cosmic rays (e.g. Ji et al., 2019), kpc-scale fountain cycles (e.g. Fraternali and Binney, 2008; Kim and Ostriker, 2018), multi-filament gas flows fueling star formation (Martin et al., 2019), galactic winds (Bordoloi et al., 2011; Burchett et al., 2020c; Huang et al., 2020) and turbulent mixing layers (Fielding et al., 2020). Furthermore, simulations have generally under-predicted observed column densities of various ions, of both high-and-low ionization potential, compared to observations of galactic halos (e.g. Stinson et al., 2012; Hummels et al., 2013; Oppenheimer et al., 2016; Liang et al., 2016; Fielding et al., 2017). Newer simulations (e.g. Oppenheimer et al., 2018; Lehner et al., 2020; Nelson et al., 2020) have more faithfully reproduced observations.

Recent simulations have shown that the CGM content, thermal structure, and kinematic properties are highly dependent on the mass and spatial resolution of simulations (e.g. van de Voort et al., 2019). High-resolution simulations that focus explicitly on the CGM, with resolved spatial scales of $\sim$100 pc and mass scales of $\sim 10^3$ $\mathrm{M}_\odot$ out to 250 kpc, have helped to ease the discrepancies with the observed column densities (Peeples et al., 2019b; Hummels et al., 2019), more naturally producing cool, denser clumps. However, it remains important to characterize the full extent of the CGM, as simulations still struggle to recreate all key aspects of the extended CGM.

For the above reasons, it has been challenging to put quantitative, physically-motivated bounds on the radial extent of the CGM. Generally, statistical QSO-galaxy samples are limited to single sightlines per galaxy due to the rarity of QSOs (but see Chen et al., 2014; Bowen et al., 2016) and thus assembling large enough samples to place constraints on the extent of the CGM is challenging work. A comprehensive $z < 1$ CGM survey requires both space and ground-based spectroscopy. The former are required to measure key far-ultraviolet transitions such as Lyα, Lyβ and metal species (e.g., C II, Si II, C IV, O VI). At low redshift, only Ca II, Na I, Fe II, and Mg I and Mg II can be observed from the ground. However, these species are confused by ionization and dust depletion and are hard to interpret by themselves. On the other hand, large samples of galaxies are accessible to ground based spectrographs to measure precise ($z_{\mathrm{err}} < 10^{-4}$) redshifts of potentially associated galaxies.

To fill this need for a comprehensive, uniform CGM study, we have completed a deep spectroscopic survey of >1000 galaxies around $z > 0.5$ QSOs, all within a few arcminutes of the background QSO sightline. The galaxies and absorption systems are blindly selected such that no preference to galaxy type or absorber optical depth is explicitly imposed. Our survey, COS and Gemini Mapping the Circumgalactic Medium (CGMCGM $\equiv$ CGM^2), was designed to address many puzzles related to the CGM, such as: (i) What is the physical state of the gaseous halo when galaxies quench their star formation? (ii) What effects do environment and merger history have on the physical state and content gaseous halo? (iii) How do galaxy properties relate to the metal content and kinematic structure of their gaseous halos and vice versa? (iv) What is the physical extent of the CGM? In this first presentation of the CGM^2 survey, we address the last question and in particular, how its spatial extent may depend upon halo mass and galaxy type. Ultimately, this study will inform how galaxy evolution and the large scale structure of the universe are connected.

CGM^2 includes a significant investment of effort from a team of University of Washington Undergraduate Astronomy majors led by CGM^2 PI Jessica Werk, the Student Quasar Absorption Diagnosticians (aka the Werk SQuAD) whose work is described in detail in the relevant sections. The outline of this chapter is as follows. Section 2.2 describes the observations and data reduction for the CGM^2 spectroscopic galaxy survey. Section 2.3 describes an analysis of the galaxies' spectroscopic and photometric properties. In Section 2.4, we discuss our initial empirical results from associating galaxies with absorbers. We then derive a characteristic physical scale for the CGM in Section 2.5 by examining the H I -galaxy clustering. Section 2.6 compares our results to previous surveys of similar design and hydrodynamical cosmological simulations. We summarize and conclude this work in Section 2.7. Throughout this analysis we adopt the Planck15 (Planck Collaboration et al., 2016) cosmology as encoded in the ASTROPY package (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018). All distances are in physical space (not co-moving) unless otherwise noted.

2.2 Observations and Data Reduction

2.2.1 QSO Sample Selection

The CGM^2 survey is built upon the COS-Halos (GO11598, GO13033; Tumlinson et al., 2013) and COS-Dwarfs (GO12248; Bordoloi et al., 2014) surveys, which used 263 orbits of COS time to observe ~ 80 QSOs with $z = 0.2 - 1.0$. These surveys were designed to probe the halo(s) of one or two foreground galaxies well inside $R_{\rm vir}$ using suitable QSO-galaxy pairs. The COS-Halos QSO catalog was selected from the SDSS DR5 quasar catalog for QSOs that are UV-bright (GALEX FUV $\lesssim 18$) and lie at $z \lesssim 1$. The target QSOs were further selected to avoid Mg II absorbers at $z > 0.4$ to avoid losing a large range of the FUV QSO spectrum to LLSs. While this criterion did not affect the original surveys, which typically targeted specific galaxies at $z < 0.4$, it does affect the CGM^2 sample because it selects against high column density systems at $z \gtrsim 0.4$.

COS-Halos galaxies were selected via photometric redshifts to target stellar masses of $M_\star \simeq 10^{10-11} M_\odot$ at $0.1 < z < 0.3$ with projected separations from a nearby QSO $\rho < 150$ kpc. The precise redshifts (and other galaxy properties) were subsequently constrained with follow-up spectroscopy (Werk et al., 2012). The COS-Dwarfs survey was designed to probe the CGM of galaxies with $\log M_\star / M_\odot \lesssim 10^{10}$ at z $\simeq 0.01 - 0.05$ with masses and redshifts derived via SDSS photometry and spectroscopy. Both surveys yield galaxies that are the closest spectroscopically-identified galaxy to each QSO sightline and were designed to avoid biases with respect to galaxy neighbors, large-scale environment, or status as a satellite of a larger halo (in cases where neighbors were known).

For the CGM^2 multislit spectroscopic follow-up, we targeted a subset of "high-value" QSO fields with our Gemini program, either those with $z_{\rm QSO} > 0.6$ or those with HST imaging available, which allows us to obtain morphology of any absorption-hosting galaxies with $z < 0.5$. To-date, we have obtained multislit galaxy spectra in 22 QSO fields as part of CGM^2 . The fields we have surveyed, and some basic properties of the background QSOs, are tabulated in Table 2.1.

COS Data Reduction

The COS spectra were taken using both the G130M and G160M gratings. The balance between the time allocated to G130M and G160M was designed to achieve S/N $\simeq 10 - 12$ per resolution element (FWHM $\simeq$ 16-18 km s^{-1}) or better over 1150-1800 Å. The reduction of the COS spectra is explained in detail in Tumlinson et al. (2013) and Bordoloi et al. (2014) and follows the same method used by Tripp et al. (2011), Meiring et al. (2011), Tumlinson et al. (2011) and Thom et al. (2012) but we provide a brief description here. The COS data were obtained from MAST[1] and processed via CALCOS (v2.12) (Holland, 2012)

---

[1] http://archive.stsci.edu

Table 2.1. QSO Fields in CGM^2

QSO	QSOLong	RA	Dec	z_{QSO}	m$_g$
J0226+0015	J022614.46+001529.7	36.56028	0.25827	0.615	17.15
J0809+4619	J080908.13+461925.6	122.2839	46.3238	0.657	16.54
J0843+4117	J084349.49+411741.6	130.9562	41.2949	0.99	17.31
J0914+2823	J091440.38+282330.6	138.66829	28.39184	0.735	17.79
J0935+0204	J093518.19+020415.5	143.82581	2.07098	0.649	16.99
J0943+0531	J094331.61+053131.4	145.88173	5.52541	0.564	17.16
J1001+5944	J100102.55+594414.3	150.2606	59.7373	0.746	16.08
J1016+4706	J101622.60+470643.3	154.09418	47.11204	0.822	17.12
J1022+0132	J102218.99+013218.8	155.57913	1.53856	0.789	16.75
J1059+1441	J105945.23+144142.9	164.9385	14.6953	0.631	16.93
J1059+2517	J105958.82+251708.8	164.9951	25.2858	0.662	17.39
J1112+3539	J111239.11+353928.2	168.16296	35.65784	0.636	17.73
J1133+0327	J113327.78+032719.1	173.36578	3.45533	0.525	17.54
J1134+2555	J113457.62+255527.9	173.7401	25.9244	0.71	16.8
J1233-0031	J123304.05-003134.1	188.26688	-0.52616	0.471	17.76
J1241+5721	J124154.02+572107.3	190.4751	57.35205	0.583	17.58
J1342-0053	J134251.60-005345.3	205.71503	-0.89592	0.326	16.92
J1419+4207	J141910.20+420746.9	214.79251	42.1297	0.873	17.04
J1437+5045	J143726.14+504555.8	219.35892	50.76551	0.783	17.57
J1553+3548	J155304.92+354828.6	238.27052	35.80795	0.722	16.46
J1555+3628	J155504.39+362848.0	238.76833	36.48001	0.714	17.76
J2345-0059	J234500.43-005936.0	356.2518	-0.99335	0.789	16.8

Note. — The full sample of QSOs included in the CGM^2 Survey: (1) QSO Short Name; (2) QSO Long Name, RA in hms and Dec in dms, all J2000; (3 & 4) RA, Dec in decimal degrees; (5) QSO redshift; (6) SDSS g-band magnitude of QSO.

Table 2.2. CGM^2 Multislit Mask Observations

QSO	Maskname	Date	Instrument	N(slits)	N(z)
J0809+4619	GN2014AQ001-01	2014-11-19	GMOS-N	31	24
J0809+4619	GN2014AQ001-02	2014-11-20	GMOS-N	27	21
J0809+4619	GN2014AQ001-03	2014-11-21	GMOS-N	29	21
J0809+4619	GN2014AQ001-04	2014-11-23	GMOS-N	17	15
J1134+2555	GN2014AQ001-05	2014-06-21	GMOS-N	29	26
J1134+2555	GN2014AQ001-06	2014-06-22	GMOS-N	29	26
J1134+2555	GN2014AQ001-07	2014-06-25	GMOS-N	24	20
J1134+2555	GN2014AQ001-08	2014-12-21	GMOS-N	16	13
J1241+5721	GN2014AQ001-09	2014-06-25	GMOS-N	27	18
J1241+5721	GN2014AQ001-10	2014-06-30	GMOS-N	24	23
J1555+3628	GN2014AQ001-11	2014-06-24	GMOS-N	33	29
J1555+3628	GN2014AQ001-12	2014-06-24	GMOS-N	26	24
J0914+2823	GN2014BLP003-01	2015-01-17	GMOS-N	45	29
J0914+2823	GN2014BLP003-02	2015-01-17	GMOS-N	44	27
J0914+2823	GN2014BLP003-03	2015-01-17	GMOS-N	45	26
J0843+4117	GN2014BLP003-04	2015-01-16	GMOS-N	39	28
J0843+4117	GN2014BLP003-05	2015-01-16	GMOS-N	39	31
J0843+4117	GN2014BLP003-06	2015-01-16	GMOS-N	39	31
J1059+1441	GN2014BLP003-08	2015-01-16	GMOS-N	45	38
J1001+5944	GN2014BLP003-10	2015-01-18	GMOS-N	45	36
J1001+5944	GN2014BLP003-11	2015-01-18	GMOS-N	45	29
J1001+5944	GN2014BLP003-12	2015-01-18	GMOS-N	41	33
J1016+4706	GN2014BLP003-13	2015-01-19	GMOS-N	42	34
J1016+4706	GN2014BLP003-14	2015-01-19	GMOS-N	40	30
J1016+4706	GN2014BLP003-15	2015-01-19	GMOS-N	42	33
J1112+3539	GN2014BLP003-19	2015-01-17	GMOS-N	49	33
J1112+3539	GN2014BLP003-20	2015-01-17	GMOS-N	46	31
J1112+3539	GN2014BLP003-21	2015-01-17	GMOS-N	43	37
J1059+2517	GN2015ALP003-01	2015-05-18	GMOS-N	47	31
J1059+2517	GN2015ALP003-02	2015-05-20	GMOS-N	43	29
J1059+2517	GN2015ALP003-03	2015-06-08	GMOS-N	42	25
J1419+4207	GN2015ALP003-04	2015-05-17	GMOS-N	45	27
J1419+4207	GN2015ALP003-05	2015-05-17	GMOS-N	46	33
J1419+4207	GN2015ALP003-06	2015-05-17	GMOS-N	43	32
J1419+4207	GN2015ALP003-07	2015-05-17	GMOS-N	41	25
J1437+5045	GN2015ALP003-08	2015-05-24	GMOS-N	40	33
J1437+5045	GN2015ALP003-09	2015-05-22	GMOS-N	43	31
J1437+5045	GN2015ALP003-10	2015-06-19	GMOS-N	42	30
J1553+3548	GN2015ALP003-11	2015-05-17	GMOS-N	45	32
J1553+3548	GN2015ALP003-12	2015-05-18	GMOS-N	45	35
J1553+3548	GN2015ALP003-13	2015-05-18	GMOS-N	41	28

Table 2. CGM^2 Multislit Mask Observations

QSO	Maskname	Date	Instrument	N(slits)	N(z)
J0943+0531	GS2014AQ002-01	2015-04-19	GMOS-S	19	13
J0943+0531	GS2014AQ002-02	2015-04-18	GMOS-S	7	3
J0943+0531	GS2014AQ002-03	2015-02-18	GMOS-S	28	22
J0943+0531	GS2014AQ002-04	2015-02-19	GMOS-S	18	9
J1133+0327	GS2014AQ002-05	2015-03-15	GMOS-S	32	27
J1133+0327	GS2014AQ002-06	2015-04-25	GMOS-S	24	18
J1133+0327	GS2014AQ002-07	2015-04-25	GMOS-S	26	21
J1133+0327	GS2014AQ002-08	2015-05-14	GMOS-S	22	17
J1233-0031	GS2014AQ002-09	2015-04-19	GMOS-S	30	19
J1233-0031	GS2014AQ002-10	2015-05-14	GMOS-S	29	24
J1342-0053	GS2014AQ002-11	2015-01-04	GMOS-S	31	18
J1342-0053	GS2014AQ002-12	2015-06-06	GMOS-S	29	22
J0226+0015	GS2014BLP004-01	2014-10-26	GMOS-S	47	40
J0226+0015	GS2014BLP004-02	2014-10-26	GMOS-S	44	32
J0226+0015	GS2014BLP004-03	2014-10-26	GMOS-S	43	29
J2345-0059	GS2014BLP004-04	2015-06-19	GMOS-S	41	25
J0935+0204	GS2014BLP004-07	2015-02-13	GMOS-S	45	26
J0935+0204	GS2014BLP004-08	2015-02-13	GMOS-S	32	20
J0935+0204	GS2014BLP004-09	2015-02-17	GMOS-S	38	25
J1022+0132	GS2015ALP004-01	2015-02-17	GMOS-S	40	26
J1022+0132	GS2015ALP004-02	2015-02-17	GMOS-S	38	30
J1022+0132	GS2015ALP004-03	2015-02-17	GMOS-S	36	27
J1022+0132	GS2015ALP004-04	2015-02-17	GMOS-S	34	23

Note. — Multislit Observation with GMOS taken as part of the CGM^2 Survey: (1) QSO short name; (2) Unique program mask name with project ID; (3) Date of mask observation; (4) Instrument, either GMOS-N or GMOS-S; (5) Number of slits placed on each mask; (6) Number of slits that yielded a reliable redshift.

with standard parameters and reference files. First, exposures taken at the same grating and CENWAVE were combined. These coadded spectra were then coadded with exposures in the same grating at different CENWAVEs. This was followed by a combination of the two grating spectra to produce a single one-dimensional (1D) trace from 1150-1800 Å. At each coadd, exposures were combined by aligning common Milky Way interstellar absorption lines. Features related to the design of COS show up in the raw data and must be removed. The photocathode grid wires cast shadows on the detector as well as other fixed-pattern noise features and were removed with a flatfield and moreover are mitigated by the use of FP SPLITS. Flat-field reference files, prepared and communicated to us by D. Massa at STScI and filtered for high-frequency noise by E. Jenkins, were used to correct these fixed pattern features. These flats do not however, correct for gain-sag depressions in the spectra created by prolonged exposure to bright geocoronal emission lines (Sahnow et al., 2011). The affected regions are flagged by the CALCOS pipeline and are rejected in our coaddition process. The 1D spectra are binned to Nyquist sampling resulting in a 1D, flat-corrected spectra with two bins per COS resolution element (FWHM $\simeq$ 18 km s^{-1}) and a S/N $\sim$ 8-10. Errors arising from counting statistics (Poisson) are propagated through each step in the calibration.

2.2.2 Spectroscopic Galaxy Survey

Survey Design

All galaxy spectra were obtained at the Gemini North and South Observatories using the GMOS instrument (Hook et al., 2004; Gimeno et al., 2016) in multi-object spectrograph mode. The GMOS observations enabled us to obtain accurate, precise redshifts for low-redshift ($0.1 \lesssim z \lesssim 1$) galaxies as faint as $g \sim 24.5$ for star-forming galaxies and $g \sim 24$ for early-type galaxies. Over 3 observing semesters from 2014−2015 at both GMOS-N and GMOS-S, we surveyed galaxies in the high-quality QSO fields listed in Table 2.1.

GMOS uses slit masks that must be cut for each field based on highly-accurate astrometry from either a 'pre-image' taken in GMOS imaging mode or derived from an accurate catalog. For most of our fields, we used Gemini GMOS pre-imaging to create the masks. However, for fields observed in 2014A, specifically J0809+4619, J0943+0531, J1133+0327, J1134+2555, J1241+5721, and J1555+3628, we used a combination of HST-WFC3 imaging for angular separations less than $1'$ from the QSO along with SDSS imaging for galaxies further from the sightline to construct a slit mask target catalog. We refer to the observational program that observed these fields as the "Rollup" program and are differentiated in Table 2 with a 'Q' in the Maskname column. Subsequent observations used Gemini-GMOS g and i band pre-imaging to create the slit mask target lists and are referred to as "Large Program" (LP) observations. From the pre-imaging, we derived astrometry as well as magnitudes from SExtractor (Bertin and Arnouts, 1996) using Gemini-calibrated photometric zero points and color corrections. We then optimized to get as many $g < 24$ galaxies into each slit mask as possible (see below).

In our GMOS programs, we used the R400 grating with $1''$ slits to balance wavelength coverage with spectral resolution. The fields from the Rollup programs listed above were dithered across two exposures per mask at central wavelengths of $\lambda\lambda 6000, 6900$Å in order to avoid losing information to the chip gaps. The LP fields have spectra consisting of three exposures per mask at three separate central wavelengths, $\lambda\lambda 6900, 7000, 7100$Å. Despite these differences, both programs yielded similar depths with total exposure times of 1 hour per mask. Our final spectra achieved a S/N of at least a few per pixel at $\lambda_{obs} = 4800$Å, the approximate wavelength of the 4000Å break at $z \sim 0.2$.

Because precise redshift determination from our spectra was the primary goal, we chose a grating to cover 4400 to 9000 Å in order to detect emission from $z \lesssim 0.5$ [OII], Hβ (for both redshifts and star formation metrics) and Hα as well as absorption from Ca H+K and Na D in passively-evolving galaxies. The R400 grating provided a spectral resolving power of $R = \lambda/\Delta\lambda \simeq 950$, which corresponds to a velocity resolution of 300 km s^{-1} per resolution element. Ultimately, this setup allowed for determining the velocity centroid to ~ 50 km s^{-1} for precise redshift determination. We note that every spectrum does not exhibit uniform wavelength coverage given the design of GMOS. Galaxies that are not placed near the center of the field have redder or bluer coverage than the range quoted above.

The slit masks were designed such that slits were placed on $\sim 85\%$ of objects within 1 arcminute of the QSO, with additional slits placed to fill the 5.5×5.5 arcminute area of the detector. Slit placement constraints meant that $\sim 40-50$ slits could be placed on one mask. We aimed to obtain spectra of $\sim$ 80-120 unique galaxies per QSO field. Some fields in the earlier Rollup programs (those observed in 2014A) had as few as two and as many as four masks per field. In the later observing campaigns, all fields (except J1059+1441, which was only observed with one mask due to problems in the mask making) have three masks per field. This is shown in Table 2. An example of the targeting strategy can be seen in Figure 2.1 for the field J0843+4117. The red circles highlight the galaxies with $z < z_{\rm QSO}$ for which we obtained reliable redshifts while the white circles were targeted for spectra but for which we did not recover a reliable redshift.

All masks were observed between June 2014 and June 2015 on either GMOS-N or GMOS-S. Table 2 provides details of each multislit mask observed. Columns 5 and 6 list the number of slits placed and the number of slits that yielded reliable redshifts, respectively. A summary of the completeness of our galaxy spectroscopic survey is shown Table 3 for each field as a function of angular separation from the central QSO and limiting magnitude. The completeness is given as the fraction of reliable, redshift-yielding spectra, with 'ZQ' > 2 of galaxies (CLASSSTAR < 0.5) by SExtractor within some angular separation to the QSO. The fields J2345-0059 and J1059+1441 have only one mask per field and thus have a lower completeness relative to the rest of the sample. In general, we are nearly 100% complete out to 2 arcminutes from the QSO for galaxies with an i-band magnitude of < 22. The completeness drops to $\sim$50% for a limiting magnitude of $i < 23$. 68% of the galaxies used in this analysis lie within 2 arcminutes of the central quasar. Details of the survey design, completeness and their impact on scientific results will be discussed in the full presentation of the CGM^2 survey, currently in preparation (Werk et al. 2021).

Table 3. Completeness

QSO	$C_{22,1'}$	$C_{22,2'}$	$C_{22,\mathrm{FOV}}$	$C_{23,1'}$	$C_{23,2'}$	$C_{23,\mathrm{FOV}}$
J0226+0015	0.57	0.79	0.47	0.34	0.41	0.23
J0809+4619	1.00	1.00	1.00	0.77	0.70	0.37
J0843+4117	1.00	1.00	0.63	0.70	0.51	0.31
J0914+2823	1.00	1.00	1.00	0.80	0.63	0.40
J0935+0204	0.93	0.86	0.42	0.33	0.40	0.20
J0943+0531	1.00	1.00	0.54	0.61	0.42	0.20
J1001+5944	1.00	1.00	0.65	0.87	0.53	0.32
J1016+4706	1.00	1.00	0.55	0.81	0.62	0.31
J1022+0132	1.00	0.97	0.79	0.59	0.51	0.35
J1059+1441	0.60	0.47	0.32	0.24	0.24	0.15
J1059+2517	1.00	1.00	0.68	0.72	0.69	0.36
J1112+3539	1.00	1.00	0.49	0.54	0.43	0.22
J1133+0327	1.00	0.95	0.48	0.68	0.46	0.23
J1134+2555	1.00	0.99	0.66	0.70	0.57	0.33
J1233-0031	1.00	0.44	0.25	0.43	0.19	0.11
J1241+5721	1.00	0.79	0.35	0.90	0.36	0.14
J1342-0053	1.00	0.58	0.29	0.54	0.30	0.14
J1419+4207	1.00	1.00	1.00	1.00	1.00	0.76
J1437+5045	1.00	1.00	0.66	1.00	0.63	0.33
J1553+3548	1.00	1.00	0.82	1.00	0.74	0.40
J1555+3628	1.00	1.00	0.62	0.67	0.53	0.24
J2345-0059	0.28	0.40	0.16	0.16	0.19	0.09
Median	1.00	1.00	0.57	0.69	0.51	0.27

Note. — Completeness of the CGM^2 galaxy catalog along with the median in each radial and magnitude limited bin. The completeness is defined as the fraction of reliable spectra, with 'ZQ' > 2, to objects designated as a galaxy (CLASSSTAR < 0.5) by SExtractor (1) QSO Field; (2) Completeness within 1' of the QSO at a limiting *i*-band magnitude of 22; (3) Completeness within 2' of the QSO at a limiting *i*-band magnitude of 22; (4) Completeness within the GMOS field of view of the QSO at a limiting *i*-band magnitude of 22; (5) Completeness within 1' of the QSO at a limiting *i*-band magnitude of 23; (6) Completeness within 2' of the QSO at a limiting *i*-band magnitude of 23; (7) Completeness within the GMOS field of view of the QSO at a limiting *i* -band magnitude of 23. The fields J2345-0059 and J1059+1441 have only one mask per field and thus have a low completeness relative to the rest of the sample.

Gemini Data Reduction

The spectra were reduced using a combination of Gemini's PyRAF package and PypeIt[2] (Prochaska et al., 2020). The initial reduction closely follows the GMOS Data Reduction Cookbook[3]. After obtaining the raw spectra from Gemini Observatory Archive, the data were organized according to field and mask. Biases were created by downloading all bias exposures of the same 2×2 binning taken within ~ 1 month from the observations if they were the same detector. Flat field and NeAr wavelength calibration exposures were taken along with each mask and were prepared using the standard methods in the cookbook. This procedure performs the bias subtraction, performs automatic slit edge finding, cuts the slits out of the image and isolates them, flat-fields each slit with the flats taken at the same central wavelength, and then performs a wavelength transformation to each slit. This results in multiple 2-D spectra. In general, the typical RMS of the wavelength solution is 0.1 pixel, which corresponds to 0.04Å given the dispersion of GMOS of 0.4Å per pixel. We then turned to PypeIt to perform sky subtraction, spectral continuum tracing, and combination of the wavelength-dithered. The spectra were then flux calibrated using a sensitivity function based on a selection of spectro-photometric flux calibration standards, BD28+4211, EG21, EG131, and Wolf1346, choosing whichever was closest on the sky. The final 1D spectra were co-added in vacuum wavelength space weighted by the inverse variance of the individual exposures.

2.3 Analysis

2.3.1 Galaxy Redshift Determination

The process of determining galaxy redshifts was done in two stages. First, each 1D extracted spectrum was passed through an automated redshift fitting code, REDROCK [4] (v0.14).

---

[2] https://github.com/pypeit/PypeIt

[3] Shaw, Richard A. 2016, GMOS Data Reduction Cookbook (Version 1.2; Tucson: National Optical Astronomy Observatory), available online at:http://ast.noao.edu/sites/default/files/GMOS_Cookbook/

[4] https://github.com/desihub/redrock

REDROCK was developed by the DESI team and uses a template fitting algorithm to generate a set of ranked best-fitting models, identifying the object's type (QSO, Star, Galaxy) and corresponding redshift.

However, much of our galaxy sample has moderate to poor S/N, and the automated REDROCK redshift guesses can thus fail catastrophically by fixating on spurious features. We constructed a method of manually vetting the REDROCK redshifts by eye using a custom GUI, VETRR[5]. Each redshift was visually assessed and assigned a quality flag, Z_Q, by one of the authors, and by at least two members of the Werk SQuAD. A Z_Q of either 0, 1, 3, 4 were assigned to each spectrum, where a $Z_Q = 0$ indicates the spectrum has a S/N that is too low to be useful. $Z_Q = 1$ denotes a good spectrum but we are not confident in identifying a redshift (i.e. has no clear absorption or emission lines). $Z_Q = 3$ are spectra with one absorption or emission line that was confidently identified. These are usually strong [OII], Hα emission or weak CaII absorption. A solitary Hα emission only falls into this category if Hβ is off the detector and the strong emission line is too narrow to be [OII], which is a marginally unresolved doublet in these data. $Z_Q = 4$ represents a spectrum for which we are most confident in the redshift, with at least two absorption or emission lines identified. $Z_Q = 2$ was not used. The fully vetted galaxy survey database that we use for our analysis contains only spectra with $Z_Q > 2$. The statistical uncertainly of our redshift identification was determined by computing the standard deviation of redshift identifications from at least three humans for a sample of 50 galaxies and is typically in the range of $\sigma_z \sim$ 50-100 km s^{-1} ($z \sim$ 0.00016-0.00030). The redshift distribution of the vetted galaxy database containing 971 galaxies with redshifts less than that of the field quasar is shown in Figure 2.2.

In addition to manually vetting the galaxy redshifts, we also visually identified a galaxy spectral type during the process of examining each 1D spectrum. If strong emission lines and weak continuum were present, we classified the galaxy as "star forming" or "SF". If only absorption lines along a detected continuum were found, we classified the galaxy as

---

[5] https://github.com/mattcwilde/vetrr

“quiescent” or “E”. If both emission and absorption lines were identified, we classified the galaxy as a combination of the two, “SF+E.” In several cases, stars were mistakenly targeted in our slit masks, and they have the identification as star. Stars are never included as part of our vetted galaxy database. Examples of galaxy spectra of each type are shown in the three panels of Figure 2.3. These spectra contain some poorly-subtracted sky lines, and telluric absorption lines at ∼7600Å, which are the sorts of spurious spectral features that can cause REDROCK to fail. The top panel displays an emission-line ‘SF’ galaxy that shows emission from several highlighted strong emission lines, the middle panel shows a combination-type, ‘SF+E’ spectrum with both [OII] emission and strong CaII absorption against a bright stellar continuum, and the bottom panel shows an example of an absorption-only spectrum, type ‘E.’

2.3.2 Galaxy Photometry and Spectral Energy Distribution Fitting

In addition to the galaxy spectroscopic catalog, we constructed a photometric galaxy catalog to derive stellar masses and galaxy star formation rates (SFRs). Our spectra are generally insufficient in signal and flux calibration to analyze them via spectral fitting codes (e.g. Cappellari, 2017). To estimate stellar masses, we used CIGALE (Noll et al., 2009; Boquien et al., 2019) to fit the spectral energy distribution and retrieve stellar mass and SFRs. We note that while there are myriad other SED fitting codes available (as well as direct color-mass relations) we chose CIGALE for its options for stellar population models, dust models, etc. across large swaths of parameter space. We also chose it to compare directly to complementary surveys such as CASBaH (Burchett et al., 2019; Prochaska et al., 2019).

Galaxy Photometry

One major challenge in constructing a photometric catalog for the CGM^2 survey was that the spectroscopic target catalog is generally fainter than available public all-sky surveys and thus the photometric coverage is not uniform in all fields. We chose to gather photometric data for every spectroscopic target in the galaxy catalog, totalling 2310 unique targets.

We created the photometric catalog by cross-matching our target to the DESI Legacy Imaging Surveys Data Release 8 (DR8) (Dey et al., 2019). The imaging survey for DESI is composed of data from three telescopes covering $\sim$ 14000 deg^2 over $-18° < \delta < +84°$ ($|b| > 18°$). These three Programs include The Beijing-Arizona Sky Survey (BASS), The DECam Legacy Survey (DECaLS), and the Mayall z-band Legacy Survey (MzLS) which provide g, r, and z band photometry to $\sim$ 23.3 mag.

The photometry is corrected for Galactic extinction. This is the deepest publicly available optical survey and provided the bulk of the photometry. In cases of overlap between the North and South catalogs, we chose the Southern DECaLS observations. We limited matches to objects with S/N > 2 and chose the closest match within 1.3 arcseconds of our targets in order to limit mismatches between our faint sources and the Legacy Survey catalogs. This gives us 1985 targets with at least one band of photometry. In addition to the g, r, and z bands, DR8 provides cross-matched WISE (Cutri et al., 2013) observations in 3.4, 4.6, 12, and 22 μm.

In order to cover a larger wavelength range to better estimate the SED, we also cross-matched our catalog with the Pan-STARRS Data Release 2 (Chambers et al., 2016) with coverage of the $grizy$ bands, utilizing the MAST cross-match service, using a 1.3 arcsecond threshold. We limited photometry to those marked as extended objects with good stack photometry and $grizy <$ 23.3, 23.2, 23.1, 22.3, 21.3., giving 393 objects with photometry in at least one band.

In addition we also queried the SDSS DR14 (Abolfathi et al., 2018) survey with $ugriz$ coverage where we restricted matches to less than $ugriz <$ 22.15, 23.13, 22.70, 22.20, 20.71, totalling to 331 targets with photometry in at least one band.

To make sure we included photometry for every object in the CGM^2 survey that may be too faint or in crowded areas for the public surveys, we included photometry from the GMOS imaging. The target selection and slit mask design was based on Gemini GMOS i-band and g-band imaging except in the case of the Rollup programs, GN-2014A-Q1 and GS-2014A-Q2, which involved a combination of HST WFC-ACS and Gemini imaging. This allowed us to

Table 4. CGM^2 Gemini GMOS Imaging

Project ID	g-t_{exp} [s]	i-t_{exp} [s]
GN-2014A-Q1	X	150
GS-2014A-Q2	X	150
GN-2014B-LP-4	450	200
GS-2014B-LP-3	450	200
GN-2015A-LP-3	450	200
GS-2015A-LP-4	450	200

Note. — Imaging observations with GMOS-N and GMOS-S taken as part of the CGM^2 Survey: (1) Project ID; (2) g-band exposure time in seconds; (3) i-band exposure time in seconds

use the Gemini images to get photometry for each field, the details of which are presented in Table 4. Two exposures in each band were combined and processed by Gemini. We obtained g and i band magnitudes of the Gemini sources using the SExtractor software. We used pre-calibrated photometric zero-points and color corrections for GMOS-N and GMOS-S for an initial pass on the Gemini photometry. In order to further calibrate the Gemini photometry, we cross-matched our m<21 sources to the g and i band SDSS photometric sources, and then bootstrapped the Gemini photometry below the SDSS magnitude limit within each field. This consisted of applying a constant magnitude offset to the Gemini sources to match the SDSS photometry in the g and i bands.

All photometry from Gemini, SDSS, and Pan-STARRS was corrected for Galactic reddening based on the values in Schlafly and Finkbeiner (2011) provided by the NASA Extragalactic Database[6] by querying our targets coordinates via ASTROQUERY[7]. We employed

---

[6] The NASA/IPAC Extragalactic Database (NED) is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology.

[7] http://dx.doi.org/10.6084/m9.figshare.805208

the SVO Filter Profile Service (Rodrigo et al., 2012) to obtain our filter transmission curves for the telescopes used in these surveys as required input to CIGALE.

CIGALE includes many models as options to include in fitting. For stellar populations, we used the Bruzual and Charlot (2003) models, assuming a Chabrier (2003) initial mass function (IMF). We chose a grid of metallicities ranging from 0.001-2.5$Z_\odot$. We used a delayed star formation history (SFH) model with an exponential burst. The e-folding time of the main stellar population models ranged from 0.1-8 Gyr. We varied the age of the oldest stars in the galaxy from 2-12 Gyr. We included an optional late burst with an e-folding time of 50 Myr and an age of 20 Myr. We varied the burst mass fraction from 0.0 or 0.1 to turn this feature on or off. Nebular emission and reprocessed dust models (Dale et al., 2014) were also included with the default values. The dust models have slopes ranging from $1 - 2.5$ and the nebular models include no active galactic nuclei.

We employed the Calzetti et al. (1994) dust attenuation law, but we also included a "bump" in the UV (see discussion in Prochaska et al. (2019)) at 217.5 nm with a FWHM of 35.6 nm. The bump amplitude is set at 1.3 and the power law slope is -0.13 (Lo Faro et al., 2017). We varied the color excess of the stellar continuum from the young population, E(B-V), from 0.12-1.98. Finally, we used a reduction factor of 0.44 to the color excess for the old population compared to the young stars.

2.3.3 Galaxy Stellar Masses and Derived Properties

After fitting each galaxy's SED, CIGALE then outputs several useful parameters including stellar mass, SFR, and the rest-frame absolute luminosity in each band. Figures 2.4 and 2.5 show the resultant mass distributions with masses spanning $M_\star \approx 10^6 - 10^{11} M_\odot$ and $\bar{M}_\star = 10^{9.3} M_\odot$ at $\bar{z} = 0.44$. In order to calculate the virial radius of the galaxies, we first calculate the halo mass using the abundance matching method of Moster et al. (2013) with the modifications used in Burchett et al. (2016). We adopt R_{200m}, the radius within which the average mass density is 200 times the mean matter density of the universe, as the virial radius ($R_{\rm vir}$) of a galaxy halo.

Figure 2.4 shows the CIGALE-derived SFRs vs stellar masses, which exhibit, as expected, a positive correlation known as the “star-forming main-sequence” (SFMS). We compare our models to the redshift dependent fit from Schreiber et al. (2015) over a range of galaxies spanning $z = [0.23, 0.63]$. This range represents the 16th and 84th percentiles of the CGM^2 galaxy catalogs redshift distribution. Typical uncertainties on photometry-derived stellar masses range from a factor of $3-5$ (e.g. Blanton and Roweis, 2007; Werk et al., 2012) corresponding to ~ 0.5 dex, and errors on CIGALE-derived SFRs are similar.

Figure 2.5 shows galaxy stellar masses vs. galaxy systemic redshift, and differentiates among our three galaxy spectral types visually derived from the GMOS spectra. Above z ~0.5, we are no longer detecting galaxies with $M_\star \lesssim 10^8$ $M_\odot$. Both SF and SF+E galaxy spectral types are preferentially distributed among lower stellar masses as seen in the marginal distributions. We find that, as expected, there are more ‘E’ type galaxies at higher galaxy stellar masses than a random distribution would predict. As shown in Figure 2.6, the CIGALE-derived color-mass diagram of our galaxy sample shows the bimodality of the star-forming and non-star-forming galaxies found in large galaxy surveys such as SDSS (e.g. Chang et al., 2015).

2.3.4 Absorber Catalog

To identify absorption features in the QSO spectra, the HST/COS UV spectra were visually inspected by members of the Werk SQuAD in a multi-step process which closely follows the procedure described in Tejos et al. (2014) and is designed to leave no QSO absorption feature unidentified. This process includes identifying all of the most ubiquitous metal ions present in QSO spectra at the wavelengths covered by the HST/COS spectra, not just the H I Lyman series. These other metal ions will be explored in future CGM^2 papers. To identify the absorption features, the Werk SQuAD used a module from the PYIGM[8] software package, IGMGUESSES . The software allows for a straightforward comparison of multiple transitions

---

[8] https://github.com/pyigm/pyigm

from different elements, as multiple lines are displayed with their expected relative intensities given by their atomic parameters simultaneously for a given redshift.

During the line identification process, the absorption lines are assigned a reliability score. ‘a’ signifies a certain feature. For example, the always-present Milky Way (MW) ISM lines at $z = 0$ fall into this reliability category. Other examples include absorption lines observed in two or more transitions of H I , or multiple metal lines that align with hydrogen lines within ± 30 km s^{-1}, and metal doublets or multiplets that show the expected relative strengths as derived from their oscillator strengths and wavelengths (Morton, 2003) and similar velocity profiles. A reliability score of ‘b’, or possible, includes single H I lines with no associated metal lines, and metal ions having only one transition within the observed wavelength range. Other cases of assigning ‘b’ values involve messy blends from absorption lines at different redshifts, weak or uncommon metal lines in an otherwise strong absorption system, and velocity offsets > 30 km s^{-1} from other ‘a’ lines. If a line did not fall into either of the previous categories we gave it a reliability score of ‘c’, or unreliable. In our analysis, we did not include any absorbers in the unreliable ’c’ category. The full catalog of identified absorbers will be presented in detail in future work.

The end result of this line identification process for all 22 QSOs is a catalog of 2914 distinct absorption components, 2071 of which have a reliability rating of certain or possible. An absorption component is defined by an absorption line or lines with a distinct central velocity (or redshift). In practice, individual components offset by < 20 km s^{-1} may not be separable in the HST/COS spectra. An absorber, or absorption system, is a set of absorption line components within $|\delta v| \approx 1000$ km s^{-1}. For example, Ly-α and Ly-β are distinct lines but would be part of the same H I component if aligned within the COS resolution velocity (redshift). This component may be grouped with other H I or metal ion components to form an absorption system.

Different absorbers lie at distinct redshifts and may physically correspond to clouds or layers of gas in the CGM of galaxies at their respective redshifts. Absorbers may also be clouds or filaments of gas in the IGM referred to as the Lyα forest, not directly associated with a nearby galaxy. The term absorber is often used interchangeably with the term "system." However, absorber distinctly does not imply an association with a galaxy and is a more empirical term. Absorption systems may have multiple components at different velocities within their assigned redshift ranges. The vast majority of our H I absorption systems cover > 3 Lyman series lines, many with multiple absorption components.

In order to retrieve more physical quantities such as column density, N_{HI}, we used the apparent optical depth method (AODM) from Savage and Sembach (1991) as encoded in the linetools[9] package. Because we have column density measurements from several Lyman Series lines in most cases, the mean 1σ uncertainties on column density is 0.17 dex for unsaturated H I lines to column densities $\simeq 10^{17.5}$ cm^{-2}.

2.4 Connecting Galaxies and Absorbers: The CGM^2 H I Survey

2.4.1 Defining CGM H I Absorption Systems

With our separately completed galaxy and absorber databases, we can now begin to connect the two as a study of the CGM. In order to construct CGM systems, we first group the individual absorption component identifications into absorption systems, or absorbers. The grouping of absorption components was done using a clustering algorithm from SKLEARN (Pedregosa et al., 2011), MEANSHIFT. This algorithm groups individual absorption components together within a window function of 1000 km s^{-1}. The resultant absorber catalog consists of groups of components we call absorption systems.

We then cross-matched the galaxies and absorption systems if the relative velocity difference of the galaxy and the velocity centroid of at least one of the components of an absorption system exhibits $|\delta v| < 500$ km s^{-1}. We chose this velocity threshold to include absorption systems that could be at or above the escape velocity of the most massive galaxies in our sample. If no absorption system is found at the redshift of a galaxy, we measure the 2σ upper limit within $\delta v = \pm 30$ km s^{-1} of the galaxies redshift using the normalized error of

---

[9] https://github.com/linetools/linetools

the quasar flux. If there was an interloping line at this redshift, we measure the AODM column density as a conservative upper limit. Our results are not sensitive to the choice in the velocity window. We find 181 systems consisting of 416 distinct components that exhibit HI column densities above our 2-σ detection threshold, giving an average of 2.3 detected components per galaxy (absorption system) within our stated velocity window. We find 2.4 average components per galaxy for a smaller window of $|\delta v| < 250$ km s^{-1}, while for a larger velocity window, we find a small decrease to 2.2 detected components per galaxy for a window of $|\delta v| < 1000$ km s^{-1}. The average column density of detected, but not saturated, components remains $10^{14.9}$ cm^{-2} in each case.

Thus, we are left with a galaxy-centric CGM survey that consists of absorption line column density measurements (or limits) around 971 galaxies with reliable redshifts that lie < 1000 km s^{-2} from the quasar. We denote these galaxy-absorber pairs as CGM systems. Figure 2.7 shows the AODM H I column densities for all 971 CGM systems as a function of galaxy systemic redshift. Saturated lines provide only lower limits to the H I column density and are shown as upward facing triangles. Non-detections are shown as 2σ upper limits, and as downward facing triangles. There is an obvious "knee" of amplitude 0.5 dex in the lowest H I column densities at a redshift of $z \sim 0.5$. This decreased sensitivity to weak H I absorption features is driven by the redshifting of Lyα out of the wavelength range of the COS G160M grating. The column density measurements at $z > 0.481$ are derived from measurements of the Lyβ absorption line and/or weaker Lyman series transitions, leading to a decrease in sensitivity for a fixed S/N. This shift in sensitivity motivates us to limit our H I analysis to $z < 0.481$, corresponding to $\lambda_{\mathrm{Ly}\alpha}(1+z) = 1800$Å, or the reddest end of the COS-G160M grating. By imposing this limit, we ensure nearly uniform sensitivity to H I column density for our CGM sample.

After limiting our CGM system survey to the aforementioned redshift, we construct the CGM^2 H I survey, consisting of 572 total galaxy-absorber pairs. The survey includes all galaxies with reliable redshifts and H I absorption systems (including the non-detections within$|\delta v| < 500$ km s^{-1}), collectively referred to as CGM systems. In the H I covering fraction analysis that follows, it is possible to have multiple galaxies at similar redshifts but at differing impact parameters that match with individual absorption systems (74 systems total). For understanding the H I extent of the CGM, we want to understand the correlation of galaxies and absorbers and thus do not limit our matches to the closest or most massive galaxy. However, in our complementary H I velocity analysis, we limit the survey to the galaxy with the smallest impact parameter, thus leaving 522 CGM systems. Future studies, depending on their specific aims, will make independent choices about how to include galaxy-absorber pairs.

In the following analyses, we will examine trends with the impact parameter, ρ, which quantifies the projected distance between the QSO and the galaxy in the rest frame of the galaxy. Figure 2.8 shows the impact parameter–redshift distribution of CGM systems. The grey curve approximates the distance to the edge of the detector in the $5.5'$ GMOS FOV, assuming the QSO is in the center, in order to highlight the survey coverage as a function of redshift. The QSO was slightly offset from the center in certain fields, either to better place guide stars or to avoid bright foreground stars; thus a few galaxies fall on or near this approximate field-size limit.

2.4.2 H I Covering Fraction: Definition and Threshold

To quantify the radial profile and extent of the CGM, we use the covering fraction f_c as a measure of the probability of the presence of H I . The covering fraction is the comparison of "hits" (H) and "misses" (M), with a hit being defined as a galaxy with a corresponding absorber at or above the detection threshold for the full ensemble (see below), while a miss occurs when the 2σ upper limit on a detection is below the threshold at the redshift of the galaxy. A system with a 2σ upper limit above the threshold is ignored altogether because this indicates that the S/N of the spectrum is not adequate for detection of the H I lines. The covering fraction is the fraction of hits versus the total number of CGM systems above the threshold in a given bin, $f_c = H/(H + M)$.

We choose a threshold of $N_{\rm HI} \geq 10^{14}$ cm^{-2} for the covering fraction calculations which is supported by previous survey work. In particular, Chen et al. (2005) find that H I column densities below $10^{13.6}$ cm^{-2} are consistent with being randomly distributed with respect to known galaxies. They also show the correlation of galaxies and absorbers does not depend sensitively on $N_{\rm HI}$ for strong absorbers $N_{\rm HI} > 10^{13.6}$ cm^{-2}. Furthermore, a value of $N_{\rm HI} \geq 10^{14}$ cm^{-2} was shown in Tejos et al. (2014) to be more highly correlated with galaxies than gas at lower column densities. Adopting a threshold of $N_{\rm HI} \geq 10^{14}$ cm^{-2} additionally provides us with a sample against which we can compare to the cumulative column density distributions of H I systems found in Danforth et al. (2016), who also use this value.

