Probing cool and warm circumgalactic gas in galaxies and clusters with large spectroscopic and imaging surveys

Abhijeet Anand

München 2022

Probing cool and warm circumgalactic gas in galaxies and clusters with large spectroscopic and imaging surveys

Abhijeet Anand

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Abhijeet Anand
aus Patna, Indien

München, den 30 May 2022

Erstgutachter: Prof. Dr. Guinevere Kauffmann
Zweitgutachter: Dr. Habil. Ariel Sánchez
Tag der mündlichen Prüfung: 19 July 2022

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन |
मा कर्मफलहेतुर्भुर्मा ते संगोऽस्त्वकर्मणि ||

|| Karmanye vadhika raste, Ma phaleshu kadachana ||
|| ma karma phala he tur bhuh, ma te sangvasta karmani ||

- Bhagavad Gita: Chapter 2, Verse 47

"To action alone hast thou a right and never at all to its fruits; let not the fruits of action be thy motive; neither let there be in thee any attachment to inaction."

To my Grandparents

and

To my family and country

Contents

List of Figures

List of Tables

List of Acronyms

A&A	Astronomy and Astrophysics
AJ	Astronomical Journal
ApJ	Astrophysical Journal
ApJL	Astrophysical Journal Letters
ApJS	Astrophysical Journal Supplement
ARA&A	Annual Review of Astronomy and Astrophysics
AGN	Active Galactic Nuclei
BCG	Brightest Cluster Galaxy
BOSS	Baryon Oscillation Spectroscopic Survey
CGM	Circumgalactic Medium
CMB	Cosmic Microwave Background
COS	Cosmic Origins Spectrograph
DESI	Dark Energy Spectroscopic Instrument
DR	Data Release
EAGLE	Evolution and Assembly of GaLaxies and their Environments
eBOSS	extended Baryon Oscillation Spectroscopic Survey
ELGs	Emission Line Galaxies
eROSITA	extended ROentgen Survey with an Imaging Telescope Array
EW	Equivalent Width
FUSE	Far Ultraviolet Spectroscopic Explorer
HETDEX	Hobby-Eberly Telescope Dark Energy Experiment
HST	Hubble Space Telescope
IllustrisTNG	Illustris The Next Generation
ICM	Intracluster Medium
IFU	Integral Field Unit
IGM	Intergalactic Medium
ISM	Interstellar Medium
JHU	Johns Hopkins University
JWST	James Webb Space Telescope
KCWI	Keck Cosmic Web Imager Integral Field Spectrograph
ΛCDM	Lambda Cold Dark Matter
LRGs	Luminous Red Galaxies
MNRAS	Monthly Notices of the Royal Astronomical Society

MPA	Max Planck Institute for Astrophysics
MUSE	Multi Unit Spectroscopic Explorer
NMF	Nonnegative Matrix Facorization
PCA	Principal Component Analysis
PFS	Subaru Prime Focus Spectrograph
PNAS	Proceedings of the National Academy of Sciences
PRL	Physical Reviews Letters
QSO	Quasi Stellar Object
ROSAT	Röntgen Satellite
SDSS	Sloan Digital Sky Survey
SFR	Star Formation Rate
sSFR	specific Star Formation Rate
SMBH	Supermassive Black Hole
UV	Ultraviolet
VLT	Very Large Telescope
WHT	William Herschel Telescope
WEAVE	WHT Enhanced Area Velocity Explorer
WHIM	Warm Hot Intergalactic Medium
WMAP	Wilkinson Microwave Anisotropy Probe
XMM	X-ray Multi-Mirror Mission

Zusammenfassung

Die gasförmige Atmosphäre um Galaxien wird als zirkumgalaktisches Medium (CGM) bezeichnet und stellt die Schnittstelle zwischen den physikalischen Prozessen auf kleinen und großen Skalen dar. Das komplexe Gleichgewicht zwischen diesen Prozessen bestimmt die Entstehung und Entwicklung von Galaxien. Einerseits wird kühles, metallarmes Gas aus dem intergalaktischen Medium (IGM) oder den kosmischen Filamenten in die Galaxien eingebracht, andererseits stoßen galaktische Ausströme von Supernovae oder dem aktiven Galaxienkern (AGN) eine große Menge an Gas und Metallen in das CGM aus, von denen ein Teil sogar aus dem Halo entweichen kann, während ein anderer Teil recycelt, vermischt und in die Galaxie zurückgeführt wird. Diese Ströme spielen eine entscheidende Rolle bei der Regulierung von Galaxienentstehung: sie bestimmen die Zeitskalen auf denen Gasschwund und Sternentstehung ablaufen, die Prozesse, die Sternentstehung unterdrücken und die Verteilung des Baryonen- und Metallhaushalts im Gas um Galaxien. Das zirkumgalaktische Gas wird in der Regel durch Absorption (von Wasserstoff oder Metallen) vor einer leuchtenden Hintergrundquelle wie einem Quasar untersucht und kann so tiefe Einblicke in die Gasströme und deren Zusammenhang mit den Eigenschaften der Galaxie und Umgebung liefern.

In dieser Arbeit charakterisiere ich die Eigenschaften des kühlen und warmen zirkumgalaktischen Gases von Galaxien, Galaxienhaufen und Quasaren mit Hilfe von Quasarabsorbern. Um das zirkumgalaktische Gas zu untersuchen, habe ich zunächst eine vollautomatische Pipeline zur Modellierung der Quasar-Kontinuumsemission entwickelt. Dann setzte ich eine angepasste Kernel-Faltungstechnik in Kombination mit einem adaptiven Signal-zu-Rausch-Kriterium ein, um dazwischen liegende Metallabsorber in ihren Spektren zu erkennen. Das Ergebnis der Pipeline ist der bisher umfangreichste Katalog von Metallabsorbern. Dieser bot die Möglichkeit, eine der umfangreichsten Studien der Absorptionslinien der galaktischen Atmosphären um verschiedene Galaxientypen herum durchzuführen. Anschließend kombiniere ich diese Absorber mit großen Stichproben spektroskopischer Aufnahmen von Galaxien und untersuche die Beschaffenheit des kalten zirkumgalaktischen Mediums bei $z \sim 0,5$. Dank der erheblichen Stichprobengröße konnte ich die Skalen der Kaltgasabsorption (verfolgt durch Mg II , einfach ionisiertes Magnesium) im CGM von sternbildenden Galaxien charakterisieren. Ich habe herausgefunden, dass die Absorption von kaltem Gas in ihrem CGM innerhalb von ~ 50 kpc von der Galaxie am stärksten ist, und dass die Häufigkeit von Absorbern in den inneren Teilen des Halos bei sternbildenden Galaxien im Vergleich zu ruhigen Galaxien bei ähnlichen Rotverschiebungen $2 - 4$ mal

höher ist. Jenseits von 50 kpc ist andererseits ein starker Rückgang der Absorptionsstärke des kalten Gases und der Detektionsrate für beide Arten von Galaxien zu beobachten, was auf eine Dichotomie der physikalischen Bedingungen im Gas hinweist. Die inneren Regionen des CGM werden hauptsächlich durch galaktische Ausströme reguliert, während das äußere CGM eng mit dem Halo der dunklen Materie verbunden ist. Galaxien mit einer höheren Sternentstehungsrate (SFR) weisen ein höheres Vorkommen von kaltem Gas auf, was auf eine enge Verbindung zu den stellaren Ausströmen schließen lässt. Andererseits stelle ich fest, dass die Bewegung des absorbierenden Gases um passive Galaxien subvirial verläuft, was bedeutet, dass das Gas gravitativ gebunden ist und wahrscheinlich nicht aus galaktischen Ausströmen stammt. Ihre geringe Sternentstehungsaktivität unterstützt diese Beobachtung zusätzlich. Daher ist der Ursprung des kühlen Gases um passive Galaxien möglicherweise auf die Akkretion aus dem intergalaktischen Medium (IGM) oder auf das Strippen des Gases von Satellitengalaxien zurückzuführen. Diese Analyse deutet darauf hin, dass der physikalische Ursprung des kühlen zirkumgalaktischen Mediums bei sternbildenden und ruhigen Galaxien sehr unterschiedlich ist und unterschiedliche Modellierungstechniken erfordert.

Als Nächstes erweitere ich die Technik der transversalen Absorptionslinien auf eine dichtere Umgebung wie einen Galaxienhaufen. Ich untersuche die Absorption von kühlem Gas in Galaxienhaufen, indem ich Quasar-Absorber sowohl im Rotverschiebungsraum als auch im projizierten kpc-Raum mit Galaxienhaufen verbinde, die in der Legacy-Imaging-Durchmusterung des Dark Energy Spectroscopic Instrument (DESI) entdeckt wurden. Ich habe herausgefunden, dass Galaxienhaufen eine beträchtliche Menge an kühlem Gas in ihren Halos beherbergen, obwohl das Medium innerhalb des Galaxienhaufens (intracluster medium, ICM) sehr heiß ist, was sich dadurch zeigt, dass das Gas hauptsächlich im Röntgenbereich emittiert. Die gesamte Mg II -Masse innerhalb des Galaxienhaufen-Halos ist $\sim 8-10$ mal höher als bei passiven Galaxien bei ähnlichen Rotverschiebungen.

Um den möglichen Zusammenhang zwischen der Absorption von kaltem Gas in Sternhaufen und der stellaren Aktivität und den Haloeigenschaften ihrer Mitglieder zu verstehen, verbinde ich diese Absorber auch mit photo$-z-$ausgewählten Galaxien aus DESI. Ich finde eine statistisch signifikante Verbindung zwischen Absorbern und Galaxien in Galaxienhaufen auf projizierten Skalen. Der mittlere projizierte Abstand zwischen Mg II -Absorbern und dem nächstgelegenen Haufenmitglied beträgt ~ 200 kpc im Vergleich zu ~ 500 kpc in zufälligen Modellen mit denselben Galaxiendichteprofilen. Ein beträchtlicher Anteil der Absorber ($\sim$ 50 Prozent) befindet sich innerhalb des Halos der dunklen Materie dieser Galaxien. Ich finde jedoch keine Korrelation zwischen der Absorptionsstärke des kühlen Gases und der Sternentstehungsrate der nächstgelegenen Galaxie im Galaxienhaufens. Dies deutet darauf hin, dass das kühle Gas in Galaxienaufen, das von den Absorbern aufgespürt wird, entweder mit Satellitengalaxien assoziiert ist, die zu massearm sind, um im DESI-Katalog enthalten zu sein, oder mit kalten Gaswolken im Intra-Galaxienhaufenmedium, die zum Teil aus Gas stammen könnten, das in der Vergangenheit von diesen Galaxienhaufen-Satelliten gestrippt wurde.

Schließlich habe ich unsere Absorber-Detektionspipeline eingesetzt, um in den Spektren von Hintergrund-Quasaren mit Mg II -Detektionen nach hochionisierten Absorbern (C IV , dreifach ionisierter Kohlenstoff) zu suchen. Durch die Kombination von Metallabsorbern in niedrigen und hohen Ionisierungszuständen untersuchte ich die räumliche Verteilung von kühlem und warmem Gas im CGM bei $z \sim 2$ unter Verwendung der größten spektroskopischen Probe von Quasaren und Absorbern. Diese Analyse zeigt große Reserven an kühlem und warmem Gas in Quasar-Halos, was darauf hindeutet, dass Quasar-Halos hochgradig mit Metall angereichert sind. Neben der Charakterisierung der allgemeinen räumlichen Verteilung von kühlem und warmem Gas um Quasare, untersuche ich auch die relative Kinematik zwischen Metallabsorbern und Quasaren. Diese Analyse zeigt, dass die Geschwindigkeitsdispersion von absorbierenden Wolken $3-4$ mal größer ist als die von sternbildenden oder ruhigen Galaxien und ähnlich der Bewegung von kühlem Gas in Galaxienhaufen. Da sich Quasare in viel kleineren Halos als Galaxienhaufen befinden, könnte dies auf ein Szenario hindeuten, bei dem metallangereichertes Gas in den Quasaren-Halos mit Ausströmen verbunden ist.