2.4.3 The Empirical H I-Galaxy connection

This section presents an empirical analysis of the H I -galaxy connection. In the following analysis we generally avoid differentiating the systems via their spectroscopic galaxy classifications due in part to the fact that we do not expect to see differences in covering fractions of H I in star forming and quiescent galaxies (Thom et al., 2012; Keeney et al., 2017) but include it in Figure 2.9 to illustrate this. Additionally, Figures 2.6 and 2.5 show that the galaxy spectral classification is correlated with the galaxy stellar mass, and thus we have relatively few galaxies with comparable masses but with differing SF classification. Future analyses that examine the metal ions present in the CGM will focus on galaxy spectroscopic type.

In Figure 2.9, we present the H I column density (left axis) and covering fraction (right axis) as a function of stellar mass for galaxies within 300 kpc (top panel) and $1.5R_{\rm vir}$ (bottom panel). The covering fraction for each bin is shown as a dotted line with the grey boxes corresponding to the 68% binomial confidence intervals. At $R < 1.5R_{\rm vir}$, f_c remains consistent with $f_c \gtrsim 0.5$ at all ranges in mass. We also notice a trend of increasing covering fraction and column density as a function of galaxy stellar mass.

The H I column densities and impact parameters are shown in Figure 2.10 as an anti-correlation of column density with increasing separation between the galaxy and QSO sightline. Figure 2.10 separates our galaxy sample into three mass bins, each containing approximately equal numbers of Galaxy-Absorber systems. The top panel contains 191 systems with $M_\star > 10^{9.9} M_\odot$. The middle panel also contains 191 systems with $10^{9.2} < 10^{10} M_\star / M_\odot < 10^{9.9}$ while the bottom panel corresponds to a sample of 190 systems with $M_\star < 10^{9.2} M_\odot$.

We find that for the highest-and-intermediate mass galaxy samples, f_c drops off monotonically but remains $\geq 50\%$ out to an impact parameter of $\rho \simeq$400 kpc, corresponding to $R \simeq 2R_{\rm vir}$ for these masses, before flattening out in the case of the intermediate mass sample. In the lowest-mass regime, we find lower ($f_c \sim 60\%$) covering fractions at the smallest separations. We also observe a shallower profile in covering fraction out to $R \simeq 2R_{\rm vir}$, beyond which the covering fraction remains elevated and consistent with $\sim 50\%$ out to 600 kpc, or $5R_{\rm vir}$.

The extended distribution seen in the lower two panels, where the covering fraction remains elevated, could imply that the galaxies are located inside (or close to) the halos of other (massive) galaxies (e.g. Burchett et al., 2016). In an upcoming CGM^2 paper, we plan to explore the impact of galaxies' environments on the properties of their CGM. Briefly, we find that 56% (243/435) of galaxies with $M_\star < 10^{10} M_\odot$ have $M_\star > 10^{10} M_\odot$ neighbors within 300 kpc and 1000 km/s of the sightline. We observe 75% (86/121) of galaxies with $M_\star > 10^{10} M_\odot$ have neighbors above this mass threshold within the same physical window, consistent with the fact that larger galaxies are more highly clustered. Furthermore, we generally see that low mass galaxies with nearby massive neighbors tend to have elevated HI covering fractions out to $\sim$ 2-3 $R_{\rm vir}$ compared to galaxies with no detected massive neighbors. In tandem, these effects support our conclusion that the elevated covering fractions out to large impact parameters for our $M_\star < 10^{10} M_\odot$ sample are consistent with environmental effects.

Our results are consistent (within 2σ) with those presented by Wakker and Savage (2009) who find that Lyα absorbers at $z < 0.017$ with equivalent width > 50 mÅ ($N_{\rm HI} \simeq 10^{13}$ cm^{-2}) have covering fractions of 100% at $\rho < 400$ kpc. Similarly, Prochaska et al. (2011) find high covering fractions out to $\rho = 300 h_{72}^{-1}$ kpc for absorbers with equivalent widths > 30 mÅ . This observed enhancement of covering fraction in the low-mass sample could also be due in part to the fact that our CGM systems were defined to have separations $|\delta v| < 500$ km s^{-1}, which is around the expected escape velocities of high-mass galaxies but is more likely to encompass unassociated absorbers in the low mass sample that trace the cosmic web. However, low-mass galaxies have smaller escape velocities and shallower potential wells, and thus would likely exhibit gas being ejected at larger velocities. There are clearly many competing effects in this mass range.

Figure 2.11 shows the covering fraction and confidence intervals for the total sample at $N_{\rm HI} \in 10^{13-15}$ cm^{-2} (blue, green, yellow). For the lowest column density threshold in the low-mass sample, the H I covering fractions remain elevated at ∼80% out to at least $4R_{\rm vir}$. This signal must be dominated by galaxy-galaxy clustering as the f_c of random incidence in a velocity window at the mean redshift is $f_c = 0.14$ for $N_{\rm HI} > 10^{13}$ cm^{-2}. (see Section 2.5). In contrast, stronger absorbers preferentially occur at smaller impact parameters. For example, it is at log $N_{\rm HI} \gtrsim 14.5$ that the covering fraction drops below 50% by $R = R_{\rm vir}$.

2.4.4 H I velocity offsets

Here, we examine the kinematics of the H I absorption in our CGM^2 sample of $z < 0.481$ absorbers that are associated with galaxies. Velocity distributions of CGM absorbers quantify the amount of material gravitationally associated with the assumed host halo. For simplicity and clarity, we trim the sample to those galaxies closest to the absorber impact parameter (ρ), thus each absorber is associated with only one galaxy. We are left with a sample of 171 unique galaxy-absorber pairs at $z < 0.48$ that were detected with a signal $> 2\sigma$.

In Figure 2.12, we show the distribution of velocity centroids of the 416 detected absorption components associated with these 181 CGM systems in the rest-frame of the galaxy systemic redshift as a function of column density. The boxes show the quartiles of distribution in velocity while the whiskers show the extent of the distribution. Outliers beyond 1.5 times the innerquartile range are displayed as points. We see that systems with $N_{\rm HI} < 10^{14.0}$ cm^{-2} have a higher median and larger spread in velocity. The component velocity centroids do not exhibit any clear trends with impact parameter. Relevant to this discussion, we recall that absorption components are required to lie within ±1000 km s^{-1} of each other. The only velocity constraint imposed with respect to the galaxy is that at least one of these absorption components must lie within ± 500 km s^{-1} of a CGM^2 galaxy systemic redshift in order to be classified as a CGM absorption system.

We find that $> 53\%$ of the detected H I components are located within ± 250 km s^{-1} of the galaxy systemic redshift, independent of H I column density. Furthermore, only 27% of all absorption components lie at $|\delta v| > 500$ km s^{-1}. Finally, approximately 83% of the total H I column density in the CGM^2 survey lies within ±250 km s^{-1} of the associated host galaxy, consistent with the results of COS-Halos (Tumlinson et al., 2013). Some absorption systems show a total extent > 500 km s^{-1}, which likely captures both bulk motions of galaxies and possible "missassociations" due to incompleteness our galaxy survey, especially at magnitudes > 23.

Were absorption components distributed uniformly in velocity space relative to the galaxies, limited only by the 1000 km s^{-1} window selection function, we would not expect to find such a high concentration of absorbers at low δv. To first order, Figure 2.12 shows that our CGM survey is capturing gas that is mostly not exceeding the escape velocity of its host halo, where the escape velocity of an $M_\star \approx 10^{9.5}\ M_\odot$ galaxy at R_{vir} is ∼250 km s^{-1}. There is very clearly a gravitational association between the galaxies and the absorption at higher column densities. The component velocity centroids do not exhibit a clear trend with impact parameter but we find that the median velocity for components at $\rho > R_{\mathrm{vir}}$ to be 233 km s^{-1} with a standard deviation of 453 km s^{-1} while components at $\rho < R_{\mathrm{vir}}$ have a median velocity of 185 km s^{-1} with a standard deviation of 354 km s^{-1}. In this unique galaxy-absorber pairing framework, 80% of the absorption components with $N_{\mathrm{HI}} > 10^{15}$ cm^{-2} lie within 1 R_{vir}. We caution that for this kinematic analysis we paired specific galaxies with absorbers on the criterion that they were the closest in physical separation from the absorption sightline. Without this bias, the trend remains, however.

Lastly, Figure 2.12 also shows the column-density dependence of the absorption component velocity distribution. Here, we see that absorption systems with $N_{\mathrm{HI}} > 10^{14}$ cm^{-2}

concentrate more strongly at low $|\delta v|$ than weaker systems, and there is a clear column density dependence to the overall velocity concentration at $|\delta|v < 250$ km s^{-1}. Counting by column density we find that 53% of the total fitted column density lies within ± 250 km s^{-1}. In contrast, we find that only 44% of the $N_{\rm HI} < 10^{14}$ cm^{-2} absorption components lie at $|\delta v| < 250$ km s^{-1}, while 67% of the $N_{\rm HI} > 10^{14.0}$ cm^{-2} absorption components lie at $|\delta v| < 250$ km s^{-1}. This concentration increases with increasing column density. For a limit of $N_{\rm HI} = 10^{15}$ cm^{-2}, the percentages of components within $|\delta v| < 250$ km s^{-1} shift to 50% and 86% for the low and high column density thresholds, respectively. Similarly, the high velocity components tend to have low column density. 27% of all absorption components lie at $|\delta v| > 500$ km s^{-1}, but less than 5% of those components have $N_{\rm HI} > 10^{14}$ cm^{-2}. Generally, there is no systematic effect preventing high column density components from appearing at higher velocities, so this trend captures an important characteristic of the CGM.

2.5 HI-Galaxy Clustering: $R^{14}_{\rm CGM}$

2.5.1 Setup

In order to quantitatively describe the radial dependence of the CGM and estimate its extent, we perform an absorber-galaxy cross-correlation analysis. We aim to measure the excess probability of detecting an absorber given the proximity of a galaxy over the proximity-agnostic average rate. Our analysis of the HI-galaxy clustering is similar to the one developed by Hennawi and Prochaska (2007) and follows more closely the analysis by Prochaska et al. (2019). We define the 3D cross-correlation function, $\xi_{ag}(r)$ as

$$\xi_{ag}(r) = \left(\frac{r}{r_0}\right)^{-\gamma}. \tag{2.1}$$

In order to determine the best fitting parameters, r_0 and γ, we define a likelihood function as

$$\mathcal{L} = \prod_i P_i^{\rm hit}(r,z) \prod_j P_j^{\rm miss}(r,z), \tag{2.2}$$

where P^{hit} is defined to be the probability of detecting one or more H I systems and P^{miss} is the probability of detecting none. This probability has both a radial and redshift dependence. An absorber is considered a "hit" if it falls within our window of $\delta v = \pm 500$ km s^{-1} of a galaxy and we measure a column density above a threshold $N_{\mathrm{HI}}^{\mathrm{thresh}}$. We define P^{miss} to be the probability of observing zero events from a Poisson distribution where the rate is the number of events expected from the average density of absorbers and our clustering term:

$$P^{\mathrm{miss}} = \exp\left(-[1+\chi_{\perp}(r)]\left\langle \frac{d\mathcal{N}}{dz} \right\rangle \delta z\right). \tag{2.3}$$

Here $\langle d\mathcal{N}/dz \rangle \delta z$ is the mean number of absorbers in a window of redshift δz. $[1+\chi_{\perp}(r)]$ represents an excess in the number of absorbers due to clustering. This boost due to clustering can be expressed in terms of the 3-dimensional correlation function as

$$\begin{aligned} \chi_{\perp}(r) &= \frac{1}{V}\int_V \xi_{ag}(r)dV \\ &\approx \frac{aH(z)}{2\delta v}\int_{-\delta v/[aH(z)]}^{-\delta v/[aH(z)]} \left(\frac{\sqrt{R_{\perp}^2 + R_{\parallel}^2}}{r_0}\right)^{-\gamma} dR_{\parallel} \end{aligned} \tag{2.4}$$

where we are integrating Equation (2.1) along the length of a cylinder of length $2\delta v/aH(z)$. Here $a = 1/(1+z)$ and $H(z)$ is the Hubble parameter. The probability of a "hit," P^{hit}, is the complement of the probability of a "miss": $P^{\mathrm{hit}} = 1 - P^{\mathrm{miss}}$. P^{hit} is equivalent to the covering fraction f_c. The covering fraction of a random sightline is then:

$$f_c = 1 - \exp\left(\left\langle \frac{d\mathcal{N}}{dz} \right\rangle \delta z\right). \tag{2.5}$$

We take $\langle d\mathcal{N}/dz \rangle$ from Danforth et al. (2016), who measured the occurrence of 5138 individual extragalactic absorption lines of H I in 82 QSO/AGN spectra at redshifts $z_{\mathrm{AGN}} < 0.85$. The occurrence rate of H I absorbers $d\mathcal{N}/dz$ is expressed with a functional form as follows:

$$\frac{d\mathcal{N}(\mathrm{N_{HI}} \geq \mathrm{N_{HI}^{thresh}}, \mathrm{z})}{dz} = C_0(1+z)^{\gamma}, \tag{2.6}$$

with $C_0 = 16$ and $\gamma = 2.3$. This measurement is valid only for absorbers with $\mathrm{N}_{\mathrm{HI}}^{\mathrm{thresh}} \geq 10^{14}$ cm^{-2}, which coincides with the threshold in our definition of "hits" and "misses."

In order to estimate r_0 and γ in our 3D correlation function, Equation (2.1), we follow a Bayesian approach. The posterior probability function can be defined as

$$p(r_0, \gamma | \{k_i, r_i, z_i\}_{i=1}^{N}) \propto p(r_0, \gamma) p(\mathbf{k}|\mathbf{r}, \mathbf{z}, r_0, \gamma). \tag{2.7}$$

where $k_i \in \{0, 1\}$ specifies whether system i is a "hit" or a "miss" and $p(\mathbf{k}|\mathbf{r}, \mathbf{z}, r_0, \gamma)$ is the likelihood function $\mathcal{L}$ defined in Equation (2.2). We define the priors as follows:

$$p(r_0) = \begin{cases} 1/10\,, & \text{if}\, 0 < r_0/\mathrm{Mpc} < 10 \\ 0\,, & \text{otherwise} \end{cases} \tag{2.8}$$

and

$$p(\gamma) = \begin{cases} \mathcal{N}(\mu = 1.6, \sigma = 1)\,, & \text{if}\, \gamma > 0 \\ 0\,, & \text{otherwise} \end{cases} \tag{2.9}$$

where r_0 is measured in h_{68}^{-1} comoving Mpc and $\mathcal{N}(\mu, \sigma)$ is the normal distribution. These priors were chosen based on physical arguments and previous results on absorber-galaxy clustering (e.g. Tejos et al., 2014). We used the Markov Chain Monte Carlo (MCMC) sampler emcee (Foreman-Mackey et al., 2013) to generate samples from the posterior probability distribution function over r_0, γ.

The data were cut as in the previous analysis with a redshift cut of $z < 0.481$ and subdivided into the three equal sized mass samples as before with identical priors used for each sample.

2.5.2 HI-Galaxy Clustering Results

Figure 2.5.3 illustrates the results of the MCMC parameter estimation. The plots on the left show the covering fraction as a function of the perpendicular separation in comoving Mpc.

The grey boxes on the bottom of the plots on the left indicate the covering fraction for sightlines taken at random, $f_c^{rand} = 1 - \exp[\langle d\mathcal{N}/dz \rangle \delta z]$ at the mean redshift of the sample, $\bar{z} = 0.34$, $\bar{z} = 0.34$ and $\bar{z} = 0.26$ for the high, intermediate and low-mass samples, respectively. The plots on the right further illustrate the marginal distributions of the posteriors of our parameters as well as indicate the median values for each. The covering fraction due a random sightline is $f_{c,\mathrm{rand}} = 0.14$.

2.5.3 Estimating R^{14}_{CGM}

In order to estimate a characteristic size of the $N_{\mathrm{HI}} > 10^{14}$ cm^{-2} CGM, R^{14}_{CGM}, we devise a method in which we use the parameters to calculate the impact parameter at which the covering fraction ($f_c = P^{\mathrm{hit}}$) exceeds 0.5. Within this impact parameter, a sightline has a greater than 50% chance of exhibiting a H I column with $N_{\mathrm{HI}} > 10^{14}\mathrm{cm}^{-2}$. We can then estimate the posterior distribution of R_{CGM} by calculating it for each sample taken from the posterior distributions of γ and r_0. Using the samples from the posterior distributions in r_0 and γ we calculate $R^{14}_{\mathrm{CGM,p}} = 346^{+57}_{-53}$ kpc ($R^{14}_{\mathrm{CGM,c}} = 463^{+76}_{-71}$ comoving kpc), $R^{14}_{\mathrm{CGM,p}} = 353^{+64}_{-50}$ kpc ($R^{14}_{\mathrm{CGM,c}} = 469^{+85}_{-66}$ comoving kpc) and $R^{14}_{\mathrm{CGM,p}} = 177^{+70}_{-65}$ kpc ($R^{14}_{\mathrm{CGM,c}} = 222^{+88}_{-81}$ comoving kpc) in order of decreasing mass samples. The extent of the CGM remains relatively similar in size (~ 350 physical kpc) except for the lowest mass sample. These correspond to $R^{14}_{\mathrm{CGM}} = 1.2^{+0.2}_{-0.2}\ R_{\mathrm{vir}}$, $R^{14}_{\mathrm{CGM}} = 2.4^{+0.4}_{-0.3}\ R_{\mathrm{vir}}$ and $R^{14}_{\mathrm{CGM}} = 1.6^{+0.6}_{-0.6}\ R_{\mathrm{vir}}$ for the mass samples in order of decreasing mass, respectively, where R_{vir} was calculated using the mean redshift and mass of each sample. These estimates are in agreement with our qualitative empirical estimates from the previous analysis. These results are summarized in Table 5.

2.6 Comparison with Other Work

2.6.1 Previous Surveys

CGM^2 contains the largest sample by a factor of 5 of CGM absorption systems within ~ 600kpc of putative host galaxies. Our increased coverage is partly due to the fact that

Table 5. Results

$M_\star$ $[M_\odot]$	$R^{14}_{\rm CGM,p}$ [kpc]	$R^{14}_{\rm CGM,c}$ [kpc]	$R^{14}_{\rm CGM}$ $[R_{\rm vir}]$	$\bar{z}$
$M_\star > 10^{9.9}$	346^{+57}_{-53}	463^{+76}_{-71}	$1.3^{+0.2}_{-0.2}$	0.34
$10^{9.2} < M_\star < 10^{9.9}$	353^{+64}_{-50}	469^{+85}_{-66}	$2.4^{+0.4}_{-0.4}$	0.33
$M_\star < 10^{9.2}$	177^{+70}_{-65}	222^{+88}_{-81}	$1.6^{+0.6}_{-0.6}$	0.26

Note. — Results of H I-galaxy clustering analysis and $R^{14}_{\rm CGM}$: (1) Stellar Mass limits of each sample; (2) $R^{14}_{\rm CGM}$ in physical kpc; (3) $R^{14}_{\rm CGM}$ in physical kpc; (4) $R^{14}_{\rm CGM}$ normalized to the average virial radius of the sample; (5) Average redshift of each sample

we probe a larger volume than previous surveys by confirming sub-L∗ galaxies out to higher redshifts. For this reason, CGM^2 presents an excellent opportunity to study the HI-traced CGM at galaxy-absorber separations that exceed $R_{\rm vir}$ for a wide range of stellar masses. Here, we compare our results to those from other observational surveys of the CGM.

Figures 2.5.3 and 2.14 and Table 5 provide a summary of our covering fraction results for our mass segregated galaxy samples. The red, green, and blue-shaded curves in Figure 2.14 correspond to the samples cut at masses of $M_{\star,cut} = 10^{9.204}, 10^{9.888} M_\odot$ as described above. This figure allows us to better compare the covering fraction profiles of each mass sample to each other. As we saw in the last section, we find a mass dependence on the extent of the CGM; the covering fraction remains higher for larger galaxy mass when looking at the physical impact parameter. When we consider the impact parameter normalized to the average virial radii of each mass sample, we find the covering fraction remains elevated to larger radii ($\gtrsim 2R_{\rm vir}$) for the intermediate mass sample, while for the other mass samples the extent of the CGM is similar ($\gtrsim 1R_{\rm vir}$). We also find that the covering fraction of the lowest mass sample never exceeds 80% while for both higher mass samples, the covering fraction is consistent with $\sim 100\%$.

Many previous studies have examined the covering fraction of H I as a function of galaxy projected separation, although most of these surveys are limited to $\rho \lesssim 300\,\mathrm{kpc}$. These surveys are all generally constructed in the same manner, with HST UV spectroscopy with COS and/or STIS and a spectroscopic galaxy catalog to connect galaxies and absorbers.

One of the largest such surveys was carried out by Tejos et al. (2014) who attempted to explore the connection between the IGM and galaxies by measuring the HI-galaxy cross-correlation at $z < 1$ at distances between ~1 and 10 Mpc. They used multiple ground based instruments to build a new spectroscopic survey of 2143 galaxies in 8 QSO fields with HST spectroscopy. They combined their catalog with existing catalogs to build a survey of ~17500 galaxies. They did not limit their galaxy survey to only the nearest galaxies but used a statistical approach to examine the cross-correlation of galaxies and absorption systems. They found that the H I Lyα forest to be divided into two main categories: a population of low column density absorbers tracing the cosmic web and a higher-column density population that traces the dark matter halos in which galaxies reside. CGM^2 has the benefit of being more sensitive to fainter galaxies closer to the QSO itself allowing us to better constrain the extent of the CGM. However, their larger galaxy sample at larger separations (which explores different scales) presents a great opportunity to compare our cross-correlation analysis. Using the same 3D cross-correlation power law, Equation 2.1, they found $r_0 = 3.8 \pm 0.2\ h_{70}^{-1}$ Mpc and $\gamma = 1.7 \pm 0.1$ in their sample of SF-galaxies. This sample is similar to our high-mass sample since we are dominated by SF galaxies and indeed, we find agreement consistent within 1σ.

A mass dependent CGM was examined in Bordoloi et al. (2018) who used the 85 galaxies in the COS-Halos and COS-Dwarfs surveys at $z \sim 0$ with $M_\star$ ranging from 8 to 11.6 $\log M_\star/M_\odot$. This sample was limited to impact parameters ($\rho < 160$ kpc). They found a mass and radius dependence of the strength of H I absorption where the equivalent width of H I increases with $M_\star$ and with decreasing impact parameter. Here we are primarily focused on the mass and radius dependence of the covering fraction, but find trends generally consistent with those observed by Bordoloi et al. (2018).

Chen et al. (2001) found Lyα with column densities $N_{\mathrm{HI}} \gtrsim 10^{14}$ cm^{-2} in 34/47 galaxies ($f_c \approx 0.7$) out to $\rho \simeq 330$ kpc. Their sample consists of $\sim L*$ HI-Galaxy pairs with $|\delta v| < 500$ km s^{-1} spanning $0.1 < z < 0.9$ ($\bar{z} = 0.36$). They also found a sharp decline thereafter. We find close agreement $f_c = 0.78^{+0.06}_{-0.89}$ (162/209) when applying the same criterion to our sample also seeing a sharp decline around 400 kpc. (Figure 2.14).

Prochaska et al. (2011) used 14 QSO sightlines with previously published equivalent width (W_0) measurements of Lyα to carry out a galaxy survey that targeted 37 $L > 0.01L_\star$ galaxies at $\bar{z} = 0.18$. They connected absorbers and galaxies with $|\delta v| < 400$ km s^{-1}, although they show it makes no qualitative difference in their results using a larger velocity window of $|\delta v| < 600$ km s^{-1}. They found covering fractions of order unity ($\approx 90\%$) for $N_{\mathrm{HI}} > 10^{13}$ cm^{-2} gas out to $\rho = 300$ kpc. Comparing our covering fractions with $N_{\mathrm{HI}} > 10^{13}$ cm^{-2} and for galaxies with $M_\star > 10^{8.55} M_\odot$ to approximate their $L > 0.01L_\star$ sample, we find $f_c = 0.85^{+0.04}_{-0.05}$ (128/150), which is roughly consistent with their value.

Wakker and Savage (2009) conducted a large survey of the HI-galaxy connection at $z \lesssim 0.017$, consisting of 76 QSO sightlines and $\sim 20\,000$ local galaxies. They found covering fractions of 77% for Lyα absorbers >50 mÅwithin $\rho < 400$ kpc and $|\delta v| < 400$ km s^{-1} of $L > 0.1L_\star$ galaxies. If we limit our sample to $L > 0.1L_\star$ galaxies and use a Lyα absorber threshold of ~50 mÅ ($N_{\mathrm{HI}} \sim 10^{13}$ cm^{-2}), we find covering fractions, $f_c = 0.89^{+0.04}_{-0.05}$ (99/111) for galaxies within 400 kpc. This discrepancy could imply that covering fractions increase with redshift.

Most recently, a large survey of H I was carried out by Keeney et al. (2018) (K18). Their survey consisted of 47 COS sightlines (COS GTO) with higher signal-to-noise (S/N $\sim$ 15-50 compared to our $\sim$ 10) QSO spectra. Using ground based telescopes, they constructed a spectroscopic galaxy database of $\sim 9,000$ galaxies with the aim of $> 90\%$ completeness to 1 Mpc down to $0.1L*$ at $z \lesssim 0.1$. Due to their higher S/N, they could consistently measure weaker absorption lines, down to $N_{\mathrm{HI}} \geq 10^{12.8}$ cm^{-2}. Leveraging the high completeness of K18 at low redshifts, they were able to measure the H I column densities out to $4R_{\mathrm{vir}}$ with enough galaxies in this range (243) to make precise statements about the radial profile of HI.

They find the covering fraction for $L < L*$ galaxies (corresponding to our low-mass sample) to feature a shallower decline than that of their $L > L*$ sample. Qualitatively, we show similar results. This difference in covering fraction behavior between high and low mass samples can readily be seen in Figures 2.11 and 2.14. The elevated covering fractions at large radii in the $L < 0.1L*$ sample imply that contributions to $N_{\mathrm{HI}} > 10^{14}$ cm^{-2} CGM gas are dominated by low-mass galaxies. This result is consistent with Prochaska et al. (2011) in relation to H I and with high-ionization metals such as O VI (Tumlinson and Fang, 2005; Pratt et al., 2018; Prochaska et al., 2019).

We note that the high mass sample drops to a very low covering fraction $f_c \simeq 0.20$ at $4R_{\mathrm{vir}}$ while in the lower mass sample, f_c remains elevated. At radii greater than $3R_{\mathrm{vir}}$, we may be limited to only higher-z high-mass galaxies due to the detector size but this detector size bias should not affect the lower mass samples.

In addition, previous low-redshift studies have found that low column density gas ($N_{\mathrm{HI}} = 10^{13-14}$ cm^{-2}) is likely uncorrelated with galaxy halos (Chen et al., 2005; Prochaska et al., 2011; Danforth et al., 2016) (but, see Tejos et al. (2014) who find that 50% of weak lines can still be correlated with galaxies on 1-10 Mpc and Tripp et al. (1998) who show that the weakest absorbers are not randomly distributed). First, low column density material exhibits high covering fractions out to 1 Mpc which we can see in Figure 2.11. Second, low column density material exhibits less velocity correspondence with the systemic velocities of galaxies nearby in projection (Tumlinson et al., 2013), as we find a median velocity difference of 233 km s^{-1} with a standard deviation 453 km s^{-1} for $\rho > 1R_{\mathrm{vir}}$ while we find a median velocity of 185 km s^{-1} with a standard deviation of 354 km s^{-1} for $\rho < 1R_{\mathrm{vir}}$. Traditionally, such low column density material is attributed to the Lyα forest, or to gas in a filament like structure (Tejos et al., 2014), physically distinct from the CGM. This was examined in greater depth by Burchett et al. (2020a) who conclusively tie the diffuse IGM to the cosmic web. They find the H I absorption signature decreases past $\rho > R_{\mathrm{vir}}$ and settles to the cosmic mean matter density.

Turning to higher redshifts, Rudie et al. (2012) use the ground based Keck Baryonic

Structure Survey; KBSS to investigate the $z \sim 2-3$ CGM surrounding 886 galaxies. Their sample contains 48 galaxies at $\bar{z} = 2.3$ within $\rho \lesssim 300$ kpc for which they measure a covering fraction of $f_c = 0.81 \pm 0.06$ for absorbers with $N_{\rm HI} > 10^{14}$ cm^{-2}. By comparison, if we choose a mass range of $10^{10.4} < M_\star/M_\odot < 10^{11}$ to approximate their mass distribution (Erb et al., 2006) we find $f_c = 0.86^{+0.08}_{-0.14}$ (18/22) which is good agreement with their results. However, looking at the extended CGM $\rho <$ 1Mpc, we find a discrepancy: we measure $f_c = 0.49^{+0.08}_{-0.08}$ (40/83) vs their 0.70 ± 0.03. These numbers are in agreement at the 2σ level, however. Interestingly, the extended CGM may show a decrease in covering fraction as the universe evolves. This phenomenon may be due to the development of virial shocks that ionize the gas, as suggested by Burchett et al. (2018). At $z = 2.3$, 300kpc $\sim 2R_{\rm vir}$ for a $7 \times 10^{10} M_\odot$ galaxy while at $z = 0.3$ (CGM^2) 300 kpc $\sim 0.8R_{\rm vir}$. Alternatively, this could simply be due to the fact that a column density of $N_{\rm HI} > 10^{14}$ cm^{-2} traces lower density peaks at high-z.

2.6.2 Comparison with Hydrodynamical Simulations

We now turn to a brief comparison with hydrodynamical simulations. van de Voort et al. (2019) simulate a roughly $z \sim 0$ $L*$ galaxy using a new refinement technique to better resolve the CGM. They find an increase in the H I column density and resultant covering fraction when the resolution is increased to resolve 1 kpc scales. We find their model to be in good agreement with our $N_{\rm HI} \geq 10^{14}$ cm^{-2} covering fraction measurements when we limit our sample to $M_\star \sim 10^{10.5} M_\odot$ (see their Figure 3). Comparing our column density measurements to van de Voort et al. (2019) and Hummels et al. (2019) we also see good agreement out to their limiting distance of 200 kpc and 100 kpc, respectively (compare Figure 2.10 to van de Voort et al. (2019) Figure 2). Our high H I covering fractions and column densities at $R < R_{\rm vir}$ are in conflict with earlier simulations that consistently underpredict the column density of low-ions in the CGM (e.g., Hummels et al., 2013; Liang et al., 2016; Stinson et al., 2012; Shen et al., 2012; Ford et al., 2013), validating the work that has gone into creating these new high resolution techniques. To understand the extent of the CGM around a diverse sample of galaxies, we encourage future simulations extending to at least $4R_{\rm vir}$, and covering a larger range of galaxy masses, down to $0.01L*$.

2.7 Summary and Conclusions

We have reported the first results from the CGM^2 survey, a comprehensive survey of the z <1 CGM at least 5× larger than previous surveys such as COS-Halos at comparable redshifts. This chapter has presented the detailed properties of the survey design and the procedures followed in the collection and processing of the data. We present an H I study that combines high-resolution HST/COS UV spectra of 22 background QSOs with Gemini/GMOS spectra of 572 foreground galaxies having stellar masses 10^6 M$_\odot$ < M$_\star$ < $10^{11.5}$ M$_\odot$ and z < 0.481. The S/N$\sim$10-12 of these COS spectra and access to the H I Lyman limit enables us to constrain the H I column densities and kinematics of associated CGM absorption, and to ultimately examine the extent of the CGM as a function of galaxy mass and physical separation from the QSO sightline.

We find that high-column density circumgalactic material is associated with galaxies at high statistical significance out to $2R_{\rm vir}$, whereas H I absorption with $N_{\rm HI} < 10^{14}$ cm^{-2} is more broadly distributed in both impact parameter and velocity space and may not be associated directly with massive galaxies. Our kinematic analysis reveals that the detected strong H I is most likely gravitationally associated with the most nearby galaxy, while weaker H I components seen at absolute velocity offsets $\gtrsim$ 500 km s^{-1} may be instead associated with extended large scale structures. We find generally good agreement between our sample and the prior studies that have examined the CGM of low-redshift galaxies out to similar and larger separations.

We define the cool CGM as the region surrounding a galaxy in which the probability of observing an absorber with $N_{\rm HI} > 10^{14.0}$ cm^{-2} is > 50%. Our column density threshold, $N_{\rm HI} > 10^{14.0}$ cm^{-2}, is motivated by previous observational work that examines the statistical, large-scale (Mpc) correlations between galaxies and QSO absorption lines. In essence, our definition of $R^{14}_{\rm CGM}$ demands that around a given galaxy, one is more likely than not to find material that has been empirically associated with galaxies. The picture that emerges is of a diffuse, CGM extending to $\sim$ 350 kpc around galaxies with stellar mass $\gtrsim 10^9 M_\odot$ at $z \lesssim 0.5$. For galaxies of lower stellar mass, the extent of the CGM is $<$ 200 kpc. At all stellar masses, the extent of this CGM exceeds a virial radius, especially for galaxies with intermediate masses where the $R^{14}_{\rm CGM}$ exceeds $2R_{\rm vir}$. Therefore, these results imply that using $R_{\rm vir}$ as a proxy for the characteristic edge of the CGM may significantly underestimate its true extent. The detailed nature of the mass dependence of the CGM will be examined in a forthcoming paper. Additional follow-up studies using CGM^2 data will consider transitions from a wide range of ionized metals and absorption-line profile analyses to characterize the ionization state, metallicity, kinematics, and mass of the CGM at low redshift.

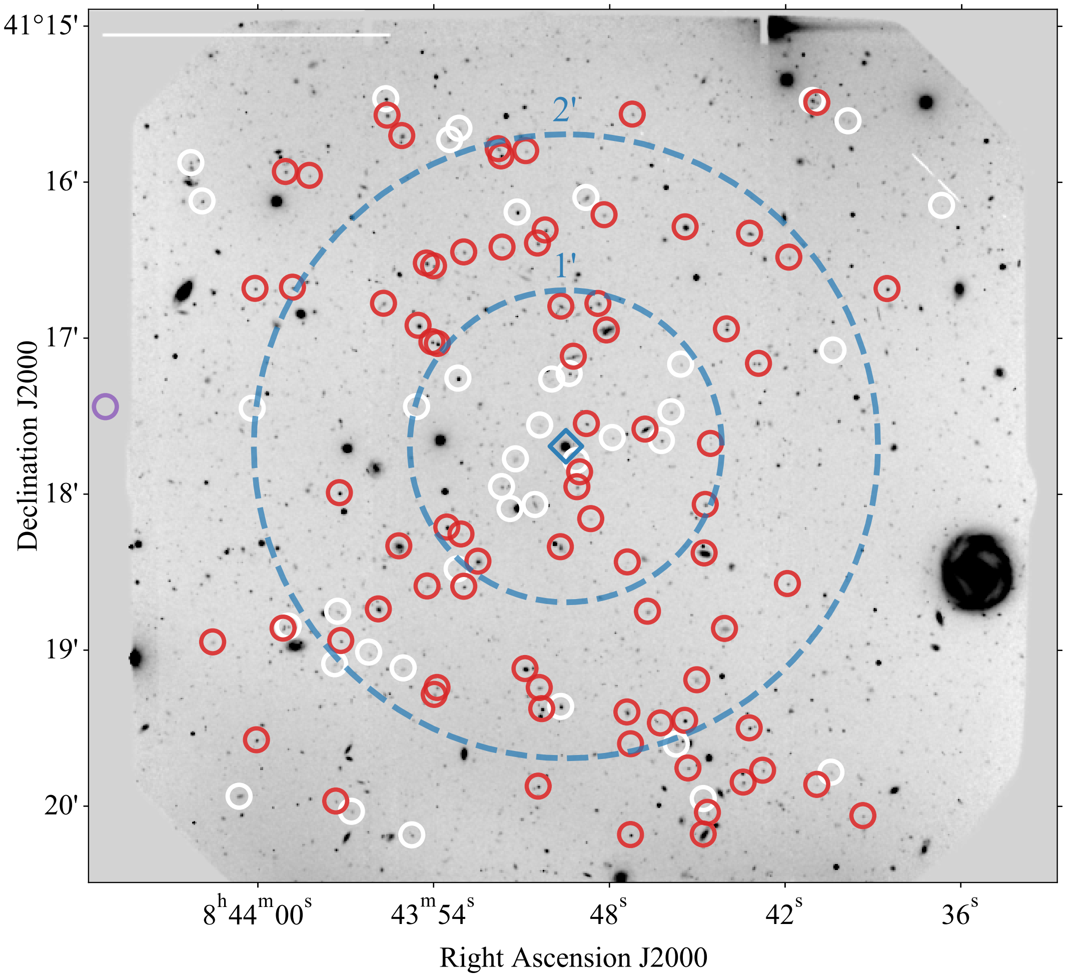


Figure 2.1: An example of the survey design and targeting strategy of CGM^2 showing the slit placements centered around QSO J0843+4117 (blue diamond) overlayed on the g-band pre-image from the Gemini-GMOS detector. The circles denote where slits were placed on the slitmasks. White circles indicate galaxies with slits whose final spectra did not yield a reliable redshift, while the red circles indicate galaxies that produced reliable redshifts. Large blue dashed circles show the one and two arcminute radii from the QSO. The purple circle just off the left of the detector is the COS-Dwarfs (Bordoloi et al., 2014) galaxy target for this field whose spectra was obtained by SDSS with an impact parameter of $\rho = 113$ kpc at $z = 0.0300$.

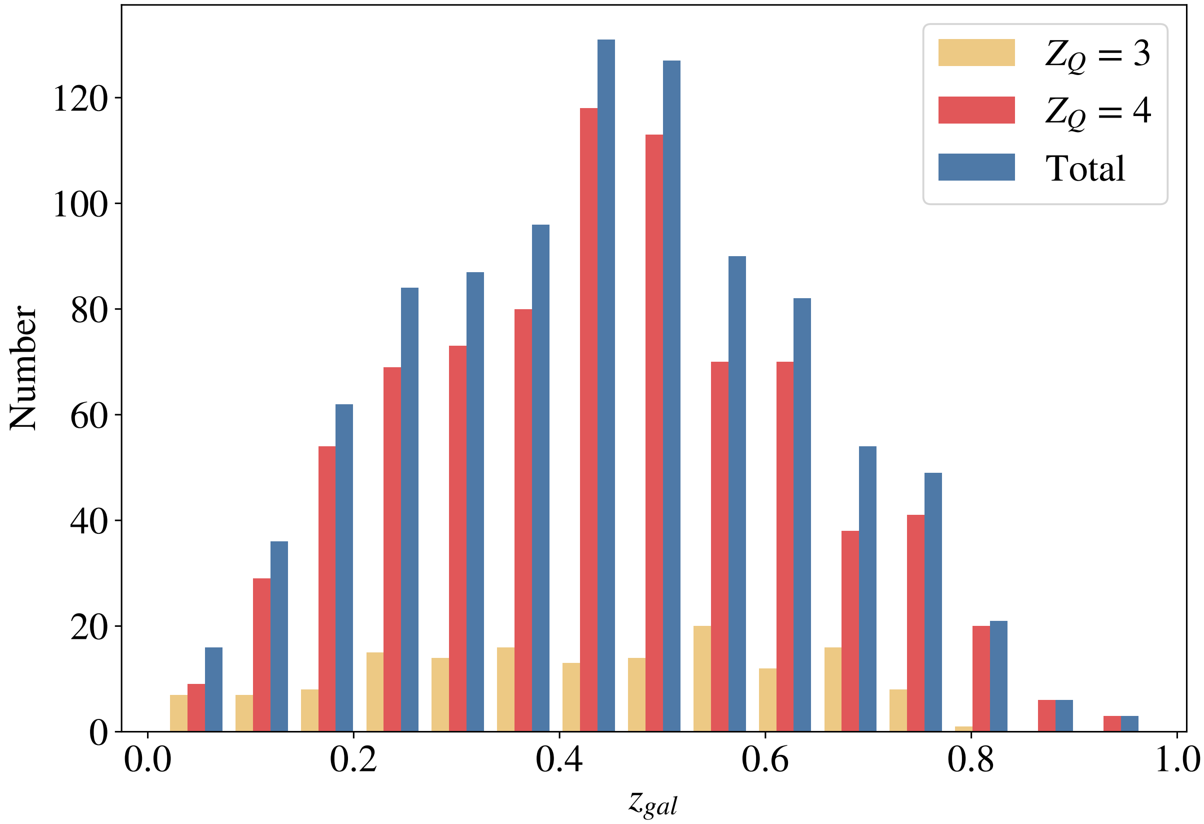


Figure 2.2: The redshift distribution of the CGM^2 galaxy catalog for galaxies with $z < z_{\rm QSO}$. The redshift reliability is encoded in yellow and red. Red represents our most reliable redshift quality flag of ‘4’ with spectra displaying more than one strong absorption or emission line. A quality flag of ‘3’ was reserved for spectra with only one strong emission line and thus a less reliable redshift designation. Approximately 85% (820 out of 971) of our spectra were given the highest reliability flag. The typical statistical uncertainly of our redshifts is $\sigma_z \sim$ 50-100 km s^{-1} ($z \simeq$ 0.00016-0.00030).

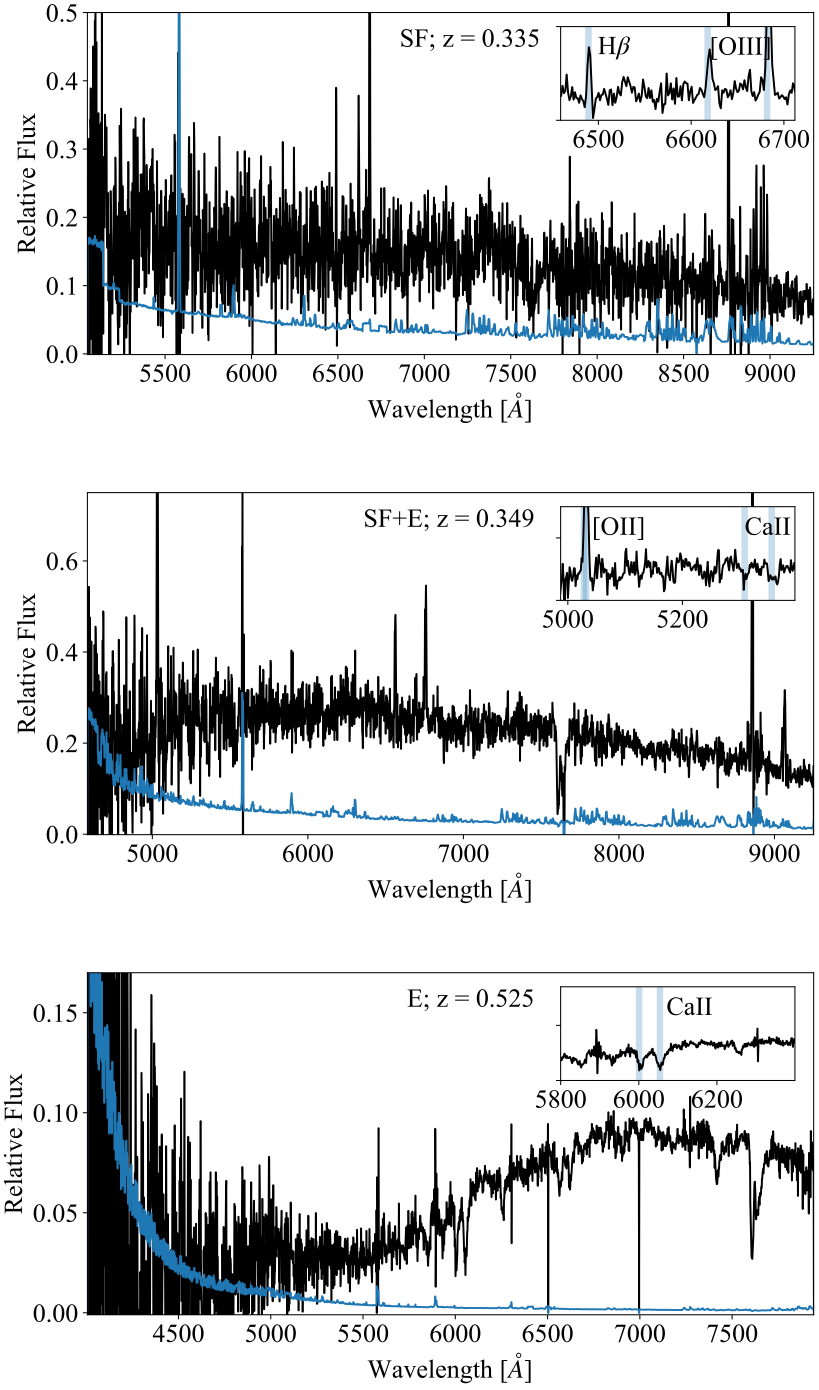
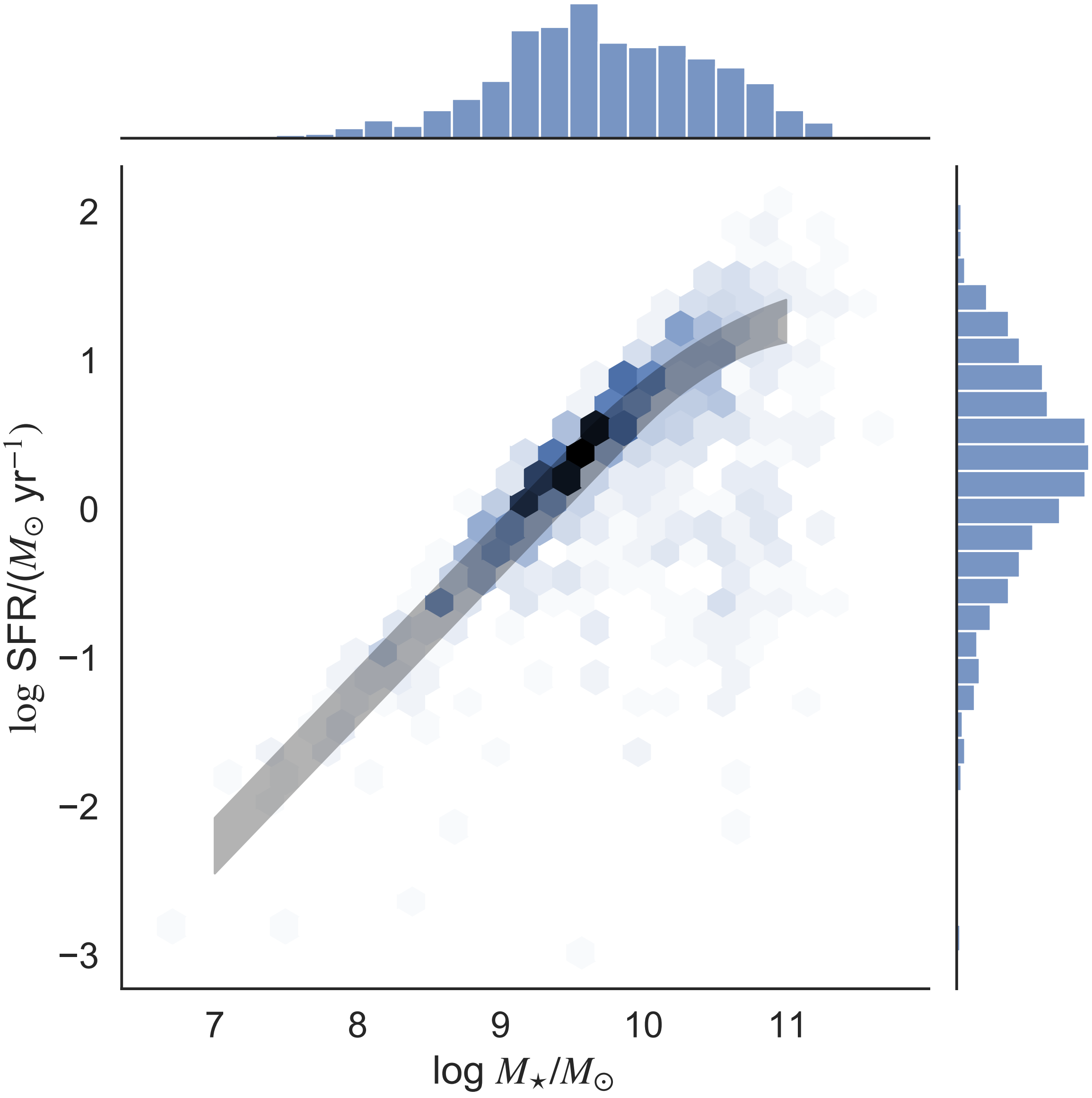


Figure 2.3: Examples of typical Gemini-GMOS spectra with a quality flag of $Z_Q = 4$, along with the error shown in blue. These spectra highlight our method of visual galaxy spectral typing. The cut out insert in each panel shows an example of the key spectral features used in redshift determination. The top spectrum is classified as star-forming galaxy, the middle panel shows a galaxy with both emission lines and absorption lines, and the bottom spectrum is an example of galaxy with an older stellar population with strong Ca H+K absorption.



.

Figure 2.4: Star formation rate (SFR) vs. stellar mass for the CGM^2 galaxy sample as estimated by CIGALE. The locus of galaxies tracks a monotonic increase of SFR with stellar mass, known informally as the "star-forming main-sequence" (SFMS). The density of galaxies in this space is indicated via shading of the hexagonal bins. The grey shaded region corresponds to the redshift dependent fit of the SFMS from Schreiber et al. (2015) of galaxies spanning $z = [0.23, 0.63]$. This range represents the 16th and 84th percentiles of the galaxy catalogs redshift distribution

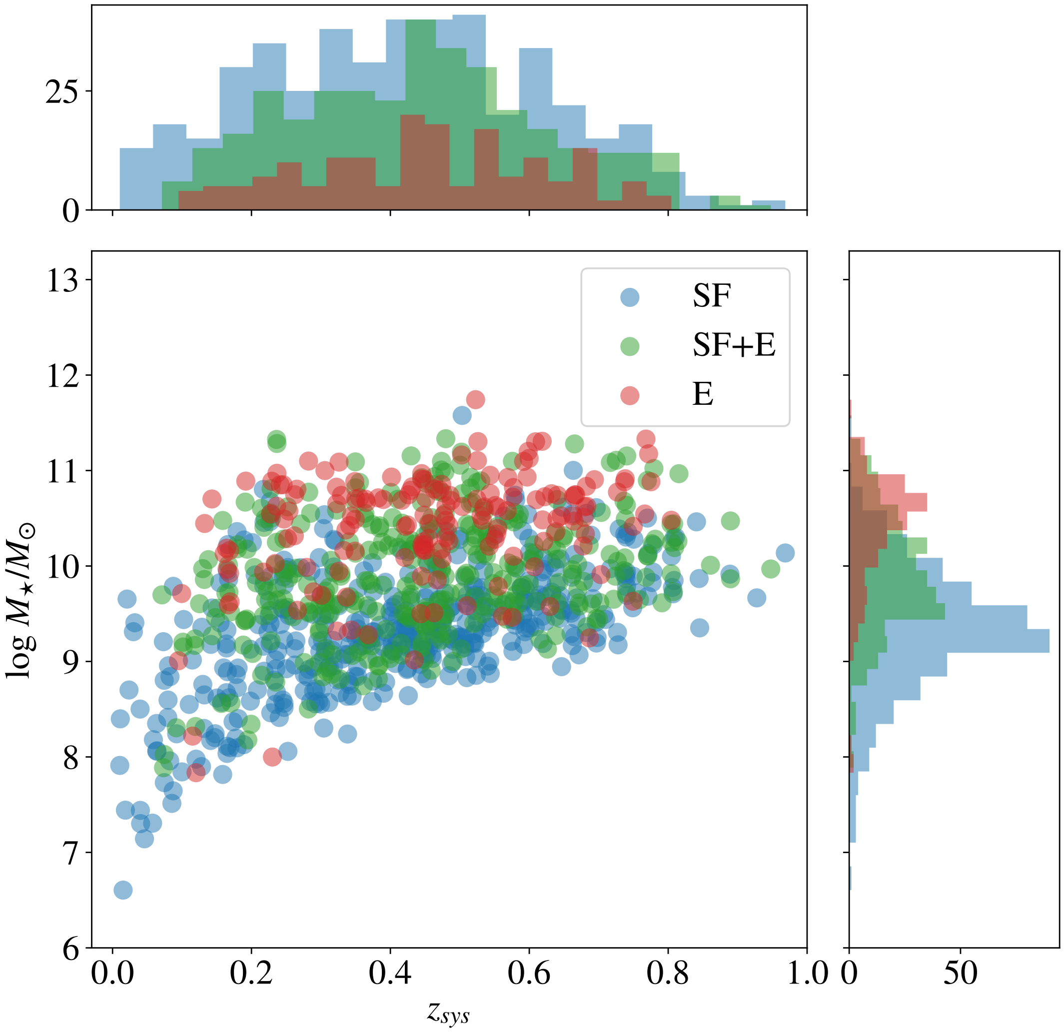


Figure 2.5: The distribution of galaxy stellar mass as a function of galaxy systemic redshift with marginal distributions on the right and top. The red, green, blue colors correspond to the galaxy spectral type determined from visual inspection of GMOS spectra. Red circles show absorption-line only, or elliptical (E) type galaxies, green circles show galaxies displaying a combination of absorption and emission lines associated with star formation (SF+E), and blue circles show emission-line only, or star forming (SF) galaxies.

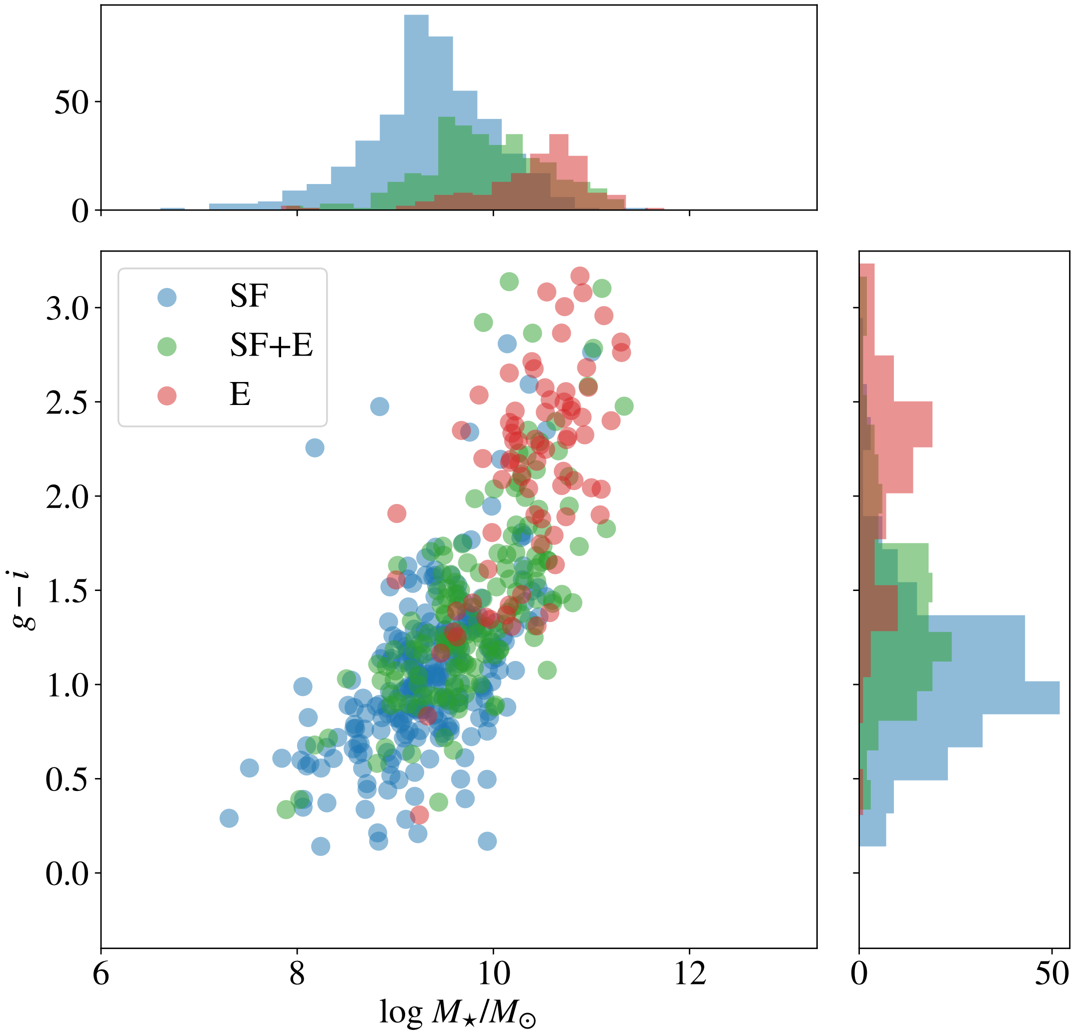


Figure 2.6: CGM^2 sample in a color-mass diagram using the $g-i$ color and the galaxy stellar mass, $M_\star$. Multi-band photometry was not available for all of the galaxy targets, only objects with both bands are shown here. The bi-modal populations of the star-forming and passive galaxies are evident. Due to the nature of our survey, we are slightly biased against faint, passive galaxies since retrieving a redshift in the case of absorption lines requires large continuum flux.

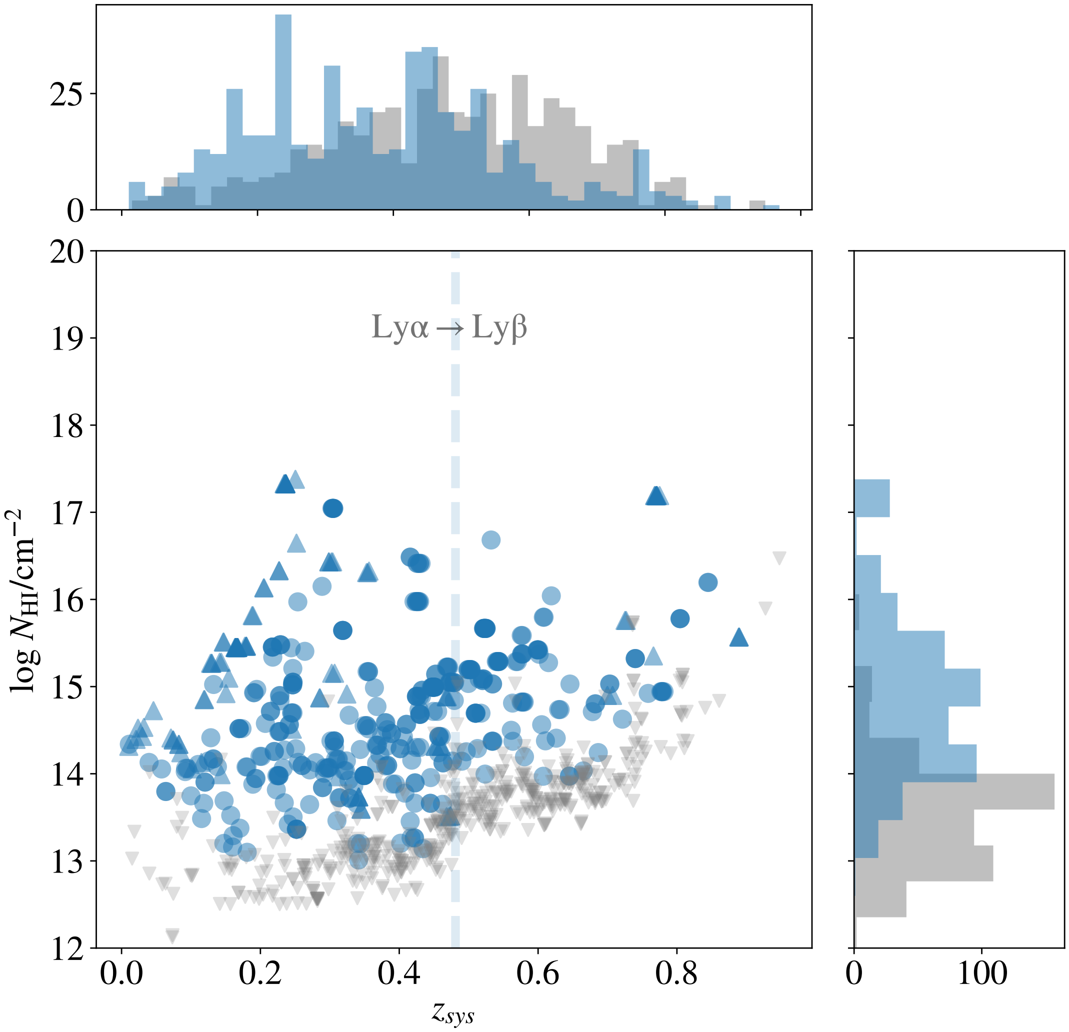


Figure 2.7: Scatter plot and marginal distributions of column densities vs. redshift for the H I systems detected in the CGM^2 survey. The mean 1σ uncertainties on column density is 0.17 dex for unsaturated H I lines to column densities $\simeq 10^{17.5}$ cm^{-2}, which is of order the size of the symbol (see Figure 2.9 for the size of our uncertainties). The measurements designated with upward triangles are saturated absorption lines and are thus lower limits while circles represent detections. Smaller gray downward facing triangles are 2σ upper limits for galaxies where no corresponding absorption was measured. The visible break in the minimum $N_{\rm HI}$ at $\gtrsim 0.48$ is shown by the vertical dashed line, and marks the redshift at which Lyα shifts out of the COS G160M bandpass and thus becomes inaccessible. There can be multiple systems at the same column density and redshift which appear as darker points in this plot. This occurs when multiple galaxies lie within $|\delta v| < 500$ km s^{-1} of the absorption systems.

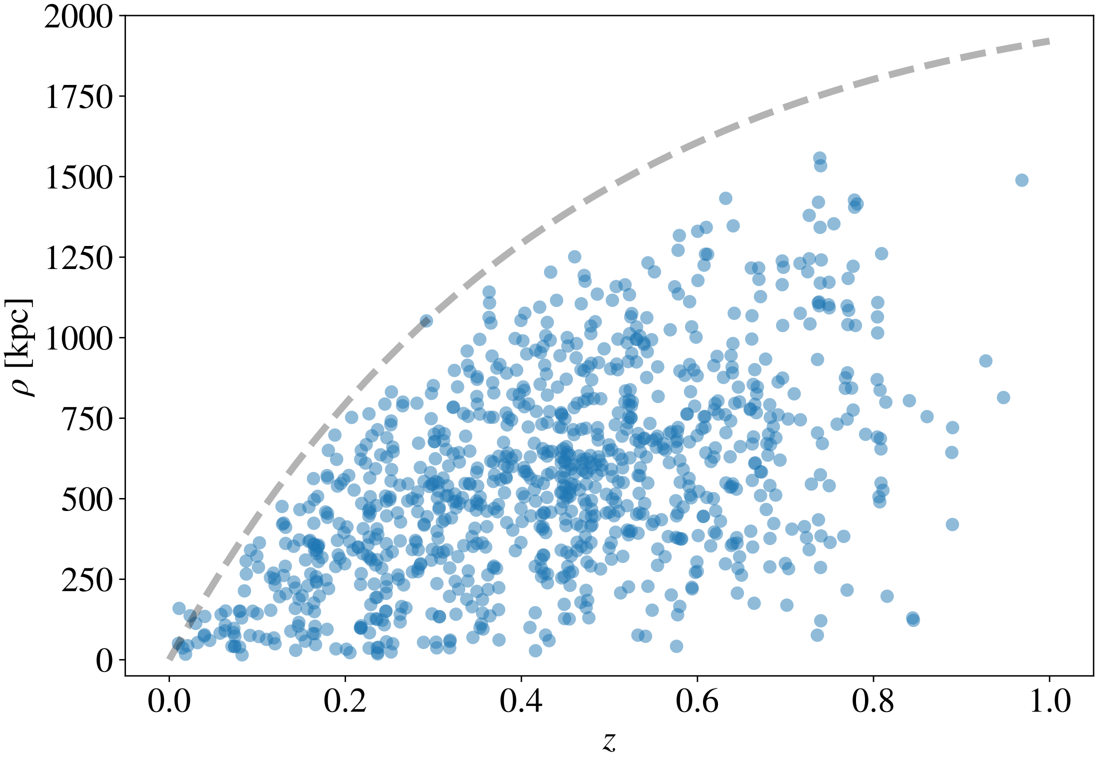


Figure 2.8: The distribution of impact parameters as a function of redshift. The grey curve approximates the distance to the edge of the detector in the 5.5′ GMOS FOV assuming the QSO is in the center to highlight the survey coverage as a function of redshift. Although most of the QSOs are centered in the FOV, in a few cases they had to be offset to avoid bright stars.

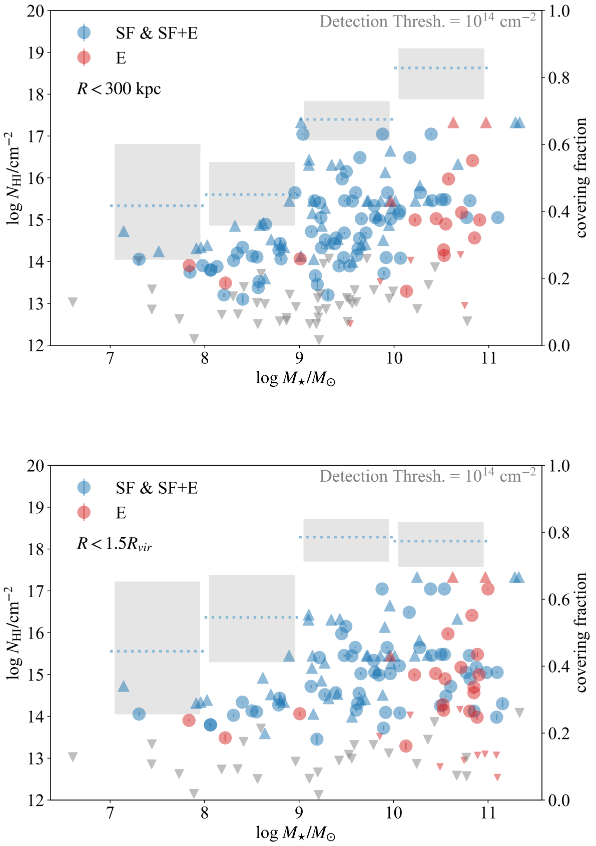


Figure 2.9: Column density as a function of mass for $\rho < 300$ kpc (top) and $R < 1.5R_{\rm vir}$ (bottom). CGM systems with saturated absorption are marked with upward facing triangles, while non-detections are displayed as lighter, downward facing triangles at their corresponding 2σ upper limits. Circles represent CGM systems with measured $N_{\rm HI}$. Measured 1σ uncertainties in the column density of the detected CGM systems are shown as lines inside the markers. Red markers indicate a spectroscopically-determined quiescent galaxy classification 'E', while blue corresponds to those galaxies with emission lines present in their spectra, classified as 'SF' and 'SF+E'. Our sample of quiescent galaxies predominately reside in the highest-mass bin. Covering fractions f_c are plotted with respect to the right axes and are calculated without differentiating spectroscopic galaxy categories. The grey boxes correspond to the binomial confidence interval of the covering fraction ($N_{\rm HI} > 10^{14}$ cm^{-2}) with the mean f_c in each bin denoted with a dotted line. The column density increases as a function of mass while the covering fraction remains greater than $f_c > 0.5$ for galaxies with masses of $M_\star = 10^{8-11} M_\odot$.

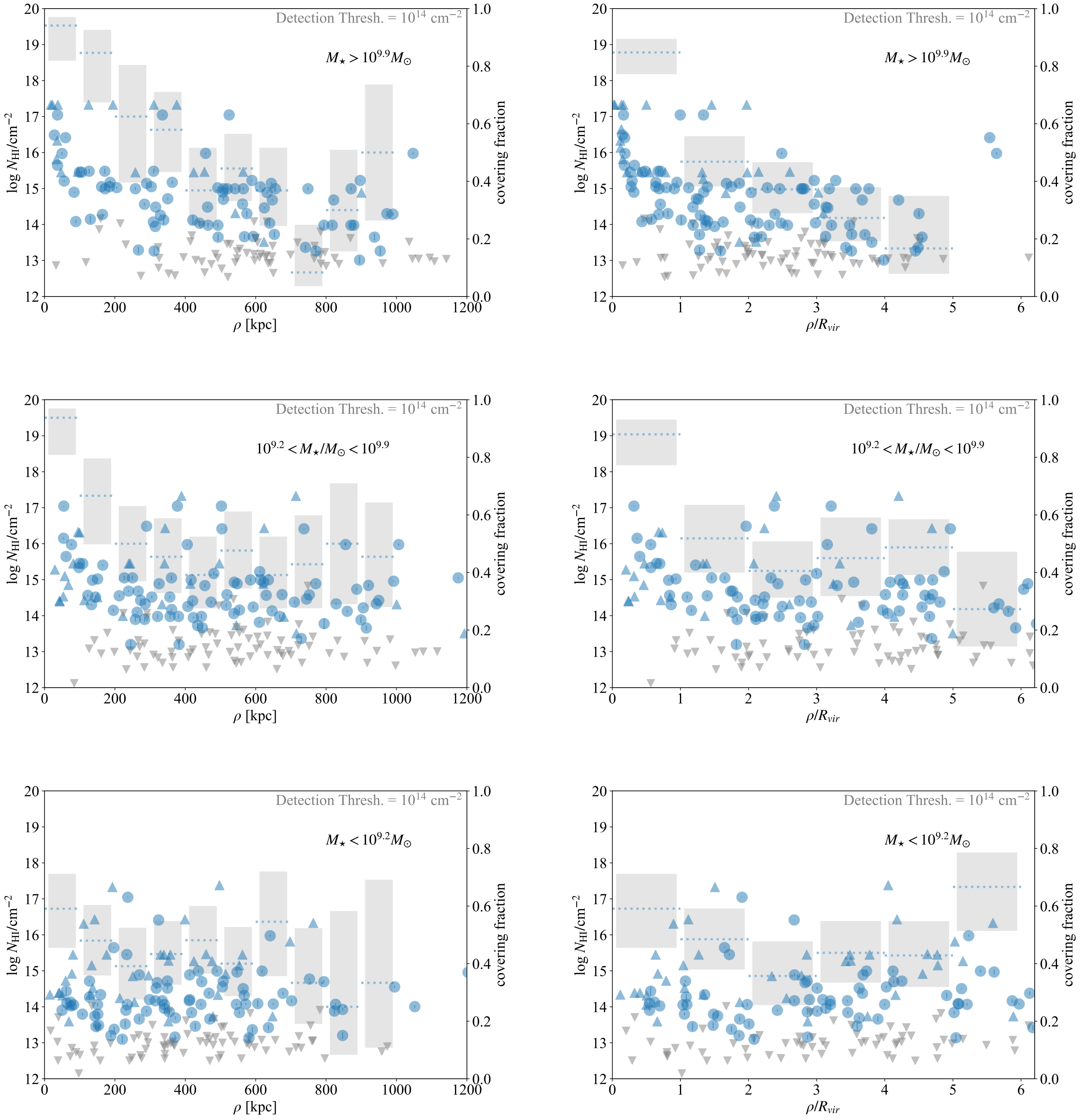


Figure 2.10: Column density $N_{\rm HI}$ as a function of impact parameter, ρ (left) and $R_{\rm vir}$ (right). The panels are arranged vertically in order of decreasing mass. Galaxy-absorber systems with dark circles represent galaxies with $N_{\rm HI}$ detections in the corresponding QSO spectrum. Systems with dark, up-arrow symbols show our inferred lower limits due to saturation in $N_{\rm HI}$. Similarly, systems with gray, down-arrow symbols denote the 2σ upper limit on $N_{\rm HI}$ (non-detections). Measured 1σ uncertainties in the column density of the detected CGM systems are shown as grey lines but are smaller than the marker size in every case. $N_{\rm HI}$ covering fraction, f_c, corresponds to the right axes in bins of 100 kpc (left) and 1 $R_{\rm vir}$ (right). The dotted lines represent f_c assuming a detection threshold of $N_{\rm HI} = 10^{14}$ cm^{-2} with the 68% binomial confidence interval shown as shaded gray regions about the mean f_c.

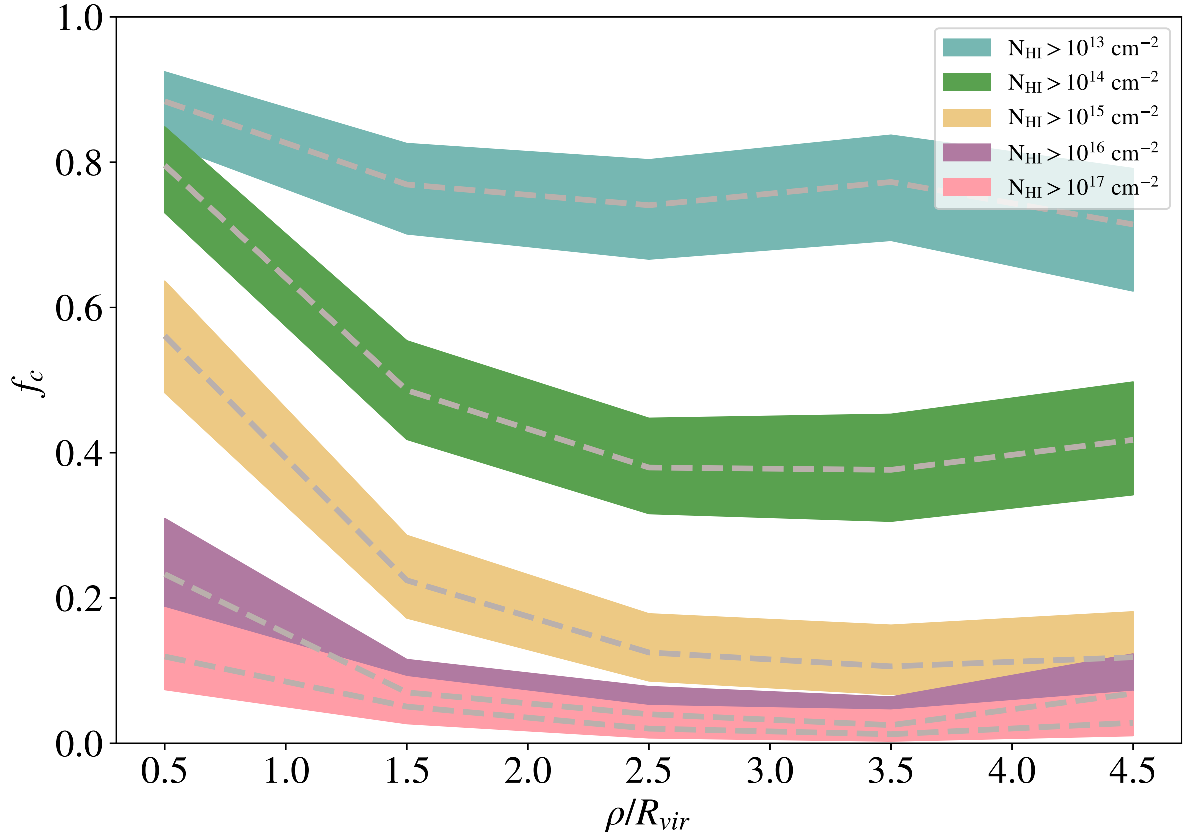


Figure 2.11: Covering fraction of H I as a function of $\rho/R_{\rm vir}$ for column density thresholds of $N_{\rm HI} > 10^{13}$ cm^{-2} (blue), $N_{\rm HI} > 10^{14}$ cm^{-2} (green), $N_{\rm HI} > 10^{15}$ cm^{-2} (yellow), $N_{\rm HI} > 10^{16}$ cm^{-2} (purple), and $N_{\rm HI} > 10^{17}$ cm^{-2} (pink). Shaded regions represent the 1-σ (68%) binomial confidence intervals. Here we connect the center of the radial bins to highlight the difference in the distributions. We see that for column densities less than 10^{14} cm^{-2} show little correlation with galaxies. The covering fraction at $R < R_{\rm vir}$ in for the highest column densities ($N_{\rm HI} > 10^{15}$ cm^{-2}) never gets higher than 0.7.

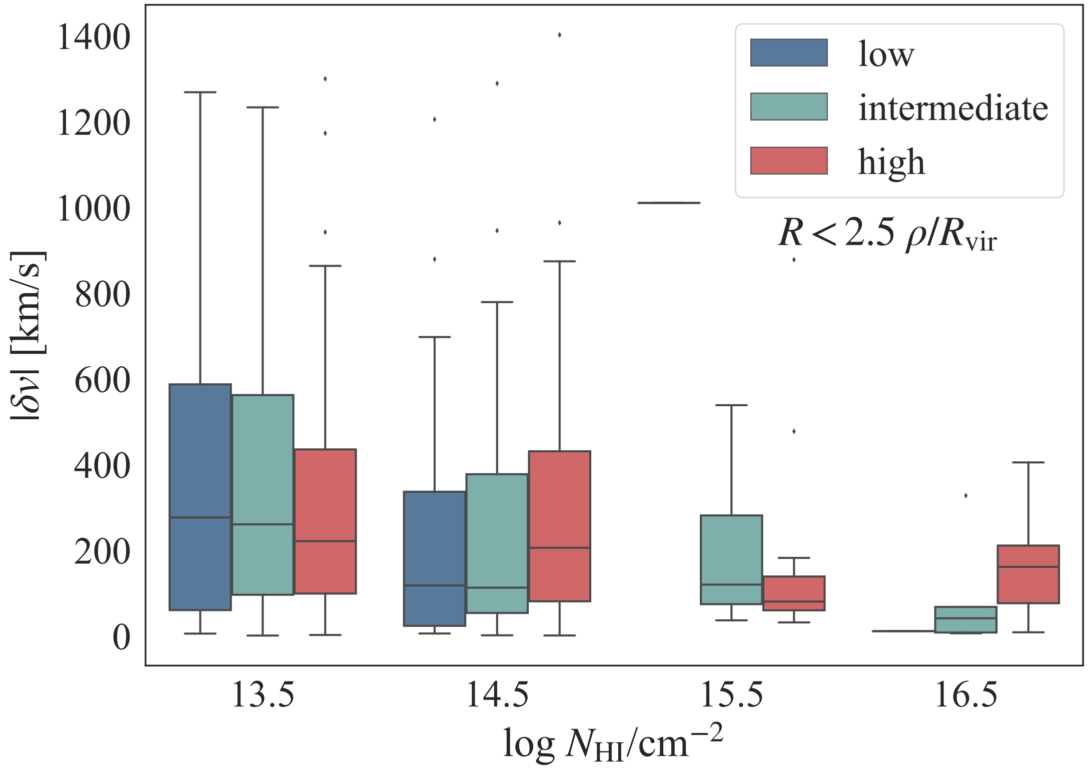


Figure 2.12: The distribution of absolute velocities as a function of $N_{\rm HI}$ displayed in a box and whisker plot for each mass sample. The boxes display the quartiles of the distribution centered at each bin spanning 1 decade in column density while the whiskers extend to show the rest of the distribution of the bins. Outliers are defined as points that lie outside 1.5 times the innerquartile range and are displayed as small diamonds. Each bin is split into a high (blue), intermediate (green) and low-mass (red) sample. We see a strong anticorrelation between velocity spread and column density. The single line seen at higher column densities in the low-mass sample indicate there are is only one point in each column density bin. This highlights how rare high-column density systems are around low-mass galaxies.

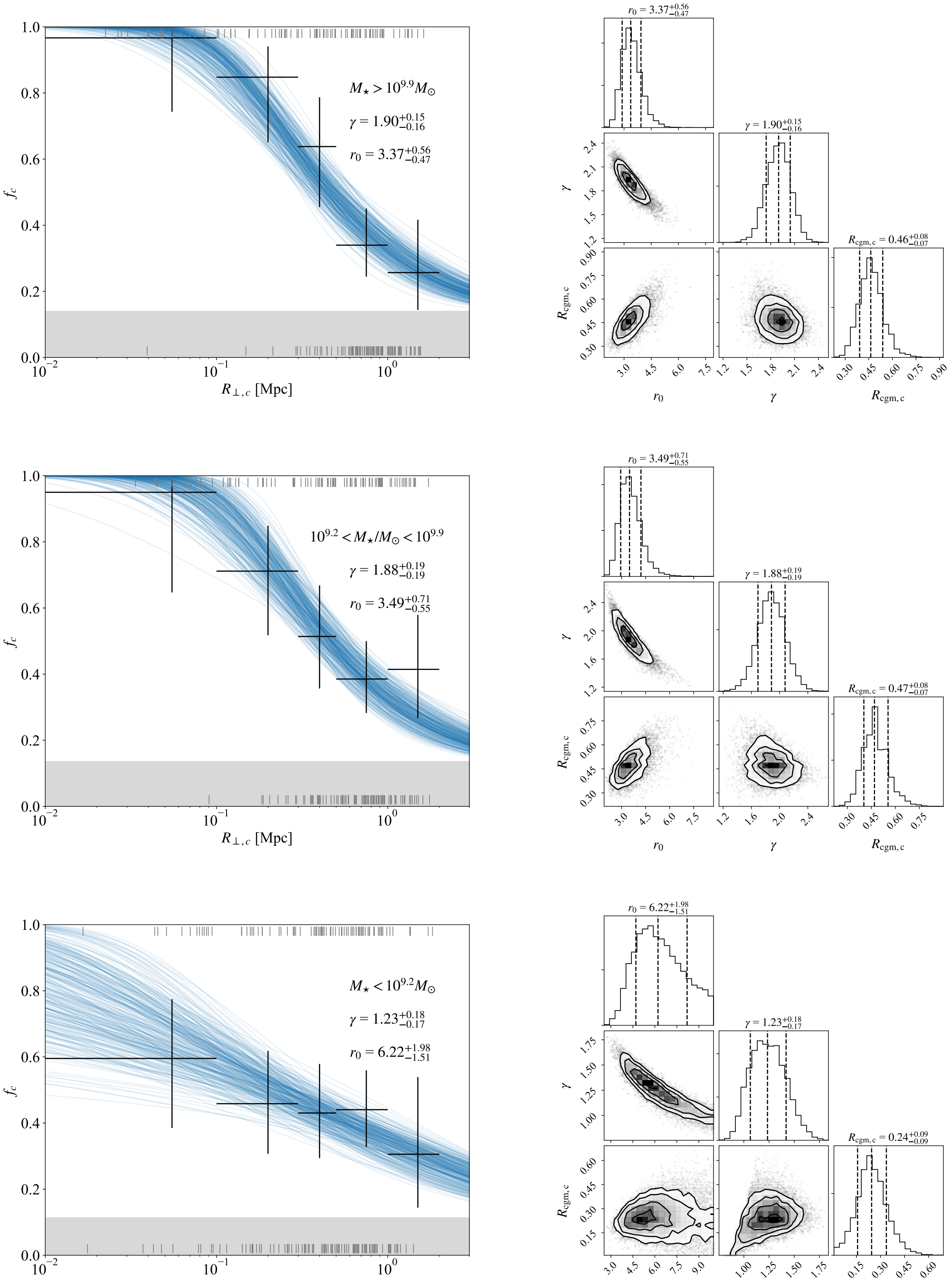


Figure 2.13: Left: covering fraction of H I gas with $N_{\rm HI} > 10^{14}$ cm^{-2} as a function of the physical impact parameter in comoving Mpc for galaxies in order of decreasing mass. The black crosses are binned evaluations of the covering fraction in each bin with 95% binomial confidence limits on covering fraction. The small grey ticks at the top and bottom of the figures indicate impact parameter for those systems that were hits (top) and misses (bottom). Beneath the black crosses are samples drawn from the posterior distribution showing the range in the γ–r_0 parameter space. The grey boxes on the bottom of the plots on the left indicate the covering fraction for sightlines taken at random, $\langle d\mathcal{N}/dz \rangle \delta z$. The plots on the right further illustrate the one and two dimensional projections of the posterior probability distributions of our parameters as well as indicate the median values for each.

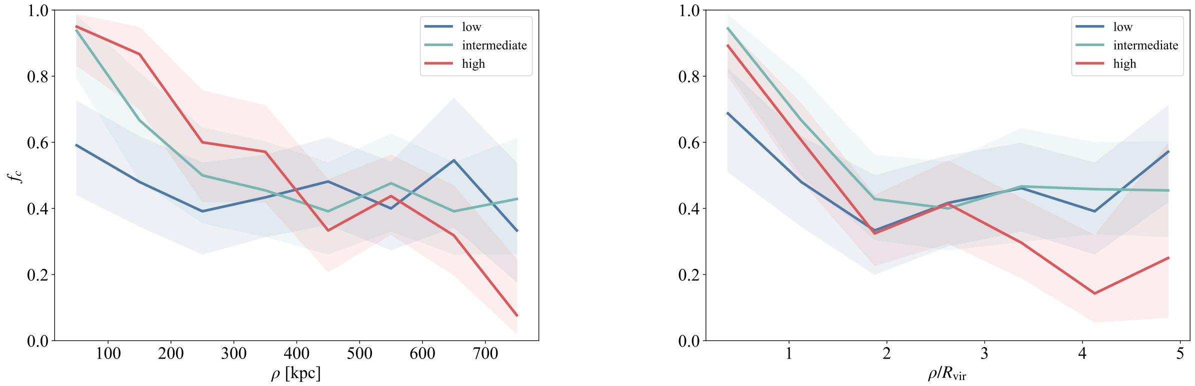


Figure 2.14: Covering fraction of H I as a function of impact parameter ρ (left) and $\rho/R_{\rm vir}$ (right) highlighting $N_{\rm HI} > 10^{14}$ cm^{-2} broken down into the three mass samples. The solid lines correspond to the covering fraction in bins of 100 physical kpc (left) and 1 $R_{\rm vir}/$ (right). The shaded region encodes the 1σ error in the covering fraction measurement. The covering fraction decreases monotonically in either measure to $\sim$ 300 kpc or $2R_{\rm vir}$. We find that the covering fraction in the lowest radial bins is consistent with $\sim$ 100% in each case except for the lowest mass sample. We note that the intermediate mass sample extends to further radius when looking at the normalized impact parameter (right) as compared to either of the other mass samples.

Chapter 3

CGM^2 + CASBAH: THE MASS DEPENDENCE OF HI-GALAXY CLUSTERING AND THE EXTENT OF THE CGM

Summary

We have examined the associations of galaxies with Lyα absorption $z < 0.48$ to explore the spatial profile of this gas and the mass dependence of the profile. Specifically, we have combined the CGM^2 and CASBaH HI and constructed a catalog of 7244 absorber-galaxy pairs around 29 QSO sightlines. The CGM^2 survey has better sampling of galaxies at low impact parameter while CASBaH samples galaxies out to 20 cMpc. This allows us to characterize the HI profile via the covering fraction as a tracer of the gas. By modelling the covering fraction as a power-law with a mass dependent length scale, we find good agreement of our clustering amplitude and power law slope parameters with previous studies such T14. We find a the clustering scale is mass dependence with a power-law slope of $\beta^{2h} = 0.08$ The single power-law model faithfully reproduces the data at the 1σ level for all galaxies $M_\star > 10^8 M_\odot$. We also model the data with an exclusionary two-component model where we use a Gaussian to model the data at smaller impact parameters and the customary power-law model at larger impact parameters. This model also faithfully reproduces the data for galaxies $M_\star > 10^8 M_\odot$. The two component model allows us to calculate the cross over radius, $R_{\rm cross}(M_\star)$, where the models are equal. $R_{\rm cross}(M_\star)$ represents a soft upper estimate of the furthest impact parameter needed to optimally fit the inner 1-halo term. We then use $R_{\rm cross}$ as an estimate of the exten of the CGM and find $R_{\rm cross}(M_\star) \approx 2 \pm 0.6 R_{\rm vir}$ for galaxies at all masses. Finally, we find no correlation between HI and galaxies for $M_\star < 10^8 M_\odot$.

3.1 Introduction

The formation and evolution of galaxies involves a complex interplay between gravitational collapse of gas from the intergalactic medium (IGM), and feedback due to stellar evolution and active galactic nuclei (AGN) that drive gaseous outflows and change the ionization state of the galaxies' gaseous halos. Together, these processes drive the "cosmic baryon cycle" which takes place largely in the region of a galaxy referred to as the circumgalactic medium (CGM). Indeed, understanding the circumgalactic medium is critical for understanding galaxy evolution as highlighted by the recent decadal survey (Peeples et al., 2019a). The extent of the gaseous CGM relative to the extent of the dark matter halo is a subject of great interest for models that aim to reproduce the properties of gaseous halos.

The existence of the CGM, first predicted by Bahcall and Spitzer (1969), was initially revealed by detection of MgII and HI absorption at subtantial impact parameters from $L*$ galaxies (Bergeron and Stasińska, 1986; Lanzetta et al., 1995; Chen et al., 2005), and subsequently higher-energy metal-line transitions such as SiIII, CIV and OVI that strongly correlate with galaxies and their global properties (e.g. Tripp et al., 2008; Prochaska et al., 2011; Tumlinson et al., 2011; Werk et al., 2013). Within 0.5 $R_{\rm vir}$ - metal line incidence – and vice-versa, Berg et al. (2022) find an 80% chance of finding a massive galaxy nearby to any metal-line absorber. The CGM of M $>$ 10^8 $M_\odot$ galaxies is now well-established to be metal-enriched (Bordoloi et al., 2014; Prochaska et al., 2017; Berg et al., 2022), and to likely extend to at least 1 $R_{\rm vir}$ Wilde et al. (2021); Borthakur (2022).

These observations are in accordance with hydrodynamical simulations of galaxy evolution which exhibit complex interactions between gravitational collapse from the cosmological large scale structure and the subsequent feedback from supernovae and AGN driven winds that heat and enrich the CGM and IGM (EAGLE Schaye et al. (2015), IllustrisTNG (Pillepich et al., 2018), SIMBA (Davé et al., 2019)), and CAMELS (Villaescusa-Navarro et al., 2022)). Yet these models still rely on prescription of the "sub-grid" physics in order to model entire galaxies (e.g. Ford et al., 2013; Hummels et al., 2013) and physical properties of the CGM are dependent of the simulation resolution (Hummels et al., 2019). Better observations if the CGM are needed to break discrepancies in these models, e.g., heating and cooling mechanisms or to develop a flexible paramteric model of the CGM (Singh et al., 2021).

The two-point correlation function between HI and galaxies has proved to be an essential tool to understand the connection of galaxies to the IGM (e.g. Chen et al., 2005; Ryan-Weber, 2006; Prochaska et al., 2011). The advantages of leveraging the clustering of these two entities over one-to-one association analyses is that it provides results for large scales

(1-10 Mpc) as well as small scales where the baryonic processes associated with the CGM inhabit. Since HI traces both enriched material from galaxies as well as primordial in-fall from the IGM, observations of the CGM, IGM and galaxies in the same volume are fundamental to testing predictions and providing a means to discern between different physical models (e.g. Fumagalli et al., 2011; Oppenheimer et al., 2012; Stinson et al., 2012; Ford et al., 2013; Hummels et al., 2013; Butsky et al., 2020; Singh et al., 2021)

Understanding the physical profile and size of the CGM sheds light on the non-linear processes of galaxy formation: at what scale do virialization, accretion and feedback transform gas around galaxies? Astronomers have long used the virial radius as an estimator for the size of galaxy halos but this is rather arbitrary and only traces the unobservable dark matter. By observing the gas profile around galaxies, we can map the gaseous halo. This boundary has other important implications. The galaxy baryon and metal budgets require a scale to integrate the total mass (Werk et al., 2013, e.g.). The quenching of satellite galaxies as they cross the threshold into the hosts gaseous halo is thought to be due to the stripping of the satellite host's CGM (Putman et al., 2021). Prochaska et al. (2011) and more recently Borthakur (2022) have shown the existence of HI outside the virial radius, thus we need a more robust estimator of the extent of the CGM.

In Wilde et al. (2021) (Paper I) we found an empirical relation between galaxy stellar mass and the extent of the CGM as indicated by HI covering fractions. In this paper, we focus on the functional forms of the mass dependence of the HI-traced CGM using 1-halo and 2-halo correlation functions. We leverage the CGM^2 which focuses on low impact parameters galaxies (< 1 Mpc) with the much larger COS Absorption Survey of Baryon Harbors (CASBaH) (< 20 Mpc) to greatly increase the absorber-galaxy sample from 543 absorber-galaxy pairs to 7244 spanning $0.003 < z < 0.48$.

The paper is structured as follows: In Section 3.2, we briefly review each of the survey and their combined properties. In Section 3.3, we cover the two ways in which we model the HI-galaxy correlation functions and cover our main results in Section 3.4. We compare our results and their context in the field in Section 3.5 and summarize our results in Section 3.6.