Abstract

The gaseous atmosphere that surrounds galaxies is known as the circumgalactic medium (CGM). The CGM is the interface between the physical processes happening at small and large scales. The complex balance between these processes drives the formation and evolution of galaxies. On the one hand, cool pristine gas is accreted from the intergalactic medium (IGM) or cosmic filaments onto galaxies. At the same time, galactic outflows produced by supernovae or the active galactic nucleus (AGN) expel a large amount of gas and metals into the CGM, some of which can even escape the halo while some recycle, mix and fall back into the galaxy. These flows play a critical role in regulating galaxy formation, such as determining the timescales of gas depletion and star formation, the processes responsible for preventing star formation in galaxies, and the distribution of the baryonic and metal budgets in the gas surrounding galaxies. The circumgalactic gas is usually studied in absorption (caused by hydrogen or metals) against a luminous background source such as a quasar. These absorbers can provide invaluable insights into the gas flows and their connection to the galactic properties and environment.

In this thesis, I characterize the properties of the cool and warm circumgalactic gas of galaxies, clusters and quasars using quasar absorbers. To study the circumgalactic gas around galaxies, I first developed a fully automated pipeline to model the quasar continuum emission. Then, I employ a matched kernel convolution technique combined with an adaptive signal-to-noise (S/N) criteria to detect intervening metal absorbers in their spectra. The pipeline resulted in the most extensive metal absorber catalogue to date. It provided an opportunity to perform one of the most extensive absorption line studies of the galactic atmosphere around different types of galaxies.

Then, I combine these absorbers with large spectroscopic samples of galaxies and investigate the nature of the cold circumgalactic medium at $z \sim 0.5$. Thanks to the large sample sizes, I could characterize the scales of cold gas absorption (traced by Mg II , singly ionized magnesium) in the CGM of star-forming galaxies. I find that cold gas absorption is higher in their CGM is strongest within ~ 50 kpc from the galaxy, and the incidence rate of absorbers in the inner parts of the halo is $2-4$ times higher for star-forming galaxies compared to quiescent galaxies at similar redshifts. However, beyond 50 kpc, I see a sharp decline in cold gas absorption strength and incidence rate for both types of galaxies, indicating a dichotomy of physical conditions in the gas. The inner regions of the CGM are regulated mainly by galactic outflows, while the outer CGM is tightly linked with the galaxy's dark matter halo. Galaxies with a higher star formation rate (SFR) have a higher

incidence of cold gas, implying a strong connection to the stellar outflows. On the other hand, I find that motion of absorbing gas is sub-virial around passive galaxies, implying that gas is gravitationally bound and unlikely to have originated from galactic outflows. Their low star formation activity further supports this observation. Therefore, the origin of cool gas around passive galaxies is possibly due to accretion from the intergalactic medium (IGM) or stripping of the gas from satellite galaxies. This analysis suggests that the physical origin of cool circumgalactic gas for star-forming versus quiescent galaxies is very different and requires different frameworks to model them.

Next, I extend the transverse absorption line technique to a more dense environment such as a galaxy cluster. I study the cool gas absorption in galaxy clusters by connecting quasar absorbers in both redshift and projected kpc space with galaxy clusters detected in the legacy imaging survey of the Dark Energy Spectroscopic Instrument (DESI). I find that clusters host a significant amount of cool gas in their haloes, despite the fact that intracluster medium (ICM) is very hot, as evidenced by the fact that the gas emits mainly in X-rays. The total Mg II mass within the cluster halo is $\sim 8 - 10$ times higher than for passive galaxies at similar redshifts.

To understand the possible connection of cold gas absorption in clusters with the stellar activity and the halo properties of its members, I also connect these absorbers with photo $-z$ selected cluster member galaxies from DESI. I find a statistically significant connection between absorbers and cluster galaxies on projected scales. The median projected distance between Mg II absorbers and the nearest cluster member is $\sim$ 200 kpc compared to $\sim$ 500 kpc in random mocks with the same galaxy density profiles. A substantial fraction of absorbers ($\sim$ 50 percent) are located within the dark matter halo of those galaxies. However, I do not see a correlation between cool gas absorption strength and the star formation rate of the closest cluster neighbour. This suggests that cool gas in clusters, as traced by these absorbers, is either associated with satellites galaxies that are too low in mass to be found in the DESI catalogue, or associated with cold gas clouds in the intracluster medium that may originate in part from gas stripped from these cluster satellites in the past.

Finally, I employed our absorber detection pipeline to search for highly ionized absorbers (C IV , triply ionized carbon) in the spectra of background quasars with Mg II detections. By combining metal absorbers in low and high ionization states, I studied the spatial distribution of cool and warm gas in the CGM at $z \sim 2$ using the largest spectroscopic sample of quasars and absorbers. This analysis reveals large cool and warm gas reservoirs in quasar haloes, indicating that quasar haloes are highly metal-enriched. In addition to characterizing the overall spatial distribution of cool and warm gas around quasars, I also study the relative kinematics between metal absorbers and quasars. This analysis shows that the velocity dispersion of absorbing clouds is $3 - 4$ times larger than star-forming or quiescent galaxies and similar to the motion of cool gas in galaxy clusters. Since quasars reside in much smaller haloes than clusters, we speculate that this may indicate a scenario where metal-enriched gas in their haloes is associated with outflows.

If you wish to make an apple pie from scratch, you must first invent the universe.

Carl Sagan

Chapter 1

Introduction

Understanding the physical processes that form galaxies and determine their evolution is one of the central goals of astronomy today. Galaxies reside at the centre of dark matter haloes, which constitute most of the mass of galactic halo. Dark matter is entirely invisible because it does not emit or absorb any form of electromagnetic wave. On the other hand, the baryons[1] constitute only a tiny fraction ($\lesssim 5$ per cent) of the total density[2] (dark matter+baryonic matter+dark energy) in the universe today. In terms of matter density, baryons constitute about ~ 15 per cent of the total matter density (dark matter+baryonic matter) in the universe (Planck Collaboration et al., 2020). Interestingly, only one-tenth of this baryonic matter is in the luminous matter. However, these baryons alone are responsible for all the light we see in our universe, from planets to stars to galaxies to the universe's large-scale structure.

Over the last hundred years, our understanding of the Universe's large scale structure has deepened. The structure formation and physical interaction between baryons and dark matter have been studied extensively with advanced telescopes and powerful numerical simulations and models. Today, we have a better picture of the processes through which the Universe's large-scale structure forms and evolves. In the next sections, we will walk through the scientific knowledge that we have accumulated over the last century and discuss what are the burning questions that remain open in the field of galaxy formation and evolution.

1.1 Structure Formation in Cosmology

The modern theory of cosmology describes the formation of structures in the universe as a result of initial fluctuations in the density field of the universe. The origin of these density perturbations has been discussed in the context of the inflationary model of our universe.

[1] In astronomy, baryons also include electrons, though electrons are members of the lepton family.

[2] In Lambda cold dark matter model, total matter/energy density, $\Omega_{\rm tot} = \Omega_{\rm DM} + \Omega_{\rm baryon} + \Omega_{\Lambda} = 1$ and the universe is assumed to be flat. Latest Planck measurements are: $\Omega_{\rm DM} = 0.2589$, $\Omega_{\rm baryon} = 0.0486$, $\Omega_{\Lambda} = 0.6911$ (Planck Collaboration et al., 2020).

It is thought to be due to quantum fluctuations of the scalar field (Guth & Pi, 1982; Starobinsky, 1982). In this scenario, the overdense regions collapse under gravity and form structures (Gamow & Teller, 1939; Lifshitz, 1946; Peebles, 1965; Silk, 1968; Peebles & Yu, 1970; Zel'dovich, 1970). The afterglow produced by the universe about $\sim 400,000$ years after the Big Bang, when matter and radiation decoupled, is visible everywhere in the sky. The radiation falls in the microwave part of the electromagnetic spectrum and is known as the cosmic microwave background or CMB, first detected in the 1960s by Penzias & Wilson (1965). The early success of Cosmic Background Explorer (COBE, Smoot et al. (1992)) in detecting temperature fluctuations in the CMB spectrum was one of the critical events in confirming the predictions of inflationary models of the universe.

In the early 1970s, several attempts were made to connect these density fluctuations to the nonlinear structures that we observe in the universe. On the one hand, Zel'dovich (1970) developed a theory of the nonlinear collapse of the initial fluctuations to form structures. At the same time, in their pioneering works, Gunn & Gott (1972) and Press & Schechter (1974) developed the spherical collapse model of structure formation from the initial density field and estimated the mass function of collapsed haloes.

The concept of 'dark matter' was first conceived by Zwicky (1933) while studying the relative motions of galaxies in dense clusters. The observation of the flat rotation curve of Andromeda (Rubin & Ford, 1970) triggered an interest in understanding the nature of dark matter (Einasto et al., 1974; Ostriker et al., 1974). In the early 80s and 90s, several groundbreaking works developed the theory of what is now known as the 'cold dark matter', in the context of structure formation in the universe (Peebles, 1982; Blumenthal et al., 1982, 1984). In the late 90s, after the discovery of the accelerated expansion of the universe (Riess et al., 1998; Perlmutter et al., 1999), the Lambda cold dark matter model (ΛCDM[3]) became the standard model of our universe (see Peebles & Ratra, 2003, for a review).

Now the latest maps of CMB by the Wilkinson Microwave Anisotropy Probe (WMAP) (Komatsu et al., 2011) and Planck Mission (Planck Collaboration et al., 2020) have constrained these fluctuation amplitudes and the cosmological parameters derived from them with very high precision.

1.2 Formation of Galaxies

Beginning with the first observations of galaxies by Messier in the 17th century (though he didn't know what they were), galaxies have been at the core of astronomy. It was not until 1923 that Hubble resolved the famous 'Shapley-Curtis debate' and concretely established the presence of galaxies beyond the Milky Way. Hubble also classified the shapes of galaxies (elliptical and spirals) in his famous 'tuning fork diagram' (Hubble, 1926). However, it was not until 40 years later, in 1962, when Eggen et al. (1962) developed the first models of galaxy formation based on the gravitational collapse of gas clouds. In the following decade

[3]Cold Dark Matter theory with a Cosmological Constant accounting for unknown Dark Energy.

Toomre & Toomre (1972) and Larson (1976) developed the theory to explain the formation of galaxies with different morphologies.

In the current framework of cosmology, the initial density fluctuations grow linearly with time until they reach a value of 1.68 times the critical density of the Universe[4]. After this point, the matter collapses under its gravity. It forms a gravitationally bound system, known as the 'dark matter halo', which is under virial equilibrium [5](Gunn & Gott, 1972; Press & Schechter, 1974). In the modern theory of galaxy formation and evolution (White & Rees, 1978), baryons accrete towards the centre of the potential well and cool to form stars via gravitational collapse (White & Frenk, 1991; Mo et al., 2010). The stellar nucleosynthesis then fuses hydrogen into heavier elements such as carbon, nitrogen and oxygen. The stellar winds or supernova (an energetic explosion when a star dies) can expel metals[6] into the interstellar medium (ISM). In addition, supernovae can also heat the ISM or kick the gas and metals out of the galactic disc into the galactic halo. On the other hand, the black hole at the galaxy's centre can accrete gas from the surroundings and grow with time. The outflows powered by a supermassive black hole (SMBH) can also heat the gas or eject it up to vast distances in the halo.

The modern understanding of these feedback processes and their implications for galaxy formation and evolution hinges upon cosmological N-body simulations of dark matter and baryonic matter within the ΛCDM framework (Navarro & White, 1994). In the last twenty years, these numerical simulations have successfully predicted the formation of galaxies with different stellar masses and star formation rates (SFR) over cosmic time (e.g., Springel et al., 2005; Genel et al., 2014; Vogelsberger et al., 2014; Schaye et al., 2015; Pillepich et al., 2018). While many simulation predictions agree well with the observations, they are often unable to reproduce the observational measurements in the galactic halo, where the complex interplay between these feedback and accretion processes constantly drives the fate of galaxies. In the next section, we will delve into the major problems in galaxy formation and evolution and how having a better hold on the processes occurring in galactic halo could be a way forward to a better understanding of the baryon cycle.

1.3 Key Problems in Galaxy Formation and Evolution

We begin with the most important problems in the field of galaxy formation and evolution and their connection to the galactic halo (defined more formally in section 1.4). Over time, these puzzles have been summarized into four fundamental problems in the field. Below we discuss each one of them briefly.

[4]The critical density is an average matter density that is required to keep the geometry of the universe flat. The universe will expand for infinite time if it maintains the critical density.