3.2 Data - Combining CGM^2 and CASBaH

Both surveys feature far-ultraviolet spectroscopy of QSOs with HST, using both the Cosmic Origins Spectrograph (COS, Green et al., 2012) and the Space Telescope Imaging Spectrograph (STIS, Woodgate et al., 1998)), and deep, ground-based optical spectroscopy of foreground galaxies in the QSO fields. CASBaH is well suited to the study of the interface between the CGM and the IGM, at scales $\gtrsim 1$ Mpc. CGM^2 provides a relatively more complete mapping of the inner CGM at scales $\lesssim 1$ Mpc.By combining the CGM^2 and CASBaH surveys, we exploit the strengths of each, which are described below. Figure (3.1) shows the distributions of galaxy redshifts and impact parameters from both surveys out to $z = 0.5$. Together, the surveys allow us to probe the CGM as it transitions into the IGM for a large sample of galaxies. In this section, we describe the data that make up each survey and how we combined them for this analysis.

3.2.1 CGM^2

The CGM^2 survey, first presented in Wilde et al. (2021), includes precise spectroscopic redshifts and bulk galaxy properties (e.g. stellar masses, M_*, and star formation rates, SFR) from a combination of Gemini GMOS spectra and deep, broadband photometry for ∼1000 galaxies in the foreground of 22 QSOs, each with S/N ≈10 HST/COS G130M+G160M spectra. By matching galaxy and absorber redshifts in ±500 km s−1 windows, the CGM^2 survey is ultimately a large collection of measurements pertaining to the CGM of $z < 1$ galaxies over a wide range of stellar masses, $10^8 \gtrsim \text{Log } M^* \, (M_\odot) \lesssim 10^{11.5}$. The details of the data acquisition and analysis are explained in detail in Wilde et al. (2021). Here we present a brief overview of the survey data relevant to the present analysis.

The CGM^2 galaxy spectra were obtained using Gemini-GMOS spectrographs on the twin Gemini North and South telescopes (Hook et al., 2004; Gimeno et al., 2016). Galaxy redshifts were inferred from the template fitting code, [1] (v0.14) and manually inspected

---

[1] https://github.com/desihub/redrock

with VETRR[2]. The typical statistical uncertainly of our redshifts is $\sigma_z \sim$ 50-100 km s^{-1} ($z \simeq$ 0.00016-0.00030). Photometry of the CGM^2 galaxy catalog was obtained from the Gemini-GMOS imaging in g and i bands as well as all available bands from DESI Legacy Imaging Surveys Data Release 8 (DR8) (Dey et al., 2019), WISE (Cutri et al., 2013), Pan-STARRS Data Release 2 (Chambers et al., 2016), and SDSS DR14 (Abolfathi et al., 2018).

The 22 QSOs included in the CGM^2 survey have HST/COS spectra were selected from the COS-Halos (GO11598, GO13033; Tumlinson et al., 2013) and COS-Dwarfs (GO12248; Bordoloi et al., 2014) surveys. In general, the CGM^2 QSO targets have either $z_{\rm QSO} > 0.6$ or available HST imaging, which permits detailed morphological analysis of absorption-hosting galaxies with $z < 0.5$. All COS spectra include both the G130M and G160M gratings, and have a S/N $\simeq 10 - 12$ per resolution element (FWHM $\simeq$ 16-18 km s^{-1}) or better over 1150-1800 Å. The COS data and their reduction are presented in detail in Tumlinson et al. (2013) and Bordoloi et al. (2014) and follows the same method used by Tripp et al. (2011), Meiring et al. (2011), Tumlinson et al. (2011) and Thom et al. (2012).

3.2.2 CASBaH

CASBaH obtained both HST/COS and HST/STIS spectra of nine QSOs at $0.92 < z_{\rm QSO} < 1.48$, its primary goal being a detection of the Ne viii doublet to constrain the warm-hot phase of galaxy halos (Burchett et al., 2019). For this reason, the CASBaH survey QSO spectra have higher S/N ratios than those of CGM^2 , $\approx 15 - 50$ per resolution element, and the targets have $z \gtrsim 1$ to probe a large enough pathlength to accumulate a statistically useful sample of Ne viii absorbers. The HST data for CASBaH consist of spectra with COS FUV G130M and G160M gratings, the COS NUV G185M and G225M gratings, and the STIS E230M echelle mode. The QSO selection criteria and data reduction are described in detail in (Burchett et al., 2019).

The CASBaH galaxy survey consists of 6701 galaxies redshifts with $z < 0.481$, appro-

---

[2] https://github.com/mattcwilde/vetrr

priate for the HI analysis presented here. The redshifts have typical uncertainties of 30 km s^{-1}.

3.2.3 Galaxy Properties

To estimate the galaxy properties for both surveys, we used CIGALE (Noll et al., 2009; Boquien et al., 2019) to fit the spectral energy distribution (SED) and retrieve stellar mass and star formation rates (SFR). We used the Bruzual and Charlot (2003) stellar population models, assuming a Chabrier (2003) initial mass function (IMF). We chose a grid of metallicities ranging from 0.001-2.5$Z_{\odot}$. A delayed star formation history (SFH) model was employed with an exponential burst. The e-folding time of the main stellar population models ranged from 0.1-8 Gyr. We varied the age of the oldest stars in the galaxy from 2-12 Gyr. We included an optional late burst with an e-folding time of 50 Myr and an age of 20 Myr. The burst mass fraction varied from 0.0 or 0.1 to turn this feature on or off. Nebular emission and reprocessed dust models (Dale et al., 2014) were also included with the default values. The dust models have slopes ranging from $1 - 2.5$ and the nebular models include no active galactic nuclei.

We employed the Calzetti et al. (1994) dust attenuation law, but we also included a "bump" in the UV (see discussion in Prochaska et al. (2019)) at 217.5 nm with a FWHM of 35.6 nm. The bump amplitude is set at 1.3 and the power law slope is -0.13 (Lo Faro et al., 2017). We varied the color excess of the stellar continuum from the young population, E(B-V), from 0.12-1.98. Finally, we used a reduction factor of 0.44 to the color excess for the old population compared to the young stars.

CIGALE then provides us with Bayesian estimates for the stellar mass and SFR for each galaxy in the combined catalog. In order to calculate the virial radius we used the abundance matching method of Moster et al. (2013) with the modifications used in Burchett et al. (2016). We adopt the convention of using $R_{\mathrm{vir}} = R_{200m}$, the radius within which the average mass density is 200 times the mean matter density of the universe, as the virial radius (R_{vir}) of a galaxy halo.

3.2.4 Combining the CGM^2 and CASBaH Surveys

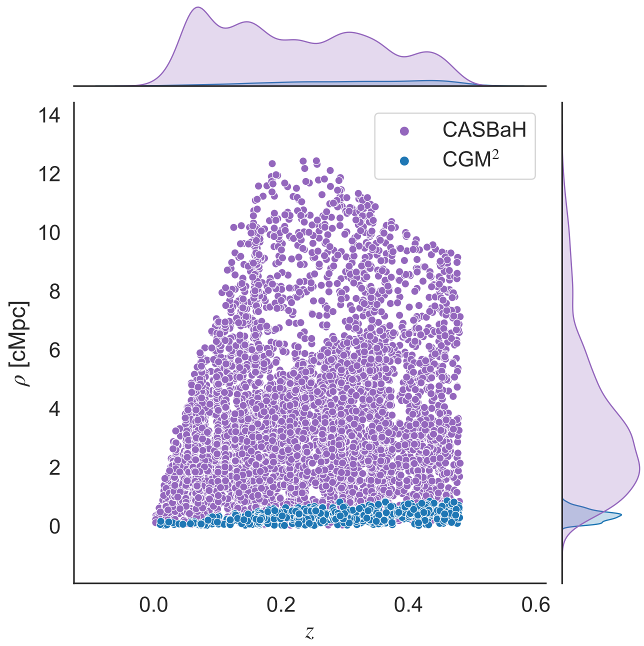
In order to combine the surveys, modifications were needed to both catalogs to ensure the same matching criteria between galaxies and absorbers. In the original CGM^2 survey, we measured the 2σ upper limit within $\delta v = \pm 30$ km s^{-1} of the galaxies redshift using the normalized error of the quasar flux when no absorption system was found within our $|\delta v| <$ 500 km s^{-1} window. In order to match the CASBaH survey, we adjusted this to a 3σ upper limit. This did not change our results in a meaningful way. The original CASBaH survey used a velocity window of $|\delta v| <$ 400 km s^{-1} to match the galaxies to absorption systems. We adjusted this for this work to $|\delta v| <$ 500 km s^{-1} to match the CGM^2 survey.

As in Paper I, we restrict our HI measurements to those less than $z < 0.481$ since at this redshift, the Lyman-α line redshifts off the G160 grating and thus, we are only sensitive to higher order transitions.

Having made these two small changes to each survey, both can be combined to give us a total survey that includes 6629 galaxies spanning $\sim 0.01 - 8$ comoving Mpc in impact parameter around 28 QSO sightlines. The distribution of impact parameter and redshift can be seen in Figure 3.1.

3.3 Modelling Absorber-Galaxy Clustering

In Paper I, we model the CGM using an absorber-galaxy cross-correlation analysis. This technique is based on modelling the covering fraction as a binomial probability distribution of detections, where we measure a column density $N_{\mathrm{HI}} > 10^{14}$ cm^{-2}. Likewise, a non-detection is the case where we do not detect gas above this threshold. The model consists of two terms: the base rate due to the random incidence of absorbers greater than this threshold and an excess above this base rate due to the clustering of galaxy-absorber pairs. This model is employed in Wilde et al. (2021) and is similar to the one developed by Hennawi and Prochaska (2007) and follows more closely the analysis by Prochaska et al. (2019). In Wilde et al. (2021), we found a mass dependence of the extent of the CGM based on breading up



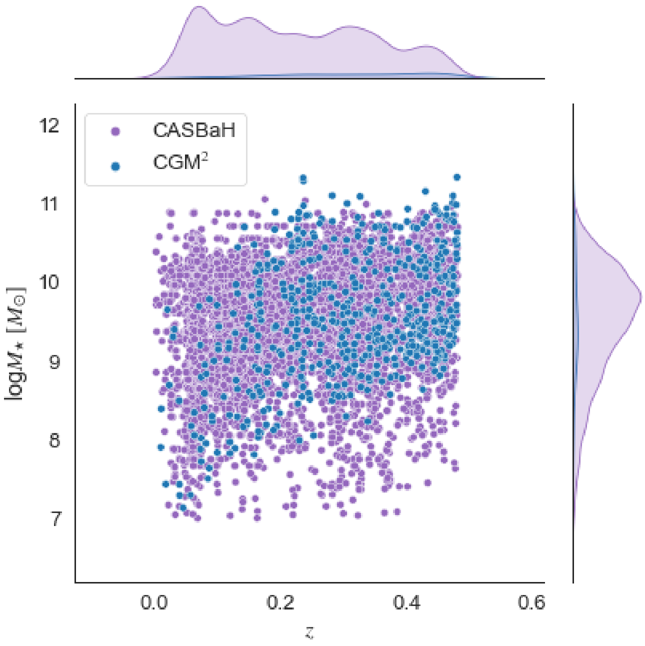


Figure 3.1: Top: Distribution of the the combined CGM^2 and CASBaH data sets in both impact parameter, ρ, and redshift. The data is roughly uniform in redshift space but we can see the relative contributions of the data sets in impact parameter space; CGM^2 is highly concentrated at lower impact parameters while CASBaH explores much greater impact parameters. Bottom: Mass distribution as a function of redshift of the data sets.

the data into three mass bins. In this work, we wish to quantify the mass dependence of the clustering as well as determine the redshift dependence given our data. The following section describes our formal model.

3.3.1 Single Power-Law Model

We define the 3D absorber-galaxy cross-correlation function, $\xi_{ag}(r)$ as

$$\xi_{ag}(r) = \left(\frac{r}{r_0}\right)^{-\gamma}. \tag{3.1}$$

To model the the mass dependence of the clustering, we add a new mass dependence to the clustering scale, r_0,

$$r_{0,m}(m) = r_0 \left(\frac{M_\star}{M_0}\right)^{\beta}. \tag{3.2}$$

As before, we examine the projected 2-D correlation function, which is obtained by integrating the 3-D correlation function over the line of sight

$$\chi_\perp(r) = \frac{1}{r_\parallel} \int_{r_\parallel} \xi_{ag}(\sqrt{r_\parallel^2 + r_\perp^2}\,)dr_\parallel. \tag{3.3}$$

For simplicity of notation, r is equivalent to $r_\perp$ in the following analysis.

In the following definitions we label the single power law clustering terms “2-halo” as the galaxy clustering method we adopt here describes the clustering of separate dark matter halos and distinguishes this method from the method we develop later in this manuscript.

We model the covering fraction, f_c, as a Poisson process,

$$f_c = 1 - \exp(-\lambda) \tag{3.4}$$

where we denote the rate of incidence as λ. We model the rate of incidence as the projected correlation function, the 2-halo term, as the excess over the probability of intersecting an absorber with $N_{\rm HI} > 10^{14}$ cm^{-2} in the redshift window,

$$\lambda = (1 + \chi_{\perp}^{2h})\ \langle d\mathcal{N}/dz \rangle \delta z. \tag{3.5}$$

In addition to parameterizing the mass dependence as in Equation (3.2), we would also like to parameterize the redshift dependence, $\langle dN/dz \rangle$. We do so as follows,

$$\frac{d\mathcal{N}(\mathrm{N_{HI}} \geq \mathrm{N_{HI}^{14}}, \mathrm{z})}{dz} = C_0(1+z)^{\alpha}, \tag{3.6}$$

where N_{HI}^{14} denotes absorbers with column densities of 10^{14} cm^{-2}, C_0 is the random rate of incidence at $z = 0$, and δz is the velocity window. We adopt a redshift window to be ±500 km s^{-1}. km s^{-1}.

Thus we have a rate of incidence of the form

$$\lambda = (1 + [\chi_{\perp}^{2h}(r, m | r_0^{2h}, \gamma^{2h}, \beta^{2h})])\ \langle d\mathcal{N}(z | C_0, \alpha)/dz \rangle\ \delta z. \tag{3.7}$$

Finally, we construct the likelihood function,

$$\mathcal{L} = \prod_i P_i^{\mathrm{hit}}(r, z | \theta) \prod_j P_j^{\mathrm{miss}}(r, z | \theta), \tag{3.8}$$

where $\theta = [r_0^{2h}, \gamma^{2h}, \beta^{2h}, C_0, \alpha]$.

As in Paper I, we apply the same Bayesian Markov Chain Monte Carlo (MCMC) sampler emcee (Foreman-Mackey et al., 2013) to generate samples from the posterior probability distribution function to estimate the parameters of interest and their distributions, using Equation (3.8).

In constructing our Bayesian model, we must choose priors. For the single power law parameters, we chose the priors based on the results of cross-correlation analysis in Tejos et al. (2014) except for our new mass dependent term, β^{2h}, which was motivated by physical arguments:

- $r_0^{2h} \sim \mathcal{N}(\mu = 3.2, \sigma = 0.3),\ r_0^{2h} > 0$

- $\gamma^{2h} \sim \mathcal{N}(\mu = 1.7, \sigma = 0.1)$, $\gamma^{2h} > 0$
- $\beta^{2h} > 0$

The priors for the redshift dependence were chosen based on the findings in Kim et al. (2021):

- $C_0 \sim \text{Lognormal}(\mu = 1.25, \sigma = 0.11)$, $C_0 > 0$
- $\alpha \sim \mathcal{N}(\mu = 0.97, \sigma = 0.87)$, $-3 < \alpha < 3$

We note that we chose to use the more recent results of Kim et al. (2021) in modelling the redshift evolution instead of that from Danforth et al. (2016), as were used in Paper I.

In Figure 3.2 we see the posterior distributions of our single power-law model, using the priors described above and $M_0 = 10^{9.5} M_\odot$. These were fit only to data with $8 < \log M_\star / M_\odot < 10.5$ as above this range there is a change in the virial radius due to the $M_\star - M_{\text{halo}}$ relation from abundance matching (Moster et al., 2013) which can be seen in Figure 3.7. Below this range we find a very flat covering fraction profile (Figure 3.5) which doesn't show a clustering signal.

3.3.2 Two-component Model

The single power-law model used in galaxy-galaxy clustering and adapted above to model the galaxy-absorber clustering makes no assumption of a CGM or overlapping gaseous halos. In the following section, we assume the existence of the CGM and use a Gaussian profile to model excess clustering signal due to the presence of the CGM. Specifically, the model is composed of a Gaussian 1-halo component combined with the 2-halo single power law and a component due to the random contribution from the IGM as used in the previous section. We employ an exclusion model where the contribution from the 2-halo term terminates at the distance it reaches the 1-halo component. This scheme, shown in Figure 3.3 also allows

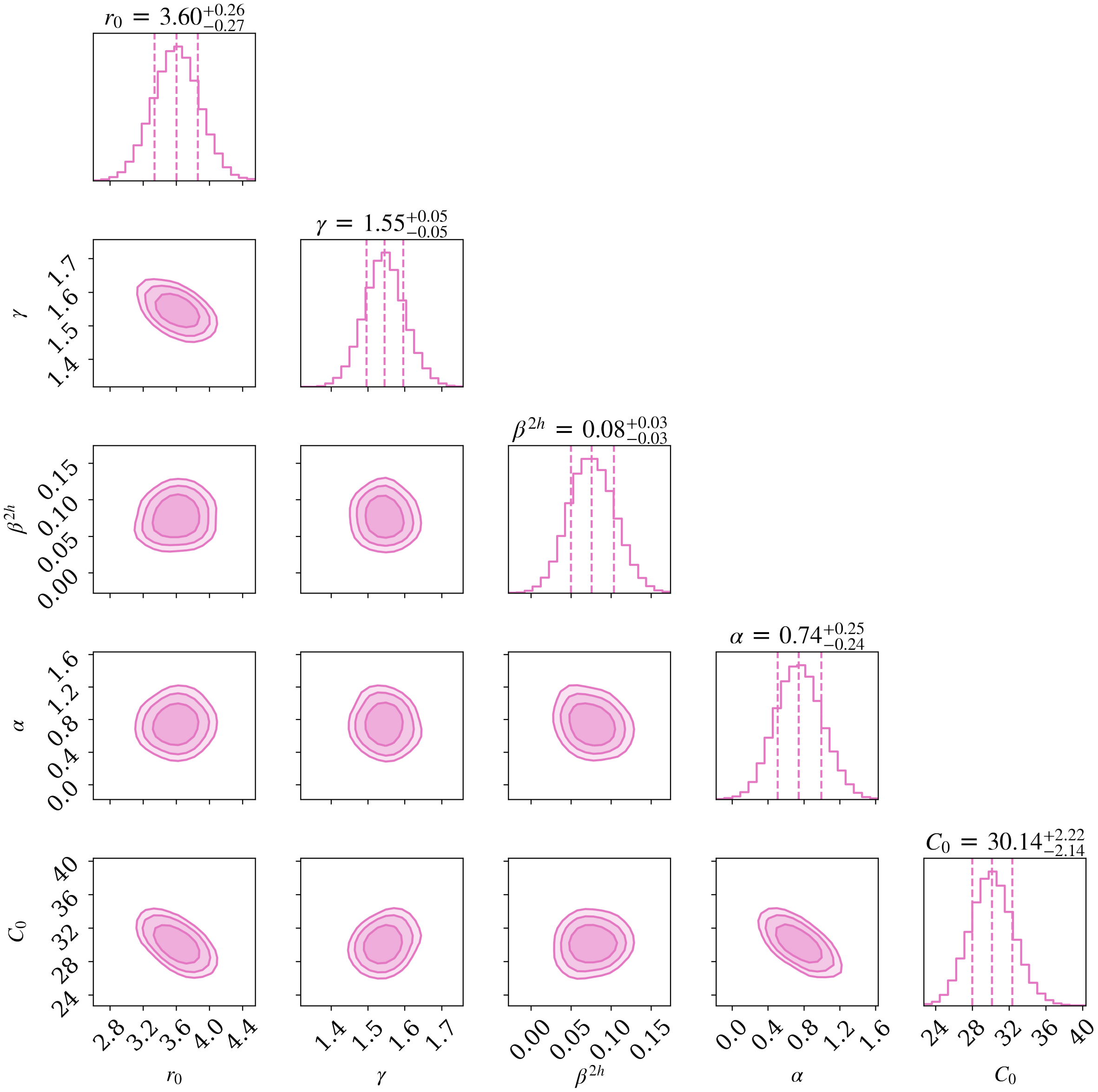


Figure 3.2: Corner plots showing the posterior parameter probabilities for the parameters in the single power-law clustering model. We find a non-zero, positive mass dependence term in the two-halo absorber-galaxy clustering, β^{2h}.

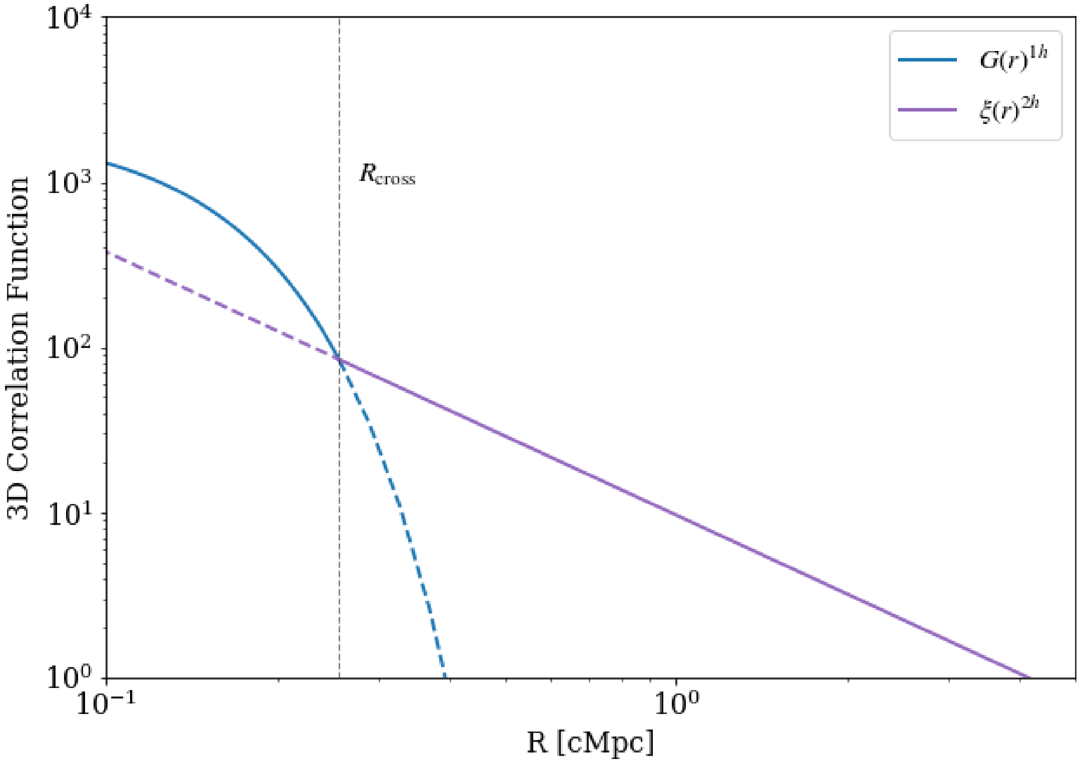


Figure 3.3: A schematic depiction of our two-component exclusion model and the determination of R_{cross}. The 2-halo component cuts off interior to R_{cross}.

us to determine a natural estimate of the extent of the CGM: the crossing point of the 1-and 2-halo components.

The model is similar to that single power-law we introduced before with a few key differences. We introduce a Gaussian one-halo term defined as:

$$G(r)^{1h} = Ae^{(r/\sigma)^2}. \tag{3.9}$$

Where the two models intersect, R_{cross}, we can solve for σ as

$$\sigma = \sqrt{\frac{1}{2}\frac{R_{\rm cross}^2}{\ln(A) + \gamma\ln(R_{\rm cross}/r_0)}}. \tag{3.10}$$

It should be noted that $R_{\rm cross}$ here is the 3-D distance and not the projected distance. In order to characterize the mass dependence of $R_{\rm cross}$ we define

$$R_{\rm cross} = \left(\frac{M_\star}{M_0}\right)^{\beta^{1h}}, \tag{3.11}$$

This parameterization allows us to compare the mass dependence of the 1-halo term, β^{1h} with that of the 2-halo term, β^{2h}.

In order to solve for the projected clustering signal, ξ, we first make some definitions to ease the notation. We use $s = r_\parallel$ in the remainder of the analysis. The integration is performed over different portions of the line of sight distance, s, corresponding to the 1 and 2-halo components. We define the line of sight crossing point $s_{\rm cross}$ as

$$s_{\rm cross} = \sqrt{\max(R_{\rm cross}^2 - r_\perp^2, 0)}, \tag{3.12}$$

and we can then integrate Equation 3.9 to $s_{\rm eval} = \min(s_{\rm cross}, s_{\rm max})$, where $s_{\rm max}$ is the maximum interval we wish to integrate over, which in our case is $[-500, 500]$ km s^{-1}. Thus we have

$$\chi(r_\perp) \propto 2\int_0^{s_{\rm eval}} G(r_\perp, s)^{1h} ds + 2\int_{s_{\rm eval}}^{s_{\rm max}} \xi(r_\perp, s)^{2h} ds \tag{3.13}$$

where the factor of 2 comes from the fact that both components are symmetric. Here we integrate the one-halo component over the more nearby regime out to $s_{\rm eval}$ and only integrate the 2-halo term beyond $s_{\rm eval}$ out to the maximum line of sight distance, thus excluding the regimes in which the models do not apply. For the two-component model, we choose fairly weak priors on unknown parameters based on physical arguments while following the same priors as described above for the parameters in the single power-law model:

- $\beta^{1h} > -3$
- $A > 0$
- $R_{\rm cross} > 0$

We can then follow the same MCMC fitting procedure described above to determine the posteriors for the parameters in this model as well as the crossing radius, $R_{\rm cross}$. These are shown in Figure 3.4. As before, we only fit data with $8 < \log M_\star/M_\odot < 10.5$ and use $M_0 = 10^{9.5}M_\odot$.

3.3.3 Model Comparison

In addition to comparing the two models to each other, Figure 3.5 compares the models to the empirical covering fraction as a function of impact parameter and mass. The data are shown in black with 1σ error bars. The single power-law model is shown in pink while the two component model is shown in purple. Both models recreate the covering fraction of the data in all mass bins except for the lowest mass bin where the clustering signal disappears. The two component model does significantly better for galaxies of $M_\star = 10^{9-10}M_\odot$ at the lowest impact parameters where the single power law model underestimates the covering fraction.

We find the 1-halo component has a stronger clustering mass dependence, $\beta^{1h} \simeq 0.14 \pm 0.07$, than the two-halo term, $\beta^{2h} \simeq 0.08 \pm 0.03$. We also find the 2-halo clustering terms in each model to be internally consistent with each other as seen in Figure 3.6

3.4 Results

3.4.1 Clustering Mass Dependence

As seen in Figure 3.2, we find the clustering parameters to be $r_0 = 3.6 \pm 0.3$, $\gamma = 1.6 \pm 0.5$. r_0 and γ are consistent with those found in Tejos et al. (2014) ($r_0 = 1.7 \pm 0.1$ and $\gamma = 3.7 \pm 0.3$. We also find a mass dependence of the absorber-galaxy clustering of $\beta^{2h} = 0.7^{+0.3}_{-0.2}$.

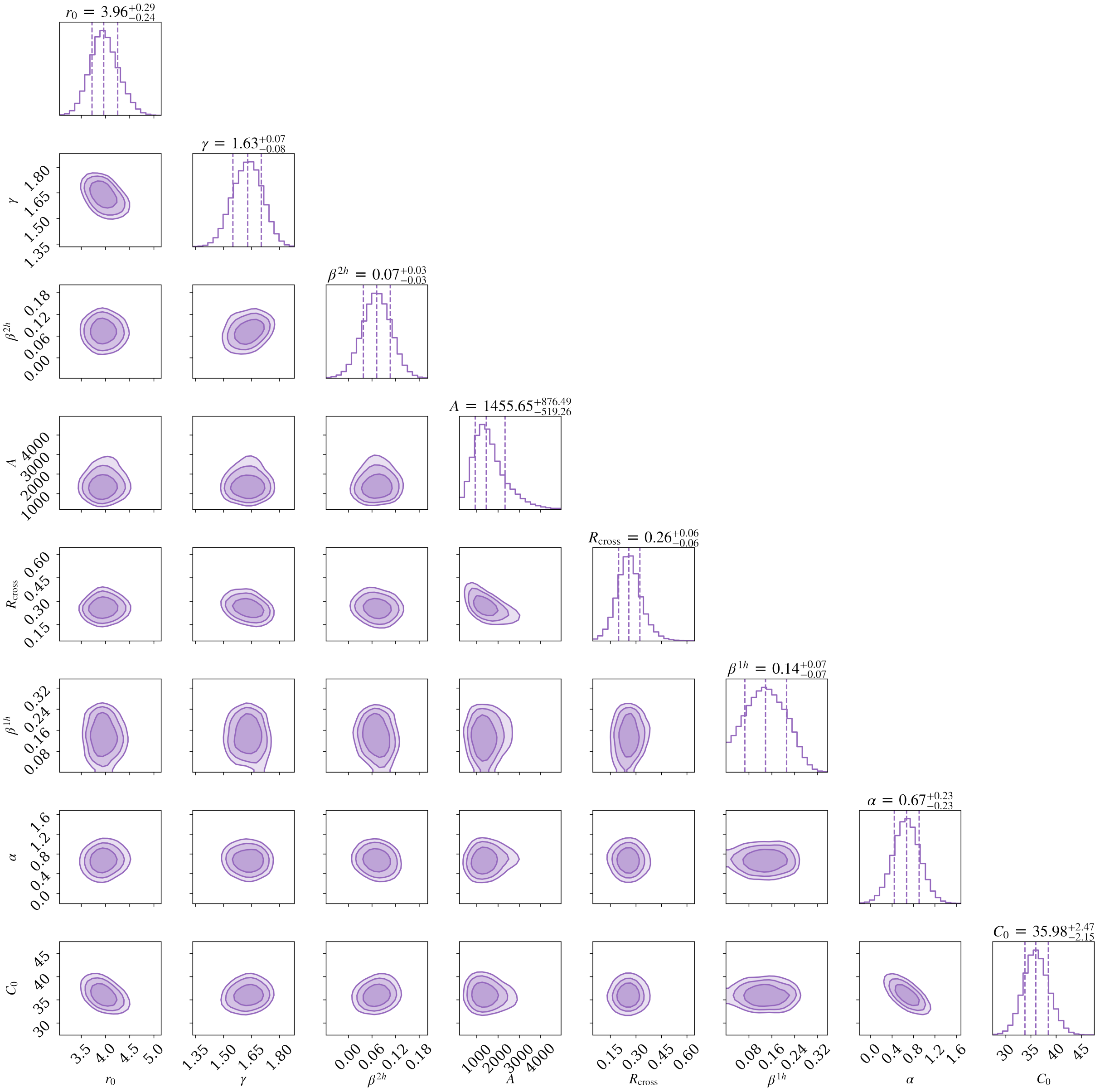


Figure 3.4: Posterior parameter probabilities for the parameters in the two-component clustering model. We again recover a non-zero, positive mass dependence term in the two-halo absorber-galaxy clustering, β^{2h} but find an even stronger one-halo CGM clustering mass dependence $\beta^{1h} \simeq 0.14 \pm 0.07$.

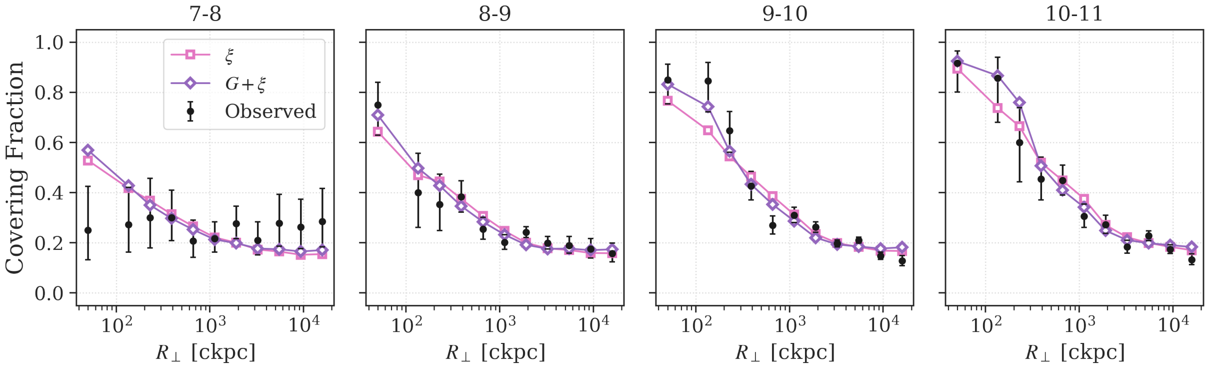


Figure 3.5: Comparison of our models to the empirical covering fraction as a function of impact parameter in comoving kpc in mass bins of $10^{7-8}M_\odot, 10^{8-9}M_\odot, 10^{9-10}M_\odot$ and $10^{10-11}M_\odot$. The data are shown in black with 1σ error bars. The single power-law model is shown in pink while the two-component model is shown in purple. Both models recreate the covering fraction of the data in all mass bins except for the lowest mass bin where the clustering signal disappears. The two-component model provides a better match to the data for galaxies of $M_\star > 10^9 M_\odot$ at the lowest impact parameters where the single power law model underestimates the covering fraction.

We find the the two component model better fits the data as can be seen in Figure 3.5. Specifically, the two component model does slightly better for galaxies of $M_\star > 10^{9-10} M_\odot$ at the lower impact parameters where the single power law model underestimates the covering fraction.

In addition, we find the two-component model reproduces the mass dependence of the 2-halo clustering term, $\beta^{2h} \simeq 0.7$ while also producing a stronger mass dependence of the 1-halo clustering term, $\beta^{1h} \simeq 0.14$.

3.4.2 Physically-Motivated Extent of the CGM

As mentioned above, using the two-component model produces an estimate of $R_{\rm cross}$, a natural metric for the extent of the CGM. This 3-D distance demarcates where the contribution to the clustering begins to be dominated by the CGM above the expected two-halo clustering

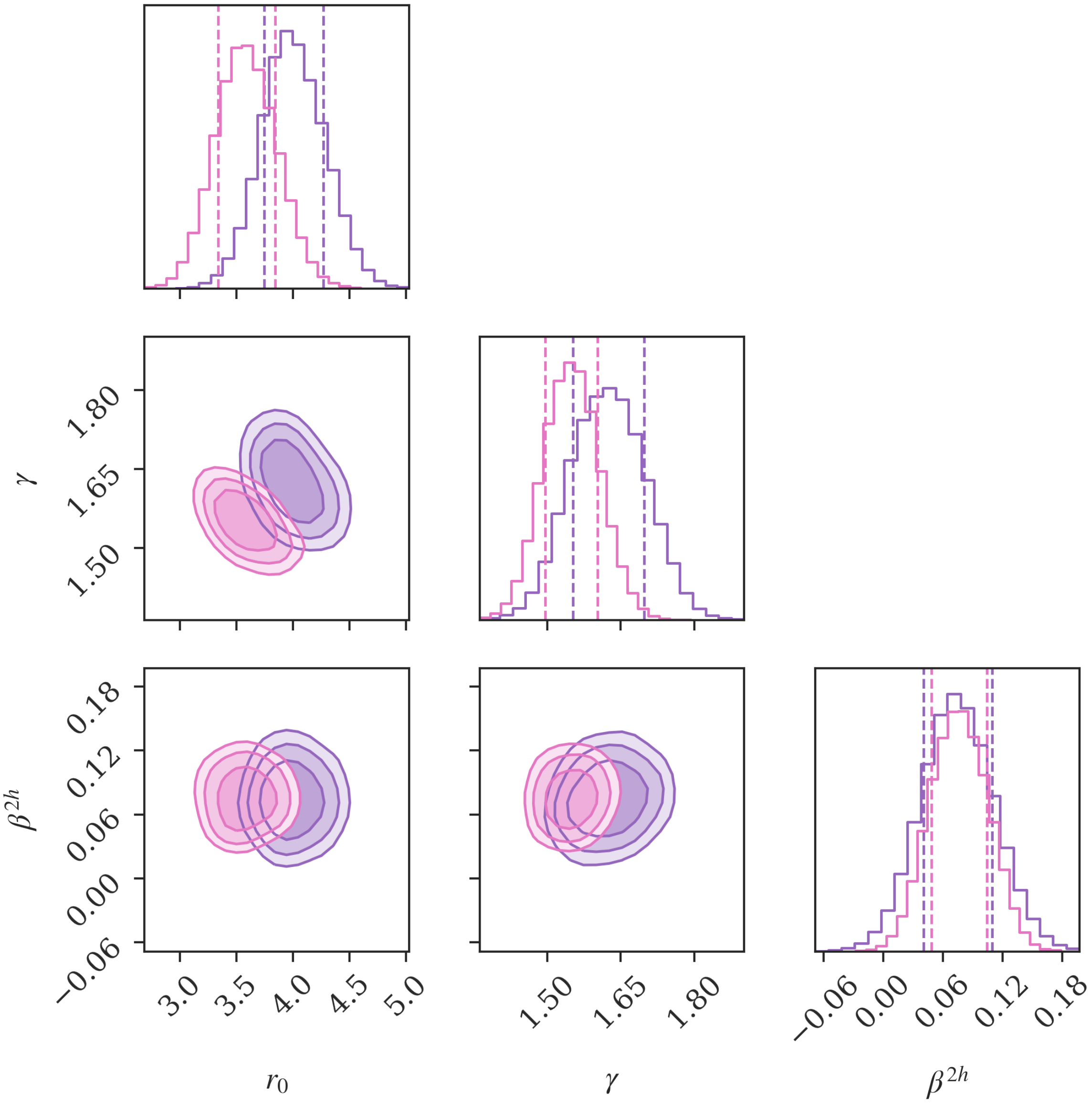


Figure 3.6: Comparison of the two-halo 3D cross correlation posteriors between the two-component model ($r_0 = 3.99^{+0.28}_{-0.24}$, $\gamma = 1.62 \pm 0.07$) and the single power-law model ($r_0 = 3.58^{+0.28}_{-0.24}$, $\gamma = 1.55 \pm 0.05$). The two models are consistent with each other within the 1σ limits and have a power-law slope consistent with the absorber-galaxy 3D cross correlation found in the literature (e.g. Tejos et al., 2014) of $\gamma = 1.7 \pm 0.1$.

due to isolated galaxy halos traced by HI. R_{cross} can be viewed as the maximum extent to which an enhancement from the CGM is needed to recreate the data. In Figure 3.7 we see R_{cross} compared with the spread in virial radii of the galaxy sample (grey filled region). The filled blue region represents the 1σ limits of the distribution in R_{cross} while the blue line denotes the median of this distribution. In addition, we compare R_{cross} with the method used in Wilde et al. (2021) (orange dashed line) of estimating R^{14}_{CGM} as the radius at which the probability of detecting HI above 10^{14} cm^{-2} surpasses 50% using the best fit parameters in our two-component model. The black crosses correspond to the values published in Paper I. The vertical dotted lines denote the mass range of $8 < \log(M_\star/M\odot) < 10.5$ that was used in our MCMC analysis. Above this range we see a change in relation of the virial radius with stellar mass and below this mass range, we find a much weaker correlation between absorbers and galaxies (see left most panel of Figure 3.5). We find excellent agreement of these three estimates of the CGM and find that R_{cross} is approximately $2R_{\mathrm{vir}}$.

In Figure 3.7, we compare the mass dependence of R_{cross} (blue line, shaded area) to the virial radius that spans the redshift of the galaxies in our sample (black shaded area). We find R_{cross} is $\sim 2 \pm 0.6 R_{\mathrm{vir}}$ for galaxies in the range $8 < \log(M_\star/M\odot) < 10.5$.

3.5 Discussion

Broadly, our results suggest that the 2-halo single power law model does an adequate job of recreating the cross correlation signal at all impact parameters and all masses $M_\star > 10^8 M_\odot$ as seen in Figure 3.5. While this does not preclude the presence of a CGM around these galaxies, it implies that the majority of HI absorption that one would associate to these galaxies on scales of $\lesssim 200\,\mathrm{kpc}$ is dominated by other, nearby galaxies.

The two-component model appears able to better recreate the data at smaller impact parameters and at higher masses as seen in Figure 3.5. For $R_{\mathrm{cross}} < 300\,\mathrm{kpc}$, one detects 52 HI systems where 46 systems are predicted with the single power-law model. Assuming Poisson statistics, the 2-halo model is consistent with the data at 1σ level, however.

The much larger number of galaxies at larger impact parameters drives the fit of the

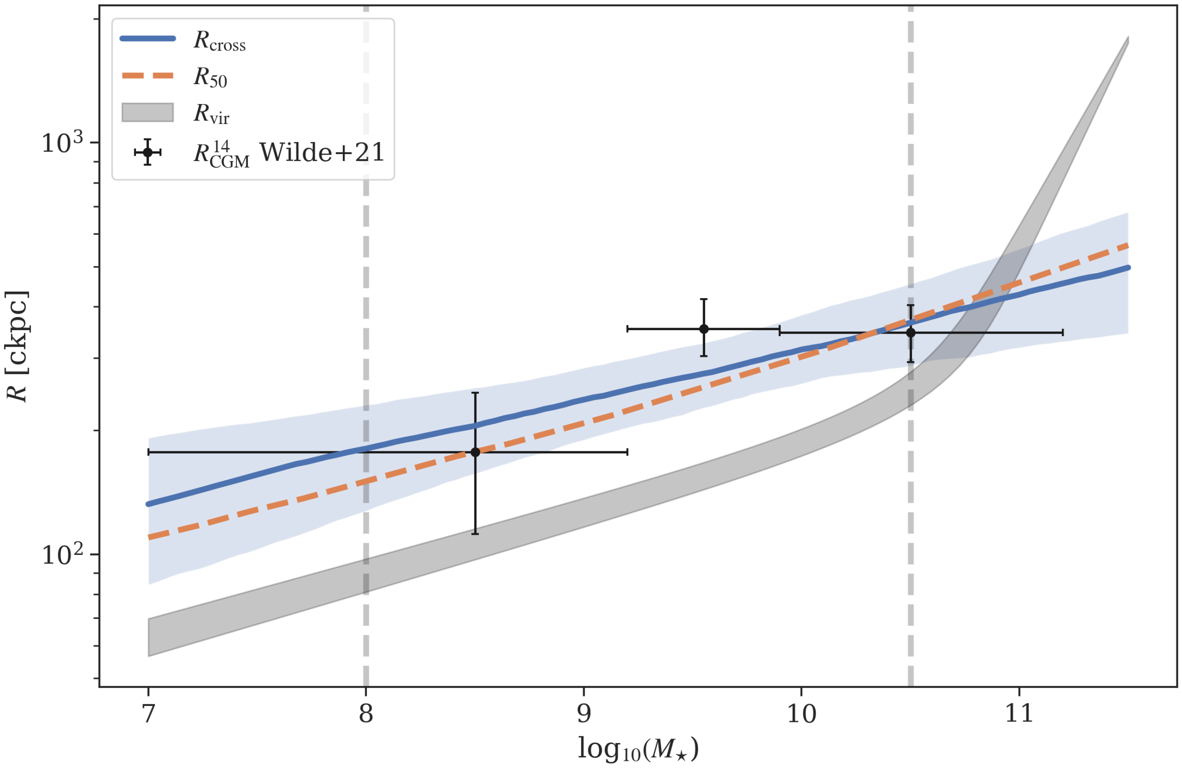


Figure 3.7: A comparison of $R_{\rm cross}$ with the spread in virial radii of the galaxy sample (grey filled region). The filled blue region represents the 1σ limits of the distribution in $R_{\rm cross}$ while the blue line denotes the median of this distribution. In addition, we compare $R_{\rm cross}$ with the method used in Wilde et al. (2021) (orange dashed line) of estimating $R^{14}_{\rm CGM}$ as the radius at which the probability of detecting HI above 10^{14} cm^{-2} surpasses 50% using the best fit parameters in our two-component model. The black crosses correspond to the values published in Paper I. The vertical dotted lines denote the mass range of $8 < \log(M_\star/M\odot) < 10.5$ that was used in our MCMC analysis. Above this range we see a change in relation of the virial radius with stellar mass and below this mass range, we find a much weaker correlation between absorbers and galaxies (see left most panel of Figure 3.5). We find excellent agreement of these three estimates of the CGM and find that $R_{\rm cross}$ is approximately $2R_{\rm vir}$.

models to the data. There is, however, a 2σ inconsistency between the models and the data at $R_\perp \sim 600$kpc in the $\log M_\star = 9 - 10 M_\odot$ mass range. This may be due to cosmic variance or the assumption that the absorber-galaxy measurements are independent and aren't correlated which would increase the scale of the error bars.

3.5.1 Comparing the mass dependence of the one and two-halo terms

Our galaxy sample includes a large number of galaxies at low (< 500 kpc) impact parameters which allows us to better model the regime in which the two-halo galaxy clustering becomes dominated by the signal of galaxies that inhabit the same dark matter halo, the one-halo term. By separating these two terms in the manner presented here, we can disentangle the large scale clustering as well as the contribution of the CGM to the 3D correlation of absorbers and galaxies.

Our analysis finds nearly identical terms for the mass dependence of the clustering at large scales, β^{2h} as well as the contribution of absorbers at random, C_0 and α. We do find a stronger mass dependence in the one-halo term, β^{1h} that at larger scales. This can be seen in Figure 3.5 where the correlation steepens in higher mass bins.