[5]When the gravitational potential energy of the halo is twice the magnitude of the total kinetic energy of its moving particles.

[6]In Astronomy, all elements except Hydrogen and Helium are referred as 'metals'.

1.3.1 How galaxies are forming stars for so long?

The gas depletion time, $t_{\rm dep} \approx M_{\rm gas}/\dot{M}_{\rm SFR}$ (i.e. the ratio of total gas mass to the rate with which stars are forming in galaxy), varies by $2-3$ times (from 6 Gyr to 2 Gyr) over three orders of magnitude difference in stellar mass ($\log\,[{\rm M/M_\odot}] = 8.5$ to $\log\,[{\rm M/M_\odot}] = 11.5$) (Whitaker et al., 2012; Peeples et al., 2014). It poses a big challenge for us because our own Milky Way has been converting gas into the stars for the last 10 billion years (Bond et al., 2013). Hence, there must be an external supply of the necessary fuel for this sustained star formation. One possible source is the galactic halo (the circumgalactic medium), which continuously feeds gas to the galactic disc (the factory where stars are born) via accretion.

1.3.2 Quenching of star formation in galaxies

Another mystery at the core of galaxy evolution is why different galaxies have had different star formation histories? What physical processes prevent some galaxies from forming stars after a while and keep them passive afterwards while other galaxies are still forming stars? The observed bimodality in the galaxy population on the $M_\star - \log {\rm sSFR}$ (stellar mass vs specific star formation rate[7]) (Schiminovich et al., 2010) plane is of particular interest where at given stellar mass, two distinct populations of galaxies can be seen (blue star-forming and passive red galaxies) based on their sSFR. The proposed solutions to the problem include control of gas supply from the CGM or IGM or continuous heating or ejection of cold gas from the ISM, thus preventing any further star formation activity (Bluck et al., 2020). However, several open questions remain poorly understood, such as what are the time scales over which galaxies quench? (Walters et al., 2022) What role does the galaxy's environment play in its quenching process?(Geha et al., 2012). Certainly, the gas flows in and out of the galactic halo are key to these questions.

1.3.3 Missing baryon problem

The standard hierarchical model of galaxy formation has successfully established that galaxies form in the centres of dark matter haloes by accreting baryonic matter (White & Rees, 1978; White & Frenk, 1991). However, galaxies contain a much lower ($\lesssim 20\,\%$) fraction of cosmic baryons ($\frac{\Omega_{\rm b}}{\Omega_{\rm m}}{\rm M_h}$)[8] in their ISM and stars combined (Afshordi et al., 2007; McGaugh, 2008; Behroozi et al., 2010). The baryons in the galactic disc alone can not account for the theoretically predicted baryon fraction for galaxies. The missing baryons are probably present in the galactic halo or the disc, and we cannot see them even with the largest telescopes. The other possibility is that baryons have been expelled far from the galactic halo into the IGM or cosmic filaments, or galaxies did not get their fair share of baryons in the first place when they were forming.

[7]specific star formation rate is defined as the ratio of star formation rate (SFR) of galaxy to its stellar mass, ${\rm sSFR}\,[{\rm yr^{-1}}] = {\rm SFR/M_\star}$

[8]The Cosmic baryon fraction, $f_{\rm b} = \frac{\Omega_b}{\Omega_m} = 0.1565$ (Planck Collaboration et al., 2020)

Cosmological simulations have predicted these missing baryons to reside in the warm-hot intergalactic medium (WHIM, $T \sim 10^5 - 10^7$ K) (Cen & Ostriker, 1999; Davé et al., 2001; Reimers, 2002; Bertone et al., 2008). Given the WHIM is a highly ionized and low-density gas, it is challenging to detect in emission[9], even with the most advanced telescopes. However, there has been tremendous success in detecting these hot filaments in recent years, particularly in absorption. Nicastro et al. (2018) detected highly ionized oxygen (O VII) absorbers associated with an overdense region of galaxies in very high S/N X-ray spectra of background quasars. These observations are consistent with cosmological predictions implying that a large fraction of missing baryons indeed resides in WHIM. In addition to absorption studies, de Graaff et al. (2019) and Tanimura et al. (2019) probed the WHIM by mapping the thermal Sunyaev-Zel'dovich (tSZ) (Sunyaev & Zeldovich, 1972) effect of SDSS galaxies and detected hot, diffuse gas in cosmic filaments. However, understanding how metals evolve at different epochs is still an ongoing research topic in the field (see Somerville & Davé, 2015; Tumlinson et al., 2017; Péroux & Howk, 2020, for reviews).

1.3.4 Missing metals problem

Even though the 'missing baryon problem' seems to have a solution, another puzzle remains unresolved. While a large fraction of the missing baryons reside in the WHIM, galaxies also lack their fair share of metals. Stellar physics tells us that heavy elements (other than hydrogen) are produced in the stars via nucleosynthesis, while more heavy metals (like gold or uranium) form via supernova. This implies that all the metals formed in stars must remain in the ISM or disc of galaxies. However, the galactic disc only contains $\lesssim 20 - 25$ per cent of the total metals produced (Peeples et al., 2014; Oppenheimer et al., 2016). Both observations (Tremonti et al., 2007; Circosta et al., 2018) and simulations (Nelson et al., 2019) have shown that metals are expelled via galactic outflows up to large distances in the galactic halo. Therefore, tracing the distribution of metals in the halo is a way forward to shed light on this problem.

1.4 The Galactic Halo: Circumgalactic Medium

The previous section described the main problems at the core of modern galaxy formation and evolution. In all four questions, the connection of the galaxy with its gaseous halo is at the forefront. Now we will focus on the nature of this gaseous envelope and how can it provide some clues to these open problems.

The circumgalactic medium (CGM) is the gaseous envelope that surrounds the galaxies (see Figure 1.1). The physical boundaries of the CGM are not well defined. However, for practical purposes, we define it as the region between the galactic disc or ISM (inner edge) and virial radius ($r_{\rm vir}$, outer boundary) of the dark matter halo. Notably, we also do not know the exact ISM boundary and the concept of $r_{\rm vir}$ is also a bit ambiguous. In

[9] Emissivity is proportional to n^2 and flux decreases as r^{-2}.

the literature, the definition of $r_{\rm vir}$ is related to the ratio of mean matter density to the universe's critical density. The critical density of the universe is defined as:

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

where Hubble parameter, $H(z)$ is the Hubble expansion at cosmological redshift z[10] and G is the gravitational constant. Now defining virial radius as the radius of a spherical halo, where the matter density is some Δ times the universe's critical density ($\rho_{\rm m} = \Delta\rho_{\rm crit}$), the total mass, M_Δ within r_Δ is given as:

$$M_\Delta = \frac{4}{3}\pi r_\Delta^3\,\Delta\rho_{\rm crit} \implies r_\Delta = \left(\frac{3M_\Delta}{4\pi\,\Delta\rho_{\rm crit}}\right)^{1/3} \tag{1.2}$$

In the literature for the CGM purposes, we take $\Delta \approx 200$ (Bryan & Norman, 1998), based on numerical simulations of dark matter halo formation; around this value of Δ, the motion of particles in the outer orbit starts to virialize. Therefore in terms of $\Delta = 200$, we can define virial radius, $r_{\rm vir} = r_{200}$ as:

$$r_{200} = \left(\frac{3M_{200}}{4\pi\,200\rho_{\rm crit}}\right)^{1/3} \tag{1.3}$$

where M_{200} can be considered as the total mass of the galaxy's dark matter halo. There are several models based on cosmological simulations of structure formation that explore the analytical relation between stellar mass and halo mass of galaxies as a function of redshift (see Moster et al., 2013, 2018; Behroozi et al., 2019, for details). We can further simplify the Eqn 1.3 in *kpc*[11] units (Fielding et al., 2020; Donahue & Voit, 2022):

$$r_{200} = r_{\rm vir} \approx 210\left(\frac{M_{200}}{10^{12}\,M_\odot}\right)^{1/3} E^{-2/3}(z)\,[\rm kpc] \tag{1.4}$$

where, $E(z) = H(z)/H(z=0) = \sqrt{\Omega_{\rm m}(1+z)^3 + (1-\Omega_{\rm m})}$; $\Omega_{\rm m} = \rho_{\rm m}/\rho_{\rm crit,0} = 0.27$ (Mo et al., 2010; Komatsu et al., 2011) is the total matter density at $z = 0$, $H(z=0) = H_{\rm o} = 67.6\,{\rm km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration et al., 2020) is the Hubble constant at present time. For typical $L^\star$ galaxies[12] with $M_{200} \sim 10^{13}\,{\rm M_\odot}$ (Cooray & Milosavljević, 2005), $r_{200} \sim 400\,{\rm kpc}$.

1.5 The Local Baryon Cycle in the CGM

Now that we have broadly defined the CGM, it is important to discuss the various physical processes happening in the CGM. These processes are critically connected to the formation

[10] The cosmological redshifts can be estimated by observing the shifts in spectral lines of atoms or molecules as $1 + z = \frac{\lambda_{\rm obs}}{\lambda_{\rm rest}}$, where $\lambda_{\rm rest}$ and $\lambda_{\rm obs}$ are the 'rest-frame' and 'observed' wavelengths of atomic transition. In cosmology, redshift is also a measure for distance of an object.

[11] 1 kpc $= 3.086 \times 10^{19}$ meters

[12] $L^\star$ is the characteristics luminosity of galaxies above which the galaxy luminosity function is dominated by exponential decline in the Schechter function.

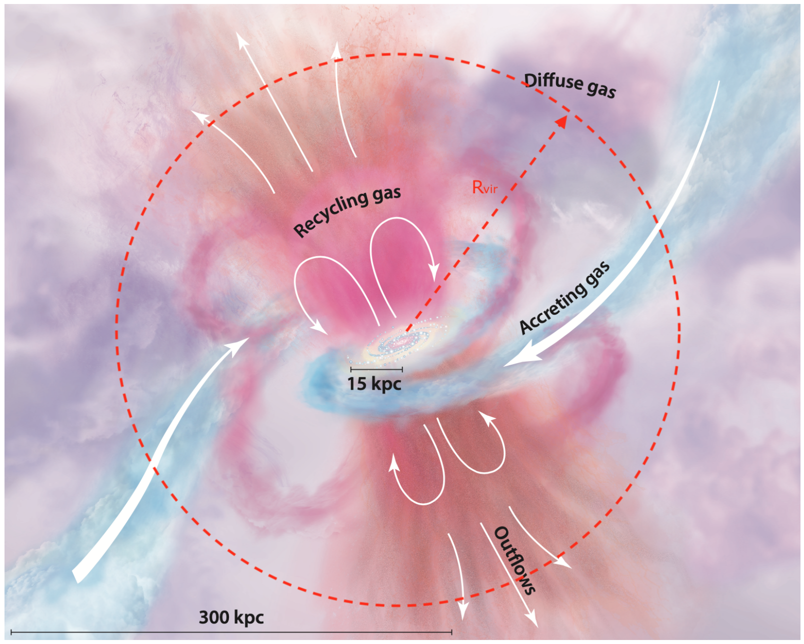

Figure 1.1: A bird's eye view of the circumgalactic medium, taken from Tumlinson et al. (2017): The galactic disc in the centre of the halo gets its cool gas supply via accretion (blue) from the gaseous halo or the cosmic filaments. The orange and pink shades show the outflowing material, powered mainly by supernovae or AGN activity of the central galaxy. Also shown in the same shade the previously ejected gas that is being recycled in the halo. Next, the purple shows the diffuse halo gas that is a cumulative result of all the physical processes over time. The CGM and galactic disc scales are also shown in the panel. The red dashed line represents the $r_{\rm vir}$ of the galaxy's dark matter halo.

and evolution of galaxies. In Figure 1.1 (Tumlinson et al., 2017), we show a bird's eye view of the galactic halo. The CGM is the main venue for cosmological gas accretion and inflows (blue) and galactic outflows (orange + pink) produced by feedback processes. These flows pass through the CGM and play a pivotal role in several key processes regulating galaxy formation.