3.5.2 Absorber-Galaxy Bias

Using our single power-law analysis, we can compute the relative absorber-galaxy bias via the standard relation $\xi_{ag} = b_a b_g \xi_{\rm DM}$, where ξ_{ag} is the absorber-galaxy correlation function (eq. 3.1), b_a and b_g are the absorber and galaxy bias respectively and $\xi_{\rm DM}$ is the dark matter 3D correlation function. We can approximate the relative bias as $(r/r_0(M))^{-\gamma} = b_a b_g \xi_{\rm DM}$ where $\xi_{\rm DM}$ is of the same form with a $\gamma = 1.62$ as calculated by Tinker et al. (2010) (hereafter, T10). Thus, we can solve for the product of the bias, taking into account our mass dependent r_0 (Eq. 3.2) and using $\gamma = 1.62$ with is consistent with the value we determine from our data, as $b_a b_g \propto (M_\star/M_0)^{\gamma\beta}$. This is shown in Figure 3.8 where we find compare our results to those in T10 as encoded in the COLOSSUS package (Diemer, 2018). This is in-part due to the fact that our sample contains a fraction of star forming (SF) to quiescent galaxies (Q) that is non-uniform in mass (Wilde et al., 2021); our galaxy sample has a greater number of SF galaxies at lower masses than at higher masses and redder, quiescent galaxies are known to be more strongly clustered (Zehavi et al., 2002; Berti et al., 2021). It is also plausible that the lower mass galaxies live in lower density regions of the cosmic web. The fact that the bias is below that of T10 bias suggests that HI is less common in these lower density regions than at higher density regions such as filaments inhabited by the higher mass galaxies Prochaska et al. (2013). An alternative explanation is that the assumption that the two tracers are linear and independent does not hold and that baryonic physics affects the correlation between galaxies and absorbers.

3.5.3 Comparison to Previous Work

One of the key aspects of this analysis is determining the mass dependence of the extent of the $N_{\rm HI} > 10^{14}$ cm^{-2} for which our model provides a direct metric, $R_{\rm cross}(M_\star)$. We compare our resulting $R_{\rm cross}(M_\star)$ to the method and results from Paper I in Figure 3.7. The result of Paper I, $R^{14}_{\rm CGM}$, which are based only on the CGM^2 survey are shown as black crosses in the mass bins they span in that paper. We also compare the method used in that paper to determine $R^{14}_{\rm CGM}$, the radius at which the probability of detecting $N_{\rm HI} > 10^{14}$ cm^{-2} is $> 50\%$, calculated with the two-component model using the combined CGM^2 + CASBaH surveys (orange dashed line). We find that our mass dependent estimate of the extent of the CGM, $R_{\rm cross}(M_\star)$ corroborates the findings of Paper I that the $N_{\rm HI} > 10^{14}$ cm^{-2} extends to approximately twice the virial radius ($\sim 2 \pm 0.6 R_{\rm vir}$).

One of the main strengths of the CGM^2+ CASBaH sample is the large number of galaxies at small projected separations (<1 Mpc). This allows us to investigate the smaller scale regime in more detail within the context of similar studies such as Tejos et al. (2014) (hereafter, T14) who uses a single power law model to measure the two-point correlation between HI and galaxies above $N_{\rm HI}$ 10^{14} cm^{-2}. In this work they break up their measurements into SF vs non-SF samples while we do not. Our sample however is dominated by the more common SF galaxies and we will compare our results to their SF sample. Comparing our cross-correlation results with T14, we find good agreement between the results in T14, $r_0^{\rm T14} = 3.8 \pm 0.2$, $\gamma = 1.7 \pm 0.1$ and the results from both models presented here, $r_0 = 3.99^{+0.28}_{-0.24}$,

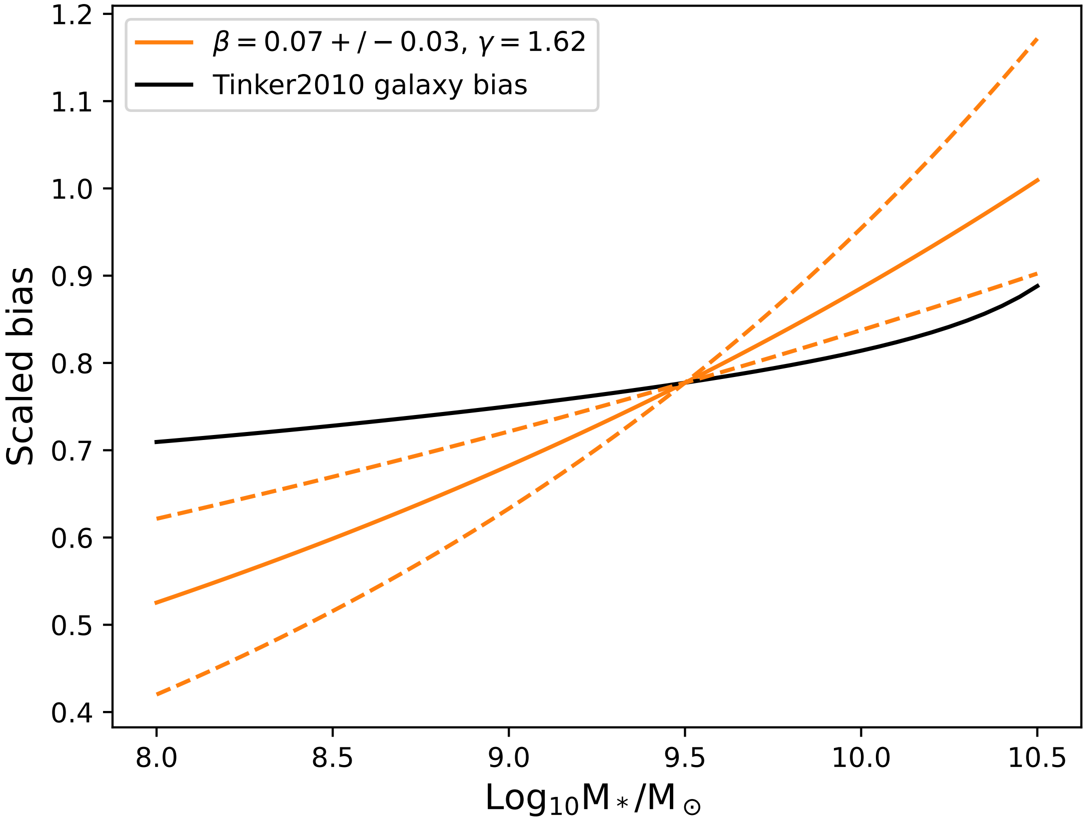


Figure 3.8: A comparison of the slopes of the relative bias as a function of mass derived from our analysis (orange) compared to Tinker et al. (2010) (T10, black). The dashed lines correspond to the ranges spanned by the 1σ limits in in β^{2h}. The relative bias, $r_0(M) \propto (M_\star/M_0)^{\gamma\beta}$, are normalized to the value of T10 at log $M_\star = 9.5$. We find a steeper mass dependence than T10 which is due in part to the proportion of SF/quiescent galaxies in our sample; the galaxies at lower mass have a greater proportion of SF galaxies which could steepen $b(M)$ due to known result that red galaxies being known to be more clustered.

$\gamma = 1.62 \pm 0.07$) and the single power-law model ($r_0 = 3.58^{+0.28}_{-0.24}$, $\gamma = 1.55 \pm 0.05$. We find a mass dependence of this cross-correlation, however as parameterized by β^{2h}.

Our results are in slightly in tension with Momose et al. (2021) who find galaxies in the $10^{9-10} M_\odot$ range dominate their HI-galaxy cross correlation signal. We find the largest mass bin sample to have the most elevated covering fractions at low impact parameter.

3.6 Summary

With this paper we have examined the associations of galaxies with Lyα absorption $z < 0.48$ to explore the spatial profile of this gas and the mass dependence of the profile. Specifically, we have combined the CGM^2 and CASBaH HI and constructed a catalog of 7244 absorber-galaxy pairs around 29 QSO sightlines. The CGM^2 survey has better sampling of galaxies at low impact parameter while CASBaH samples galaxies out to 20 cMpc. This allows us to characterize the HI profile via the covering fraction as a tracer of the gas.

1. By modelling the covering fraction as a power-law with a mass dependent length scale, we find good agreement of our clustering amplitude and power law slope parameters with previous studies such T14.

2. We find a the clustering scale is mass dependence with a power-law slope of $\beta^{2h} = 0.08$.

3. From this power-law slope we can compare the slope of our absorber-galaxy bias to that of Tinker et al. (2010) and find a steeper bias as a function of galaxy mass.

4. The single power-law model faithfully reproduces the data at the 1σ level for all galaxies $M_\star > 10^8 M_\odot$. While this does not preclude the presence of a CGM around these galaxies, it implies that the majority of HI absorption that one would associate to these galaxies on scales of $\lesssim 200\,\mathrm{kpc}$ is dominated by other, nearby galaxies.

5. We also model the data with an exclusionary two-component model where we use a Gaussian to model the data at smaller impact parameters and the customary power-law model at larger impact parameters. This model also faithfully reproduces the data for galaxies $M_\star > 10^8 M_\odot$.

6. The two component model allows us to calculate the cross over radius, $R_{\text{cross}}(M_\star)$, where the models are equal. $R_{\text{cross}}(M_\star)$ represents a soft upper estimate of the furthest impact parameter needed to optimally fit the inner 1-halo term. We then use R_{cross} as an estimate of the exten of the CGM and find $R_{\text{cross}}(M_\star) \approx 2 \pm 0.6 R_{\text{vir}}$ for galaxies at all masses.

7. Finally, we find no correlation between HI and galaxies for $M_\star < 10^8 M_\odot$.

Chapter 4

SDSS DR17: THE COSMIC SLIME VALUE ADDED CATALOG

Summary

A tenet of modern cosmology is the existence of the "cosmic web", a vast filamentary structure formed via the collapse of matter due to gravity. This structure is ubiquitous in cosmological simulations yet challenging to observe due to its diffuse nature. Recently, a new technique was developed which is inspired by the growth and movement of Physarum polycephalum slime mold (Burchett et al., 2020b), to map the cosmic web of a low redshift subsample of the SDSS spectroscopic galaxy catalog. This model, the Monte Carlo Physarum Machine (MCPM) was shown to faithfully reconstruct the cosmic web. Here, we make great strides in improving the methods used in MCPM and apply them to a larger cosmological volume using the classic SDSS and eBOSS spectroscopic surveys. We have constructed a catalog of galaxy positions with estimates for the cosmic density. In addition, we include the 3D density cubes of these volumes. These data products were released as part of Sloan Digital Sky Survey Data Release 17 (Abdurro'uf et al., 2022) and are now publicly available. We present the input catalogs as well as the methodology in constructing these data products. We also highlight exciting potential applications to cosmology, the intergalactic and circumgalactic medium and gravitational wave progenitor detection.

4.1 Introduction

The cosmic web is an established prediction of ΛCDM and is readily identifiable in simulations, where the underlying density distribution is known. However, we are primarily motivated by the problem of unveiling the latent large scale structure in the observational realm, where the structure is traced by galaxies and absorption signatures of the intergalactic medium. However, mapping the actual dark matter distribution remains a challenge as tracers such as quasar absorption lines are limited by the scarcity of quasars in the universe.

Of paramount concern in galaxy evolution science is the impact of a galaxy's environment on its morphology and star formation activity. Correlations between environmental metrics and galaxy properties, such as morphology (Dressler, 1980, e.g.), color (Abell, 1965, e.g.), and star formation (Peng et al., 2010) have been known about for many decades, but the physical mechanisms and their relative importance remain heavily pursued problems. Galaxy-environment analyses typically fall along one of two paths: local environment-centric or large-scale environment-centric. In the former, one employs an environmental density metric, such as a nearest-neighbor distance or density within some aperture (Kauffmann et al., 2004; Peng et al., 2010), or galaxies are associated with a local group or cluster environment (Yang et al., 2007; Berlind et al., 2016) and galaxy properties are studied with respect to the properties of the group or cluster (Carollo et al., 2013; Catinella et al., 2013).

The latter path is less straightforward, as one must infer the large scale structure from tracers, typically the galaxies themselves, and correlate galaxies back to that structure in some way. Various methods have been devised to reconstruct the cosmic web from discrete tracers. Libeskind et al. (2018) reviewed a number of these, and we refer to the reader to this valuable resource for an overview of the techniques employed and comparisons between them. Essentially, once the underlying density field is inferred, one can correlate galaxy properties with this density field (an approach one can directly employ with the catalog described here) or attempt to geometrically relate a galaxy's position to the structure identified, e.g., the distance to a filament. One should appreciate that filament identification (DisPerSE, Luber et al. (2019); Tempel et al. (2014), etc.), whether from a density field or some other methodology, is a separate problem from inference of the field itself.

Studies of galaxy-cosmic web dependencies report somewhat mixed results. Kuutma et al. (2017) find higher elliptical-to-spiral ratio and decreasing Star Formation Rate (SFR) towards filament spines. Similarly, Crone Odekon et al. (2018) report that, at fixed stellar mass, galaxies closer to filaments or in higher density environments are more deficient in HI. These large-scale effects on galaxies have also been investigated in the cosmological simulation domain. Codis et al. (2018) measure the spin-filament alignment in IllustrisTNG (Vogelsberger et al., 2014) and find a strong dependence on spin alignment with galaxy mass. Pasha et al. (2022) find that the collapse of large scale structure into sheets at higher redshifts ($z \sim 3$) can create shocks that explain quenching in dwarf galaxies similar to the effect seen in nearby dwarf galaxies that exist in the presence of clusters and groups.

Aside from the galaxies themselves, one may also study the intergalactic medium in context with the cosmic web environment. Wakker et al. (2015) measured the Lyα absorption in quasar spectra probing a foreground visually identified filament, finding increasing absorber equivalent width and linewidth with decreasing projected distance to the center of the filament. With a larger archival sample of QSOs and filaments, Bouma et al. (2021) find similar results, with Lyα absorbers showing both greater incidence and column density at small projected distance and velocity offsets from filaments first identified by Courtois et al. (2013). In the first application of the reconstruction framework we use here, Burchett et al. (2020a) analyze the Lyα optical depth as a function of cosmic web density probed by QSO sightlines and find three distinct regimes: a void regime at low matter overdensity with no detected absorption, an onset of absorption in the outer skins of filaments with monotonically increasing optical depth, and the highest-density regime where the absorption no longer increases with local density but rather turns over and declines at the highest densities.

In this manuscript, we employ the novel method first introduced in Burchett et al. (2020a) which is based on the movements of the organism Physarum polycephalum slime mold to map the cosmic density field. This model implicitly traces the Cosmic Web structure by efficiently finding optimal pathways between galaxies, which are known to trace filaments. We apply our model to two large catalogs of galaxies, the NASA Sloan Atlas (NSA) (Blanton et al., 2011) and the catalogs of Luminous Red Galaxies (LRGs) from SDSS-IV Extended Baryon Oscillation Spectroscopic Survey (Bautista et al., 2018). Our method faithfully reconstructs the cosmic matter density of the cosmic web for all points in the observational footprints, allowing the study of this dark matter distribution with any objects of interest in the footprints. We have released this data as part of the SDSS Data Release 17 (DR17) as a Value Added Catalog (VAC) that is open access for the communities use.

Unless stated otherwise, we adopt the Planck15 (Planck Collaboration et al., 2016) cosmology as encoded in the ASTROPY package (Astropy Collaboration et al., 2013; Price-Whelan et al., 2018).

4.2 Data

We need a three dimensional map as well as the mass of a galaxy field in order to infer the density field. To this end, we use spectroscopic surveys that provide us with precise redshifts as well combined with value-added catalogs that provide us with galaxy masses. We use two primary catalogs for our galaxy positions, the NASA-Sloan Atlas (NSA, or NSA/SDSS) for galaxies with $z < 0.1$ and the Large Scale Structure catalogs from Sloan Digital Sky Survey (SDSS) for galaxies at higher redshifts. These two catalogs each offer advantages and disadvantages and are described below. Note that no data used in this VAC came from DR17.

In addition to the two galaxy catalogs, we need to map our MCPM density estimates to the cosmic matter density ($\rho_m/\langle\rho_m\rangle$). We employ the dark matter only Bolshoi-Plank ΛCDM (BP) simulation as described below.

4.2.1 NASA Sloan Atlas

The NASA Sloan atlas (NSA) is a value added catalog constructed from reprocessed SDSS *ugriz* photometry combined with Galaxy Evolution Explorer (GALEX) photometry in the ultraviolet. It was designed to improve upon the standard SDSS sky subtraction pipeline (Blanton et al., 2011). We use the most recent version of this catalog, nsa_v1_0_1.fits, which contains galaxies out to $z = 0.15$. For our purposes, we cut the catalog to those galaxies with $z = 0.1$ resulting in a catalog of 325321 galaxies. We will often refer to this catalog in this paper as simply “NSA/SDSS” to distinguish it from the other catalogs that come from BOSS.

4.2.2 LRG catalogs

For the higher redshift portion of our catalog, we use a sample of Luminous Red Galaxies (LRG) from the Baryon Oscillation Spectroscopic Survey (BOSS), part of the Sloan Digital Sky Survey (SDSS) III project, which at the time of its release provided the largest survey of galaxy redshifts available, in terms of both the number of galaxy redshifts measured by a single survey, and the effective cosmological volume covered. We chose to use the LRG catalogs as tracers of the DM density as these catalogs are more complete at these redshifts than using a more general galaxy catalog. The LRG sample of from BOSS comes from the Large scale structure catalogues provided by the team and are broken into Northern and Southern Galactic caps (LRG-NGC and LRG-SGC, respectively) (Ross et al., 2011; Ho et al., 2012; Ross et al., 2012). We use the LOWZ catalogs which provide a sample of LRGs to $z \lesssim 0.5$ and are found in these files[1]: galaxy_DR12v5_LOWZ_North.fits.gz and galaxy_DR12v5_LOWZ_South.fits.gz. We chose to use the LRG catalogs as tracers of the DM density as these catalogs are more complete at these redshifts than using a more general galaxy catalog. The procedure to create this catalog is mostly based on Reid et al. (2016) with modifications to the redshift failure and systematic corrections described in Bautista et al. (2018).

4.2.3 Mass Determination

For the LRG galaxies we used the Firefly VAC. The Firefly VAC[2] (Comparat et al., 2017) provides galaxy properties of all SDSS, BOSS and eBOSS spectra with the FIREFLY fitting routine (Wilkinson et al., 2017) (v1_0_4 for DR14 and v1_1_1 for DR16) was combined with the the stellar population models of Maraston and Strömbäck (2011). The Firefly catalog includes light- and mass-weighted stellar population properties (age and metallicity), E(B-V) values and most crucially to this work, stellar mass for all galaxies in the catalog. We used the DR14 catalog to determine masses for the galaxies found in this file,sdss_eboss_firefly-dr14.fits.

The lower redshift NSA/SDSS, catalog contains many galaxies that are spatially resolved and require more careful photometric analysis (Blanton et al., 2011, e.g.). The most recent version of this catalog provides elliptical Petrosian aperture photometry which is more accu-

---

[1] The files can be found here: https://www.sdss.org/dr14/spectro/lss/

[2] https://www.sdss.org/dr16/spectro/eboss-firefly-value-added-catalog/

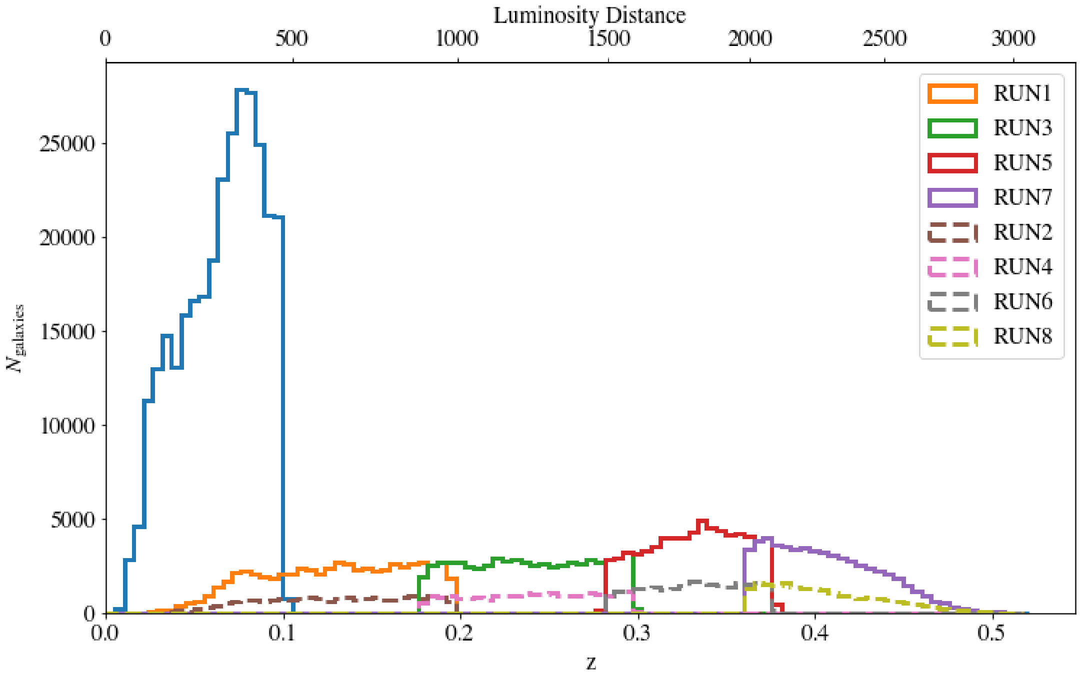


Figure 4.1: Distribution of the galaxy redshifts for the NSA/SDSS (blue, solid) and LRG-NGC (multi-colored, solid) and LRG-SGC (multi-colored, dashed) data sets that were used to reconstruct the cosmic density map. The NSA/SDSS catalog includes all galaxies out to $z = 0.1$ and is denoted as RUN1 in the the MCMP VAC. The LRG catalogs extend to higher redshifts but only include the rarer LRGs, hence the lower galaxy count. This figure also shows the slicing scheme used to self consistently fit the MCMP model in subsets of redshift as the density of galaxies decreases with luminosity distance.

rate than the standard SDSS pipeline. We adopt the Petrosian mass as our estimate of the galaxy stellar mass for this sample.

4.2.4 Bolshoi-Planck Simulations

To calibrate our MCPM density estimates to the cosmic matter density, we use the dark matter only Bolshoi-Plank ΛCDM (BP) simulation (Klypin et al., 2016; Rodríguez-Puebla et al., 2016). The BP simulation uses 2048^3 particles in a volume of $250\mathrm{h}^{-1}$ Mpc^3 and based on the 2013 Planck (Planck Collaboration et al., 2014) cosmological parameters and com-

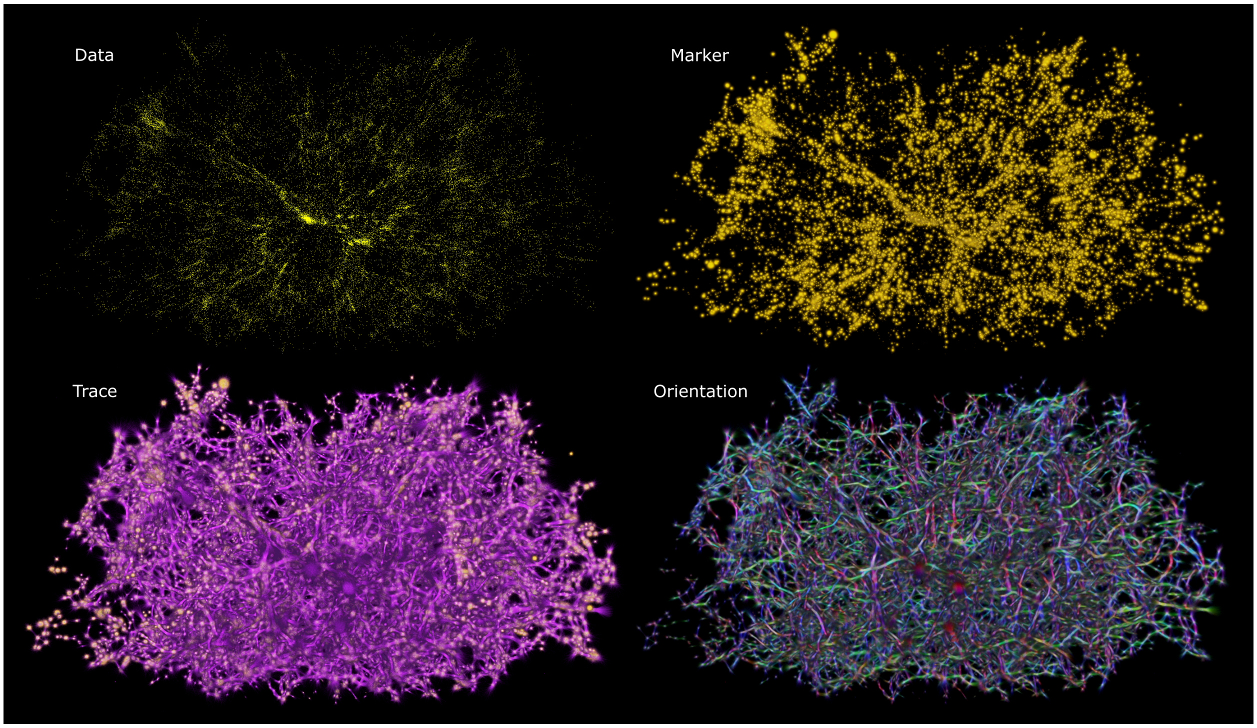


Figure 4.2: Overview of MCPM's operating modalities, demonstrated on the $0.018 < z < 0.038$ sample of SDSS galaxies. In reading order: input data points and the marker concentration emitted by the data (yellow), reconstructed trace field f_{T} (purple), corresponding orientation field f_{O} (XYZ directions mapped to RGB colors).

patible with the Planck 2015 parameters (Planck Collaboration et al., 2016). We implement a Gaussian smoothing kernel to the simulation to smooth over scales of 0.25 Mpc h^{-1} (Lee et al., 2017; Goh et al., 2019). We also employ a halo catalog produced using the Rockstar algorithm (Behroozi et al., 2012).

4.3 Methodology

4.3.1 The MCPM algorithm

The VAC has been produced with the Monte Carlo Physarum Machine (MCPM) algorithm implemented in the Polyphorm software[3]. MCPM has been first used in Burchett et al.

---

[3] https://github.com/CreativeCodingLab/Polyphorm

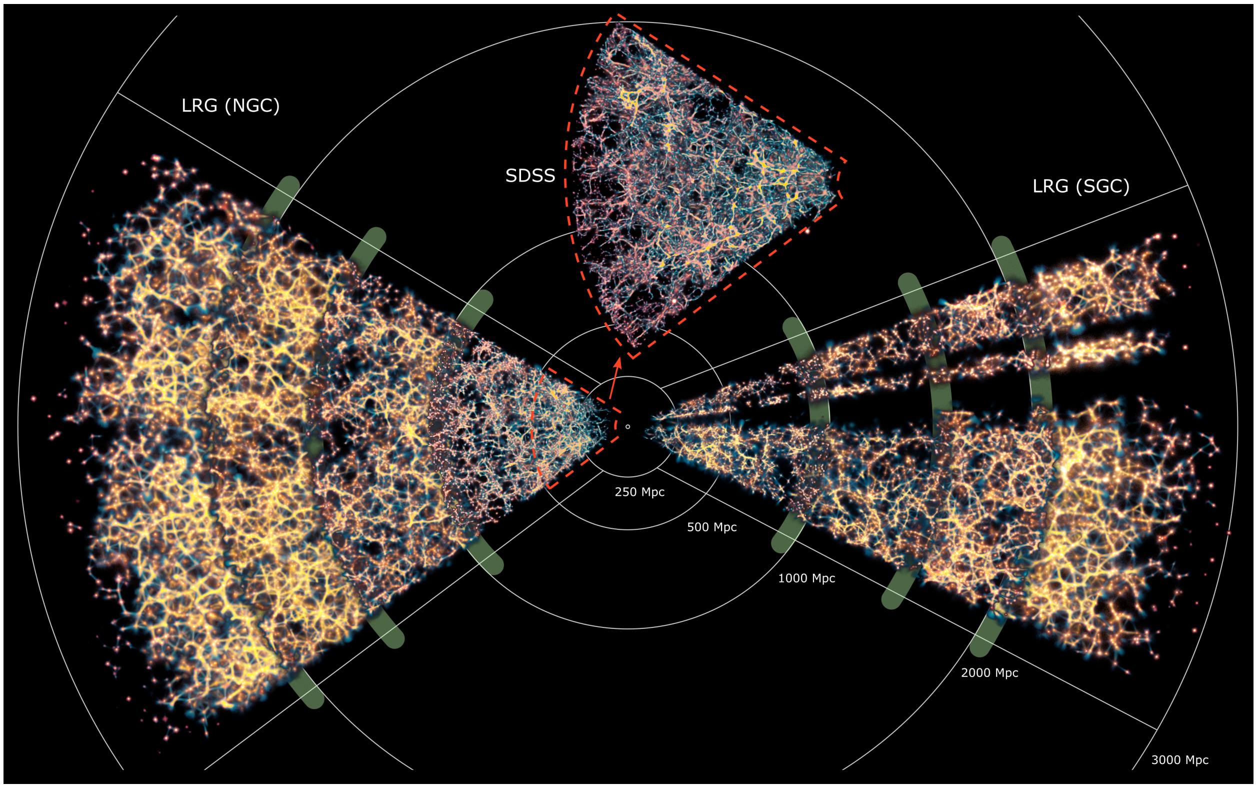


Figure 4.3: Diagram of all the data products and their spatial relations. The green bands highlight regions of overlapping LRG slices. The SDSS portion of the data is magnified to visualize the higher amount of recovered structure owing to the denser observations.

(2020a) to reconstruct a 3D density field estimate of the large scale structure spanning 37.6k SDSS galaxies within the $0.018 < z < 0.038$ range. We provide a brief summary of the model here; for the full description please refer to Elek et al. (2021a).

MCPM is a massively parallel agent-based model inspired by the growth patterns of Physarum polycephalum slime mold. Its main modalities are visualized in Fig. 4.2. Using a swarm of millions of particle-like agents, MCPM iteratively traces the network structures implicit in the input data: dark matter halos or galaxies represented as a weighted 3D point cloud. In proportion to their mass, the data points emit a virtual marker which the agents navigate towards at every iteration. The key innovation of this model is the probabilistic navigation of the agents, which significantly increases the complexity of the structures which

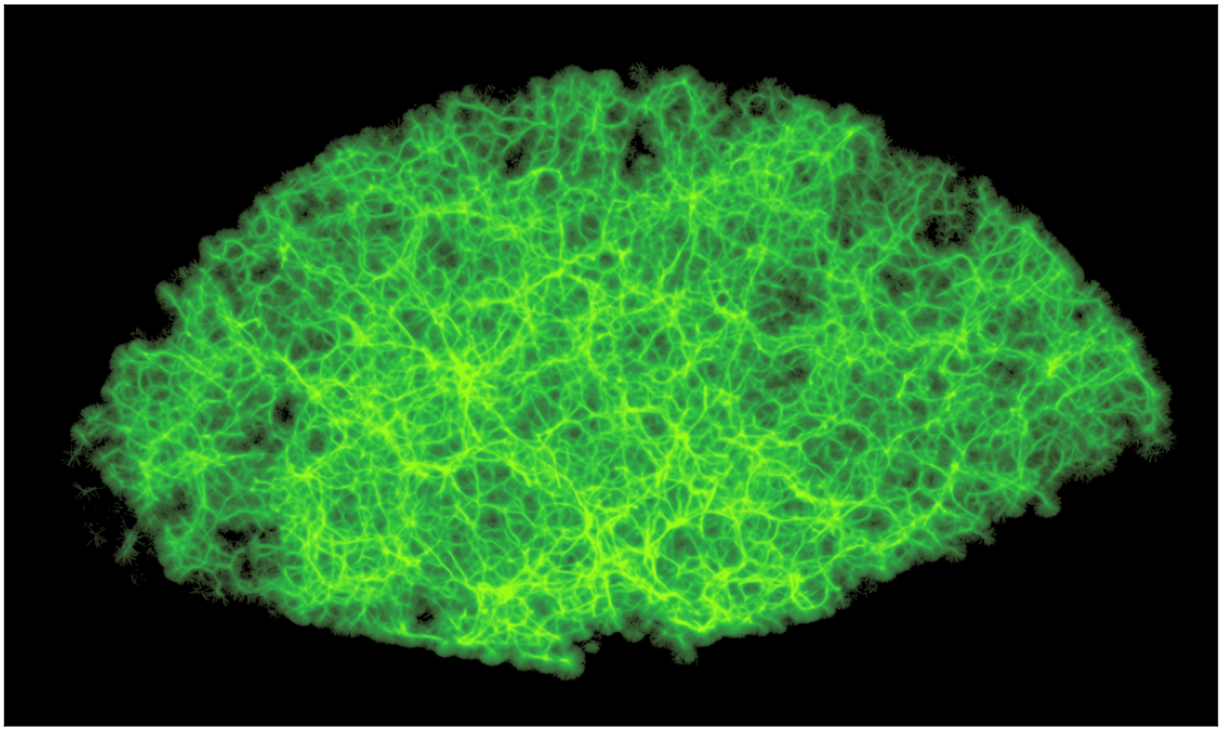


Figure 4.4: Zoom in of a slice of the $0.018 < z < 0.1$ SDSS MCPM's reconstruction of the cosmic web.

the model is able to generate.

MCPM produces two main quantities: the trace field and the orientation field. The trace field $f_\mathrm{T} : \mathbb{R}^3 \rightarrow \mathbb{R}_+$ accumulates the superimposed trajectories of all active agents, and represents the reconstructed LSS density field (after statistical standardization, see Sec. 4.3.5). The orientation field $f_\mathrm{O} : \mathbb{R}^3 \rightarrow \mathbb{R}^3_+$ records the averaged unsigned directions of the agents, and serves as a clustering criterion in our FoG compensation step (Sec. 4.3.6). Both of these are robust (i.e., stable in time) Monte-Carlo estimates of the equilibrium agent distributions.

Compared to our earlier applications of the MCPM model (Burchett et al., 2020a; Simha et al., 2020) we introduce a few methodological and implementation changes aimed at improving the quality of the fits:

1. Linear accumulation of f_T and f_O values instead of the original exponential floating window averaging. The latter is used for the supervised part of fitting, i.e., when exploring different MCPM configurations. After finding the optimal dataset-specific set of model parameters we switch to linear averaging which dramatically reduces the solution variance.

2. To avoid numerical errors we increase the numerical precision from fp16 to fp32 for both f_T and f_O. This slows the software down by 10-20%, which is acceptable for maintaining interactivity during fitting.

3. We redesigned the agent rerouting step. Rerouting is invoked when an agent encounters no data for too many subsequent steps, indicating either a boundary of the dataset or a large void. Our original rerouting assigned such an agent to a random location in space; currently, we reposition it to the location of a random data point. This change leads to a significant decrease of background noise and effectively increases the dynamic range of the obtained solutions for both f_T and f_O.

4.3.2 MCPM fit to Bolshoi-Planck

In order to calibrate MCPM's hyperparameters, we fit the model to two snapshots of the Bolshoi-Planck simulation dataset (at $z = 0$ and $z = 0.5$, both containing roughly 16M halos extracted with the Rockstar algorithm). Some of the parameter estimates carried over from our previous work (Burchett et al., 2020a), namely the sensing angle at 20 deg, moving angle at 10 deg, moving distance at 0.1 Mpc and persistence of 0.9 (now adjusted to 0.92 due to the finer granularity of halos used here). We focused on constraining the remaining critical parameters: sampling exponent (which controls the acuity of obtained structures, especially filaments) and sensing distance (which determines the scale of the structures, such as mean segment length and by transition the diameter of loops, voids etc).

Using the procedure described in Burchett et al. (2020a) we matched the MCPM fits to the ground truth densities in Bolshoi-Planck. We determined the optimal sampling exponent to be 2.5 at $z = 0$ and 2.2 at $z = 0.5$, which aligns with the observation that the LSS at higher redshifts be less condensed. For the sensing distance, the optimal value was found to be 2.37 Mpc. To obtain these values, we balanced two key criteria: fit to the halos (as defined by the fitness function, cf. Elek et al. (2021b)), and the monotonicity of the obtained overdensity mapping as shown in Figure 4.6.

The above values of sampling exponent and sensing distance pose the lower limits for the values actually used to fit the observational data. This is due to the significantly lower spatial density of data points in the galaxy catalogs (relative to Bolshoi-Planck Simulations), which need to be compensated for by proportionally increasing the two parameters.

In Figure 4.5 we demonstrate that MCPM reconstructs not just the halos that we feed into it but the cosmic structure, including filaments and voids.

4.3.3 Fit to NASA-Sloan Atlas

The first component of the VAC is based on the MCPM fit to the NASA-Sloan Atlas catalog for $0 < z < 0.1$, which contains roughly 325k galaxies in luminosity distances between 44 and

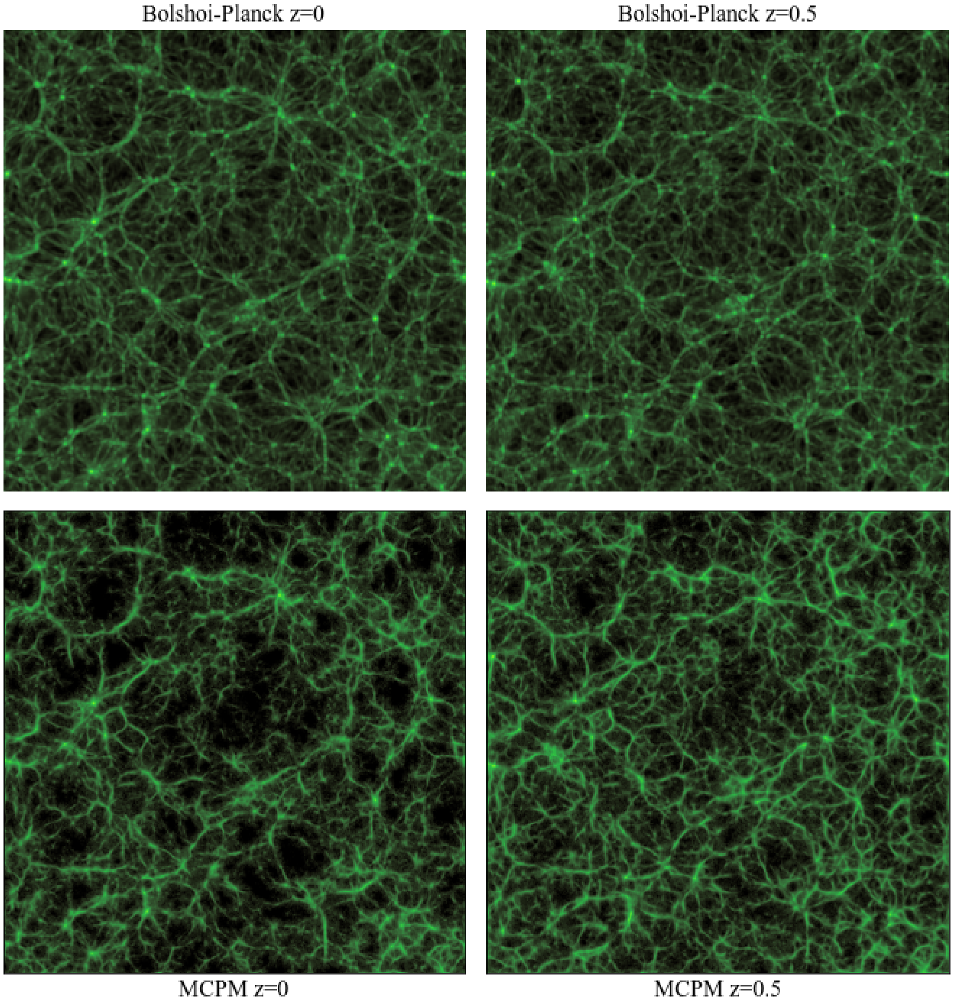


Figure 4.5: Comparison of the Bolshoi-Planck simulations (top row) at redshifts of $z = 0.0$ (left) and $z = 0.5$ (right) to the MCPM trace of the simulations (bottom row). We see MCPM faithfully reconstructs not only the galaxy halo population but also the cosmic structure.

476 Mpc. Similar to the BP dark matter halos we treat the galaxies as 3D point attractors, in this case weighted by their stellar masses.

The fits are based on the hyperparameters calibrated on the BP simulations. Further, to reflect the lower spatial density of the galaxies in comparison to the halos, we adjust the two critical parameters of MCPM: sampling exponent to 3.5 and sensing distance to 5.2. These values were obtained by maximizing the model's fitness function within intervals of the parametric space which yield a well connected reconstruction, determined in a semi-supervised way through in Polyphorm.

To verify the consistency of the fit across different z values, we have split the SDSS catalog into 3 overlapping slices (44-270 Mpc / 250-370 Mpc / 350-476 Mpc, each containing about 120k galaxies) and fitted them separately by only adjusting the sensing distance parameter. The resulting optimal values (Figure 4.7) follow a linear trend, implying that the spatial density of galaxies decreases in corresponding proportion. However, the obtained variation of sensing distance (3.8–5.6) is well within the ability of the model to perform a consistent fit using a single value of the parameter. Therefore, we opt for a single fit to the entire catalog using the aforementioned sensing distance value of 5.2.

4.3.4 Fit to LRG Catalogs

The procedure of fitting to the LRG NGC and SGC catalogs is identical with the SDSS data: using the sampling exponent of 3.5 and the BP-calibrated values for the other hyperparameters, we continued increasing sensing distance until reaching an optimal fit.

Due to the much lower spatial density of LRG observations compared to SDSS, the optimal values of sensing distance end up being considerably higher (Figure 4.3). Also, unlike SDSS, the LRG galaxies span a significantly longer range of redshifts. The consequence is nearly a two-fold growth of the optimal sensing distance value across the catalog's redshift range (Figure 4.7). Therefore to construct the VAC, we split the LRG galaxies into 4 overlapping 'slices' of approximately equal numbers of galaxies (about 70k per slice for NGC, about 25k per slice for SGC) and fit each separately. The resulting distance intervals are 0-1000 Mpc ($z \approx 0 - 0.2$), 900-1600 Mpc ($z \approx 0.18 - 0.3$), 1500-2100 Mpc ($z \approx 0.28 - 0.38$), and 2000-3000 Mpc ($z \approx 0.36 - 0.51$).

The visualization of all thus obtained density slices and their spatial relations is provided in Figure 4.3. An added benefit if this approach is the higher resolution of each slice we can afford. This is desirable again due to the massive redshift range of the LRG data.

4.3.5 Statistical Standardization & Mapping

The MCPM densities fit to each survey slice do not reflect any physical density, rather they are the density of agents in the fit. In order to translate the MCPM density to cosmic overdensity, we standardize each distribution to the MCPM fit of the simulation so that a mapping between MCPM and cosmic overdensity can be applied. The MCPM fits to the galaxy surveys differ from the fits to the BP simulations in that they suffer from luminosity selection functions and are thus much sparser. This affects the lowest density regime of the density distribution. In order to account for this effect, we used the Wasserstein distance[4] or the "Earth Movers Distance" to calculate the stretch and shift values such that the distribution of MCPM densities of the surveys could be linearly transformed to best fit the BP-MCPM fit. That is TargetDist = stretch×SurveyDist + shift. Where the TargetDist is the BP-MCPM density distribution and SurveyDist is the density distribution of each survey slice. The benefit of this method is that we can impose a lower limit on the density distributions to only take into account the higher density wing of the distribution corresponding to densities that contain structure and avoid the empty space in the survey fits.

In order to retrieve the cosmic matter density, $\rho_m/\langle\rho_m\rangle$, we must map the MCPM trace density to that of the BP simulations at each redshift. That is, we fit the BP simulations using the MCPM algorithm and then apply a mapping from MCPM density to cosmic matter density. This mapping was achieved by sampling the MCPM fits in bins of equal density and then determining the density from the BP simulations at the same location. This is shown by the multi-colored stripes in Figure 4.8. We then determine the median (and 1σ limits) of each MCPM density bin. The median densities in each bin were then used to create a mapping

---

[4] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.wasserstein$_d$*istance.html*

function. We chose to base our mapping function on the rectified linear activation function (ReLU) where the inflection point was determined by the maximum change of the median of the bins. On the right-hand side of flat part of the function, we fit a cubic polynomial to the data, creating a piece-wise continuous mapping function. This method was chosen over other methods such as simply interpolating the bins, or using a spline function because the behavior for densities below or above the density in the MCPM fits is well defined. Our method is illustrated in Figure 4.8 where the thicker black line shows the mapping function applied to the $z = 0$ simulation, the thinner black lines show the 1σ limits of our mapping, which correspond to ±0.5 dex in log cosmic over density, $\rho_m/\langle\rho_m\rangle$.

4.3.6 Correction for Redshift Space Distortions

As MCPM operates in 3D space, applying the algorithm necessitates attaching physical distances to the input dataset. Although distance measurements via more direct methods (e.g., tip of red giant branch or Type Ia supernovae) (Tully et al., 2016) may be available for a small subset of the galaxies (and therefore tracers of the underlying density field), we must primarily assume distances concordant with the Hubble flow. Thus, we initially attach to each galaxy the luminosity distance given the adopted cosmology and galaxy redshift. Clearly, denser environments such as galaxy groups and clusters will include galaxies with large peculiar velocities. These peculiar velocities will result in redshift space distortions (RSDs), or 'fingers of god' (FoG), if adopted directly. For example, a typical velocity dispersion for a $> 10^{14} M_\odot$ galaxy cluster (~ 1000 km/s) would propagate to a systematic error in the distance by assuming pure Hubble flow of > 10 Mpc. This issue plagues our low-redshift SDSS sample significantly more than the LRG samples for two key reasons: 1) Low-mass galaxies are much more abundant and likely to be observed at low z in the magnitude-limited SDSS, which results in many objects composing false apparent structures along the direction pointing away from (and towards) the observer. 2) High-mass galaxies, which will dominate the samples at progressively higher redshifts, preferentially reside as central galaxies in their local environments (Lan et al., 2016). Therefore, these galaxy samples will be less subject to the systematic error in cosmological distance than our lowest redshift sample. Thus, we employ an RSD correction for the $z < 0.1$ SDSS galaxy sample that we detail here.