As described in section 1.3.1, galaxies can not form stars for so long unless their gas supply is replenished continuously via accretion of cold, pristine gas from the IGM or cosmic filaments (Kereš et al., 2005; Dekel et al., 2009; Tumlinson et al., 2013; Borthakur et al., 2015). These accretion flows necessarily pass through the CGM (see cartoon in Figure 1.1), and by constraining the CGM properties, we can understand the nature of these accretion flows. The solution to the quenching mechanism is also connected to the gas flows in the galaxy's CGM. In section 1.3.2, we discussed the very different star formation histories of galaxies. Again, the resolution involves regulating and controlling the gas supply to the disc, where stars are formed. Either the cold gas is being heated or expelled out of the disc by the powerful supernovae driven winds or AGN outflows (Gabor et al., 2010; Choi et al., 2015; Suresh et al., 2017), preventing the collapse of molecular clouds (the birthplace of stars), or the galaxy didn't get enough fuel via accretion (Kereš et al., 2005). What are the timescales over which the cold gas is heated and expelled? How far can the ejected gas travel before being recycled back to the ISM? What is the average temperature and density structure of the galactic halo? These are a few challenging questions we are struggling with (Somerville & Davé, 2015; Naab & Ostriker, 2017; Péroux & Howk, 2020; Donahue & Voit, 2022). Whatever is the dominant mechanism (though it is probably a combination of all), this must involve the interaction with the CGM. Finally, the missing baryon and metal problems (see section 1.3.3 and 1.3.4) can be resolved by studying the gas content of the CGM. In the CGM, the gas with different densities and temperatures can co-exist up to vast distances. We will return to the multiphase nature of the CGM in section 1.7.

In the previous sections, we have illustrated how fundamental the complex interaction between these gas flows is to the fate of galaxies and how they drive their physical properties and evolution. Therefore our understanding of galaxy formation is itself limited by our current knowledge of the CGM. However, extensive sky surveys performed with the ground and space-based telescopes such as the Sloan Digital Sky Survey (SDSS), Keck, Very Large Telescope (VLT), and the Hubble Space Telescope (HST) combined with state-of-art numerical simulations have significantly deepened our understanding of the CGM over the past two decades. In the following sections, we will walk through the methodologies and surveys that have helped decipher the evolution of the CGM properties around galaxies at different epochs. Our goal is to present a brief synthesis of important discoveries and results on the gaseous haloes of galaxies.

1.6 CGM in Observations

A wide range of observational tools is available to study this invisible gas around galaxies. The gaseous envelope of galaxies can be analysed with the electromagnetic spectrum ranging from high-energy X-ray observations to ultraviolet (UV) to optical and sub-millimetre wavelengths. Combining these observations provides a holistic picture of different phases of the CGM. Below we summarize the observational techniques.

1.6.1 Quasar Absorption Line Spectroscopy

One of the most powerful tools to study the CGM of galaxies is so-called 'transverse absorption - line studies'. The basic idea is the following: the light coming from a bright distant quasar gets absorbed by the medium present between it and the observer (see Figure 1.2). Quasars are the most luminous and powerful active galactic nuclei (AGN) in the universe (luminosity, $L > 10^{44}\,erg\,s^{-1}$); therefore, they can be observed up to very far distances. The accretion disc fueled by continuous feeding of matter to the powerful central black hole gives rise to several emission features characteristic of the quasar (see Inayoshi et al., 2020, for a recent review). The intrinsic spectral shape of the accretion disk is a power law. By observing a quasar spectrum, we can look for flux decrements (absorption dips) relative to the quasar's intrinsic continuum features. Those absorption dips are at lower

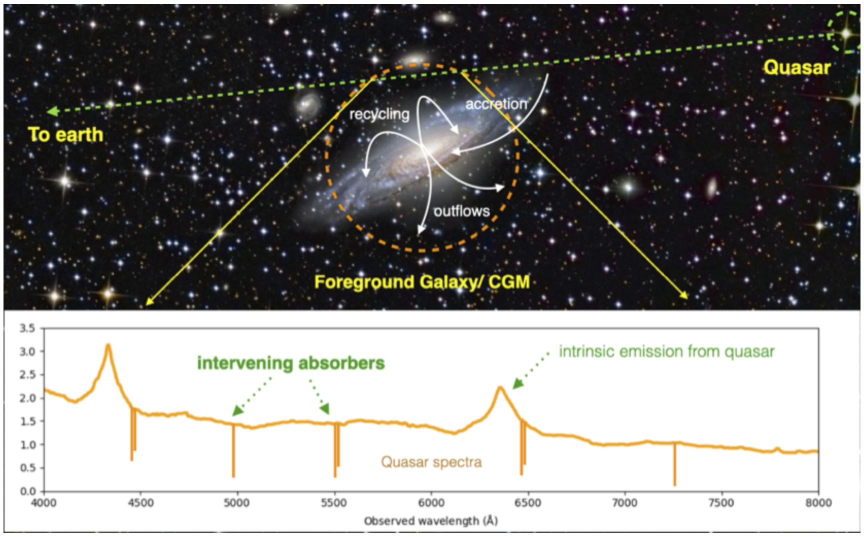

Figure 1.2: The basic setup of an absorber-CGM cross-correlation study, where the CGM is seen in absorption against a bright background quasar.

redshifts (distance) than the quasar itself, by definition. The method has its merit because absorption depends exponentially on the density of the matter present, therefore can be sensitive to very low column densities and trace a broad range of temperature and densities. In addition, absorber detection limit does not depend on the distance or luminosity of the foreground objects. In contrast, emission flux declines as $\propto r^{-2}$ and depends on the density as $\propto n^2$, therefore emission studies are limited to certain distances and densities, even with the most advanced telescopes. The only downside to the absorption line studies is that we are often limited by one projected pencil-beam sightline per galaxy. Hence our ability to completely characterize the CGM of even one galaxy is very challenging with this method. However, in recent years, using multiple sightlines passing through our galaxy, studies have constrained the properties of Milky Way's CGM (Zheng et al., 2019). With the recent advances in spectroscopic surveys, the quasar-galaxy pair studies have been extended up to $z \sim 4$ (Prochaska et al., 2014; Turner et al., 2016), including thousands of galaxies and absorbing systems (Zhu & Ménard, 2013a). Before we delve into what we have learnt from absorption line studies, we discuss the basic concepts of absorption line spectroscopy.

1.6.1.1 Basic Physics of Absorption Lines

Suppose we measure the spectral flux density, F_ν as a function of frequency ν, of the bright background continuum source (e.g. Quasar), then this flux at the observer can be given as

$$F_\nu = F_\mathrm{o}\, exp(-\tau_\nu) \tag{1.5}$$

where F_o is the intrinsic flux of the object in the absence of any absorption and τ_ν is the optical depth[13] of the intervening matter. The optical depth is related to the column density (N) of absorbing material through[14]

$$\tau_\nu = \frac{\pi e^2}{m_e c} f_{osc} \, N \, \phi(\nu) \tag{1.6}$$

where $N = \int n \, ds$ is the column density of the absorbing matter, n is density, ds is the path length of light through the medium, f_{osc} is the quantum mechanical oscillator strength of the atomic transition, e, m_e are the charge and mass of electron, c is the speed of light. $\phi(\nu)$ is the normalized line profile[15], such that $\int \phi(\nu)\, d\nu = 1$. For more detailed derivations see Ch 7 of Draine (2011). However, column density is also related to a directly observed quantity, called *equivalent width*, which is defined as

$$EW = \int_{\lambda_1}^{\lambda_2} \Big(1 - \frac{F_\lambda}{F_o}\Big) d\lambda = \int_{\lambda_1}^{\lambda_2} (1 - e^{-\tau_\nu}) d\lambda \tag{1.7}$$

where λ_1 and λ_2 are the regions where absorption feature is located, such that outside this region the function, $1 - \frac{F_\lambda}{F_o} \to 0$. Intuitively, EW measures absorption strength in terms of missing flux from the intrinsic continuum; therefore can be measured even for unresolved absorbers. EW is related to the absorber's column density via the *curve of growth*, $EW(N)$ function. Generally, our goal is to invert this function and measure N at a given EW, though it is challenging because it depends on the shape of the intrinsic line profile (see Eqn 1.6). Now, if the absorbers have a Gaussian velocity distribution with optical depth at the line-centre, τ_o and dispersion σ_V, the optical depth can be expressed as

$$\tau_\nu = \tau_o \, exp(-\Delta v^2/b^2); \quad b = \sqrt{2}\sigma_V \tag{1.8}$$

where τ_o is given by

$$\tau_o = \sqrt{\pi} \frac{e^2}{m_e c} \frac{N f \lambda}{b} \tag{1.9}$$

and $\Delta v = c\frac{\nu_o - \nu}{\nu_o}$, is the frequency shift from the line-centre due to particle motion. If the velocity dispersion is purely due to the thermal motion then we can measure the kinetic temperature of the absorbers using b parameter as

$$T = \frac{M_{atom} b^2}{2k_B} \implies T\,[\mathrm{K}]/10^4 = \Big(\frac{\mathrm{b\,[km\,s^{-1}]}}{12.90}\Big)^2 \mathrm{M_{atom}\,[amu]} \tag{1.10}$$

[13]It's a measure of how opaque a medium is and depends on the attenuation coefficient (κ_ν) and path that light traverses in the cloud (ds) as $d\tau_\nu = \kappa_\nu \, ds$. Larger the path, the larger the optical depth.

[14]In this equation, stimulated emission is neglected, so not applicable if we have a population inversion. For more details see Eqn 9.5 of Draine (2011).

[15]The most commonly used line profiles are Gaussian or Voigt profiles. The latter is a convolution of Gaussian (characterizing the thermal broadening of the line) with a Lorentzian profile (describing the collisional broadening of the line).

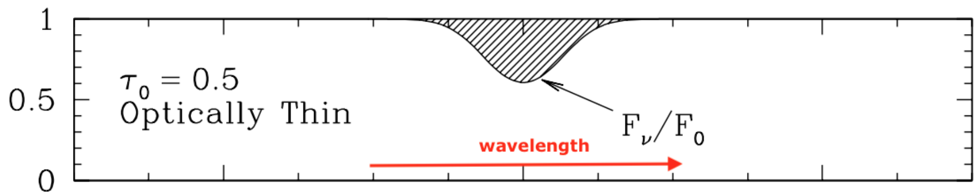

Figure 1.3: An example of absorption line profile in optically thin regime, taken from Draine (2011).

In the optically thin regime (see Figure 1.3), $\tau_\nu \ll 1$, Eqn 1.7 can be solved analytically and EW and N can be approximated as

$$EW \approx \sqrt{\pi}\frac{b}{c}\,\tau_o\lambda \implies N\,[\mathrm{cm}^{-2}] \approx 1.13 \times 10^{20} \times \frac{EW}{f\lambda^2} \tag{1.11}$$

where, we have replaced τ_o from Eqn 1.9 and used the usual values for electronic constants, and both EW and λ are in Å units. In optically thick ($\tau_\nu > 1$), regions Eqn 1.11 will only give a lower limit on column density. From here, it is trivial to estimate the surface mass density of absorbers, by just multiplying the column density by the mass of the absorber.

Sometimes, it is difficult to detect weak systems in individual quasar spectra, given the signal-to-noise (S/N) limits. In that case, we can then 'stack' thousands of spectra in the rest-frame of foreground objects to enhance the spectral S/N and detect those faint absorption signals. However, this comes with the cost of averaging out the properties of the individual absorbers. Nevertheless, in recent years, there have several stacking analyses with quasars from the Sloan Digital Sky Survey (SDSS) and the Dark Energy Spectroscopic Instrument (DESI) surveys (see Zhu et al., 2014, 2015; Lan & Mo, 2018). These studies have detected several weak lines in the halo of galaxies and studied their properties as a function of the galaxy's stellar and halo properties.

On the other hand, instead of using a quasar as the background source, we can observe the CGM against the background light of its host galaxy, the so-called 'Down the Barrel' technique. This is particularly useful for disentangling the nature of gas flows, such as accretion (inflows) or outflows. There has been substantial progress in this direction in recent years. Studies have revealed significant covering factors of metals along minor axes of galaxies, indicating the metal enrichment of the CGM via galactic outflows (Bordoloi et al., 2011; Rubin et al., 2014).