A key feature of MCPM is that the cosmic web reconstruction converges to an equilibrium state but is a dynamical system nonetheless, and the adopted ‘densities’ are simply aggregated trajectories of the millions of agents seeking out efficient pathways between galaxy tracers. MCPM also outputs the components of an aggregated three dimensional agent velocity vector for each cell in the volume. We use these velocities to identify RSDs, as the agent velocities producing them will be preferentially oriented perpendicular to the plane of the sky along the line of sight and will be clustered in their celestial coordinates. We select points in the MCPM cube by orientation as follows: We convert each input galaxy’s location in the MCPM-output cube to its equivalent celestial coordinates and find the three components of a unit radial vector parallel to the line of sight in Cartesian space to match the coordinate system of the MCPM velocity vectors. We then calculate the dot product between the aggregated velocity vector at each galaxy’s position in the cube with the unit radial vector and assign the result to that galaxy. Galaxies within an RSD structure (FoG), having either parallel or antiparallel velocity vectors to the unit radial vector, should not have dot product absolute values close to zero. Therefore, we filter out galaxies with dot product absolute values less than 10, chosen upon inspecting the distribution of galaxy dot product values as a conservative cut. To identify galaxy positions with similar velocity orientation and projected location on the sky, we then employ the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm as implemented in the scikit-learn[5] python package, feeding it the sky coordinates and redshift. For this step, we further filter the galaxy catalog by mass to those with $M_* > 10^{10}\ M_\odot$, as the completeness of SDSS declines for less massive galaxies at the upper end of our redshift range ($z \sim 0.1$). DBSCAN operates by locating high-density cores in the data, which are the beginnings of the clusters, and the algorithm searches out from these cores, adding points until no more points are found in

---

[5] https://scikit-learn.org/stable/

within some distance tolerance (in whatever space the data occupy). This algorithm contains a number of advantages over other possible choices, including scalability, compatibility with non-flat geometries, and the feature that certain points may not be included in any cluster (they are deemed 'noise'). Two key parameters for DBSCAN are the distance tolerance (eps) and the mininum number points to be considered a core in the data (min_samples). We chose min_samples=3 as a minimum number of galaxies (e.g., such as in a group or cluster) that might form a false RSD structure (FoG) in the MCPM model. We chose a value of eps=2 upon experimenting with several values through visual inspection, to balance the inclusion of FoGs (which are readily identified by eye) containing a relatively small number of galaxies with minimizing false identification of filaments not oriented antiparallel to the plane of the sky as RSD structures. Figure 4.9 shows the resulting clusters identified by DBSCAN in a slice in declination of our galaxy catalog, with galaxies belonging to the same having the same color.

Having the output clusters identified by DBSCAN, we then find the velocity range spanned by galaxy redshifts within each cluster. This is adopted as the full-width-half-maximum (FWHM) of the velocity distribution (v_{FWHM}). For clusters with $v_{\mathrm{FWHM}} > 300$ km/s, we adopt new redshifts for the associated galaxies to be commensurate with more realistic physical distance separations inferred from a simple luminosity distance based on the redshift; this procedure is as follows. Assuming the cluster members are bound to the same virialized structure, we convert the velocity FWHM to a velocity dispersion by the relation:

$$\sigma_v = \frac{v_{\mathrm{FWHM}}}{2\sqrt{\ln 2}} \tag{4.1}$$

This velocity dispersion is then used to infer a virial radius, R_{200}, of the cluster:

$$R_{200}^{\mathrm{infer}} = \frac{\sigma_v}{4/3\pi G\Delta_{200}\rho_{\mathrm{crit}}} \tag{4.2}$$

Where Δ_{200} and ρ_{crit} are the overdensity and critical density, respectively. We then adopt new redshifts (solely for the purpose of feeding MCPM) about the median redshift of the cluster members by sampling from a normal distribution with standard deviation corresponding to the change in redshift that would result in luminosity distance difference equal to the inferred R_{200}. Finally, we convert these galaxy coordinates and adopted redshifts to 3D Cartesian space via luminosity distances based on the new redshifts; these then serve as inputs to MCPM.

4.4 Data Products

4.4.1 The Catalog

The final value added catalog contains the positions and redshifts as well as stellar mass of the galaxies in the NASA-Sloan Atlas and the eBOSS Firefly Value-Added Catalog. We include a column, MASS_SOURCE, to indicate which catalog was used to estimate the mass. The MCPM algorithm uses the galaxy mass to build the matter density field. The primary field of interest here is MATTERDENS, the matter density field at the location of a given galaxy. This is derived from fits of MCPM models in 3D volumes, which are then mapped to the cosmological matter density (relative to the mean matter density) using MCPM fits to the Bolshoi-Planck simulations. The CATALOGID is a combination of PLATE-MJD-FIBERID. A unique identifier is the combination of CATALOGID and MCPM_RUN. Objects with the same value of MCPM_RUN were fitted with the MCPM model simultaneously. The data were sliced in redshift to yield samples producing self-consistent large-scale structures over the volume in each slice. MCPM_RUN = 0 correspond to $0.01 < z < 0.1$ SDSS galaxies with masses from the NASA/Sloan Atlas. Samples of LOWZ LRGs are marked 1-2 ($z < 0.20$), 3-4 ($0.18 < z < 0.30$), 5-6 ($0.28 < z < 0.38$), and 7-8 ($0.36 < z < 0.51$); each pair (e.g., 3-4), corresponds to the NGC/SGC samples in some redshift slice, with odd and even numbers for NGC and SGC, respectively. The data model for the catalog is described in Table (4.1).

Table 4.1. Data Model

Name	Type	Unit	Description
catalogID	char[13]		Combination of PLATE-MJD-FIBERID
plate	int32		Plate number
mjd	int32		MJD of observation
fiberid	int32		Fiber identification number
ra	float64	deg	Right ascension of fiber, J2000
dec	float64	deg	Declination of fiber, J2000
z	float32		Best redshift
massSource	char[7]		Source of the mass determination (nsa or firefly)
mcpmRun	int8		Index of galaxy sample fitted simultaneously with MCPM
mstars	float64	$M_{\odot}$	Stellar mass
matterDens	float32		log10 of the ratio of the matter density/mean matter density

Note. — Schema for the MCPM Value-Added Catalog, v1.0.0 as found in slimeMold_galaxy_catalog_v1_0_0.fits.

4.4.2 3D Density Cube

In addition to the VAC, which contains the density at the location of each galaxy, we offer the full 3D density field of the relevant volumes, available at http://astronomy.nmsu.edu/jnb/slimeVAC/. These may be queried using our custom package, pyslime[6]. The data will unzip to a directory which may be opened by pyslime. This will enable the user to query the overdensity at arbitrary points in the cube, allowing the study of voids and filamentary structure outside the local environment of the galaxy field.

4.5 Discussion

We focused on producing density fields, instead of classifying/indexing filaments, which will be produced in a forthcoming data product.

4.5.1 Comparison to Peng et al. (2010)

We can additionally validate our model by comparing our findings to that of other surveys, Although Burchett et al. (2020a) has demonstrated the efficacy of our model, we present comparisons to other studies, leveraging our deeper and larger surveys.

(Peng et al., 2010) used a method based on the 5th nearest galaxy neighbors to estimate the cosmic density and compared this galaxy stellar mass and SFR and the quenched fraction of galaxies. Burchett et al. (2020a) illustrate that the MCPM method of computing cosmic density qualitatively matched the results (Peng et al. (2010), see Figure 5 & 6). In Figure 4.11, we demonstrate the improvement in signal gained with the NSA/SDSS sample as the increase in the number of galaxies is significant and we again qualitatively reproduce their density-stellar mass-sSFR relations.

---

[6] https://github.com/jnburchett/pyslime

4.5.2 Potential Applications

Our primary aim in this manuscript is simply to showcase the dataset and describe its construction. There are, however, many exciting applications for this dataset that are well beyond the scope of this publication. Here, we list four general areas of application:

- Galaxy evolution in the cosmic web: Figure 4.11 highlights one direct application of the galaxy-density catalog to study the possible impacts of a galaxy's location within the cosmic web on its evolution. However, a vast amount of galaxy properties measured and inferred from both multiwavelength photometry and spectroscopy have been cataloged for SDSS galaxies () (many also released as VACs; e.g., Salim et al., 2016); via straightforward crossmatching with our catalog, myriad galaxy-environment analyses may be readily conducted.

- Void finding: In the linear regime, the sizes of voids and their correlation statistics are sensitive to cosmology, particularly dark energy (Pisani et al., 2015). Although the most of the analyses we have alluded to thus so far focus on the denser regions of the cosmic web, namely filaments and nodes, our density cubes naturally also include the underdense regions. Simple centroiding and clustering algorithms may be readily applied to these density fields to directly identify and characterize the voids, which in turn may be used as inputs for cosmological parameter estimating using, e.g., the Alcock-Paczynski effect (Alcock and Paczynski, 1979).

- The intergalactic medium: Cosmological simulations and analytical models alike predict a rich multiphase structure to the thermodynamic properties of the intergalactic gas permeating the cosmic web (Cen and Ostriker, 1999; Davé and Tripp, 2001; Tepper-García et al., 2012). In addition to the physical states of gas resulting from large-scale structure formation (Bertschinger, 1985; Molnar et al., 2009), energetic feedback from the galaxies themselves might extend well beyond the virial radius, which is often adopted as a fiducial extent of a galaxy's halo (Finlator and Davé, 2008; Schaye et al., 2015; Nelson et al., 2019). Burchett et al. (2020a) used HST-observed background quasar sightlines through the MCPM reconstructed volume to find a relationship between cosmic web density and Lyα optical density. Similar analysis could and should be done leveraging other absorption tracers, such as Mg II.

- Multimessenger transient followup: Transient phenomena such as gravitational waves and fast radio bursts are typically detected with imprecise localization, with scales of minutes or degrees on the sky (Chen and Holz, 2016; CHIME/FRB Collaboration et al., 2019). Space-based and ground-based facilities around the world then follow up these detections to identify and characterize the sources (e.g., Coulter et al., 2017). As extragalactic sources are statistically more likely to be found within the large-scale structure, transient observers could employ our reconstructed density field of the cosmic web in followup imaging campaigns to prioritize pointings toward regions of the sky most likely to contain the source counterparts.

4.5.3 Known Limitations

The VAC volumes have the usual luminosity function systematics that are present in the underlying SDSS and LRG catalogs. Specifically, the density of galaxies is greater at lower redshifts. This is reflected in the trace as well and can be seen in Figure 4.3 in the SDSS data as well as each slice of the LRG catalogs. This presents itself as an increased density at the lower redshift end of the volume. However, the mean matter density at the low and high redshift ends of each volume are consistent.

Some sub-optimality of the model fit arising from the fact that the optimal sensing distance should grow linearly, at least according to the data in Figure 4.7. Whereas the catalog is a piece-wise constant approximation of this.

Do to the differing sensing distance in each slice, there is a small discontinuity of the MCPM densities extracted from the overlaps between the LRG slices. Therefor, any comparison of densities should be done on a slice-by-slice basis and comparisons any quantities based on the density done on differing redshift slices should be avoided.

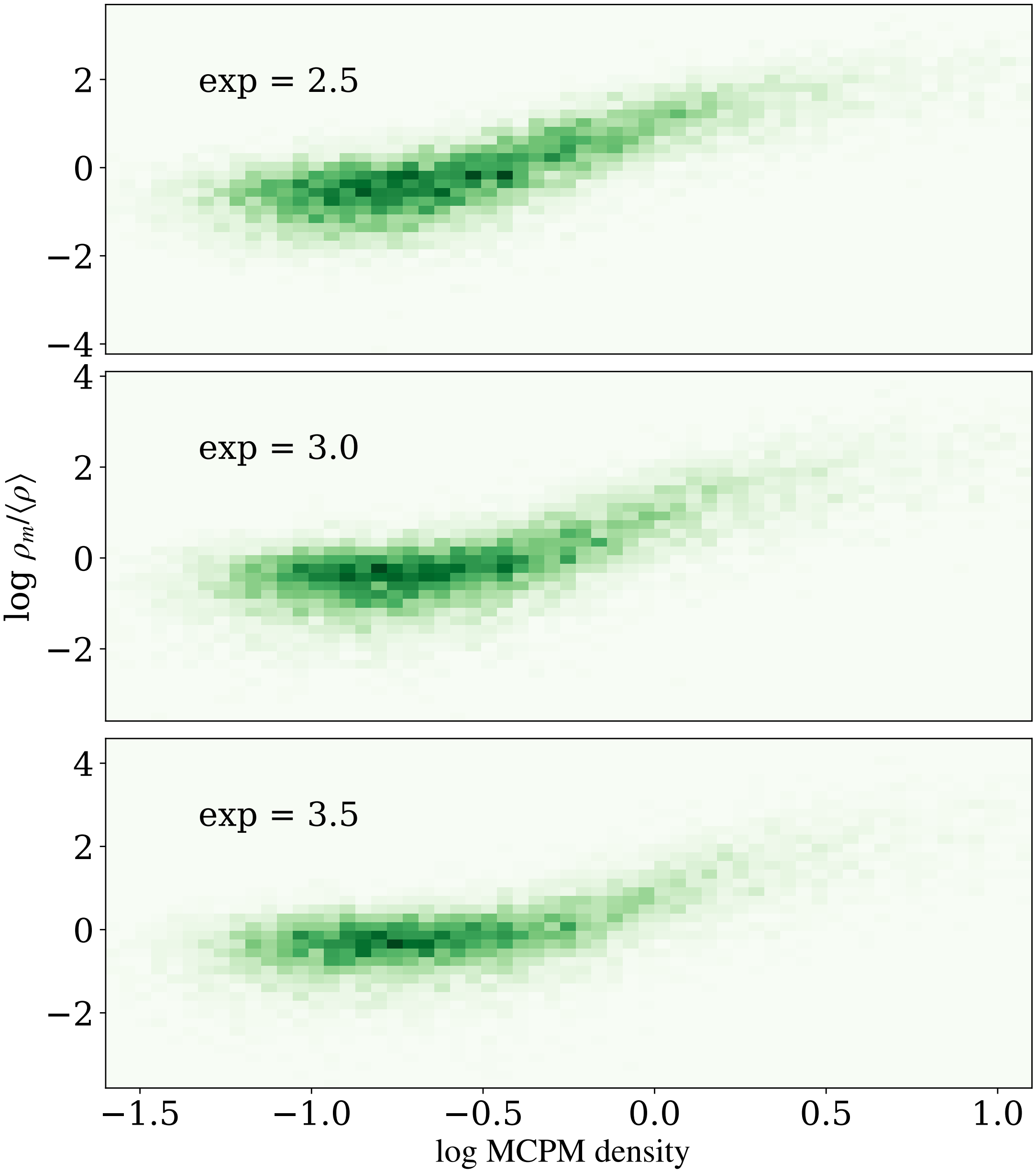


Figure 4.6: Comparison of different sampling exponents in an increasing order from top to bottom. We find that a sampling exponent of 2.5 produces the most linear mapping between the MCPM densities and the cosmic matter densities from the BP simulations, especially at lower densities where previous versions of the MCPM have generally failed to recover the lowest density structures. (see Figure 10 in Burchett et al. (2020a)).

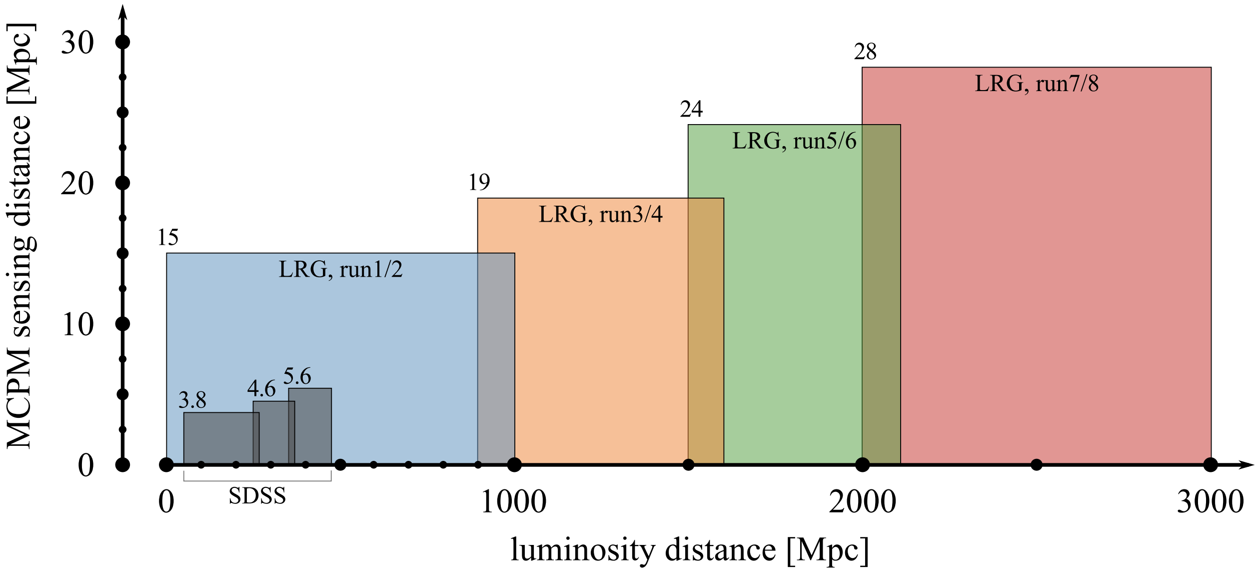


Figure 4.7: Plot of MCPM agents' sensing distance (the main feature scaling parameter) as read out from the best fits for the LRG data, radially sliced into 4 runs at overlapping luminous distance intervals. For comparison we also show the best-fit sensing distances for 3 SDSS slices, which manifest a similar linear growth as we observe in the LRG data.

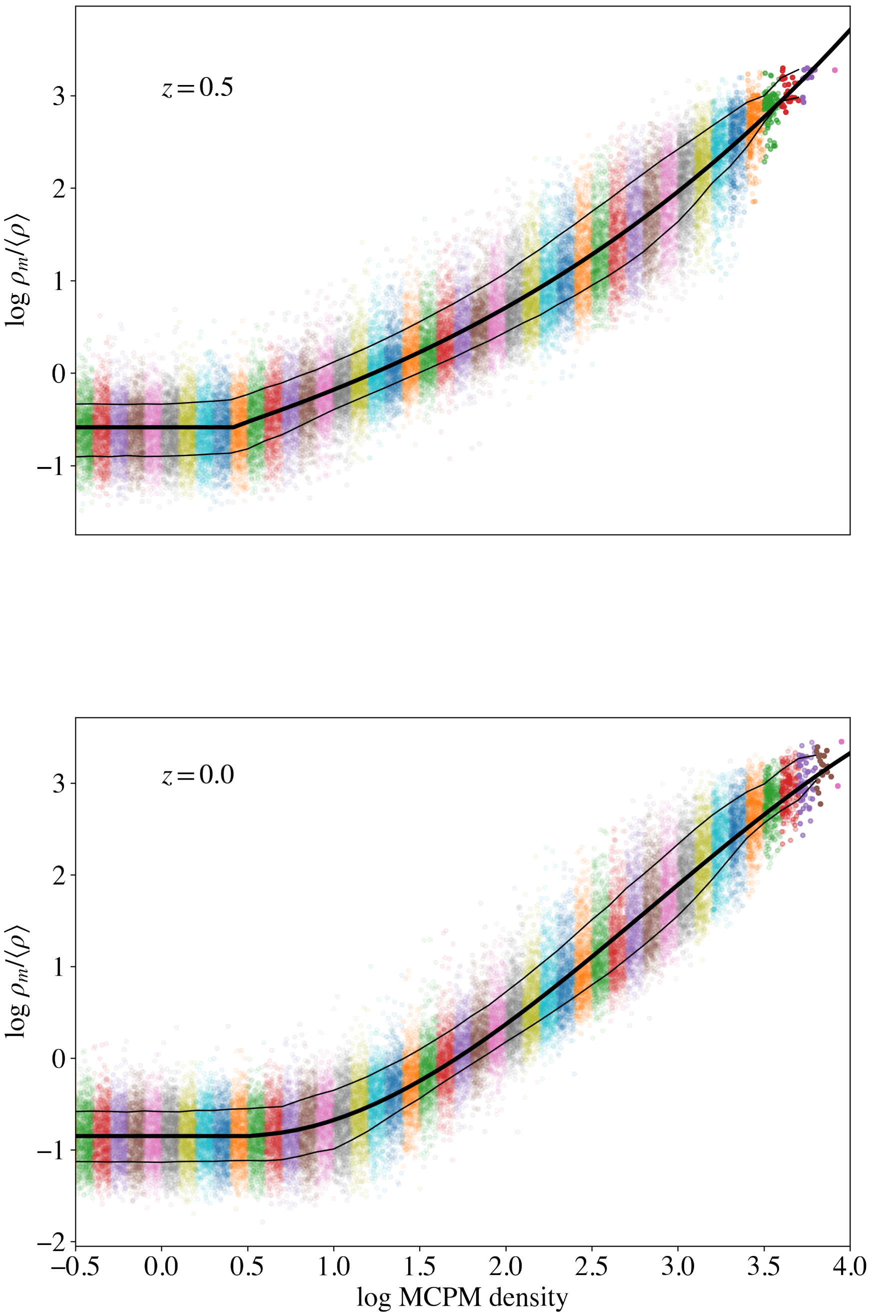


Figure 4.8: Mapping of the MCPM derived density to the cosmic matter density from the BP simulation. The MCPM densities were binned evenly in MCPM space in bins of 0.1 dex as demarcated by the colored stripes. The custom ReLu mapping function fit to the bins (thick black line) and 1σ limits (thinner black lines) are plotted on top of the data. This mapping function provides a translation from the MCPM density to the cosmic overdensity.

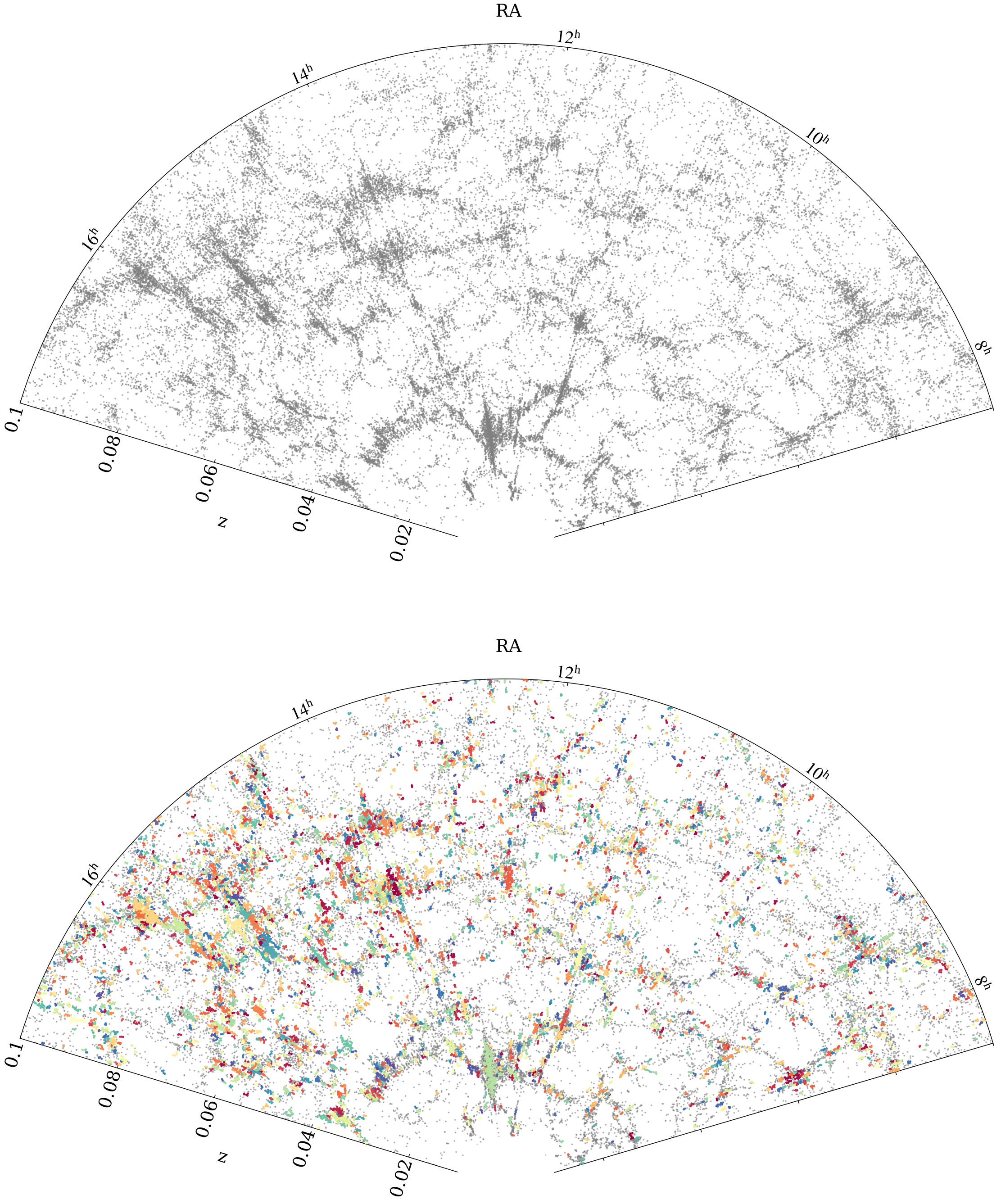


Figure 4.9: A slice in declination of our input galaxy catalog (grey points, top). RSD structures identified by DBSCAN shown in various colors overlayed on the original points (bottom).

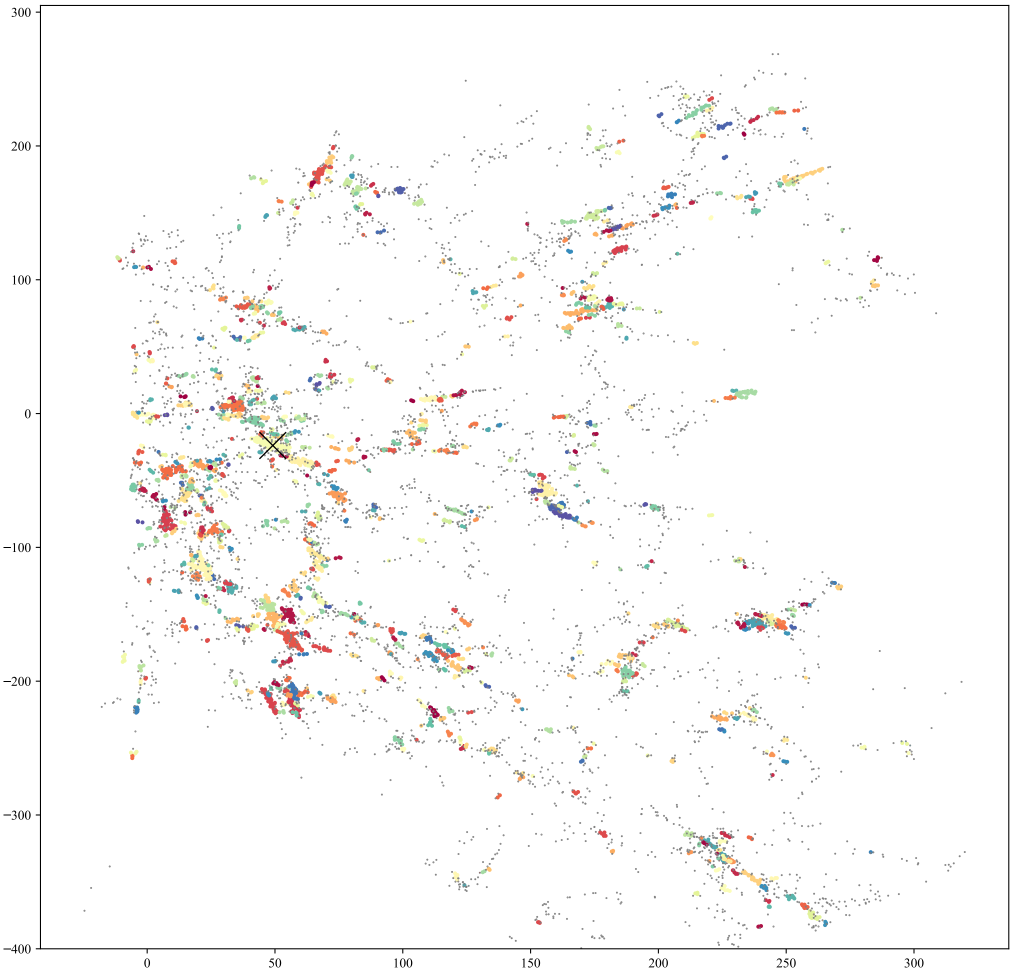


Figure 4.10: A slice in right ascension of our input galaxy catalog (grey points) with RSD structures identified by DBSCAN shown in various colors.

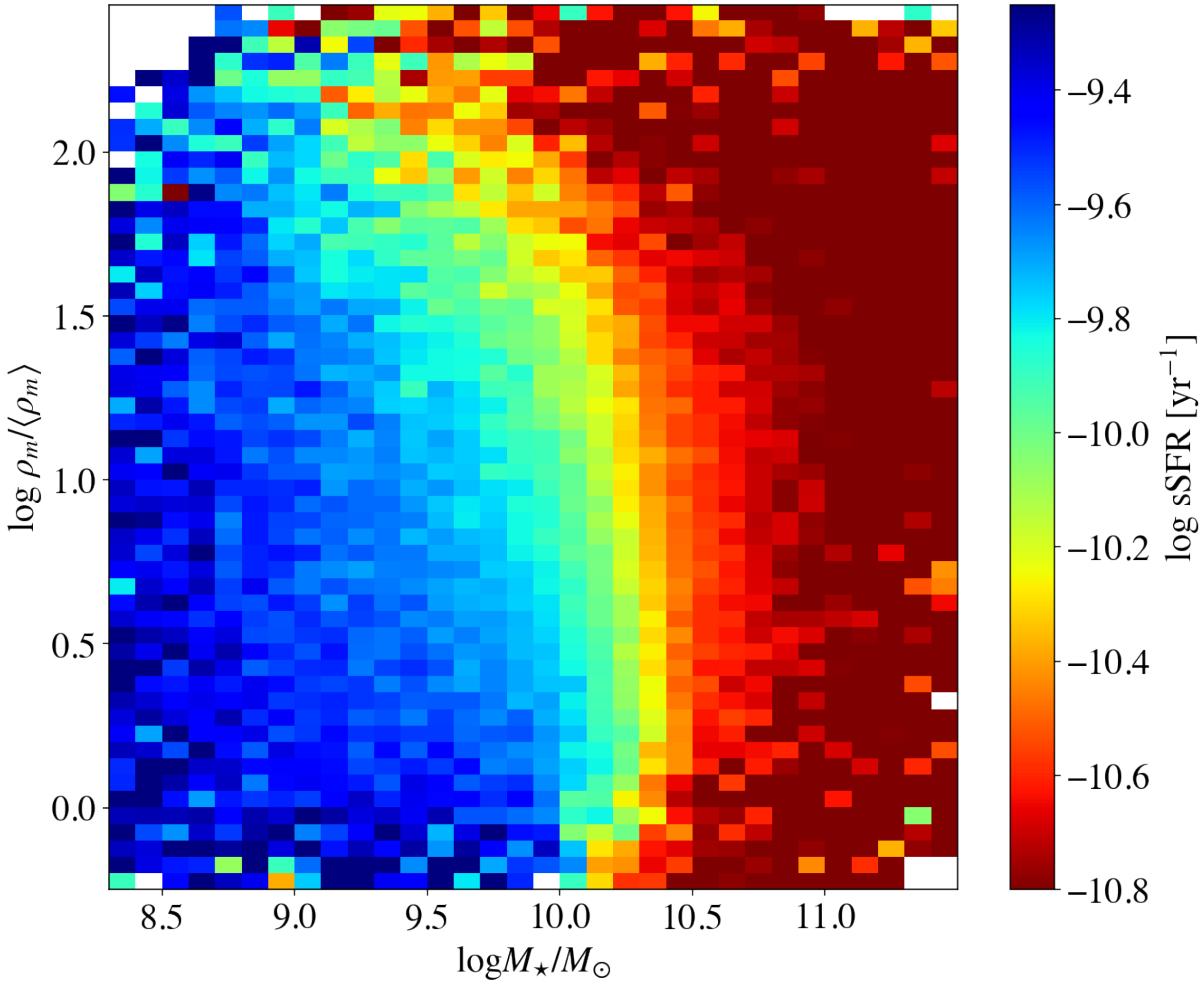


Figure 4.11: The dependence of star formation activity on galaxy environment and stellar mass for the galaxies within the NSA/SDSS volume ($z < .0.1$). The color coding denotes the fraction of "red" galaxies in the population within each mass/environment bin, where the environmental density (slime mold trace density) is determined from our MCPM cosmic web reconstruction algorithm. A comparison with Figure 6 of Peng et al. (2010) shows a similarly increasing red fraction as functions of both mass at fixed density and density at fixed mass.

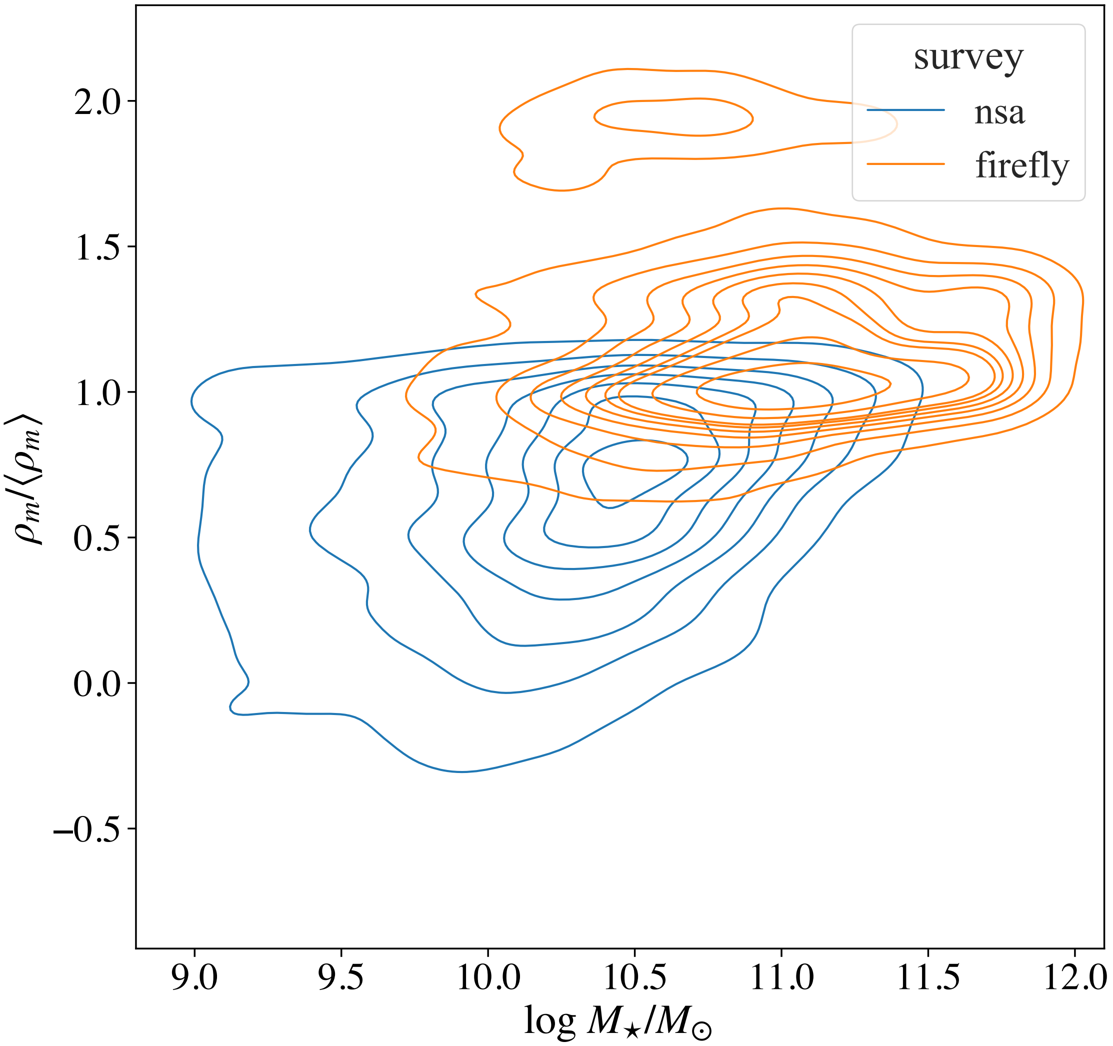


Figure 4.12: Comparison of the galaxy stellar mass vs cosmic matter overdensity in each survey. The contours are created via a kernel density estimate of the density of galaxies in the galaxy mass-overdensity space. We break it into the two subsamples: the lower redshift and lower mass NSA/SDSS catalog (blue) and the higher mass, higher redshift LRG catalogs with masses derived from the firefly VAC. This highlights the differnt structures of the cosmic web traced by the two types of galaxy catalogs. The LRG catalogs, which only include higher mass LRGs, trace higher density structure while the NSA/SDSS catalog traces lower density structures.

Chapter 5

DISCUSSION & CONCLUSION

5.1 Understanding the Circumgalactic Medium is Critical for Understanding Galaxy Evolution

There remain many unsolved challenges in coming to a holistic picture of galaxy evolution. How galaxies quench and stay quenched remains a critical blind spot in our understanding of galaxy evolution. Similarly, the mechanism that allows lower mass galaxies to continue to form stars despite the lack of a reservoir in the ISM to sustain this formation are still unclear. There has been recent progress on accounting for all of the galaxies baryons, but the cycle of the baryons is still not completely understood. Accretion from the IGM, feedback from both stars and AGN, outflows all play a role and all take place in the CGM. In Chapter 1, we describe how CGM is thought to be responsible, or at least serve as an important observable in deciphering the multitude of models.

5.1.1 The CGM^2 Survey

An understanding of the size of the CGM is critical to answering some of these outstanding problems. The size of the CGM sets the scale for accounting for the baryonic mass in the universe. Environmental quenching is thought to be due to the disruption of the CGM by external gaseous halos, either via ram pressure stripping or strangulation. Both of these phenomena depend on the size of the atmospheres of the constituent galaxies.

In addition, the radial profile of the gas around galaxies can shed light on the inflow and feedback mechanisms of the galaxies. It is therefor imperative to have statistically significant observations of this gas in the CGM.

In Chapter 2 (Wilde et al., 2021), we introduced a new survey, CGM^2 that promises to be a valuable resource in answering the pressing questions related to the CGM. The survey consists of 22 $z \sim 1$ QSO sightlines observed with HST-COS. Every foreground absorption line has been identified in the spectra. These are combined with 976 galaxy spectra obtained with Gemini GMOS North and South. The galaxy spectra allow us to provide accurate redshift measurements of each galaxy. We match the galaxies to the absorbers that lie within ±500 km s^{-1} of the systemic redshift of each galaxy. In cases where no absorption was detected we place upper limits on the column density of the absorption signal. In order to obtain mass and SFH estimates, we used CIGALE to fit the SEDs of each galaxy which requires photometric data. We collected photometry for every galaxy from SDSS, the DESI Legacy Surveys and Pan-STARRS. One major challenge with our spectroscopic galaxies was that they were much deeper than many of the imaging surveys. We supplemented the photometry with imaging from the Gemini GMOS instruments. Thus we have a survey of 976 absorber-galaxy pairs with estimates for stellar mass, SFR and column density for a multitude of ions including CIII, CIV, OVI in addition to HI. Tchernyshyov et al. (2022) has investigated the profile of OVI and the mass dependence around the galaxies in CGM^2 . In future work, we wish to investigate the kinematics of the metals.

Focusing on the HI, we cut the survey to galaxies with $z < 0.48$, where the Lyα absorption lines redshift out of the range of the COS detectors. We are left with a sample of 572 HI-galaxy absorber pairs. We find a strong mass dependence in the column density and covering fraction for gas within 1.5 R_{vir} or 300kpc and with column densities $N_{\mathrm{HI}} \gtrsim 10^{14}$ cm^{-2}. This mass dependence has also been observed in Bordoloi et al. (2018).

By examining the empirical covering fractions as a function of impact parameter and mass, we show the covering fraction declines as a function of impact parameter. This trend is weakest for our lowest mass galaxy sample $\lesssim 10^9 M_\odot$. We also find that the covering fraction of the lowest mass sample never exceeds 80% while for both higher mass samples, the covering fraction is consistent with $\sim 100\%$. Using a single power law clustering model, we examined the radial profile in three mass bins. $< 10^9 M_\odot$, $10^9 < M_* < 10^{10}$ and $> 10^{10} M_\odot$. Using a probabilistic argument, we define a metric, R^{14}_{CGM}, which demands that around a given galaxy, one is more likely than not to find material that has been empirically associated with galaxies. The picture that emerges is of a diffuse, CGM extending to ~ 350 kpc around galaxies with stellar mass $\gtrsim 10^9 M_\odot$ at $z \lesssim 0.5$. For galaxies of lower stellar mass, the extent of the CGM is < 200 kpc. At all stellar masses, the extent of this CGM exceeds a virial radius, especially for galaxies with intermediate masses where the R^{14}_{CGM} exceeds $2R_{\mathrm{vir}}$.

Therefore, these results imply that using $R_{\rm vir}$ as a proxy for the characteristic edge of the CGM may significantly underestimate its true extent.

Broadly, we find that high-column density circumgalactic material is associated with galaxies at high statistical significance out to $2R_{\rm vir}$, whereas H1 absorption with $N_{\rm HI} < 10^{14}$ cm^{-2} is more broadly distributed in both impact parameter and velocity space and may not be associated directly with massive galaxies. Our kinematic analysis reveals that the detected strong H1 is most likely gravitationally associated with the most nearby galaxy, while weaker H1 components seen at absolute velocity offsets $\gtrsim 500$ km s^{-1} may be instead associated with extended large scale structures. We find generally good agreement between our sample and the prior studies that have examined the CGM of low-redshift galaxies out to similar and larger separations (Chen et al., 2005; Wakker and Savage, 2009; Prochaska et al., 2011; Werk et al., 2014; Tejos et al., 2014; Keeney et al., 2018, e.g.).

5.1.2 Characterizing the Galactic Atmospheric Boundary

In order to improve upon the results presented in Chapter 2, we combined the HI survey in CGM^2 with that of CASBaH. This sample gives us an order of magnitude galaxies with greater sampling at larger impact parameter which allows us to examine the region around galaxies where the CGM transitions to the IGM. We build upon the analysis of HI-galaxy correlation in Paper I by including more sophisticated models.

The first model uses the standard two point correlation power law, where we parameterize the clustering scale with a power law in mass. This allows us to model the full mass range in addition to the radial profile of the HI. We find clustering parameters that are consistent with Tejos et al. (2014) with the addition of a non-zero, positive clustering scale slope. This implies the HI-galaxy clustering increases for galaxies in the range $8 < logM_* < 10.5$. We use this clustering to compare the absorber galaxy bias as a function of stellar mass to that of the galaxy-halo bias found in Tinker et al. (2010). We find that HI is a less biased tracer at lower mass galaxies and is more biased than galaxies at higher masses. This could be explained by the fact that our high mass galaxies include a higher fraction of quiescent galaxies, which are known to cluster more strongly (Berti et al., 2021). However, our high mass sample is not dominated by quiescent galaxies and this relation could imply the low mass galaxies exist in lower density environments such as near voids or tendrils in the cosmic web than the higher mass galaxies which exist in higher density environments (Prochaska et al., 2013).

In order to examine the effects of the CGM and leverage the high number of galaxies at small impact parameters in the CGM^2 survey, we include a two-component model in the analysis. This model consists of a Gaussian 1-halo component and a single power law model 2-halo component. These two components are exclusionary in the sense that the 2-halo model doesn't contribute to the clustering below $R_{\rm cross}$, the impact parameter that the two models cross. In addition, we parameterize $R_{\rm cross}$ and the power law clustering scale as power laws in mass. This allows us to model both the inner and outer portions of the model with mass. We recover the same mass dependence for the outer 2-halo model as the single power law model. However, the inner mass dependence is $\sim 2x$ stronger. The HI at low impact parameters $R_\perp < R_{\rm cross}$ is more strongly mass dependent. This model is also consistent with the data at all masses $> 10^8 M_\odot$. Another benefit of this model is that we can use $R_{\rm cross}$ as a estimate for the extent of the CGM. $R_{\rm cross}$ represents the furthest impact parameter that the data require an elevated clustering signal.

We find that $R_{\rm cross}$ $2 \pm 0.6 R_{\rm vir}$ for galaxies with stellar masses below $M_* < 10^{10.5} M_\odot$. This is consistent with the results in Paper I and imply that the virial radius may underestimate the boundary of gas surrounding galaxies. This has important implications for satellites, who may have their CGM impacted by host halos CGM at distances that are larger than previously thought. In addition, this enhancement of gas around galaxies to larger distances implies a larger fuel reserve for galaxy star formation. It also paints a picture of an environmental dependence on the clustering of HI around galaxies.

5.2 Reconstructing the Cosmic Web

Of paramount concern in galaxy evolution science is the impact of a galaxy's environment on its morphology and star formation activity. Correlations between environmental metrics and galaxy properties, such as morphology (Dressler, 1980, e.g.), color (Abell, 1965, e.g.), and star formation (Peng et al., 2010) have been known about for many decades, but the physical mechanisms and their relative importance remain heavily pursued problems.

In order to address the effects that environment have on the CGM and thus galaxy evolution, we created a public data product, consisting of a galaxy catalog and 3D data cubes that were released as part of the Sloan Digital Sky Survey Data Release 17 (Abdurro'uf et al., 2022) and are now publicly available. we employ the novel method first introduced in Burchett et al. (2020a) which is based on the movements of the organism Physarum polycephalum slime mold to map the cosmic density field. This model implicitly traces the Cosmic Web structure by efficiently finding optimal pathways between galaxies, which are known to trace filaments. We apply our model to two large catalogs of galaxies, the NASA Sloan Atlas (NSA) (Blanton et al., 2011) and the catalogs of Luminous Red Galaxies (LRGs) from SDSS-IV Extended Baryon Oscillation Spectroscopic Survey (Bautista et al., 2018). Our method faithfully reconstructs the cosmic matter density of the cosmic web for all points in the observational footprints, allowing the study of this dark matter distribution with any objects of interest in the footprints.

5.3 Ongoing & Future Work

In future work, it would be valuable to repeat the methodology of fitting the the extent of the CGM presented in Chapter 3 with the metal line species already identified in the CGM^2 and CASBaH surveys. Comparing the low-ions tracing the cool gas such as SiII, SiIII, CIII, etc. with warmer gas such as the OVI. The metals emanate from the galaxies themselves and could paint a fuller picture of the extent of the CGM and dependence galaxy mass has on the extent of warm-hot vs cool gas.