1.6.2 CGM in Emission

In contrast to absorption line studies, in emission, we map the photons emitted from the gaseous component of CGM. As discussed in sections 1.6.1, the flux of emitted photons declines rapidly with distance ($\propto r^{-2}$) and emissivity is directly proportional to the density of the medium ($\propto n^2$). Given the diffuse nature of the CGM, it is incredibly challenging to construct an emission map of the galactic halo, even with the largest telescopes. The Milky Way and other nearby galaxies have been mapped with H I 21 cm emission in radio waves and have revealed an extensive reservoir of neutral hydrogen in their haloes (see Kalberla & Kerp, 2009, for review).

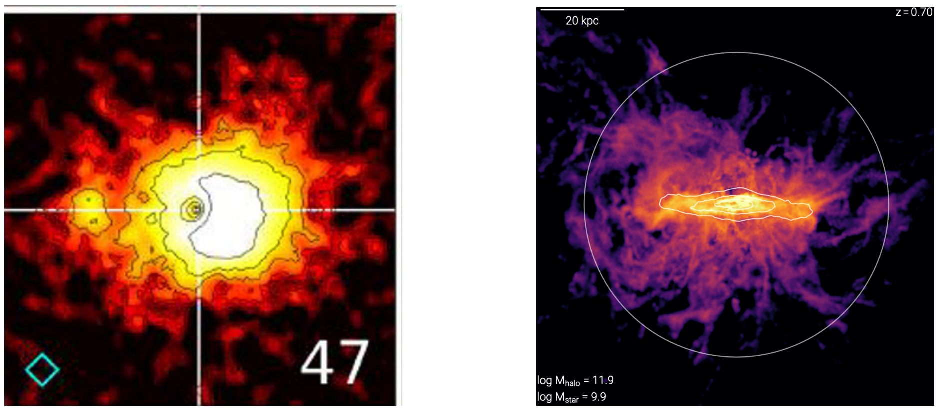

Figure 1.4: **Left:** The Lyα emission from galactic halo, taken from Arrigoni Battaia et al. (2019). **Right:** The CGM in TNG50 simulations, taken from Nelson et al. (2020).

On the other hand, with the space-based X-ray telescopes (e.g., XMM-Newton, ROSAT and Chandra), the CGM has been studied in X-ray, revealing its hot and diffuse content (Anderson et al., 2016; Kaaret et al., 2020; Das et al., 2021b).

Moreover, with the development of integral field units (IFUs), mapping the halo of galaxies and quasars in emission at high redshifts is now possible. In recent years, several studies have successfully mapped the Lyα[16] emission from the halo of high redshift quasars and galaxies ($z \sim 3-6$, see an example in Figure 1.4, left) with VLT's Multi Unit Spectroscopic Explorer (MUSE) (Cantalupo et al., 2014; Farina et al., 2017; Wisotzki et al., 2018; Arrigoni Battaia et al., 2019). There have also been advances in the emission maps of metals such C IV , Mg II in high redshift galaxies and quasars (Arrigoni Battaia et al., 2015; Burchett et al., 2021) with MUSE and Keck Cosmic Web Imager (KCWI). In the upcoming years, the James Webb Space Telescope (JWST) and the Hobby-Eberly Telescope Dark Energy Experiment (HETDEX) will construct emission maps of the CGM up to high redshifts. Combining absorption line studies with emission maps will be the key goal in the near future to constrain the physical properties of CGM.

1.6.3 CGM in Numerical Simulations

Another field that has taken off is the numerical simulation of galaxy formation and evolution. With advances in physical models of gaseous haloes combined with observation constraints, the CGM has been at the core of next-generation computer simulations. In recent years, there have been several key developments in revealing the multiphase nature

[16] A spectral line in neutral hydrogen atom when electron transitions from the first excited state ($n = 2$) to ground state ($n = 1$), where n is the principal quantum number. The emitted photon has a wavelength of 1215.7 Å. The line can be seen in optical at $z > 2$.

of this gas across a wide range of galactic properties at a broad range of redshifts (see Figure 1.4, right for an example of CGM emission in TNG50 simulation, taken from Nelson et al. (2020)). The key successes and challenges in simulating the CGM have been reviewed in Somerville & Davé (2015) and Naab & Ostriker (2017).

Studies with very high-resolution numerical simulations, have reproduced the observed fraction of hydrogen and metals around different types of galaxies (Oppenheimer et al., 2016; Fielding et al., 2020; Nelson et al., 2020, 2021). Idealized hydrodynamics simulations have revealed the co-existence of cool ($T \sim 10^{4-4.5}$ K) and warm ($T \gtrsim 10^6$ K) gas within the virial radius of the dark matter halo (Nelson et al., 2020; Lochhaas et al., 2021). The simulation predictions and results also agree reasonably well with the observations (Oppenheimer et al., 2016; Nelson et al., 2021). However, there are still many question that remain unanswered, given the limited resolution of current simulations. The 'subgrid models'[17] used to model SFR, feedback processes, formation history, and ISM physics of the galaxies and non-convergence of gas properties are also not very well understood and make simulations more complex and challenging (see Naab & Ostriker, 2017; van de Voort et al., 2019, for details). Multiple physical models have been proposed to explain the results, such as (i) formation of cool gas from hot, diffuse gas due to thermal instabilities (McCourt et al., 2012; Sharma et al., 2012) or (ii) mixing of hot gas via galactic outflows (Fielding et al., 2020) or formation of cold filaments via condensation of hot gas driven by AGN outflows (Qiu et al., 2020).

However, these simulations have provided us with deep insights into the nature of this complex multiphase gas. In the near future, one key goal is to perform a detailed comparative study between simulations and observations to push our understanding further in this field. In the next sections, we focus on quasar absorption line studies only as that is the main topic of this thesis.

1.7 What do these absorbers tell us?

As described earlier, absorption line studies are currently one of the most valuable techniques to study the gaseous envelopes of galaxies up to very high redshifts. In this section, we will discuss how absorption measurements can be used to get constraints on the physical properties of the CGM, such as its temperature, density and kinematics.

The quasar absorbers detected in the foreground galaxies, the CGM or the IGM, can trace different phases, densities and temperatures of the gas. In Figure 1.5 (taken from Tumlinson et al. (2017)), we show the phase diagram ($T-n$, temperature vs density) of the halo gas (within $r_{\rm vir}$) at $z = 0$ from the EAGLE[18] simulation (Oppenheimer et al., 2016). This diagram is created using Photoionization Equilibrium (PIE, needed to account for low and intermediate ions) and Collisional Ionization Equilibrium (CIE, needed for highly

[17] Given the limited resolution of numerical simulations, it is generally impossible to model many small-scale physical processes using first principles. Therefore simulations implement sub-grid models for those processes based on the theoretical and numerical framework of those problems.

[18] Evolution and Assembly of GaLaxies and their Environments

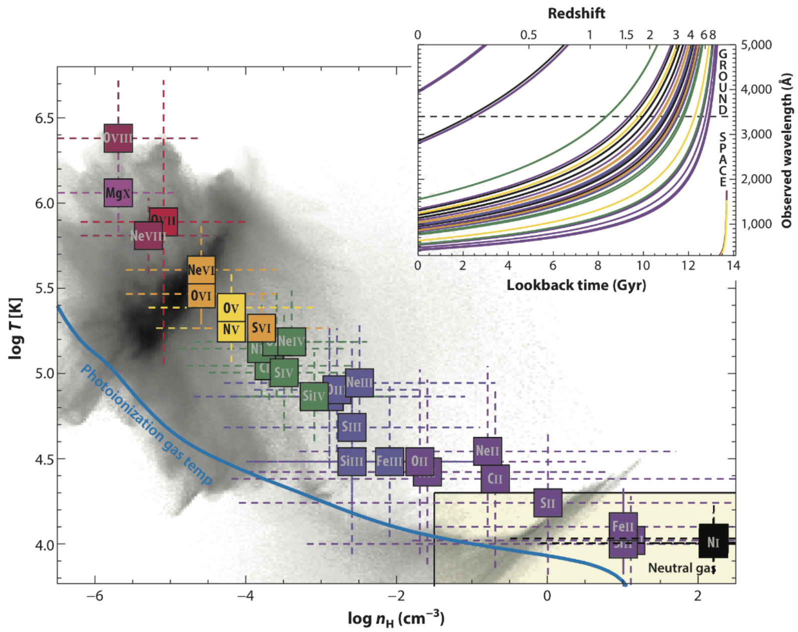

Figure 1.5: The metal absorption lines present in the CGM are detected with a broad range of wavelengths ($19 < \lambda_{\rm rest} < 6000\,\text{Å}$, X-ray to optical). The curve is calculated within $r_{\rm vir}$ of $z = 0$ galaxy in the EAGLE simulation. The vertical axis is the temperature of gas-producing the absorption line. The horizontal axis is the density of neutral hydrogen in the medium. The Figure is taken from Tumlinson et al. (2017). The inset shows the observed wavelengths of absorbers detected with space and ground-based telescopes.

ionized absorbers) models implemented in the CGM simulations. As we can see, there are several metal absorbers present in the CGM, tracing very different temperatures and densities. The metals in high ionization states (e.g Ne VIII, O VII, O VIII) are tracers of diffuse hot phase ($T > 10^6$ K) of the CGM, while metals in low ionization states (e.g, Mg II , Fe II , C II) trace the dense cool phase ($T < 10^{4.5}$ K) of the CGM. Suppose we could observe all the metal species with all their kinematic properties. In that case, it's possible to put strong constraints on the metallicity and baryonic budget of the CGM around all types of galaxies. However, as the inset shows, these transitions occur at a wide range of wavelengths; therefore, we have restricted access to them in redshift space with ground and space based telescopes. Nevertheless, it is pretty clear that the CGM of galaxies is multiphase and hosts a wide variety of gas densities and temperatures with very different spatial distributions, implying the complexity of the co-occurring physical processes.

After the pioneering first works of absorber-CGM cross-correlation studies of Bergeron (1986) and Steidel & Sargent (1992), the field has grown tremendously in the last 30 years. The most commonly observed metals, such as Mg, C, N, O, and Si ranging from optical to UV to X-ray, have provided a plethora of information about the gaseous component of the CGM. These studies have investigated the physical relationships between galaxies and the CGM, covering different types of galaxies over a broad range of redshifts (Steidel et al.,

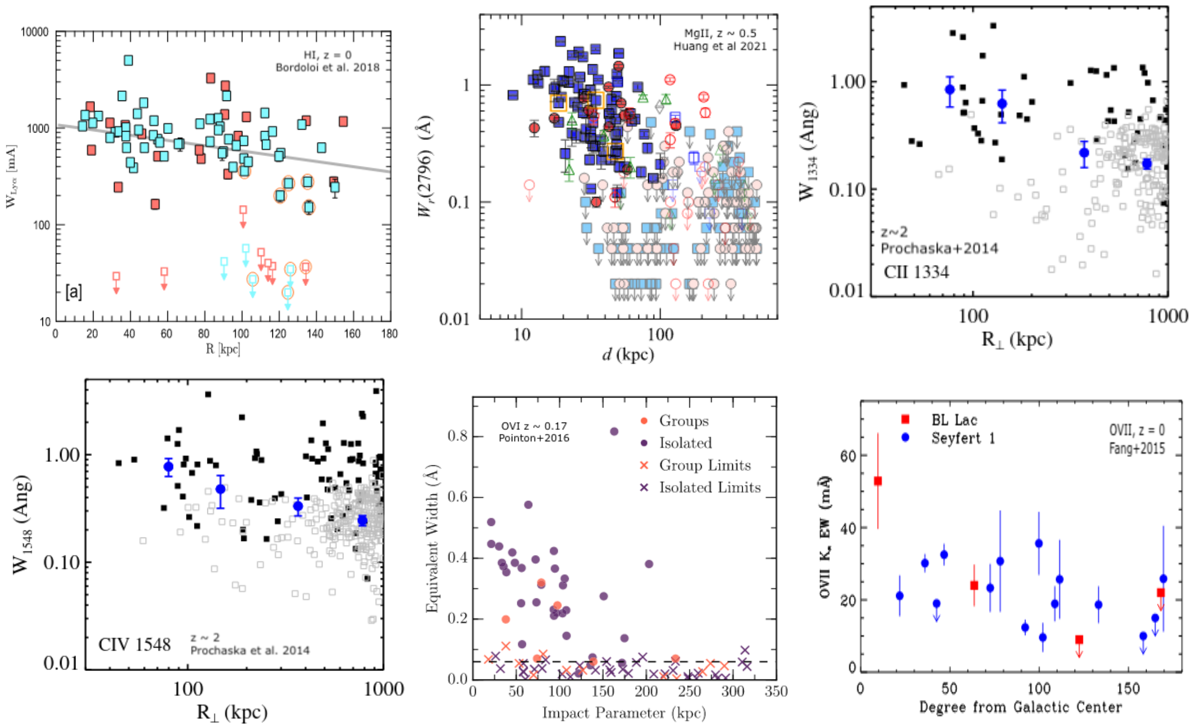

Figure 1.6: Rest Equivalent Widths of hydrogen and metal absorbers in the CGM of galaxies as a function of projected distance or impact parameter. The atoms and ions trace different phases of the CGM. The neutral hydrogen, H I (top left, Bordoloi et al. (2018)) and low ionized metals, Mg II (top middle, Huang et al. (2021)) and C II (top right, Prochaska et al. (2014)) trace cool phase ($T \sim 10^{4-4.5}$ K) of the CGM. On the other hand, metals in high ionization state, C IV (bottom left, Prochaska et al. (2014), tracing $T > 10^5$ K), O VI (bottom middle, Pointon et al. (2017)) and O VII (bottom right, Fang et al. (2015)) trace warm-hot phase ($T > 10^6$ K) of the CGM in galaxies.

1994; Churchill et al., 2005; Chen et al., 2010; Nielsen et al., 2013). Starting with the CGM of galaxies in the Local Group (Lehner et al., 2015), quasar absorption spectroscopy has been extended to the nearby galaxies up to $z \lesssim 0.1$ (Keeney et al., 2017), galaxies at $0.1 < z < 1.5$ (Steidel et al., 1994; Bouché et al., 2006; Chen et al., 2010; Nielsen et al., 2013; Zhu et al., 2014; Chen, 2017; Burchett et al., 2019; Lan, 2020), gaseous haloes at $2 < z < 3.5$ (Steidel et al., 2010; Rudie et al., 2012; Pieri et al., 2014; Lehner et al., 2014; Nielsen et al., 2020) and the CGM at high redshift $z > 3.5$ (Matejek & Simcoe, 2012; Prochaska et al., 2014; Turner et al., 2016; Chen et al., 2016). These studies have found that the amount of gas varies strongly with distance from galaxies, and can depend on galaxy stellar mass, star formation history, color and shape (Bordoloi et al., 2014; Borthakur et al., 2016; Lopez et al., 2018; Lan & Mo, 2018; Rubin et al., 2018).

In Figure 1.6, we compile the rest equivalent width distribution of neutral hydrogen and several ionized metal absorbers from published surveys. These observations range from UV to optical to X-ray over the past decade. Using HST/COS - Halos and Dwarfs surveys, Bordoloi et al. (2018) characterized the spatial extent of H I (tracer of $T \sim 10^4$ K gas) in the CGM of $z = 0$ galaxies with Lyα absorbers detected in background quasars. They found that the rest EW is anti-correlated with the impact parameter (top left panel) and correlated with the stellar mass of galaxies. In a recent work on the CGM of quiescent galaxies at $z \sim 0.5$, Huang et al. (2021) found a similar decreasing trend of EW with impact parameter for Mg II absorbers (tracer of $T \sim 10^4$ K, top middle panel). They also characterized the dependence of EW on halo radius, SFR and stellar mass of galaxies. Extending absorber-CGM cross-correlation studies to $z \sim 2$ quasars, Prochaska et al. (2014) detected large reservoirs of metal-enriched gas (C II and C IV , top right and bottom left panels) in the CGM of quasars. Note that the anti-correlation of EW with the impact parameter is not very clear. By analyzing HST/COS and archival Far Ultraviolet Spectroscopic Explorer (FUSE) data for O VI absorption lines Pointon et al. (2017) detected strong absorbers at small impact parameters in the CGM of $z \sim 0.17$ galaxies (bottom middle), indicating the hot gas presence in the halo. On the hotter phase of the CGM gas, with XMM-Newton data of O VII absorbers at $z = 0$, Fang et al. (2015) characterized the spatial distribution of $T \sim 10^6$ K gas in the CGM (bottom right). They didn't see a strong anti-correlation between EW and impact parameter, indicating that hot gas is more diffuse in the CGM of nearby galaxies. Pieri et al. (2014) stacked several hundred thousand Ly α absorbers at $2.4 < z < 3.1$ to probe the CGM at high redshift. They estimated the average neutral hydrogen column densities in those systems and found evidence for metallicities near the solar value. These observations have shown that both cold and hot gas are spatially extended up to $200 - 300$ kpc in the galactic halo, hinting that the CGM of galaxies is multiphase in nature. In the next section, we will discuss two metal absorbers namely, Mg II (tracer of $T \sim 10^{4-4.5}$ K gas) and C IV (tracer of $T \sim 10^{5-5.5}$ K gas) in more detail, as they are the most relevant absorbers for this thesis. The basic line parameters of these two absorbers are compiled in Table 1.1.

1.7.1 Atomic Properties of Mg II and C IV

1.7.1.1 Mg II : Mg$^+$

Among all the detected metals, the most extensively examined is magnesium (Mg). It is the ninth most abundant element in the universe (Ishigaki et al., 2012; Hasselquist et al., 2021). The nucleosynthesis in massive stars produces large quantities of Mg, which is expelled into the ISM or the CGM via stellar winds or the supernovae, enriching the overall Mg content in galaxies. The most commonly detected ion in the spectra of background quasars is Mg II (singly ionized Mg atom). The electronic transition between the ground and the first excited state (fine structure splitting) of this ion results in a resonant doublet line with rest-frame wavelengths 2796 Å and 2803 Å.

We show the quantum mechanical energy diagram of this transition in the left panel of Figure 1.7. Given the relatively high abundance of Mg in the universe and the large oscillator strength of this transition ($f_{2796} = 0.608$, $f_{2803} = 0.305$, Kelleher & Podobedova (2008)), it is one of the strongest absorption features that ground-based optical telescopes can detect at modest redshifts ($0.3 \lesssim z \lesssim 2.5$). Based on PIE models and broadening of the line, Mg II absorbers are known to trace low-ionization cold gas ($T \sim 10^{4-4.5}$ K) in the CGM and the intergalactic medium (IGM). The doublet nature with large line separation ($\Delta\lambda = 7.17$ Å) makes it easier to detect in quasar spectra with automated pipelines. The doublet ratio varies from 2 (for unsaturated lines) to 1 (saturated case), and saturation starts for column densities, $\log\,[\mathrm{N/cm^2}] > 12.5$, that corresponds to $EW_{2796} \gtrsim 0.1$ Å. Basic spectral features are also compiled in Table 1.1

1.7.1.2 C IV : C^{3+}

Carbon is one of the most abundant elements after oxygen in the universe (Henning & Salama, 1998). It can be easily detected in quasar spectra as C IV (triply ionized carbon atom). The C IV line is also a doublet with rest-frame wavelengths 1550 Å and 1548 Å.

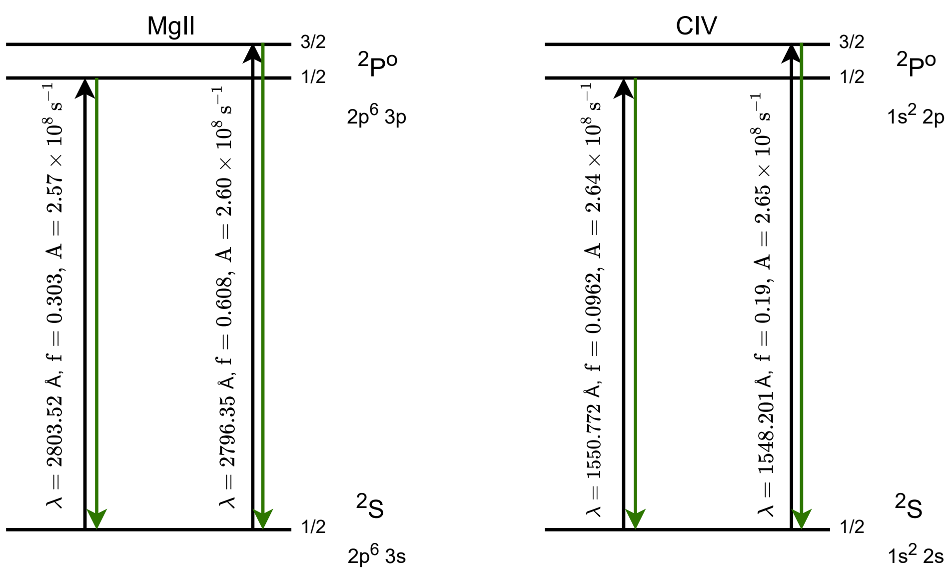

Figure 1.7: **Left:** Energy-level diagram of Mg II $\lambda\lambda\,2796, 2803$ doublet. **Left:** Energy-level diagram of C IV $\lambda\lambda\,1548, 1550$ doublet. The diagram is created based on data compiled by National Institute of Standards and Technology (NIST) Atomic Spectra Data system (Kramida et al., 2021).

We show the quantum mechanical energy diagram for C IV line in the right Figure 1.7. Given its high abundance and the large oscillator strength ($f_{1548} = 0.19$, $f_{1550} = 0.096$, Kelleher & Podobedova (2008)), it can easily be detected with ground-based optical telescopes at $1.5 \lesssim z \lesssim 4.5$). Based on CIE models and broadening of the line, C IV absorbers trace high-ionization warm gas ($T \sim 10^{5-5.5}$ K) in the CGM and the intergalactic medium (IGM). The doublet nature (though the lines are narrowly separated $\Delta\lambda = 2.57$ Å) enables us to search automatically for them in quasar spectra. For C IV also, the doublet ratio

Table 1.1: Line parameters of Mg II and C IV doublets, The rest-frame wavelengths (λ) are in Å. The oscillator strengths (f) are taken from National Institute of Standards and Technology (NIST) atomic data compiled by Kramida et al. (2021).

Ion	λ_1 [Å]	λ_2 [Å]	$\Delta\lambda$ [Å]	f_{λ_1}	f_{λ_1}
Mg II (Mg^+)	2796.35	2803.52	7.17	0.608	0.305
C IV (C^{3+})	1548.20	1550.77	2.57	0.192	0.096

varies from 2 (for unsaturated lines) to 1 (saturated case). C IV starts saturating at column densities, $\log\,[\mathrm{N/cm^2}] \gtrsim 14$, that corresponds to the equivalent width, $EW_{1548} \gtrsim 0.4$ Å (Puech et al., 2018). We compile the basic atomic data for C IV in Table 1.1.

1.8 Mg II Surveys

Several Mg II surveys have been performed to constrain their distribution and evolution in the universe as a function of redshift. Combining the earliest works by Sargent et al. (1988) and Steidel & Sargent (1992) with more recent large surveys of Mg II with current telescopes such as SDSS and Keck, it has been established that dN/dW (number of absorber per unit equivalent width) of strong Mg II systems ($\mathrm{EW}^{2796}_{\mathrm{rest}} \geq 0.3$ Å) can be best described by an exponential distribution (Figure 1.8, left) between redshifts $0.3 < z < 2.5$ (Nestor et al., 2005; Zhu & Ménard, 2013a). Similarly, the differential number of absorbers per unit redshift (dN/dz) can be best-fit by a combination of power-law and exponential function over a broad range of redshifts (Zhu & Ménard, 2013a). The rest equivalent width of Mg II absorbers and dN/dz increase with redshift up to $z \sim 1.5$ and then declines at higher redshifts (Figure 1.8, right, Matejek & Simcoe (2012); Zhu & Ménard (2013a)), following a very similar trend as the cosmic star formation history (Madau & Dickinson, 2014). These findings have triggered an investigation into the connection of Mg II absorbers with global star formation processes.

On the other hand, for weak systems ($\mathrm{EW}^{2796}_{\mathrm{rest}} \leq 0.3$ Å), the number of absorbers per unit equivalent width (dN/dW) can only be described by a power-law (Churchill et al., 1999; Narayanan et al., 2007). In contrast to the strong systems, the dN/dz evolves weakly with redshift for weak absorbers (Steidel & Sargent, 1992; Churchill et al., 1999; Narayanan et al., 2007). These differences in the properties of weak and strong absorbers indicate that different astrophysical processes might be responsible for driving the evolution of absorber systems with different strengths, and overall cold gas content evolves with the environment and age of the universe.