The two point correlation function used in that analysis has also been the mainstay for cosmological studies for years (Peebles, 1980) yet is not sensitive to the complexity found in the cosmic web. However, it would prove a useful "zeroth-order" comparison of the two-point correlation of the slime mold cosmic web reconstruction to the simulations that were used to calibrate it. The MCPM has been optimized to recreate a visual reconstruction of the cosmic web along with a linear (in log scale) correspondence to the dark matter density. A two point correlation of the density could provide another metric by which to optimize the algorithm.

In addition to analyzing the faithfulness of the reproduction, we are working on running other density recreation methods on the surveys such as the 5th nearest neighbor technique used by Peng et al. (2010). This would allow us more directly compare our galaxy-environment dependence with their findings.

Methodologically, the MCPM is currently tied to very specific hardware up which it must be run. We are in the process of making a portable and open source version of the Physarum machine that can be run on generic GPU clusters.

Connecting the cosmic web to the CGM, future studies should incorporate the densities provided by slime mold to the CGM^2 and CASBaH surveys to better place their environmental dependence's in context.

5.4 Concluding Remarks

In conclusion, we anticipate a bright future for the study of the circumgalactic medium and its connection to the large scale environment. In this work, we have presented a new survey and a method of reconstructing said large scale environment that should suite these purposes. There remain many questions related to galaxy evolution. However, we believe we have contributed one of the more basic properties to an astronomical object: the size. Having this in hand can push forward future work that will hope to solve these puzzles.

BIBLIOGRAPHY

Abdurro'uf, Accetta, K., Aerts, C., Silva Aguirre, V., Ahumada, R., Ajgaonkar, N., Filiz Ak, N., Alam, S., Allende Prieto, C., Almeida, A., Anders, F., Anderson, S. F., Andrews, B. H., Anguiano, B., Aquino-Ortíz, E., Aragón-Salamanca, A., Argudo-Fernández, M., Ata, M., Aubert, M., Avila-Reese, V., Badenes, C., Barbá, R. H., Barger, K., Barrera-Ballesteros, J. K., Beaton, R. L., Beers, T. C., Belfiore, F., Bender, C. F., Bernardi, M., Bershady, M. A., Beutler, F., Bidin, C. M., Bird, J. C., Bizyaev, D., Blanc, G. A., Blanton, M. R., Boardman, N. F., Bolton, A. S., Boquien, M., Borissova, J., Bovy, J., Brandt, W. N., Brown, J., Brownstein, J. R., Brusa, M., Buchner, J., Bundy, K., Burchett, J. N., Bureau, M., Burgasser, A., Cabang, T. K., Campbell, S., Cappellari, M., Carlberg, J. K., Wanderley, F. C., Carrera, R., Cash, J., Chen, Y.-P., Chen, W.-H., Cherinka, B., Chiappini, C., Choi, P. D., Chojnowski, S. D., Chung, H., Clerc, N., Cohen, R. E., Comerford, J. M., Comparat, J., da Costa, L., Covey, K., Crane, J. D., Cruz-Gonzalez, I., Culhane, C., Cunha, K., Dai, Y. S., Damke, G., Darling, J., Davidson, James W., J., Davies, R., Dawson, K., De Lee, N., Diamond-Stanic, A. M., Cano-Díaz, M., Sánchez, H. D., Donor, J., Duckworth, C., Dwelly, T., Eisenstein, D. J., Elsworth, Y. P., Emsellem, E., Eracleous, M., Escoffier, S., Fan, X., Farr, E., Feng, S., Fernández-Trincado, J. G., Feuillet, D., Filipp, A., Fillingham, S. P., Frinchaboy, P. M., Fromenteau, S., Galbany, L., García, R. A., García-Hernández, D. A., Ge, J., Geisler, D., Gelfand, J., Géron, T., Gibson, B. J., Goddy, J., Godoy-Rivera, D., Grabowski, K., Green, P. J., Greener, M., Grier, C. J., Griffith, E., Guo, H., Guy, J., Hadjara, M., Harding, P., Hasselquist, S., Hayes, C. R., Hearty, F., Hernández, J., Hill, L., Hogg, D. W., Holtzman, J. A., Horta, D., Hsieh, B.-C., Hsu, C.-H., Hsu, Y.-H., Huber, D., Huertas-Company, M., Hutchinson, B., Hwang, H. S., Ibarra-Medel, H. J., Chitham, J. I., Ilha, G. S., Imig, J., Jaekle, W., Jayasinghe, T., Ji, X., Johnson, J. A., Jones, A., Jönsson, H., Katkov, I., Khalatyan, Arman, D., Kinemuchi, K., Kisku, S., Knapen, J. H., Kneib, J.-P., Kollmeier, J. A., Kong, M., Kounkel, M., Kreckel, K., Krishnarao, D., Lacerna, I., Lane, R. R., Langgin, R., Lavender, R., Law, D. R., Lazarz, D., Leung, H. W., Leung, H.-H., Lewis, H. M., Li, C., Li, R., Lian, J., Liang, F.-H., Lin, L., Lin, Y.-T., Lin, S., Lintott, C., Long, D., Longa-Peña, P., López-Cobá, C., Lu, S., Lundgren, B. F., Luo, Y., Mackereth, J. T., de la Macorra, A., Mahadevan, S., Majewski, S. R., Manchado, A., Mandeville, T., Maraston, C., Margalef-Bentabol, B., Masseron, T., Masters, K. L., Mathur, S., McDermid, R. M., Mckay, M., Merloni, A., Merrifield, M., Meszaros, S., Miglio, A., Di Mille, F., Minniti, D., Minsley, R., Monachesi, A., Moon, J., Mosser, B., Mulchaey, J., Muna, D., Muñoz, R. R., Myers, A. D., Myers, N., Nadathur, S., Nair, P., Nandra, K., Neumann, J., Newman, J. A., Nidever, D. L., Nikakhtar, F., Nitschelm, C., O'Connell, J. E., Garma-Oehmichen, L., Luan Souza de Oliveira, G., Olney, R., Oravetz, D., Ortigoza-Urdaneta, M., Osorio, Y., Otter, J., Pace, Z. J., Padilla, N., Pan, K., Pan, H.-A., Parikh, T., Parker, J., Peirani, S., Peña Ramírez, K., Penny, S., Percival, W. J., Perez-Fournon, I., Pinsonneault, M., Poidevin, F., Poovelil, V. J., Price-Whelan, A. M., Bárbara de Andrade Queiroz, A., Raddick, M. J., Ray, A., Rembold, S. B., Riddle, N., Riffel, R. A., Riffel, R., Rix, H.-W., Robin, A. C., Rodríguez-Puebla, A., Roman-Lopes, A., Román-Zúñiga, C., Rose, B., Ross, A. J., Rossi, G., Rubin, K. H. R., Salvato, M., Sánchez, S. F., Sánchez-Gallego, J. R., Sanderson, R., Santana Rojas, F. A., Sarceno, E., Sarmiento, R., Sayres, C., Sazonova, E., Schaefer, A. L., Schiavon, R., Schlegel, D. J., Schneider, D. P., Schultheis, M., Schwope, A., Serenelli, A., Serna, J., Shao, Z., Shapiro, G., Sharma, A., Shen, Y., Shetrone, M., Shu, Y., Simon, J. D., Skrutskie, M. F., Smethurst, R., Smith, V., Sobeck, J., Spoo, T., Sprague, D., Stark, D. V., Stassun, K. G., Steinmetz, M., Stello, D., Stone-Martinez, A., Storchi-Bergmann, T., Stringfellow, G. S., Stutz, A., Su, Y.-C., Taghizadeh-Popp, M., Talbot, M. S., Tayar, J., Telles, E., Teske, J., Thakar, A., Theissen, C., Tkachenko, A., Thomas, D., Tojeiro, R., Hernandez Toledo, H., Troup, N. W., Trump, J. R., Trussler, J., Turner, J., Tuttle, S., Unda-Sanzana, E., Vázquez-Mata, J. A., Valentini, M., Valenzuela, O., Vargas-González, J., Vargas-Magaña, M., Alfaro, P. V., Villanova, S., Vincenzo, F., Wake, D., Warfield, J. T., Washington, J. D., Weaver, B. A., Weijmans, A.-M., Weinberg, D. H., Weiss, A., Westfall, K. B., Wild, V., Wilde, M. C., Wilson, J. C., Wilson, R. F., Wilson, M., Wolf, J., Wood-Vasey, W. M., Yan, R., Zamora, O., Zasowski, G., Zhang, K., Zhao, C., Zheng, Z., Zheng, Z., and Zhu, K. (2022). The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar, and APOGEE-2 Data. ApJS, 259(2):35.

Abell, G. O. (1965). Clustering of Galaxies. ARAA, 3:1.

Abolfathi, B., Aguado, D. S., Aguilar, G., Allende Prieto, C., Almeida, A., Ananna, T. T., Anders, F., Anderson, S. F., Andrews, B. H., Anguiano, B., Aragón-Salamanca, A., Argudo-Fernández, M., Armengaud, E., Ata, M., Aubourg, E., Avila-Reese, V., Badenes, C., Bailey, S., Balland, C., Barger, K. A., Barrera-Ballesteros, J., Bartosz, C., Bastien, F., Bates, D., Baumgarten, F., Bautista, J., Beaton, R., Beers, T. C., Belfiore, F., Bender, C. F., Bernardi, M., Bershady, M. A., Beutler, F., Bird, J. C., Bizyaev, D., Blanc, G. A., Blanton, M. R., Blomqvist, M., Bolton, A. S., Boquien, M., Borissova, J., Bovy, J., Bradna Diaz, C. A., Brandt, W. N., Brinkmann, J., Brownstein, J. R., Bundy, K., Burgasser, A. J., Burtin, E., Busca, N. G., Cañas, C. I., Cano-Díaz, M., Cappellari, M., Carrera, R., Casey, A. R., Cervantes Sodi, B., Chen, Y., Cherinka, B., Chiappini, C., Choi, P. D., Chojnowski, D., Chuang, C.-H., Chung, H., Clerc, N., Cohen, R. E., Comerford, J. M., Comparat, J., Correa do Nascimento, J., da Costa, L., Cousinou, M.-C., Covey, K., Crane, J. D., Cruz-Gonzalez, I., Cunha, K., da Silva Ilha, G., Damke, G. J., Darling, J., Davidson, James W., J., Dawson, K., de Icaza Lizaola, M. A. C., de la Macorra, A., de la Torre, S., De Lee, N., de Sainte Agathe, V., Deconto Machado, A., Dell'Agli, F., Delubac, T., Diamond-Stanic, A. M., Donor, J., Downes, J. J., Drory, N., du Mas des Bourboux, H., Duckworth, C. J., Dwelly, T., Dyer, J., Ebelke, G., Davis Eigenbrot, A., Eisenstein, D. J., Elsworth, Y. P., Emsellem, E., Eracleous, M., Erfanianfar, G., Escoffier, S., Fan, X., Fernández Alvar, E., Fernandez-Trincado, J. G., Fernand o Cirolini, R., Feuillet, D., Finoguenov, A., Fleming,

S. W., Font-Ribera, A., Freischlad, G., Frinchaboy, P., Fu, H., Gómez Maqueo Chew, Y., Galbany, L., García Pérez, A. E., Garcia-Dias, R., García-Hernández, D. A., Garma Oehmichen, L. A., Gaulme, P., Gelfand , J., Gil-Marín, H., Gillespie, B. A., Goddard, D., González Hernández, J. I., Gonzalez-Perez, V., Grabowski, K., Green, P. J., Grier, C. J., Gueguen, A., Guo, H., Guy, J., Hagen, A., Hall, P., Harding, P., Hasselquist, S., Hawley, S., Hayes, C. R., Hearty, F., Hekker, S., Hernand ez, J., Hernandez Toledo, H., Hogg, D. W., Holley-Bockelmann, K., Holtzman, J. A., Hou, J., Hsieh, B.-C., Hunt, J. A. S., Hutchinson, T. A., Hwang, H. S., Jimenez Angel, C. E., Johnson, J. A., Jones, A., Jönsson, H., Jullo, E., Khan, F. S., Kinemuchi, K., Kirkby, D., Kirkpatrick, Charles C., I., Kitaura, F.-S., Knapp, G. R., Kneib, J.-P., Kollmeier, J. A., Lacerna, I., Lane, R. R., Lang, D., Law, D. R., Le Goff, J.-M., Lee, Y.-B., Li, H., Li, C., Lian, J., Liang, Y., Lima, M., Lin, L., Long, D., Lucatello, S., Lundgren, B., Mackereth, J. T., MacLeod, C. L., Mahadevan, S., Maia, M. A. G., Majewski, S., Manchado, A., Maraston, C., Mariappan, V., Marques-Chaves, R., Masseron, T., Masters, K. L., McDermid, R. M., McGreer, I. D., Melendez, M., Meneses-Goytia, S., Merloni, A., Merrifield, M. R., Meszaros, S., Meza, A., Minchev, I., Minniti, D., Mueller, E.-M., Muller-Sanchez, F., Muna, D., Muñoz, R. R., Myers, A. D., Nair, P., Nand ra, K., Ness, M., Newman, J. A., Nichol, R. C., Nidever, D. L., Nitschelm, C., Noterdaeme, P., O'Connell, J., Oelkers, R. J., Oravetz, A., Oravetz, D., Ortíz, E. A., Osorio, Y., Pace, Z., Padilla, N., Palanque-Delabrouille, N., Palicio, P. A., Pan, H.-A., Pan, K., Parikh, T., Pâris, I., Park, C., Peirani, S., Pellejero-Ibanez, M., Penny, S., Percival, W. J., Perez-Fournon, I., Petitjean, P., Pieri, M. M., Pinsonneault, M., Pisani, A., Prada, F., Prakash, A., Queiroz, A. B. d. A., Raddick, M. J., Raichoor, A., Barboza Rembold, S., Richstein, H., Riffel, R. A., Riffel, R., Rix, H.-W., Robin, A. C., Rodríguez Torres, S., Román-Zúñiga, C., Ross, A. J., Rossi, G., Ruan, J., Ruggeri, R., Ruiz, J., Salvato, M., Sánchez, A. G., Sánchez, S. F., Sanchez Almeida, J., Sánchez-Gallego, J. R., Santana Rojas, F. A., Santiago, B. X., Schiavon, R. P., Schimoia, J. S., Schlafly, E., Schlegel, D., Schneider, D. P., Schuster, W. J., Schwope, A., Seo, H.-J., Serenelli, A., Shen, S., Shen, Y., Shetrone, M., Shull, M., Silva Aguirre, V., Simon, J. D., Skrutskie, M., Slosar, A., Smethurst, R., Smith, V., Sobeck, J., Somers, G., Souter, B. J., Souto, D., Spindler, A., Stark, D. V., Stassun, K., Steinmetz, M., Stello, D., Storchi-Bergmann, T., Streblyanska, A., Stringfellow, G. S., Suárez, G., Sun, J., Szigeti, L., Taghizadeh-Popp, M., Talbot, M. S., Tang, B., Tao, C., Tayar, J., Tembe, M., Teske, J., Thakar, A. R., Thomas, D., Tissera, P., Tojeiro, R., Tremonti, C., Troup, N. W., Urry, M., Valenzuela, O., van den Bosch, R., Vargas-González, J., Vargas-Magaña, M., Vazquez, J. A., Villanova, S., Vogt, N., Wake, D., Wang, Y., Weaver, B. A., Weijmans, A.-M., Weinberg, D. H., Westfall, K. B., Whelan, D. G., Wilcots, E., Wild, V., Williams, R. A., Wilson, J., Wood-Vasey, W. M., Wylezalek, D., Xiao, T., Yan, R., Yang, M., Ybarra, J. E., Yèche, C., Zakamska, N., Zamora, O., Zarrouk, P., Zasowski, G., Zhang, K., Zhao, C., Zhao, G.-B., Zheng, Z., Zheng, Z., Zhou, Z.-M., Zhu, G., Zinn, J. C., and Zou, H. (2018). The Fourteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the Extended Baryon Oscillation Spectroscopic Survey and from the Second Phase of the Apache Point Observatory Galactic Evolution Experiment. ApJS, 235(2):42.

Alam, S., Ata, M., Bailey, S., Beutler, F., Bizyaev, D., Blazek, J. A., Bolton, A. S., Brownstein, J. R., Burden, A., Chuang, C.-H., Comparat, J., Cuesta, A. J., Dawson, K. S., Eisenstein, D. J., Escoffier, S., Gil-Marín, H., Grieb, J. N., Hand, N., Ho, S., Kinemuchi, K., Kirkby, D., Kitaura, F., Malanushenko, E., Malanushenko, V., Maraston, C., McBride, C. K., Nichol, R. C., Olmstead, M. D., Oravetz, D., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Pellejero-Ibanez, M., Percival, W. J., Petitjean, P., Prada, F., Price-Whelan, A. M., Reid, B. A., Rodríguez-Torres, S. A., Roe, N. A., Ross, A. J., Ross, N. P., Rossi, G., Rubiño-Martín, J. A., Saito, S., Salazar-Albornoz, S., Samushia, L., Sánchez, A. G., Satpathy, S., Schlegel, D. J., Schneider, D. P., Scóccola, C. G., Seo, H.-J., Sheldon, E. S., Simmons, A., Slosar, A., Strauss, M. A., Swanson, M. E. C., Thomas, D., Tinker, J. L., Tojeiro, R., Magaña, M. V., Vazquez, J. A., Verde, L., Wake, D. A., Wang, Y., Weinberg, D. H., White, M., Wood-Vasey, W. M., Yèche, C., Zehavi, I., Zhai, Z., and Zhao, G.-B. (2017). The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. MNRAS, 470(3):2617–2652.

Alcock, C. and Paczynski, B. (1979). An evolution free test for non-zero cosmological constant. Nature, 281:358.

Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., Hopkins, P. F., Quataert, E., and Murray, N. (2017). The cosmic baryon cycle and galaxy mass assembly in the FIRE simulations. MNRAS, 470(4):4698–4719.

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Unther, H. M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J. E. H., Singer, L., Fox, R., Weaver, B. A., Zabalza, V., Edwards, Z. I., Azalee Bostroem, K., Burke, D. J., Casey, A. R., Crawford, S. M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P. L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., and Streicher, O. (2013). Astropy: A community Python package for astronomy. A&A, 558:A33.

Bahcall, J. N. and Spitzer, Jr., L. (1969). Absorption Lines Produced by Galactic Halos. ApJ, 156:L63.

Bautista, J. E., Vargas-Magaña, M., Dawson, K. S., Percival, W. J., Brinkmann, J., Brownstein, J., Camacho, B., Comparat, J., Gil-Marín, H., Mueller, E.-M., Newman, J. A., Prakash, A., Ross, A. J., Schneider, D. P., Seo, H.-J., Tinker, J., Tojeiro, R., Zhai, Z., and Zhao, G.-B. (2018). The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations at Redshift of 0.72 with the DR14 Luminous Red Galaxy Sample. ApJ, 863(1):110.

Behroozi, P. S., Wechsler, R. H., and Conroy, C. (2013). On the Lack of Evolution in Galaxy Star Formation Efficiency. ApJ, 762(2):L31.

Behroozi, P. S., Wechsler, R. H., and Wu, H.-Y. (2012). THE ROCKSTAR PHASE-SPACE TEMPORAL HALO FINDER AND THE VELOCITY OFFSETS OF CLUSTER CORES. The Astrophysical Journal, 762(2):109.

Berg, M. A., Lehner, N., Howk, J. C., O'Meara, J. M., Schaye, J., Straka, L. A., Cooksey, K. L., Tripp, T. M., Prochaska, J. X., Oppenheimer, B. D., Johnson, S. D., Muzahid, S., Bordoloi, R., Werk, J. K., Fox, A. J., Katz, N., Wendt, M., Peeples, M. S., Ribaudo, J., and Tumlinson, J. (2022). The Bimodal Absorption System Imaging Campaign (BASIC) I. A Dual Population of Low-metallicity Absorbers at z < 1. arXiv e-prints, page arXiv:2204.13229.

Bergeron, J., Cristiani, S., and Shaver, P. A. (1992). Discovery of Z about 1 galaxies causing quasar absorption lines. A&A, 257:417–424.

Bergeron, J. and Stasińska, G. (1986). Absorption line systems in QSO spectra : properties derived from observations and from photoionization models. A&A, 169:1–3.

Berlind, A. A., Florez, J., Calderon, V., Sinha, M., Moffett, A. J., Eckert, K. D., Kannappan, S., Stark, D., Baker, A., and RESOLVE Team (2016). Mocking the ECO and RESOLVE Surveys: Probing the Environmental Dependencies of Galaxy Properties. In American Astronomical Society Meeting Abstracts #227, volume 227 of American Astronomical Society Meeting Abstracts, page 311.04.

Berti, A. M., Coil, A. L., Hearin, A. P., and Behroozi, P. S. (2021). Main-sequence Scatter is Real: The Joint Dependence of Galaxy Clustering on Star Formation and Stellar Mass. AJ, 161(1):49.

Bertin, E. and Arnouts, S. (1996). SExtractor: Software for source extraction. A&AS, 117:393–404.

Bertschinger, E. (1985). Self-similar secondary infall and accretion in an Einstein-de Sitter universe. ApJS, 58:39–65.

Blanton, M. R., Kazin, E., Muna, D., Weaver, B. A., and Price-Whelan, A. (2011). Improved Background Subtraction for the Sloan Digital Sky Survey Images. AJ, 142(1):31.

Blanton, M. R. and Moustakas, J. (2009). Physical Properties and Environments of Nearby Galaxies. ARA&A, 47(1):159–210.

Blanton, M. R. and Roweis, S. (2007). K-Corrections and Filter Transformations in the Ultraviolet, Optical, and Near-Infrared. AJ, 133(2):734–754.

Boksenberg, A. and Sargent, W. L. W. (1978). The existence of Ca II absorption lines in the spectrum of the quasar 3C 232 due to the galaxy NGC 3067. ApJ, 220:42–46.

Boquien, M., Burgarella, D., Roehlly, Y., Buat, V., Ciesla, L., Corre, D., Inoue, A. K., and Salas, H. (2019). CIGALE: a python Code Investigating GALaxy Emission. A&A, 622:A103.

Bordoloi, R., Lilly, S. J., Knobel, C., Bolzonella, M., Kampczyk, P., Carollo, C. M., Iovino, A., Zucca, E., Contini, T., Kneib, J. P., Le Fevre, O., Mainieri, V., Renzini, A., Scodeggio, M., Zamorani, G., Balestra, I., Bardelli, S., Bongiorno, A., Caputi, K., Cucciati, O., de la Torre, S., de Ravel, L., Garilli, B., Kovač, K., Lamareille, F., Le Borgne, J. F., Le Brun, V., Maier, C., Mignoli, M., Pello, R., Peng, Y., Perez Montero, E., Presotto, V., Scarlata, C., Silverman, J., Tanaka, M., Tasca, L., Tresse, L., Vergani, D., Barnes, L., Cappi, A., Cimatti, A., Coppa, G., Diener, C., Franzetti, P., Koekemoer, A., López-Sanjuan, C., McCracken, H. J., Moresco, M., Nair, P., Oesch, P., Pozzetti, L., and Welikala, N. (2011). The Radial and Azimuthal Profiles of Mg II Absorption around 0.5 < z < 0.9 zCOSMOS Galaxies of Different Colors, Masses, and Environments. ApJ, 743(1):10.

Bordoloi, R., Prochaska, J. X., Tumlinson, J., Werk, J. K., Tripp, T. M., and Burchett, J. N.

(2018). On the CGM Fundamental Plane: The Halo Mass Dependency of Circumgalactic H I. ApJ, 864(2):132.

Bordoloi, R., Tumlinson, J., Werk, J. K., Oppenheimer, B. D., Peeples, M. S., Prochaska, J. X., Tripp, T. M., Katz, N., Davé, R., Fox, A. J., Thom, C., Ford, A. B., Weinberg, D. H., Burchett, J. N., and Kollmeier, J. A. (2014). The COS-Dwarfs Survey: The Carbon Reservoir around Sub-L* Galaxies. ApJ, 796(2):136.

Borthakur, S. (2022). How are Lyα Absorbers in the Cosmic Web Related to Gas-rich Galaxies? ApJ, 924(2):123.

Borthakur, S., Heckman, T., Tumlinson, J., Bordoloi, R., Thom, C., Catinella, B., Schiminovich, D., Davé, R., Kauffmann, G., Moran, S. M., and Saintonge, A. (2015). Connection between the Circumgalactic Medium and the Interstellar Medium of Galaxies: Results from the COS-GASS Survey. ApJ, 813(1):46.

Bouma, S. J. D., Richter, P., and Wendt, M. (2021). The relation between Lyα absorbers and local galaxy filaments. A&A, 647:A166.

Bowen, D. V., Blades, J. C., and Pettini, M. (1995). The Distribution of Metal-absorbing High-Velocity Clouds in the Galaxy. ApJ, 448:662.

Bowen, D. V., Chelouche, D., Jenkins, E. B., Tripp, T. M., Pettini, M., York, D. G., and Frye, B. L. (2016). The Structure of the Circumgalactic Medium of Galaxies: Cool Accretion Inflow Around NGC 1097. ApJ, 826(1):50.

Bowen, D. V., Pettini, M., and Blades, J. C. (2002). Lyα Absorption around Nearby Galaxies. ApJ, 580(1):169–193.

Bowen, D. V., Pettini, M., Penston, M. V., and Blades, C. (1991). A survey of interstellar Ca II absorption in the haloes of low-redshift galaxies. MNRAS, 249:145.

Bregman, J. N. (2007). The search for the missing baryons at low redshift. Annual Review of Astronomy and Astrophysics, 45(1):221–259.

Bruzual, G. and Charlot, S. (2003). Stellar population synthesis at the resolution of 2003. MNRAS, 344(4):1000–1028.

Bryan, G. L. and Norman, M. L. (1998). Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. ApJ, 495(1):80–99.

Burchett, J. N., Elek, O., Tejos, N., Prochaska, J. X., Tripp, T. M., Bordoloi, R., and Forbes, A. G. (2020a). Revealing the Dark Threads of the Cosmic Web. ApJ, 891(2):L35.

Burchett, J. N., Elek, O., Tejos, N., Tripp, T. M., Bordoloi, R., and Forbes, A. G. (2020b). Revealing the dark threads of the cosmic web. ApJL, 891:L35.

Burchett, J. N., Rubin, K. H. R., Prochaska, J. X., Coil, A. L., Rickards Vaught, R., and Hennawi, J. F. (2020c). Circumgalactic Mg II Emission from an Isotropic Starburst Galaxy Outflow Mapped by KCWI. arXiv e-prints, page arXiv:2005.03017.

Burchett, J. N., Tripp, T. M., Bordoloi, R., Werk, J. K., Prochaska, J. X., Tumlinson, J., Willmer, C. N. A., O'Meara, J., and Katz, N. (2016). A Deep Search for Faint Galaxies Associated with Very Low Redshift C IV Absorbers. III. The Mass- and Environment-dependent Circumgalactic Medium. ApJ, 832:124.

Burchett, J. N., Tripp, T. M., Prochaska, J. X., Werk, J. K., Tumlinson, J., Howk, J. C., Willmer, C. N. A., Lehner, N., Meiring, J. D., Bowen, D. V., Bordoloi, R., Peeples, M. S., Jenkins, E. B., O'Meara, J. M., Tejos, N., and Katz, N. (2019). The COS Absorption Survey of Baryon Harbors (CASBaH): Warm-Hot Circumgalactic Gas Reservoirs Traced by Ne VIII Absorption. ApJ, 877(2):L20.

Burchett, J. N., Tripp, T. M., Wang, Q. D., Willmer, C. N. A., Bowen, D. V., and Jenkins, E. B. (2018). Warm-hot gas in X-ray bright galaxy clusters and the H I-deficient circumgalactic medium in dense environments. MNRAS, 475(2):2067–2085.

Butsky, I. S., Fielding, D. B., Hayward, C. C., Hummels, C. B., Quinn, T. R., and Werk, J. K. (2020). The Impact of Cosmic Rays on Thermal Instability in the Circumgalactic Medium. ApJ, 903(2):77.

Calzetti, D., Kinney, A. L., and Storchi-Bergmann, T. (1994). Dust Extinction of the Stellar Continua in Starburst Galaxies: The Ultraviolet and Optical Extinction Law. ApJ, 429:582.

Cappellari, M. (2017). Improving the full spectrum fitting method: accurate convolution with Gauss-Hermite functions. MNRAS, 466(1):798–811.

Carollo, C. M., Cibinel, A., Lilly, S. J., Miniati, F., Norberg, P., Silverman, J. D., van Gorkom, J., Cameron, E., Finoguenov, A., Peng, Y., Pipino, A., and Rudick, C. S. (2013). The Zurich Environmental Study of Galaxies in Groups along the Cosmic Web. I. Which Environment Affects Galaxy Evolution? ApJ, 776:71.

Catinella, B., Schiminovich, D., Cortese, L., Fabello, S., Hummels, C. B., Moran, S. M., Lemonias, J. J., Cooper, A. P., Wu, R., Heckman, T. M., and Wang, J. (2013). The GALEX Arecibo SDSS Survey - VIII. Final data release. The effect of group environment on the gas content of massive galaxies. MNRAS, 436:34–70.

Cen, R. and Ostriker, J. P. (1999). Where Are the Baryons? ApJ, 514:1–6.

Chabrier, G. (2003). Galactic Stellar and Substellar Initial Mass Function. PASP, 115:763–795.

Chambers, K. C., Magnier, E. A., Metcalfe, N., Flewelling, H. A., Huber, M. E., Waters, C. Z., Denneau, L., Draper, P. W., Farrow, D., Finkbeiner, D. P., Holmberg, C., Koppenhoefer, J., Price, P. A., Rest, A., Saglia, R. P., Schlafly, E. F., Smartt, S. J., Sweeney, W., Wainscoat, R. J., Burgett, W. S., Chastel, S., Grav, T., Heasley, J. N., Hodapp, K. W., Jedicke, R., Kaiser, N., Kudritzki, R. P., Luppino, G. A., Lupton, R. H., Monet, D. G., Morgan, J. S., Onaka, P. M., Shiao, B., Stubbs, C. W., Tonry, J. L., White, R., Bañados, E., Bell, E. F., Bender, R., Bernard, E. J., Boegner, M., Boffi, F., Botticella, M. T., Calamida, A., Casertano, S., Chen, W. P., Chen, X., Cole, S., Deacon, N., Frenk, C., Fitzsimmons, A., Gezari, S., Gibbs, V., Goessl, C., Goggia, T., Gourgue, R., Goldman, B., Grant, P., Grebel, E. K., Hambly, N. C., Hasinger, G., Heavens, A. F., Heckman, T. M., Henderson, R., Henning, T., Holman, M., Hopp, U., Ip, W. H., Isani, S., Jackson, M., Keyes, C. D., Koekemoer, A. M., Kotak, R., Le, D., Liska, D., Long, K. S., Lucey, J. R., Liu, M., Martin, N. F., Masci, G., McLean, B., Mindel, E., Misra, P., Morganson, E., Murphy, D. N. A., Obaika, A., Narayan, G., Nieto-Santisteban, M. A., Norberg, P., Peacock, J. A., Pier, E. A., Postman, M., Primak, N., Rae, C., Rai, A., Riess, A., Riffeser, A., Rix, H. W., Röser, S., Russel, R., Rutz, L., Schilbach, E., Schultz, A. S. B., Scolnic, D., Strolger, L., Szalay, A., Seitz, S., Small, E., Smith, K. W., Soderblom, D. R., Taylor, P., Thomson, R., Taylor, A. N., Thakar, A. R., Thiel, J., Thilker, D., Unger, D., Urata, Y., Valenti, J., Wagner, J., Walder, T., Walter, F., Watters, S. P., Werner, S., Wood-Vasey, W. M., and Wyse, R. (2016). The Pan-STARRS1 Surveys. arXiv e-prints, page arXiv:1612.05560.

Chang, Y.-Y., van der Wel, A., da Cunha, E., and Rix, H.-W. (2015). Stellar Masses and Star Formation Rates for 1M Galaxies from SDSS+WISE. ApJS, 219(1):8.

Chen, H.-W., Gauthier, J.-R., Sharon, K., Johnson, S. D., Nair, P., and Liang, C. J. (2014). Spatially resolved velocity maps of halo gas around two intermediate-redshift galaxies. MNRAS, 438:1435–1450.

Chen, H.-W., Lanzetta, K. M., Webb, J. K., and Barcons, X. (1998). The Gaseous Extent of Galaxies and the Origin of Lyα Absorption Systems. III. Hubble Space Telescope Imaging of Lyα-absorbing Galaxies at z < 1. ApJ, 498(1):77–94.

Chen, H.-W., Lanzetta, K. M., Webb, J. K., and Barcons, X. (2001). The Gaseous Extent of Galaxies and the Origin of Lyα Absorption Systems. V. Optical and Near-Infrared Photometry of Lyα-absorbing Galaxies at z1. ApJ, 559:654–674.

Chen, H.-W., Prochaska, J. X., Weiner, B. J., Mulchaey, J. S., and Williger, G. M. (2005). Probing the Intergalactic Medium-Galaxy Connection toward PKS 0405-123. II. A Cross-Correlation Study of Lyα Absorbers and Galaxies at z<0.5. ApJ, 629(1):L25–L28.

Chen, H.-W. and Tinker, J. L. (2008). The Baryon Content of Dark Matter Halos: Empirical Constraints from Mg II Absorbers. ApJ, 687(2):745–756.

Chen, H.-W., Zahedy, F. S., Johnson, S. D., Pierce, R. M., Huang, Y.-H., Weiner, B. J., and Gauthier, J.-R. (2018). Characterizing circumgalactic gas around massive ellipticals at z̃ 0.4 - I. Initial results. MNRAS, 479(2):2547–2563.

Chen, H.-Y. and Holz, D. E. (2016). Finding the One: Identifying the Host Galaxies of Gravitational-Wave Sources. arXiv e-prints, page arXiv:1612.01471.

CHIME/FRB Collaboration, Andersen, B. C., Bandura, K., Bhardwaj, M., Boubel, P., Boyce, M. M., Boyle, P. J., Brar, C., Cassanelli, T., Chawla, P., Cubranic, D., Deng, M., Dobbs, M., Fandino, M., Fonseca, E., Gaensler, B. M., Gilbert, A. J., Giri, U., Good, D. C., Halpern, M., Hill, A. S., Hinshaw, G., Höfer, C., Josephy, A., Kaspi, V. M., Kothes, R., Landecker, T. L., Lang, D. A., Li, D. Z., Lin, H. H., Masui, K. W., Mena-Parra, J., Merryfield, M., Mckinven, R., Michilli, D., Milutinovic, N., Naidu, A., Newburgh, L. B., Ng, C., Patel, C., Pen, U., Pinsonneault-Marotte, T., Pleunis, Z., Rafiei-Ravandi, M., Rahman, M., Ransom, S. M., Renard, A., Scholz, P., Siegel, S. R., Singh, S., Smith, K. M., Stairs, I. H., Tendulkar, S. P., Tretyakov, I., Vanderlinde, K., Yadav, P., and Zwaniga, A. V. (2019). CHIME/FRB Discovery of Eight New Repeating Fast Radio Burst Sources. ApJ, 885(1):L24.

Christensen, C. R., Davé, R., Governato, F., Pontzen, A., Brooks, A., Munshi, F., Quinn, T., and Wadsley, J. (2016). In-N-Out: The Gas Cycle from Dwarfs to Spiral Galaxies. ApJ, 824(1):57.

Codis, S., Jindal, A., Chisari, N. E., Vibert, D., Dubois, Y., Pichon, C., and Devriendt, J. (2018). Galaxy orientation with the cosmic web across cosmic time. MNRAS, 481(4):4753–4774.

Cole, S. and Lacey, C. (1996). The structure of dark matter haloes in hierarchical clustering models. Monthly Notices of the Royal Astronomical Society, 281(2):716–736.

Comparat, J., Maraston, C., Goddard, D., Gonzalez-Perez, V., Lian, J., Meneses-Goytia, S., Thomas, D., Brownstein, J. R., Tojeiro, R., Finoguenov, A., Merloni, A., Prada, F., Salvato, M., Zhu, G. B., Zou, H., and Brinkmann, J. (2017). Stellar population properties for 2 million galaxies from SDSS DR14 and DEEP2 DR4 from full spectral fitting. arXiv e-prints, page arXiv:1711.06575.

Coulter, D. A., Foley, R. J., Kilpatrick, C. D., Drout, M. R., Piro, A. L., Shappee, B. J., Siebert, M. R., Simon, J. D., Ulloa, N., Kasen, D., Madore, B. F., Murguia-Berthier, A., Pan, Y. C., Prochaska, J. X., Ramirez-Ruiz, E., Rest, A., and Rojas-Bravo, C. (2017). Swope Supernova Survey 2017a (SSS17a), the optical counterpart to a gravitational wave source. Science, 358(6370):1556–1558.

Courtois, H. M., Pomarède, D., Tully, R. B., Hoffman, Y., and Courtois, D. (2013). Cosmography of the Local Universe. AJ, 146(3):69.

Crone Odekon, M., Hallenbeck, G., Haynes, M. P., Koopmann, R. A., Phi, A., and Wolfe, P.-F. (2018). The Effect of Filaments and Tendrils on the H I Content of Galaxies. ApJ, 852(2):142.

Cutri, R. M., Wright, E. L., Conrow, T., Fowler, J. W., Eisenhardt, P. R. M., Grillmair, C., Kirkpatrick, J. D., Masci, F., McCallon, H. L., Wheelock, S. L., Fajardo-Acosta, S., Yan, L., Benford, D., Harbut, M., Jarrett, T., Lake, S., Leisawitz, D., Ressler, M. E., Stanford,

S. A., Tsai, C. W., Liu, F., Helou, G., Mainzer, A., Gettings, D., Gonzalez, A., Hoffman, D., Marsh, K. A., Padgett, D., Skrutskie, M. F., Beck, R. P., Papin, M., and Wittman, M. (2013). Explanatory Supplement to the AllWISE Data Release Products. Technical report.

Dale, D. A., Helou, G., Magdis, G. E., Armus, L., Díaz-Santos, T., and Shi, Y. (2014). A Two-parameter Model for the Infrared/Submillimeter/Radio Spectral Energy Distributions of Galaxies and Active Galactic Nuclei. ApJ, 784(1):83.

Danforth, C. W., Keeney, B. A., Tilton, E. M., Shull, J. M., Stocke, J. T., Stevans, M., Pieri, M. M., Savage, B. D., France, K., Syphers, D., Smith, B. D., Green, J. C., Froning, C., Penton, S. V., and Osterman, S. N. (2016). An HST/COS Survey of the Low-redshift Intergalactic Medium. I. Survey, Methodology, and Overall Results. ApJ, 817:111.

Davé, R., Anglés-Alcázar, D., Narayanan, D., Li, Q., Rafieferantsoa, M. H., and Appleby, S. (2019). SIMBA: Cosmological simulations with black hole growth and feedback. MNRAS, 486(2):2827–2849.

Davé, R., Oppenheimer, B. D., Katz, N., Kollmeier, J. A., and Weinberg, D. H. (2010). The intergalactic medium over the last 10 billion years - I. Lyα absorption and physical conditions. MNRAS, 408(4):2051–2070.

Davé, R. and Tripp, T. M. (2001). The statistical and physical properties of the low-redshift LY forest observed with the hubble space Telescope/STIS. ApJ, 553:528–537.

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., Walker, A. R., Valdes, F., Weaver, B. A., Yèche, C., Zou, H., Zhou, X., Abareshi, B., Abbott, T. M. C., Abolfathi, B., Aguilera, C., Alam, S., Allen, L., Alvarez, A., Annis, J., Ansarinejad, B., Aubert, M., Beechert, J., Bell, E. F., BenZvi, S. Y., Beutler, F., Bielby, R. M., Bolton, A. S., Briceño, C., Buckley-Geer, E. J., Butler, K., Calamida, A., Carlberg, R. G., Carter, P., Casas, R., Castander, F. J., Choi, Y., Comparat, J., Cukanovaite, E., Delubac, T., DeVries, K., Dey, S., Dhungana, G., Dickinson, M., Ding, Z., Donaldson, J. B., Duan, Y., Duckworth, C. J., Eftekharzadeh, S., Eisenstein, D. J., Etourneau, T., Fagrelius, P. A., Farihi, J., Fitzpatrick, M., Font-Ribera, A., Fulmer, L., Gänsicke, B. T., Gaztanaga, E., George, K., Gerdes, D. W., Gontcho, S. G. A., Gorgoni, C., Green, G., Guy, J., Harmer, D., Hernand ez, M., Honscheid, K., Huang, L. W., James, D. J., Jannuzi, B. T., Jiang, L., Joyce, R., Karcher, A., Karkar, S., Kehoe, R., Kneib, J.-P., Kueter-Young, A., Lan, T.-W., Lauer, T. R., Le Guillou, L., Le Van Suu, A., Lee, J. H., Lesser, M., Perreault Levasseur, L., Li, T. S., Mann, J. L., Marshall, R., Martínez-Vázquez, C. E., Martini, P., du Mas des Bourboux, H., McManus, S., Meier, T. G., Ménard, B., Metcalfe, N., Muñoz-Gutiérrez, A., Najita, J., Napier, K., Narayan, G., Newman, J. A., Nie, J., Nord, B., Norman, D. J., Olsen, K. A. G., Paat, A., Palanque-Delabrouille, N., Peng, X., Poppett, C. L., Poremba, M. R., Prakash, A., Rabinowitz, D., Raichoor, A., Rezaie, M., Robertson, A. N., Roe, N. A., Ross, A. J., Ross, N. P., Rudnick, G., Safonova, S., Saha, A., Sánchez, F. J., Savary, E., Schweiker, H., Scott, A., Seo, H.-J., Shan, H., Silva, D. R., Slepian, Z., Soto, C., Sprayberry, D., Staten, R., Stillman, C. M., Stupak, R. J., Summers, D. L., Sien Tie, S., Tirado, H., Vargas-Magaña, M., Vivas, A. K., Wechsler, R. H., Williams, D., Yang, J., Yang, Q., Yapici, T., Zaritsky, D., Zenteno, A., Zhang, K., Zhang, T., Zhou, R., and Zhou, Z. (2019). Overview of the DESI Legacy Imaging Surveys. AJ, 157(5):168.

Diemer, B. (2018). COLOSSUS: A Python Toolkit for Cosmology, Large-scale Structure, and Dark Matter Halos. ApJS, 239(2):35.

Diemer, B. and Kravtsov, A. V. (2014). Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate. ApJ, 789(1):1.

Donahue, M. and Voit, G. M. (2022). Baryon cycles in the biggest galaxies. Phys. Rep., 973:1–109.

Dressler, A. (1980). Galaxy morphology in rich clusters - implications for the formation and evolution of galaxies. ApJ, 236:351–365.

Elek, O., Burchett, J. N., Prochaska, J. X., and Forbes, A. G. (2021a). Monte carlo physarum machine: Characteristics of pattern formation in continuous stochastic transport networks. Artificial Life.

Elek, O., Burchett, J. N., Prochaska, J. X., and Forbes, A. G. (2021b). Polyphorm: Structural analysis of cosmological datasets via interactive physarum polycephalum visualization. IEEE Transactions on Visualization and Computer Graphics, 27(2):806–816.

Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., Reddy, N. A., and Adelberger, K. L. (2006). The Stellar, Gas, and Dynamical Masses of Star-forming Galaxies at z $\sim$2. ApJ, 646(1):107–132.

Fielding, D., Quataert, E., McCourt, M., and Thompson, T. A. (2017). The impact of star formation feedback on the circumgalactic medium. MNRAS, 466(4):3810–3826.

Fielding, D. B., Ostriker, E. C., Bryan, G. L., and Jermyn, A. S. (2020). Multiphase Gas and the Fractal Nature of Radiative Turbulent Mixing Layers. ApJ, 894(2):L24.

Finlator, K. and Davé, R. (2008). The origin of the galaxy mass-metallicity relation and implications for galactic outflows. MNRAS, 385(4):2181–2204.

Ford, A. B., Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., and Weinberg, D. H. (2013). Hydrogen and metal line absorption around low-redshift galaxies in cosmological hydrodynamic simulations. MNRAS, 432(1):89–112.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). emcee: The mcmc hammer. PASP, 125:306–312.

Fraternali, F. and Binney, J. J. (2008). Accretion of gas on to nearby spiral galaxies. MNRAS, 386:935–944.

Fumagalli, M., Prochaska, J. X., Kasen, D., Dekel, A., Ceverino, D., and Primack, J. R. (2011). Absorption-line systems in simulated galaxies fed by cold streams. MNRAS, 418(3):1796–1821.

Geha, M., Blanton, M. R., Yan, R., and Tinker, J. L. (2012). A Stellar Mass Threshold for Quenching of Field Galaxies. ApJ, 757(1):85.

Gimeno, G., Roth, K., Chiboucas, K., Hibon, P., Boucher, L., White, J., Rippa, M., Labrie, K., Turner, J., Hanna, K., Lazo, M., Pérez, G., Rogers, R., Rojas, R., Placco, V., and Murowinski, R. (2016). On-sky commissioning of Hamamatsu CCDs in GMOS-S. In Evans, C. J., Simard, L., and Takami, H., editors, Ground-based and Airborne Instrumentation for Astronomy VI, volume 9908, pages 872 – 885. International Society for Optics and Photonics, SPIE.

Goh, T., Primack, J., Lee, C. T., Aragon-Calvo, M., Hellinger, D., Behroozi, P., Rodriguez-Puebla, A., Eckholm, E., and Johnston, K. (2019). Dark matter halo properties versus local density and cosmic web location. MNRAS, 483(2):2101–2122.

Green, J. C., Froning, C. S., Osterman, S., Ebbets, D., Heap, S. H., Leitherer, C., Linsky, J. L., Savage, B. D., Sembach, K., Shull, J. M., Siegmund, O. H. W., Snow, T. P., Spencer, J., Stern, S. A., Stocke, J., Welsh, B., Béland, S., Burgh, E. B., Danforth, C., France, K., Keeney, B., McPhate, J., Penton, S. V., Andrews, J., Brownsberger, K., Morse, J., and Wilkinson, E. (2012). The Cosmic Origins Spectrograph. ApJ, 744(1):60.

Haislmaier, K. J., Tripp, T. M., Katz, N., Prochaska, J. X., Burchett, J. N., O'Meara, J. M., and Werk, J. K. (2020). The COS Absorption Survey of Baryon Harbors: Unveiling the Physical Conditions of Circumgalactic Gas through Multiphase Bayesian Ionization Modeling. MNRAS.