Many studies have also investigated the connection of these absorbers with the CGM of galaxies as functions of galactic properties such as stellar mass, halo mass, SFR, luminosity, redshift and environment. However, the origin of Mg II absorbers in the CGM remains an open question, though observations have pointed toward several competing mechanisms. By cross-correlating Mg II absorbers detected in early SDSS data (DR3) with luminous red

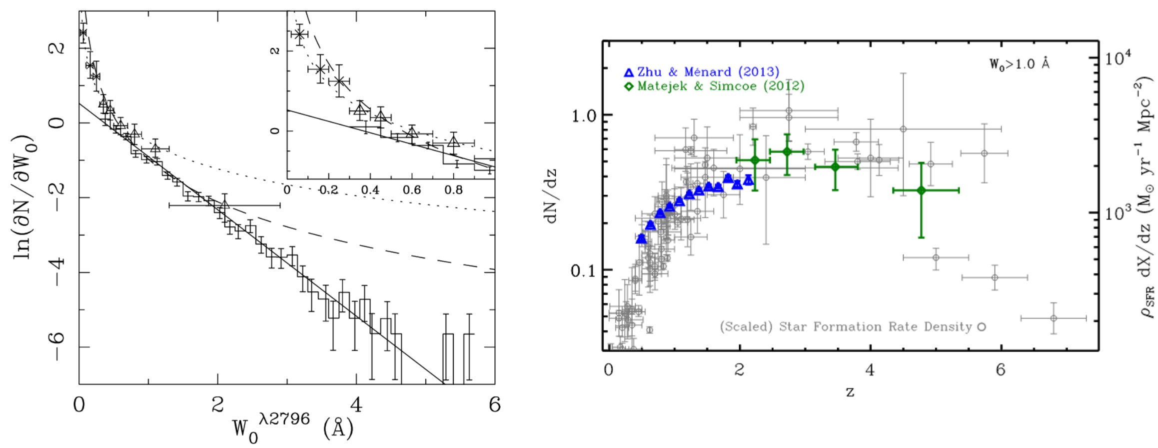

Figure 1.8: **Left:** Rest equivalent width (EW_{2796}) distribution of Mg II absorbers (taken from Nestor et al. (2005)). The solid black line is the best-fit exponential function, and the dashed line is the power-law from Steidel & Sargent (1992) (triangles are also from the same work). The crosses are from Churchill et al. (1999). The power law describes the weak absorbers very well, while exponential is better for strong systems. **Right:** The incidence rate of Mg II absorber per unit redshift for strong Mg II absorbers ($EW_{2796} > 1$ Å) as a function of redshift (taken from Zhu & Ménard (2013a)). Also shown is the scaled cosmic star formation rate using data compiled in Madau & Dickinson (2014).

galaxies (LRGs)[19] from the same data release, Bouché et al. (2006) studied the correlation between Mg II EW and halo mass of galaxies. They found a strong anti-correlation between these two quantities, indicating that Mg II clouds are not virialized in LRG's haloes. In contrast, Churchill et al. (2013) found no such anti-correlation between these two quantities and argued in support of a self-similar CGM across a wide range of halo mass.

There have also been statistical studies investigating the nature of the CGM around star-forming versus quiescent galaxies (Bordoloi et al., 2011; Ménard et al., 2011; Zhu et al., 2014; Peek et al., 2015; Huang et al., 2016; Lan, 2020; Huang et al., 2021). The sample sizes have varied from a few hundreds (in high resolution spectra, e.g., Churchill et al. (2020)) to several thousands (in large spectroscopic surveys, e.g., Zhu & Ménard (2013a)) for metal absorbers. Zhu et al. (2014) stacked SDSS DR7 quasars (Schneider et al., 2010) in the rest-frame of LRGs from the Baryonic Oscillation Spectroscopic Survey (BOSS) program of SDSS DR11 (Dawson et al., 2013). They found a slope change in mean equivalent width and surface mass density of Mg II absorbers (EW_{2796}) at a projected distance, $D_{\rm proj} \sim 1$ Mpc (see Figure 1.9, left). They modelled this slope change with a 2-halo term to account for contribution of multiple haloes to the metal absorption. The galaxy's halo dominates the gas properties at smaller distances, and neighbouring haloes contribute to

[19]The LRGs have very low or no star formation activity, and a significant fraction of these galaxies do not show strong emission lines in their spectra. The presence of old stellar populations makes them look optically red. They are typically very luminous and massive ($M_{\rm halo} \sim 10^{13.5}$ M$_\odot$) and can be detected up to $z \sim 0.5$ (Padmanabhan et al., 2012).

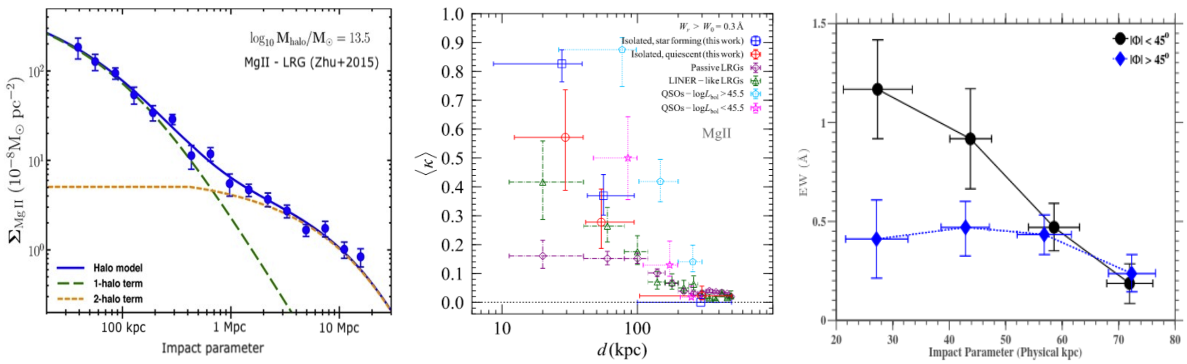

Figure 1.9: Mg II surface mass density around SDSS LRGs (left, Zhu et al. (2015)), covering fraction around different types of galaxies (middle, Huang et al. (2021)) and rest equivalent width (right, Bordoloi et al. (2011)) as a function of impact parameter from central galaxy.

the observed distribution at large distances. Recently, Zu (2021) performed a more careful analysis using the redshift-distortion method with the latest DR16 data for galaxies and quasars and found a similar slope change. Extending this cross-correlation to the extended Baryonic Oscillation Spectroscopic Survey (eBOSS, Dawson et al. (2016)) emission-line galaxies (ELGs)[20], Lan & Mo (2018), found that mean equivalent width of Mg II absorbers is $2-3$ times higher around ELGs compared to LRGs within $\sim 50-100$ kpc. The mean Mg II absorption in the CGM of ELGs positively correlates with their SFR, indicating a connection to the stellar outflows. They also estimated the total mass of cool gas traced by Mg II absorbers around star-forming and quiescent galaxies and found ~ 2 times more neutral hydrogen in ELGs than LRGs within $r_{\rm vir}$. It is observed that the mean covering factor[21] of Mg II absorbers varies strongly with galaxy type. The Mg II absorber have higher covering fractions around star-forming galaxies than passive or quiescent galaxies (Huang et al., 2021) (see Figure 1.9, middle). Similarly, Zibetti et al. (2005, 2007) performed image stacking to study the photometric properties of Mg II systems in SDSS and found that weak and strong absorbers are originated in different types of galaxies. Although the cool CGM (traced by Mg II absorbers) of galaxies has been studied extensively, few studies have explored the nature and origin of Mg II absorbers in galaxy clusters (Lopez et al., 2008; Padilla et al., 2009; Lee et al., 2021) or groups (Nielsen et al., 2018; Dutta et al., 2020). Lopez et al. (2008) found that the number density of strong absorbers ($EW_{2796} > 1$ Å) is higher around clusters than in field galaxies, possibly due to the overdensity of galaxies. Similar results for covering fractions have been found (Dutta et al., 2021b). However, we have not explored the impact of galaxy environment on its CGM in great detail yet.

To explore the possible origin of Mg II absorbers in the galactic halo, Bordoloi et al.

[20] The characteristic features of ELG spectra are strong gas emission lines such as [O II] λ3727 [O III] λ5007 and Hβ, due to high star formation activity. They are less massive than LRGs (Raichoor et al., 2021). Their typical halo mass is $M_{\rm haloa} \sim 10^{12}\,{\rm M}_\odot$.

[21] The probability of detecting Mg II absorbers in a given sightline that passes through the galactic atmosphere.

(2011) measured Mg II equivalent widths along major and minor axes of galaxies using VLT's zCOSMOS survey. They found an enhanced absorption along the minor axis of the galaxies (see Figure 1.9 right), indicating that Mg II is associated with outflows, as outflowing material is mainly aligned with the minor axis of galaxies where it feels less resistance (Péroux & Howk, 2020). Gas metallicity is also higher along the minor axes of galaxies, pointing towards a galactic outflow origin of Mg II absorbers in star-forming or emission-line galaxies (ELGs) (Péroux et al., 2020a). On the other hand, spectroscopic samples of absorbers and galaxies also allow the characterization of line-of-sight relative kinematics of galaxies and absorbing clouds (Tremonti et al., 2007). For example, using samples of quiescent and star-forming galaxies from BOSS and eBOSS, Lan & Mo (2018) showed that Mg II absorbers have dispersion similar to the virial velocity of dark matter halo for star-forming galaxies. At the same time, the motion is suppressed in passive galaxies, indicating different mechanisms at work. The consensus now is that the origin of cool circumgalactic gas is very different in star-forming and passive galaxies, and different theoretical frameworks would be needed to model their CGM.

1.9 C IV Surveys

Previously, large C IV surveys have found that equivalent widths or column densities follow an exponential or a power-law distributions (with slopes varying from $\alpha \approx -1.8$ to -2.3) (Songaila, 2001; D'Odorico et al., 2010; Cooksey et al., 2010, 2013; Boksenberg & Sargent, 2015). In addition, studies have shown that the C IV mass density relative to the universe's critical density, $\Omega_{\mathrm{C\,IV}}$ increases with redshift (see Figure 1.10, left) from $z = 5 \rightarrow 0$ (D'Odorico et al., 2010; Cooksey et al., 2013). Moreover, the number of C IV systems per unit comoving path length (dN/dX) also increases from $z = 5.5 \rightarrow 0$, though $\Omega_{\mathrm{C\,IV}}$ increases steeper than dN/dX (Cooksey et al., 2010; Simcoe et al., 2011; Burchett et al., 2015; Codoreanu et al., 2018; Hasan et al., 2020).

It is important to constrain the cosmic C IV densities in the IGM over a broad range of redshift. In addition, C IV absorption in quasar spectra also allows us to study the warm phase in the CGM of galaxies and quasars. The optical survey of galaxies at $z > 1.5$ is minimal because galaxies appear very faint at far distances and fall below the flux sensitivity of even the largest telescopes[22]. However, with UV spectroscopy from HST observations, there have been several studies constraining the C IV properties in the local universe.

In an early work, Chen et al. (2001) studied the C IV absorber distribution in $z < 1$ galaxies. Their analysis revealed a large reservoir of C IV absorbers up to ~ 200 kpc in the galactic halo. Using HST's Cosmic Origin Spectrograph (COS)-Halos survey, Bordoloi et al. (2014) analyzed the C IV properties as a function of galaxy type at $z < 0.1$. They

[22] Also, the most exciting and valuable spectral lines in galaxies are at optical or near-infrared (NIR), which redshifts further to infrared (IR) at these redshifts. They are challenging to be seen from the ground-based telescopes because the earth's atmosphere blocks most of the infrared. However, the James Webb Space Telescope (JWST) will detect many galaxies at these redshifts with its infrared spectrometer.