Hennawi, J. F. and Prochaska, J. X. (2007). Quasars Probing Quasars. II. The Anisotropic Clustering of Optically Thick Absorbers around Quasars. ApJ, 655(2):735–748.

Ho, S., Cuesta, A., Seo, H.-J., de Putter, R., Ross, A. J., White, M., Padmanabhan, N., Saito, S., Schlegel, D. J., Schlafly, E., Seljak, U., Hernández-Monteagudo, C., Sánchez, A. G., Percival, W. J., Blanton, M., Skibba, R., Schneider, D., Reid, B., Mena, O., Viel, M., Eisenstein, D. J., Prada, F., Weaver, B. A., Bahcall, N., Bizyaev, D., Brewinton, H., Brinkman, J., Nicolaci da Costa, L., Gott, J. R., Malanushenko, E., Malanushenko, V., Nichol, B., Oravetz, D., Pan, K., Palanque-Delabrouille, N., Ross, N. P., Simmons, A., de Simoni, F., Snedden, S., and Yeche, C. (2012). Clustering of Sloan Digital Sky Survey III Photometric Luminous Galaxies: The Measurement, Systematics, and Cosmological Implications. ApJ, 761(1):14.

Holland, S. T. e. (2012). Cosmic Origins Spectrograph Instrument Handbook for Cycle 21 v. 5.0.

Hook, I. M., Jørgensen, I., Allington-Smith, J. R., Davies, R. L., Metcalfe, N., Murowinski, R. G., and Crampton, D. (2004). The Gemini-North Multi-Object Spectrograph: Performance in Imaging, Long-Slit, and Multi-Object Spectroscopic Modes. PASP, 116(819):425–440.

Huang, S., Katz, N., Scannapieco, E., Cottle, J., Davé, R., Weinberg, D. H., Peeples, M. S., and Brüggen, M. (2020). A new model for including galactic winds in simulations of galaxy formation - I. Introducing the Physically Evolved Winds (PhEW) model. MNRAS, 497(3):2586–2604.

Hummels, C. B., Bryan, G. L., Smith, B. D., and Turk, M. J. (2013). Constraints on hydrodynamical subgrid models from quasar absorption line studies of the simulated circumgalactic medium. MNRAS, 430(3):1548–1565.

Hummels, C. B., Smith, B. D., Hopkins, P. F., O'Shea, B. W., Silvia, D. W., Werk, J. K., Lehner, N., Wise, J. H., Collins, D. C., and Butsky, I. S. (2019). The Impact of Enhanced Halo Resolution on the Simulated Circumgalactic Medium. ApJ, 882(2):156.

Ji, S., Chan, T. K., Hummels, C. B., Hopkins, P. F., Stern, J., Kereš, D., Quataert, E., Faucher-Giguère, C.-A., and Murray, N. (2019). Properties of the Circumgalactic Medium in Cosmic Ray-Dominated Galaxy Halos. arXiv e-prints, page arXiv:1909.00003.

Johnson, S. D., Chen, H.-W., and Mulchaey, J. S. (2015). On the possible environmental effect in distributing heavy elements beyond individual gaseous haloes. MNRAS, 449(3):3263–3273.

Johnson, S. D., Chen, H.-W., Mulchaey, J. S., Schaye, J., and Straka, L. A. (2017). The Extent of Chemically Enriched Gas around Star-forming Dwarf Galaxies. ApJ, 850(1):L10.

Kauffmann, G., White, S. D. M., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, C., and Brinkmann, J. (2004). The environmental dependence of the relations between stellar mass, structure, star formation and nuclear activity in galaxies. MNRAS, 353:713–731.

Keeney, B. A., Stocke, J. T., Danforth, C. W., Shull, J. M., Pratt, C. T., Froning, C. S., Green, J. C., Penton, S. V., and Savage, B. D. (2017). Characterizing the Circumgalactic Medium of Nearby Galaxies with HST/COS and HST/STIS Absorption-line Spectroscopy. II. Methods and Models. ApJS, 230(1):6.

Keeney, B. A., Stocke, J. T., Pratt, C. T., Davis, J. D., Syphers, D., Danforth, C. W., Shull, J. M., Froning, C. S., Green, J. C., Penton, S. V., and Savage, B. D. (2018). A Galaxy Redshift Survey Near HST/COS AGN Sight Lines. ApJS, 237:11.

Kereš, D., Katz, N., Weinberg, D. H., and Davé, R. (2005). How do galaxies get their gas? MNRAS, 363(1):2–28.

Kim, C.-G. and Ostriker, E. C. (2018). Numerical Simulations of Multiphase Winds and Fountains from Star-forming Galactic Disks. I. Solar Neighborhood TIGRESS Model. ApJ, 853(2):173.

Kim, T. S., Wakker, B. P., Nasir, F., Carswell, R. F., Savage, B. D., Bolton, J. S., Fox, A. J., Viel, M., Haehnelt, M. G., Charlton, J. C., and Rosenwasser, B. E. (2021). The evolution of the low-density H I> intergalactic medium from z = 3.6 to 0: data, transmitted flux, and H I> column density,. MNRAS, 501(4):5811–5833.

Klypin, A., Yepes, G., Gottlöber, S., Prada, F., and Heß, S. (2016). MultiDark simulations: the story of dark matter halo concentrations and density profiles. MNRAS, 457(4):4340–4359.

Kuutma, T., Tamm, A., and Tempel, E. (2017). From voids to filaments: environmental transformations of galaxies in the SDSS. A&A, 600:L6.

Lan, T.-W., Ménard, B., and Mo, H. (2016). The galaxy luminosity function in groups and clusters: the faint-end upturn and the connection to the field luminosity function. MNRAS, 459:3998–4019.

Lanzetta, K. M., Bowen, D. V., Tytler, D., and Webb, J. K. (1995). The Gaseous Extent of Galaxies and the Origin of Lyman-Alpha Absorption Systems: A Survey of Galaxies in the Fields of Hubble Space Telescope Spectroscopic Target QSOs. ApJ, 442:538.

Lea, S. M., Silk, J., Kellogg, E., and Murray, S. (1973). Thermal-Bremsstrahlung Interpretation of Cluster X-Ray Sources. ApJ, 184:L105.

Lee, C. T., Primack, J. R., Behroozi, P., Rodríguez-Puebla, A., Hellinger, D., and Dekel, A. (2017). Properties of dark matter haloes as a function of local environment density. MNRAS, 466(4):3834–3858.

Lehner, N., Berek, S. C., Howk, J. C., Wakker, B. P., Tumlinson, J., Jenkins, E. B., Prochaska, J. X., Augustin, R., Ji, S., Faucher-Giguere, C.-A., Hafen, Z., Peeples, M. S., Barger, K. A., Berg, M. A., Bordoloi, R., Brown, T. M., Fox, A. J., Gilbert, K. M., Guhathakurta, P., Kalirai, J. S., Lockman, F. J., O'Meara, J. M., Pisano, D. J., Ribaudo, J., and Werk, J. K. (2020). Project AMIGA: The Circumgalactic Medium of Andromeda. arXiv e-prints, page arXiv:2002.07818.

Liang, L., Durier, F., Babul, A., Davé, R., Oppenheimer, B. D., Katz, N., Fardal, M., and Quinn, T. (2016). The growth and enrichment of intragroup gas. MNRAS, 456(4):4266–4290.

Libeskind, N. I., van de Weygaert, R., Cautun, M., Falck, B., Tempel, E., Abel, T., Alpaslan, M., Aragón-Calvo, M. A., Forero-Romero, J. E., Gonzalez, R., Gottlöber, S., Hahn, O., Hellwing, W. A., Hoffman, Y., Jones, B. J. T., Kitaura, F., Knebe, A., Manti, S., Neyrinck, M., Nuza, S. E., Padilla, N., Platen, E., Ramachandra, N., Robotham, A., Saar, E., Shandarin, S., Steinmetz, M., Stoica, R. S., Sousbie, T., and Yepes, G. (2018). Tracing the cosmic web. MNRAS, 473(1):1195–1217.

Lo Faro, B., Buat, V., Roehlly, Y., Alvarez-Marquez, J., Burgarella, D., Silva, L., and Efstathiou, A. (2017). Characterizing the UV-to-NIR shape of the dust attenuation curve of IR luminous galaxies up to z $\sim$ 2. MNRAS, 472(2):1372–1391.

Luber, N., van Gorkom, J. H., Hess, K. M., Pisano, D. J., Fernández, X., and Momjian, E. (2019). Large-scale Structure in CHILES Using DisPerSE. AJ, 157(6):254.

Maraston, C. and Strömbäck, G. (2011). Stellar population models at high spectral resolution. MNRAS, 418(4):2785–2811.

Martin, D. C., O'Sullivan, D., Matuszewski, M., Hamden, E., Dekel, A., Lapiner, S., Morrissey, P., Neill, J. D., Cantalupo, S., Prochaska, J. X., Steidel, C., Trainor, R., Moore, A., Ceverino, D., Primack, J., and Rizzi, L. (2019). Multi-filament gas inflows fuelling young star-forming galaxies. Nature Astronomy, page 356.

McQuinn, M. (2016). The evolution of the intergalactic medium. Annual Review of Astronomy and Astrophysics, 54(1):313–362.

McQuinn, M. and Werk, J. K. (2018). Implications of the Large O VI Columns around Low-redshift L $_*$ Galaxies. ApJ, 852(1):33.

Meiring, J. D., Tripp, T. M., Prochaska, J. X., Tumlinson, J., Werk, J., Jenkins, E. B., Thom, C., O'Meara, J. M., and Sembach, K. R. (2011). The First Observations of Low-redshift Damped Lyα Systems with the Cosmic Origins Spectrograph. ApJ, 732(1):35.

Molnar, S. M., Hearn, N., Haiman, Z., Bryan, G., Evrard, A. E., and Lake, G. (2009). Accretion Shocks in Clusters of Galaxies and Their SZ Signature from Cosmological Simulations. ApJ, 696:1640–1656.

Momose, R., Shimasaku, K., Kashikawa, N., Nagamine, K., Shimizu, I., Nakajima, K., Terao, Y., Kusakabe, H., Ando, M., Motohara, K., and Spitler, L. (2021). Environmental dependence of galactic properties traced by lyα forest absorption: Diversity among galaxy populations. The Astrophysical Journal, 909(2):117.

More, S., Diemer, B., and Kravtsov, A. V. (2015). The Splashback Radius as a Physical Halo Boundary and the Growth of Halo Mass. ApJ, 810(1):36.

Morris, S. L., Weymann, R. J., Dressler, A., McCarthy, P. J., Smith, B. A., Terrile, R. J., Giovanelli, R., and Irwin, M. (1993). The Environment of Lyman- alpha Absorbers in the Sight Line toward 3C 273. ApJ, 419:524.

Morton, D. C. (2003). Atomic Data for Resonance Absorption Lines. III. Wavelengths Longward of the Lyman Limit for the Elements Hydrogen to Gallium. ApJS, 149(1):205–238.

Moster, B. P., Naab, T., and White, S. D. M. (2013). Galactic star formation and accretion histories from matching galaxies to dark matter haloes. MNRAS, 428:3121–3138.

Münch, G. and Zirin, H. (1961). Interstellar Matter at Large Distances from the Galactic Plane. ApJ, 133:11.

Muzahid, S., Srianand, R., and Petitjean, P. (2011). Revisiting the He II to H I ratio in the intergalactic medium. MNRAS, 410(4):2193–2202.

Nelson, D., Pillepich, A., Springel, V., Pakmor, R., Weinberger, R., Genel, S., Torrey, P., Vogelsberger, M., Marinacci, F., and Hernquist, L. (2019). First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback. MNRAS, 490(3):3234–3261.

Nelson, D., Sharma, P., Pillepich, A., Springel, V., Pakmor, R., Weinberger, R., Vogelsberger, M., Marinacci, F., and Hernquist, L. (2020). Resolving small-scale cold circumgalactic gas in TNG50. arXiv e-prints, page arXiv:2005.09654.

Nielsen, N. M., Churchill, C. W., and Kacprzak, G. G. (2013). MAGIICAT II. General Characteristics of the Mg II Absorbing Circumgalactic Medium. ApJ, 776(2):115.

Noll, S., Burgarella, D., Giovannoli, E., Buat, V., Marcillac, D., and Muñoz-Mateos, J. C. (2009). Analysis of galaxy spectral energy distributions from far-UV to far-IR with CIGALE: studying a SINGS test sample. A&A, 507:1793–1813.

Oppenheimer, B. D., Crain, R. A., Schaye, J., Rahmati, A., Richings, A. J., Trayford, J. W., Tumlinson, J., Bower, R. G., Schaller, M., and Theuns, T. (2016). Bimodality of low-redshift circumgalactic O VI in non-equilibrium EAGLE zoom simulations. MNRAS, 460(2):2157–2179.

Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., and Weinberg, D. H. (2012). The intergalactic medium over the last 10 billion years - II. Metal-line absorption and physical conditions. MNRAS, 420(1):829–859.

Oppenheimer, B. D., Schaye, J., Crain, R. A., Werk, J. K., and Richings, A. J. (2018). The multiphase circumgalactic medium traced by low metal ions in EAGLE zoom simulations. MNRAS, 481(1):835–859.

Pasha, I., Mandelker, N., van den Bosch, F. C., Springel, V., and van de Voort, F. (2022). Quenching in Cosmic Sheets: Tracing the Impact of Large Scale Structure Collapse on the Evolution of Dwarf Galaxies. arXiv e-prints, page arXiv:2204.04097.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in python. J. Mach. Learn. Res., 12:2825–2830.

Peebles, P. J. E. (1980). Nature of the matter distribution now and at Z = 1000. Phys. Scr, 21(5):720–724.

Peeples, M., Behroozi, P., Bordoloi, R., Brooks, A., Bullock, J. S., Burchett, J. N., Chen, H.-W., Chisholm, J., Christensen, C., Coil, A., Corlies, L., Diamond-Stanic, A., Donahue, M., Faucher-Giguère, C.-A., Ferguson, H., Fielding, D., Fox, A. J., French, D. M., Furlanetto, S. R., Gennaro, M., Gilbert, K. M., Hamden, E., Hathi, N., Hayes, M., Henry, A., Howk, J. C., Hummels, C., Kereš, D., Kirby, E., Koekemoer, A. M., Lan, T.-W., Lanz, L., Law, D. R., Lehner, N., Lotz, J. M., McQuinn, K., McQuinn, M., Munshi, F., Oh, S. P., O'Meara, J. M., O'Shea, B. W., Pacifici, C., Peek, J. E. G., Postman, M., Prescott, M., Putman, M., Quataert, E., Rafelski, M., Ribaudo, J., Rowlands, K., Rubin, K., Salmon, B., Scarlata, C., Shapley, A. E., Simons, R., Snyder, G. F., Stern, J., Strom, A. L., Tollerud, E., Tremblay, G., Tripp, T. M., Tumlinson, J., Tuttle, S., Bosch, F. C. v. d., Voit, G. M., Wang, Q. D., Werk, J. K., Williams, B. F., Zaritsky, D., and Zheng, Y. (2019a). Understanding the circumgalactic medium is critical for understanding galaxy evolution. BAAS, 51(3):368.

Peeples, M. S., Corlies, L., Tumlinson, J., O'Shea, B. W., Lehner, N., O'Meara, J. M., Howk, J. C., Earl, N., Smith, B. D., Wise, J. H., and Hummels, C. B. (2019b). Figuring Out Gas & Galaxies in Enzo (FOGGIE). I. Resolving Simulated Circumgalactic Absorption at $2 \leq z \leq 2.5$. ApJ, 873(2):129.

Peeples, M. S., Werk, J. K., Tumlinson, J., Oppenheimer, B. D., Prochaska, J. X., Katz, N., and Weinberg, D. H. (2014). A Budget and Accounting of Metals at z $\sim$0: Results from the COS-Halos Survey. ApJ, 786(1):54.

Peng, Y.-j., Lilly, S. J., Kovač, K., Bolzonella, M., Pozzetti, L., Renzini, A., Zamorani, G., Ilbert, O., Knobel, C., Iovino, A., Maier, C., Cucciati, O., Tasca, L., Carollo, C. M., Silverman, J., Kampczyk, P., de Ravel, L., Sanders, D., Scoville, N., Contini, T., Mainieri, V., Scodeggio, M., Kneib, J.-P., Le Fèvre, O., Bardelli, S., Bongiorno, A., Caputi, K., Coppa, G., de la Torre, S., Franzetti, P., Garilli, B., Lamareille, F., Le Borgne, J.-F., Le Brun, V., Mignoli, M., Perez Montero, E., Pello, R., Ricciardelli, E., Tanaka, M., Tresse, L., Vergani, D., Welikala, N., Zucca, E., Oesch, P., Abbas, U., Barnes, L., Bordoloi, R., Bottini, D., Cappi, A., Cassata, P., Cimatti, A., Fumana, M., Hasinger, G., Koekemoer, A., Leauthaud, A., Maccagni, D., Marinoni, C., McCracken, H., Memeo, P., Meneux, B., Nair, P., Porciani, C., Presotto, V., and Scaramella, R. (2010). Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function. ApJ, 721(1):193–221.

Pillepich, A., Springel, V., Nelson, D., Genel, S., Naiman, J., Pakmor, R., Hernquist, L., Torrey, P., Vogelsberger, M., Weinberger, R., and Marinacci, F. (2018). Simulating galaxy formation with the IllustrisTNG model. MNRAS, 473(3):4077–4106.

Pisani, A., Sutter, P. M., Hamaus, N., Alizadeh, E., Biswas, R., Wandelt, B. D., and Hirata, C. M. (2015). Counting voids to probe dark energy. Phys. Rev. D, 92(8):083531.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Battaner, E., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bobin, J., Bock, J. J., Bonaldi, A., Bond, J. R., Borrill, J., Bouchet, F. R., Bridges, M., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cappellini, B., Cardoso, J. F., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chen, X., Chiang, H. C., Chiang, L. Y., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Delouis, J. M., Désert, F. X., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Gaier, T. C., Galeotta, S., Galli, S., Ganga, K., Giard, M., Giardino, G., Giraud-Héraud, Y., Gjerløw, E., González-Nuevo, J., Górski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Haissinski, J., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D., Henrot-Versillé, S., Hernández-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hou, Z., Hovest, W., Huffenberger, K. M., Jaffe, A. H., Jaffe, T. R., Jewell, J., Jones, W. C., Juvela, M., Keihänen, E., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J. M., Lasenby, A., Lattanzi, M., Laureijs, R. J., Lawrence, C. R., Leach, S., Leahy, J. P., Leonardi, R., León-Tavares, J., Lesgourgues, J., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Caniego, M., Lubin, P. M., Macías-Pérez, J. F., Maffei, B., Maino, D., Mandolesi, N., Maris, M., Marshall, D. J., Martin, P. G., Martínez-González, E., Masi, S., Massardi, M., Matarrese, S., Matthai, F., Mazzotta, P., Meinhold, P. R., Melchiorri, A., Melin, J. B., Mendes, L., Menegoni, E., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., O'Dwyer, I. J., Osborne, S., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paladini, R., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, D., Pearson, T. J., Peiris, H. V., Perdereau, O., Perotto, L., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Platania, P., Pointecouteau, E., Polenta, G., Ponthieu, N., Popa, L., Poutanen, T., Pratt, G. W., Prézeau, G., Prunet, S., Puget, J. L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Ricciardi, S., Riller, T., Ristorcelli, I., Rocha, G., Rosset, C., Roudier, G., Rowan-Robinson, M., Rubiño-Martín, J. A., Rusholme, B., Sandri, M., Santos, D., Savelainen, M., Savini, G., Scott, D., Seiffert, M. D., Shellard, E. P. S., Spencer, L. D., Starck, J. L., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sureau, F., Sutton, D., Suur-Uski, A. S., Sygnet, J. F., Tauber, J. A., Tavagnacco, D., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Tucci, M., Tuovinen, J., Türler, M., Umana, G., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Vittorio, N., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Wilkinson, A., Yvon, D., Zacchei, A., and Zonca, A. (2014). Planck 2013 results. XVI. Cosmological parameters. A&A, 571:A16.

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Catalano, A., Challinor, A., Chamballu, A., Chary, R. R., Chiang, H. C., Chluba, J., Christensen, P. R., Church, S., Clements, D. L., Colombi, S., Colombo, L. P. L., Combet, C., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Danese, L., Davies, R. D., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Désert, F. X., Di Valentino, E., Dickinson, C., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Doré, O., Douspis, M., Ducout, A., Dunkley, J., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Farhang, M., Fergusson, J., Finelli, F., Forni, O., Frailis, M., Fraisse, A. A., Franceschi, E., Frejsel, A., Galeotta, S., Galli, S., Ganga, K., Gauthier, C., Gerbino, M., Ghosh, T., Giard, M., Giraud-Héraud, Y., Giusarma, E., Gjerløw, E., González-Nuevo, J., Górski, K. M., Gratton, S., Gregorio, A., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Hansen, F. K., Hanson, D., Harrison, D. L., Helou, G., Henrot-Versillé, S., Hernández-Monteagudo, C., Herranz, D., Hildebrand t, S. R., Hivon, E., Hobson, M., Holmes, W. A., Hornstrup, A., Hovest, W., Huang, Z., Huffenberger, K. M., Hurier, G., Jaffe, A. H., Jaffe, T. R., Jones, W. C., Juvela, M., Keihänen, E., Keskitalo, R., Kisner, T. S., Kneissl, R., Knoche, J., Knox, L., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Leahy, J. P., Leonardi, R., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Linden-Vørnle, M., López-Caniego, M., Lubin, P. M., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marchini, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Masi, S., Matarrese, S., McGehee, P., Meinhold, P. R., Melchiorri, A., Melin, J. B., Mendes, L., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Nørgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Oxborrow, C. A., Paci, F., Pagano, L., Pajot, F., Paladini, R., Paoletti, D., Partridge, B., Pasian, F., Patanchon, G., Pearson, T. J., Perdereau, O., Perotto, L., Perrotta, F., Pettorino, V., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Popa, L., Pratt, G. W., Prézeau, G., Prunet, S., Puget, J. L., Rachen, J. P., Reach, W. T., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Ristorcelli, I., Rocha, G., Rosset, C., Rossetti, M., Roudier, G., Rouillé d'Orfeuil, B., Rowan-Robinson, M., Rubiño-Martín, J. A., Rusholme, B., Said, N., Salvatelli, V., Salvati, L., Sandri, M., Santos, D., Savelainen, M., Savini, G., Scott, D., Seiffert, M. D., Serra, P., Shellard, E. P. S., Spencer, L. D., Spinelli, M., Stolyarov, V., Stompor, R., Sudiwala, R., Sunyaev, R., Sutton, D., Suur-Uski, A. S., Sygnet, J. F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Trombetti, T., Tucci, M., Tuovinen, J., Türler, M., Umana, G., Valenziano, L., Valiviita, J., Van Tent, F., Vielva, P., Villa, F., Wade, L. A., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Wilkinson, A., Yvon, D., Zacchei, A., and Zonca, A. (2016). Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13.

Pratt, C. T., Stocke, J. T., Keeney, B. A., and Danforth, C. W. (2018). The Spread of

Metals into the Low-redshift Intergalactic Medium. ApJ, 855:18.

Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Paper Contributors, P., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., Coordination Committee, A., Ardelean, C., Babej, T., Bach, Y. P., Bachetti, M., Bakanov, A. V., Bamford, S. P., Barentsen, G., Barmby, P., Baumbach, A., Berry, K. L., Biscani, F., Boquien, M., Bostroem, K. A., Bouma, L. G., Brammer, G. B., Bray, E. M., Breytenbach, H., Buddelmeijer, H., Burke, D. J., Calderone, G., Cano Rodríguez, J. L., Cara, M., Cardoso, J. V. M., Cheedella, S., Copin, Y., Corrales, L., Crichton, D., D'Avella, D., Deil, C., Depagne, É., Dietrich, J. P., Donath, A., Droettboom, M., Earl, N., Erben, T., Fabbro, S., Ferreira, L. A., Finethy, T., Fox, R. T., Garrison, L. H., Gibbons, S. L. J., Goldstein, D. A., Gommers, R., Greco, J. P., Greenfield, P., Groener, A. M., Grollier, F., Hagen, A., Hirst, P., Homeier, D., Horton, A. J., Hosseinzadeh, G., Hu, L., Hunkeler, J. S., Ivezić, Ž., Jain, A., Jenness, T., Kanarek, G., Kendrew, S., Kern, N. S., Kerzendorf, W. E., Khvalko, A., King, J., Kirkby, D., Kulkarni, A. M., Kumar, A., Lee, A., Lenz, D., Littlefair, S. P., Ma, Z., Macleod, D. M., Mastropietro, M., McCully, C., Montagnac, S., Morris, B. M., Mueller, M., Mumford, S. J., Muna, D., Murphy, N. A., Nelson, S., Nguyen, G. H., Ninan, J. P., Nöthe, M., Ogaz, S., Oh, S., Parejko, J. K., Parley, N., Pascual, S., Patil, R., Patil, A. A., Plunkett, A. L., Prochaska, J. X., Rastogi, T., Reddy Janga, V., Sabater, J., Sakurikar, P., Seifert, M., Sherbert, L. E., Sherwood-Taylor, H., Shih, A. Y., Sick, J., Silbiger, M. T., Singanamalla, S., Singer, L. P., Sladen, P. H., Sooley, K. A., Sornarajah, S., Streicher, O., Teuben, P., Thomas, S. W., Tremblay, G. R., Turner, J. E. H., Terrón, V., van Kerkwijk, M. H., de la Vega, A., Watkins, L. L., Weaver, B. A., Whitmore, J. B., Woillez, J., Zabalza, V., and Contributors, A. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. AJ, 156:123.

Prochaska, J. X., Burchett, J. N., Tripp, T. M., Werk, J. K., Willmer, C. N. A., Howk,

J. C., Lange, S., Tejos, N., Meiring, J. D., Tumlinson, J., Lehner, N., Ford, A. B., and Davé, R. (2019). The COS absorption survey of baryon harbors: The galaxy database and cross-correlation analysis of o vi systems. The Astrophysical Journal Supplement Series, 243(2):24.

Prochaska, J. X., Hennawi, J. F., Lee, K.-G., Cantalupo, S., Bovy, J., Djorgovski, S. G., Ellison, S. L., Lau, M. W., Martin, C. L., Myers, A., Rubin, K. H. R., and Simcoe, R. A. (2013). Quasars Probing Quasars. VI. Excess H I Absorption within One Proper Mpc of z ‿2 Quasars. ApJ, 776(2):136.

Prochaska, J. X., Hennawi, J. F., Westfall, K. B., Cooke, R. J., Wang, F., Hsyu, T., Davies, F. B., and Farina, E. P. (2020). PypeIt: The Python Spectroscopic Data Reduction Pipeline. arXiv e-prints, page arXiv:2005.06505.

Prochaska, J. X., Weiner, B., Chen, H.-W., Mulchaey, J., and Cooksey, K. (2011). Probing the Intergalactic Medium/Galaxy Connection. V. On the Origin of Lyα and O VI Absorption at z 0.2. ApJ, 740:91.

Prochaska, J. X., Werk, J. K., Worseck, G., Tripp, T. M., Tumlinson, J., Burchett, J. N., Fox, A. J., Fumagalli, M., Lehner, N., Peeples, M. S., and Tejos, N. (2017). The COS-Halos Survey: Metallicities in the Low-redshift Circumgalactic Medium. ApJ, 837(2):169.

Putman, M. E., Zheng, Y., Price-Whelan, A. M., Grcevich, J., Johnson, A. C., Tollerud, E., and Peek, J. E. G. (2021). The Gas Content and Stripping of Local Group Dwarf Galaxies. ApJ, 913(1):53.

Rauch, M. (1998). The lyman alpha forest in the spectra of quasistellar objects. Annual Review of Astronomy and Astrophysics, 36(1):267–316.

Reid, B., Ho, S., Padmanabhan, N., Percival, W. J., Tinker, J., Tojeiro, R., White, M., Eisenstein, D. J., Maraston, C., Ross, A. J., Sánchez, A. G., Schlegel, D., Sheldon, E., Strauss, M. A., Thomas, D., Wake, D., Beutler, F., Bizyaev, D., Bolton, A. S., Brownstein,

J. R., Chuang, C.-H., Dawson, K., Harding, P., Kitaura, F.-S., Leauthaud, A., Masters, K., McBride, C. K., More, S., Olmstead, M. D., Oravetz, D., Nuza, S. E., Pan, K., Parejko, J., Pforr, J., Prada, F., Rodríguez-Torres, S., Salazar-Albornoz, S., Samushia, L., Schneider, D. P., Scóccola, C. G., Simmons, A., and Vargas-Magana, M. (2016). SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large-scale structure catalogues. MNRAS, 455(2):1553–1573.

Rodrigo, C., Solano, E., and Bayo, A. (2012). SVO Filter Profile Service Version 1.0. Technical report.

Rodríguez-Puebla, A., Behroozi, P., Primack, J., Klypin, A., Lee, C., and Hellinger, D. (2016). Halo and subhalo demographics with Planck cosmological parameters: Bolshoi-Planck and MultiDark-Planck simulations. MNRAS, 462(1):893–916.

Ross, A. J., Ho, S., Cuesta, A. J., Tojeiro, R., Percival, W. J., Wake, D., Masters, K. L., Nichol, R. C., Myers, A. D., de Simoni, F., Seo, H. J., Hernández-Monteagudo, C., Crittenden, R., Blanton, M., Brinkmann, J., da Costa, L. A. N., Guo, H., Kazin, E., Maia, M. A. G., Maraston, C., Padmanabhan, N., Prada, F., Ramos, B., Sanchez, A., Schlafly, E. F., Schlegel, D. J., Schneider, D. P., Skibba, R., Thomas, D., Weaver, B. A., White, M., and Zehavi, I. (2011). Ameliorating systematic uncertainties in the angular clustering of galaxies: a study using the SDSS-III. MNRAS, 417(2):1350–1373.

Ross, A. J., Percival, W. J., Sánchez, A. G., Samushia, L., Ho, S., Kazin, E., Manera, M., Reid, B., White, M., Tojeiro, R., McBride, C. K., Xu, X., Wake, D. A., Strauss, M. A., Montesano, F., Swanson, M. E. C., Bailey, S., Bolton, A. S., Dorta, A. M., Eisenstein, D. J., Guo, H., Hamilton, J.-C., Nichol, R. C., Padmanabhan, N., Prada, F., Schlegel, D. J., Magaña, M. V., Zehavi, I., Blanton, M., Bizyaev, D., Brewington, H., Cuesta, A. J., Malanushenko, E., Malanushenko, V., Oravetz, D., Parejko, J., Pan, K., Schneider, D. P., Shelden, A., Simmons, A., Snedden, S., and Zhao, G.-b. (2012). The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: analysis of potential systematics. MNRAS, 424(1):564–590.

Rubin, V. C., Ford, W. K., J., and Thonnard, N. (1980). Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). ApJ, 238:471–487.

Rudie, G. C., Steidel, C. C., Pettini, M., Trainor, R. F., Strom, A. L., Hummels, C. B., Reddy, N. A., and Shapley, A. E. (2019). The Column Density, Kinematics, and Thermal State of Metal-Bearing Gas within the Virial Radius of z$\sim$2 Star-Forming Galaxies in the Keck Baryonic Structure Survey. arXiv e-prints, page arXiv:1903.00004.

Rudie, G. C., Steidel, C. C., Trainor, R. F., Rakic, O., Bogosavljević, M., Pettini, M., Reddy, N., Shapley, A. E., Erb, D. K., and Law, D. R. (2012). THE GASEOUS ENVIRONMENT OF HIGH-zGALAXIES: PRECISION MEASUREMENTS OF NEUTRAL HYDROGEN IN THE CIRCUMGALACTIC MEDIUM OFz$\sim$ 2-3 GALAXIES IN THE KECK BARYONIC STRUCTURE SURVEY. The Astrophysical Journal, 750(1):67.

Ryan-Weber, E. V. (2006). Cross-correlation of Lyman α absorbers with gas-rich galaxies. MNRAS, 367(3):1251–1260.

Sahnow, D. J., Oliveira, C., Aloisi, A., Hodge, P. E., Massa, D., Osten, R., Proffitt, C., Bostroem, A., McPhate, J. B., Béland, S., Osterman, S. N., and Penton, S. V. (2011). Gain sag in the FUV detector of the Cosmic Origins Spectrograph. Instrument Science Report COS 2011-05.

Salim, S., Lee, J. C., Janowiecki, S., da Cunha, E., Dickinson, M., Boquien, M., Burgarella, D., Salzer, J. J., and Charlot, S. (2016). GALEX-SDSS-WISE Legacy Catalog (GSWLC): Star Formation Rates, Stellar Masses, and Dust Attenuations of 700,000 Low-redshift Galaxies. ApJS, 227(1):2.

Sargent, W. L. W., Young, P. J., Boksenberg, A., and Tytler, D. (1980). The distribution of Lyman-alpha absorption lines in the spectra of six QSOs: evidence for an intergalactic origin. ApJS, 42:41–81.

Savage, B. D. and Sembach, K. R. (1991). The Analysis of Apparent Optical Depth Profiles for Interstellar Absorption Lines. ApJ, 379:245.

Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D. M., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., and Trayford, J. (2015). The EAGLE project: simulating the evolution and assembly of galaxies and their environments. MNRAS, 446:521–554.

Schaye, J., Crain, R. A., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., Jenkins, A., Rosas-Guevara, Y. M., White, S. D. M., Baes, M., Booth, C. M., Camps, P., Navarro, J. F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P. A., and Trayford, J. (2015). The EAGLE project: simulating the evolution and assembly of galaxies and their environments. MNRAS, 446(1):521–554.

Schlafly, E. F. and Finkbeiner, D. P. (2011). MEASURING REDDENING WITH SLOAN DIGITAL SKY SURVEY STELLAR SPECTRA AND RECALIBRATING SFD. The Astrophysical Journal, 737(2):103.

Schreiber, C., Pannella, M., Elbaz, D., Béthermin, M., Inami, H., Dickinson, M., Magnelli, B., Wang, T., Aussel, H., Daddi, E., Juneau, S., Shu, X., Sargent, M. T., Buat, V., Faber, S. M., Ferguson, H. C., Giavalisco, M., Koekemoer, A. M., Magdis, G., Morrison, G. E., Papovich, C., Santini, P., and Scott, D. (2015). The Herschel view of the dominant mode of galaxy growth from z = 4 to the present day. A&A, 575:A74.

Shen, S., Madau, P., Aguirre, A., Guedes, J., Mayer, L., and Wadsley, J. (2012). The Origin of Metals in the Circumgalactic Medium of Massive Galaxies at z = 3. ApJ, 760(1):50.

Shull, J. M., Smith, B. D., and Danforth, C. W. (2012). The Baryon Census in a Multiphase Intergalactic Medium: 30% of the Baryons May Still be Missing. ApJ, 759(1):23.

Simha, S., Burchett, J. N., Prochaska, J. X., Chittidi, J. S., Elek, O., Tejos, N., Jorgenson, R., Bannister, K. W., Bhandari, S., Day, C. K., Deller, A. T., Forbes, A. G., Macquart, J.-P., Ryder, S. D., and Shannon, R. M. (2020). Disentangling the Cosmic Web toward FRB 190608. ApJ, 901(2):134.

Singh, P., Voit, G. M., and Nath, B. B. (2021). Constraints on precipitation-limited hot haloes from massive galaxies to galaxy clusters. MNRAS, 501(2):2467–2477.

Stern, J., Fielding, D., Faucher-Giguère, C.-A., and Quataert, E. (2019). Cooling flow solutions for the circumgalactic medium. MNRAS, 488(2):2549–2572.

Stinson, G. S., Brook, C., Prochaska, J. X., Hennawi, J., Shen, S., Wadsley, J., Pontzen, A., Couchman, H. M. P., Quinn, T., Macciò, A. V., and Gibson, B. K. (2012). MAGICC haloes: confronting simulations with observations of the circumgalactic medium at z=0. MNRAS, 425(2):1270–1277.

Stocke, J. T., Keeney, B. A., Danforth, C. W., Shull, J. M., Froning, C. S., Green, J. C., Penton, S. V., and Savage, B. D. (2013). Characterizing the Circumgalactic Medium of Nearby Galaxies with HST/COS and HST/STIS Absorption-line Spectroscopy. ApJ, 763(2):148.

Tchernyshyov, K., Werk, J. K., Wilde, M. C., Prochaska, J. X., Tripp, T. M., Burchett, J. N., Bordoloi, R., Howk, J. C., Lehner, N., O'Meara, J. M., Tejos, N., and Tumlinson, J. (2022). The CGM^2 Survey: Circumgalactic O VI from Dwarf to Massive Star-forming Galaxies. ApJ, 927(2):147.

Tejos, N., Morris, S. L., Finn, C. W., Crighton, N. H. M., Bechtold, J., Jannuzi, B. T., Schaye, J., Theuns, T., Altay, G., Le Fèvre, O., Ryan-Weber, E., and Davé, R. (2014).

On the connection between the intergalactic medium and galaxies: the H I-galaxy cross-correlation at z 1. MNRAS, 437:2017–2075.

Tempel, E., Stoica, R. S., Martínez, V. J., Liivamägi, L. J., Castellan, G., and Saar, E. (2014). Detecting filamentary pattern in the cosmic web: a catalogue of filaments for the SDSS. MNRAS, 438:3465–3482.

Tepper-García, T., Richter, P., Schaye, J., Booth, C. M., Dalla Vecchia, C., and Theuns, T. (2012). Absorption signatures of warm-hot gas at low redshift: broad H I Ly absorbers. MNRAS, 425:1640–1663.

Thom, C., Tumlinson, J., Werk, J. K., Prochaska, J. X., Oppenheimer, B. D., Peeples, M. S., Tripp, T. M., Katz, N. S., OMeara, J. M., Ford, A. B., Davé, R., Sembach, K. R., and Weinberg, D. H. (2012). NOT DEAD YET: COOL CIRCUMGALACTIC GAS IN THE HALOS OF EARLY-TYPE GALAXIES. The Astrophysical Journal, 758(2):L41.

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., Klypin, A., Warren, M. S., Yepes, G., and Gottlöber, S. (2010). The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. ApJ, 724(2):878–886.

Tolman, R. C. (1934). Effect of inhomogeneity on cosmological models. Proceedings of the National Academy of Sciences, 20(3):169–176.

Tripp, T. M., Lu, L., and Savage, B. D. (1998). The Relationship between Galaxies and Low-Redshift Weak Lyα Absorbers in the Directions of H1821+643 and PG 1116+215. ApJ, 508(1):200–231.

Tripp, T. M., Meiring, J. D., Prochaska, J. X., Willmer, C. N. A., Howk, J. C., Werk, J. K., Jenkins, E. B., Bowen, D. V., Lehner, N., Sembach, K. R., Thom, C., and Tumlinson, J. (2011). The Hidden Mass and Large Spatial Extent of a Post-Starburst Galaxy Outflow. Science, 334(6058):952.

Tripp, T. M., Sembach, K. R., Bowen, D. V., Savage, B. D., Jenkins, E. B., Lehner, N., and Richter, P. (2008). A High-Resolution Survey of Low-Redshift QSO Absorption Lines: Statistics and Physical Conditions of O VI Absorbers. ApJS, 177(1):39–102.

Tully, R. B., Courtois, H. M., and Sorce, J. G. (2016). Cosmicflows-3. AJ, 152(2):50.

Tumlinson, J. and Fang, T. (2005). Hot Baryons and the Distribution of Metals in the Intergalactic Medium. ApJ, 623(2):L97–L100.

Tumlinson, J., Peeples, M. S., and Werk, J. K. (2017). The circumgalactic medium. Annual Review of Astronomy and Astrophysics, 55(1):389–432.

Tumlinson, J., Thom, C., Werk, J. K., Prochaska, J. X., Tripp, T. M., Katz, N., Davé, R., Oppenheimer, B. D., Meiring, J. D., Ford, A. B., O'Meara, J. M., Peeples, M. S., Sembach, K. R., and Weinberg, D. H. (2013). The COS-Halos Survey: Rationale, Design, and a Census of Circumgalactic Neutral Hydrogen. ApJ, 777(1):59.

Tumlinson, J., Werk, J. K., Thom, C., Meiring, J. D., Prochaska, J. X., Tripp, T. M., O'Meara, J. M., Okrochkov, M., and Sembach, K. R. (2011). Multiphase Gas in Galaxy Halos: The O VI Lyman-limit System toward J1009+0713. ApJ, 733(2):111.

van de Voort, F., Springel, V., Mandelker, N., van den Bosch, F. C., and Pakmor, R. (2019). Cosmological simulations of the circumgalactic medium with 1 kpc resolution: enhanced H I column densities. MNRAS, 482(1):L85–L89.

Villaescusa-Navarro, F., Genel, S., Anglés-Alcázar, D., Thiele, L., Dave, R., Narayanan, D., Nicola, A., Li, Y., Villanueva-Domingo, P., Wandelt, B., Spergel, D. N., Somerville, R. S., Zorrilla Matilla, J. M., Mohammad, F. G., Hassan, S., Shao, H., Wadekar, D., Eickenberg, M., Wong, K. W. K., Contardo, G., Jo, Y., Moser, E., Lau, E. T., Machado Poletti Valle, L. F., Perez, L. A., Nagai, D., Battaglia, N., and Vogelsberger, M. (2022). The CAMELS Multifield Data Set: Learning the Universe's Fundamental Parameters with Artificial Intelligence. ApJS, 259(2):61.

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., and Hernquist, L. (2014). Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. MNRAS, 444(2):1518–1547.

Voit, G. M. (2019). Ambient Column Densities of Highly Ionized Oxygen in Precipitation-limited Circumgalactic Media. ApJ, 880(2):139.

Wakker, B. P., Hernandez, A. K., French, D. M., Kim, T.-S., Oppenheimer, B. D., and Savage, B. D. (2015). Nearby Galaxy Filaments and the Ly-alpha Forest: Confronting Simulations and the UV Background with Observations. ApJ, 814(1):40.

Wakker, B. P. and Savage, B. D. (2009). The Relationship Between Intergalactic H I/O VI and Nearby (z 0.017) Galaxies. ApJS, 182:378–467.

Werk, J. K., Prochaska, J. X., Thom, C., Tumlinson, J., Tripp, T. M., O'Meara, J. M., and Meiring, J. D. (2012). The COS-Halos Survey: Keck LRIS and Magellan MagE Optical Spectroscopy. ApJS, 198(1):3.

Werk, J. K., Prochaska, J. X., Thom, C., Tumlinson, J., Tripp, T. M., O'Meara, J. M., and Peeples, M. S. (2013). The COS-Halos Survey: An Empirical Description of Metal-line Absorption in the Low-redshift Circumgalactic Medium. ApJS, 204(2):17.

Werk, J. K., Prochaska, J. X., Tumlinson, J., Peeples, M. S., Tripp, T. M., Fox, A. J., Lehner, N., Thom, C., O'Meara, J. M., Ford, A. B., Bordoloi, R., Katz, N., Tejos, N., Oppenheimer, B. D., Davé, R., and Weinberg, D. H. (2014). The COS-Halos Survey: Physical Conditions and Baryonic Mass in the Low-redshift Circumgalactic Medium. ApJ, 792(1):8.

White, S. D. M. and Rees, M. J. (1978). Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering. MNRAS, 183:341–358.

Wilde, M. C., Werk, J. K., Burchett, J. N., Prochaska, J. X., Tchernyshyov, K., Tripp, T. M., Tejos, N., Lehner, N., Bordoloi, R., O'Meara, J. M., and Tumlinson, J. (2021). CGM^2 I: The Extent of the Circumgalactic Medium Traced by Neutral Hydrogen. ApJ, 912(1):9.

Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D., and Parikh, T. (2017). FIREFLY (Fitting IteRativEly For Likelihood analYsis): a full spectral fitting code. MNRAS, 472(4):4297–4326.

Woodgate, B. E., Kimble, R. A., Bowers, C. W., Kraemer, S., Kaiser, M. E., Danks, A. C., Grady, J. F., Loiacono, J. J., Brumfield, M., Feinberg, L., Gull, T. R., Heap, S. R., Maran, S. P., Lindler, D., Hood, D., Meyer, W., Vanhouten, C., Argabright, V., Franka, S., Bybee, R., Dorn, D., Bottema, M., Woodruff, R., Michika, D., Sullivan, J., Hetlinger, J., Ludtke, C., Stocker, R., Delamere, A., Rose, D., Becker, I., Garner, H., Timothy, J. G., Blouke, M., Joseph, C. L., Hartig, G., Green, R. F., Jenkins, E. B., Linsky, J. L., Hutchings, J. B., Moos, H. W., Boggess, A., Roesler, F., and Weistrop, D. (1998). The Space Telescope Imaging Spectrograph Design. PASP, 110(752):1183–1204.

Yang, X., Mo, H. J., van den Bosch, F. C., Pasquali, A., Li, C., and Barden, M. (2007). Galaxy groups in the SDSS DR4. i. the catalog and basic properties. ApJ, 671:153–170.

Zehavi, I., Blanton, M. R., Frieman, J. A., Weinberg, D. H., Mo, H. J., Strauss, M. A., Anderson, S. F., Annis, J., Bahcall, N. A., Bernardi, M., Briggs, J. W., Brinkmann, J., Burles, S., Carey, L., Castander, F. J., Connolly, A. J., Csabai, I., Dalcanton, J. J., Dodelson, S., Doi, M., Eisenstein, D., Evans, M. L., Finkbeiner, D. P., Friedman, S., Fukugita, M., Gunn, J. E., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Kent, S., Knapp, G. R., Kron, R., Kunszt, P., Lamb, D. Q., Leger, R. F., Long, D. C., Loveday, J., Lupton, R. H., McKay, T., Meiksin, A., Merrelli, A., Munn, J. A., Narayanan, V., Newcomb, M., Nichol, R. C., Owen, R., Peoples, J., Pope, A., Rockosi, C. M., Schlegel, D., Schneider, D. P., Scoccimarro, R., Sheth, R. K., Siegmund, W., Smee, S., Snir, Y., Stebbins, A., Stoughton, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Tegmark, M., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Yanny, B., and York, D. G. (2002). Galaxy Clustering in Early Sloan Digital Sky Survey Redshift Data. ApJ, 571(1):172–190.