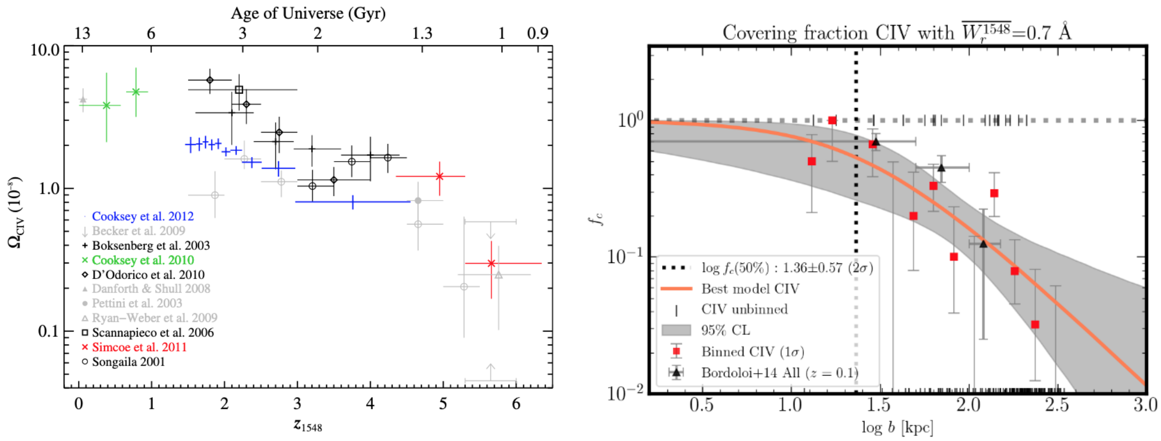

Figure 1.10: **Left:** $\Omega_{\text{C IV}}$ as a function of redshift (taken from (Cooksey et al., 2013)). The plot also includes data from several previous studies (shown in different colours). The C IV mass density increases smoothly from $z = 4$ to $z = 1.5$. **Right:** Covering fraction of C IV absorbers in the CGM of galaxies at $z \sim 1 - 1.5$, taken from Schroetter et al. (2021). The solid line is the best-fit model based on a Bayesian logistic regression method developed by authors. The black triangles are the results for $z = 0$ galaxies from Bordoloi et al. (2014).

found a higher covering factor (~ 2 times) of C IV absorbers around star-forming galaxies than red galaxies up to $D_{\text{proj}} \lesssim 0.5\, r_{\text{vir}}$. Comparison of these low redshift C IV absorption with simulations (Davé et al., 2011, 2013) shows that models based on energy-driven galactic winds can reproduce the C IV properties in the CGM fairly well. In addition, using the same COS-Halos Burchett et al. (2016) characterized the nature of C IV absorbers in low redshift ($0.0015 < z < 0.015$) galaxies and groups. They found that C IV absorbers are frequently detected in galaxies with $M_\star > 10^{9.5}\,\mathrm{M}_\odot$, while detection is lower in galaxies that reside in groups.

Extending these studies to high redshifts, to compare metal absorber properties in the CGM, Schroetter et al. (2021) estimated covering fractions of C IV and Mg II absorbers using the MUSE MEGAFLOW survey for $z \sim 1.2$ galaxies (see Figure 1.10, right). They also developed a Bayesian logistic regression method and modelled metal covering fraction as a function of projected distance from galactic center. Their analysis showed that C IV halo sizes are slightly smaller than Mg II haloes, though the error bars are large due to small sample sizes. Given the limited sample sizes of high redshift galaxies, studies have explored the C IV cross-correlation with foreground quasars. The AGN activity powered by a supermassive black hole (SMBH) at the centre of quasars can expel metals with very high velocities (several thousands $\mathrm{km\,s^{-1}}$) up to large distances (Perrotta et al., 2019; Nelson et al., 2020). In addition, they can also ionize the halo gas (Hennawi & Prochaska, 2007; Farina et al., 2013). Hence, quasar haloes are expected to have large reservoirs of warm metal gas. Previous studies have looked into these and tried to constrain the metal distribution in quasar haloes.

Studies have shown that the doublet ratio (EW_{1548}/EW_{1550}) of C IV absorbers decreases with redshift, indicating an overall increase in mean C IV density over time (Péroux et al., 2004; Songaila, 2005). On the other hand, the cross-correlation studies of C IV absorbers with $z \sim 2$ quasars have revealed a large reservoir of metal gas in their CGM (see bottom left panel of Figure 1.6), where the covering factor is as high as $60 - 80\,\%$ within the $r_{\rm vir}$ of quasar halo (Prochaska et al., 2014; Landoni et al., 2016). Based on quasar - C IV two-point cross-correlation analysis and bias estimation, it has been predicted that C IV absorbers are predominantly tracing the massive $M_{\rm halo} = 10^{12-12.5}\,{\rm M}_\odot$ haloes at $z \sim 2$ (Vikas et al., 2013; Prochaska et al., 2014; Gontcho A Gontcho et al., 2018). In addition, Finlator et al. (2020) found that the galaxy-absorber cross-correlation function ($\xi_{\rm gal-abs}(r)$) decreases with distance from the galaxy and positively correlates with the luminosity of $z > 5$ galaxies.

1.10 Goals of this thesis

This thesis aims to expand our understanding of cool and warm gas in the CGM and its connection to galaxies' properties and environment using the latest and largest imaging and spectroscopic data of quasars and galaxies. I will interpret the findings in the context of the CGM and its connection to galaxy formation and evolution.

The thesis is based on two published papers (Anand et al., 2021, 2022) and one *in preparation* work. I organize this thesis as follows: In the first part of Chapter 2, I will develop a novel automated method to estimate the quasar continuum and search for Mg II doublets in their spectra. I will then run our pipeline on the final sample of quasars from the SDSS Data Release 16 (DR16) and compile the largest and most extensive Mg II catalogue to date. After analyzing the properties of individual absorbers, I will perform a galaxy-absorber cross-correlation study in the second half of the chapter and present the most extensive CGM-Mg II cross-correlation study so far. Using this very high S/N CGM-galaxy correlation, I will also characterize the scales in the spatial distribution of cool gas traced by Mg II absorbers in the CGM of star-forming and passive galaxies between redshift $0.5 < z < 1$ in SDSS. I will then search for clues to the origin of these absorbers by studying their kinematics and connection to the galactic properties.

Chapter 3 will extend this cool gas cross-correlation to the galaxy clusters and present the most extensive search for Mg II absorbers in $z \sim 0.5$ clusters so far. I will study the spatial distribution of cool gas traced by weak and strong Mg II absorbers in the halo of clusters. To understand the origin of cool gas in a cluster environment, I will also study the association between Mg II absorber and cluster member galaxy properties. Finally, I will constrain the relative kinematics of absorbers in cluster halo and compare them with the motion of cool gas around galaxies.

Chapter 4 will search for the C IV absorbers in quasar spectra using our automated pipeline developed in chapter 2. The goal will be to compile an extensive Mg II selected C IV absorber catalogue based on the latest SDSS DR16 quasar sample. I will characterize a few statistical properties of C IV absorbers. Then I will cross-correlate C IV and Mg II absorbers to the CGM of quasars at $z \sim 2$ and study the cool (Mg II absorbers) and warm

(C IV absorbers) phase of the CGM. By studying their properties and kinematics in quasar haloes and comparing them with galaxies at low redshifts, I will explore the differences in absorber properties. Overall, this thesis will be an effort to present a holistic picture of this multiphase CGM in galaxies and clusters at a wide range of redshifts.

Light brings us the news of the Universe.

William Bragg

Chapter 2

The cool circumgalactic medium in the Sloan Digital Sky Survey (SDSS) galaxies

The contents of this chapter is based on Anand et al. (2021) published in Monthly Notices of Royal Astronomical Society, 504, 65

2.1 Introduction

As described in the previous chapter, galaxy-absorber pair studies have been at the centre of our understanding of the nature of the circumgalactic medium (CGM). Since the first discovery of a galaxy-metal absorber pair by Bergeron (1986), the quasar absorption line studies have provided us invaluable insights into the gas flows around galaxies. More recently, there have been several galaxy-absorber pair studies investigating the physical relationships between galaxies and the CGM, covering different types of galaxies over a broad range of redshifts (Steidel et al., 1994; Churchill et al., 2005; Chen et al., 2010; Nielsen et al., 2013). These studies have found that the amount of gas varies strongly with distance from galaxies, and its properties and environment (Bordoloi et al., 2014; Borthakur et al., 2016; Lopez et al., 2018; Lan & Mo, 2018; Rubin et al., 2018).

One of the most examined metal lines in the CGM of galaxies is the Mg II $\lambda\lambda 2796, 2803$ doublet. It is one of the strongest absorption features that can be detected by ground based optical telescopes at modest redshifts ($0.3 \lesssim z \lesssim 2.5$). Mg II absorbers are tracers of low-ionization cold gas ($\lesssim 10^4$ K) in the CGM and in the intergalactic medium (IGM). Several Mg II surveys have been performed to constrain the temperature and density profile of low-ionization gas in the CGM along with the statistical properties of weak ($\mathrm{EW}_{\mathrm{rest}}^{2796} \leq 0.3$ Å) and strong ($\mathrm{EW}_{\mathrm{rest}}^{2796} > 0.3$ Å) Mg II absorbers (Weymann et al., 1979; Tytler et al., 1987; Sargent et al., 1988; Caulet, 1989; Steidel & Sargent, 1992; Churchill et al., 1999, 2000; York et al., 2006; Quider et al., 2011; Zhu & Ménard, 2013b).

On the other hand, most similar to our present work, Zhu et al. (2014) performed a detailed statistical galaxy-absorber pair study with DR7 Mg II absorbers (Zhu & Ménard, 2013a) and Luminous Red Galaxies (LRGs) from DR11 of SDSS (Dawson et al., 2013). They estimated average cool surface densities as traced by Mg II absorbers around massive galaxies out to projected radii of 10 Mpc. They observed a change of slope on scales of 1 Mpc, consistent with the expected gas distribution in the parent halo together with gas outside the halo. Extending this analysis, Lan et al. (2014) found that both LRGs and ELGs have high covering fractions of cold gas ($\sim$ 1 percent) even at impact parameters of 500 kpc.

Recently, Lan & Mo (2018) have proceeded with a larger dataset of several thousand galaxy-quasar pairs using the Extended baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) in the Sloan Digital Sky Survey IV (SDSS-IV; Blanton et al. 2017) finding an anisotropic metal absorption distribution around emission line galaxies (ELGs). They also observed the amount of cool gas to be different around ELGs and LRGs. There have also been statistical studies investigating the nature of the CGM around star-forming versus quiescent galaxies (Bordoloi et al., 2011; Ménard et al., 2011; Peek et al., 2015; Huang et al., 2016; Lan, 2020; Huang et al., 2021). It is observed that the mean covering fraction of Mg II absorbers varies strongly with galaxy type. Apart from these galaxy-CGM pair studies, these large spectroscopic surveys also enable us to perform clustering and statistical cross-correlations studies exclusively with absorbers (Nestor et al., 2005, 2008; Quider et al., 2011; Zhu & Ménard, 2013b).

The distribution of velocity separations between galaxies and absorbers allows the kinematics of cold gas around galaxies to be investigated (Tremonti et al., 2007). For example, the clustering of Mg II systems around BOSS LRGs ($\mathrm{M}_\star \sim 10^{11.5}\,\mathrm{M}_\odot$) shows an excess of Mg II up to $R_p = 20$ Mpc, as well as relative velocities of $\Delta \mathrm{v} \sim 10000\ \mathrm{km\,s^{-1}}$within a projected distance of $\lesssim 800$ kpc (Kauffmann et al., 2017). The implication is that cool circumgalactic gas can originate in either supernovae or supermassive black hole driven outflows, as well as due to infall and accretion.

Since the first light of Sloan Digital Sky Survey (SDSS; York et al. 2000) more than a million quasar spectra have been observed, and these can be searched for intervening absorber systems. Several detection (both automated and visual) algorithms have been developed to find Mg II absorption lines in SDSS quasars (Nestor et al., 2005; York et al., 2006; Bouché et al., 2006; Prochter et al., 2006; Lundgren et al., 2009; Quider et al., 2011; Zhu & Ménard, 2013a). With the continuous release of ever larger datasets, it is increasingly important to develop efficient automated pipelines to detect absorber systems in background sources, and thereby study gas absorption as a function of galaxy properties.

In this chapter, we develop an automated continuum estimation and absorption detection pipeline using the approach of Zhu & Ménard (2013a) as our starting point. We run our pipeline on the full DR16 quasar sample and study the statistical variation of absorbers as well as galaxy-absorber pairs to measure the physical properties of CGM gas as a function of galaxy mass, star formation activity, impact parameter, and redshift.

The chapter is divided into five sections: Section 2.2 introduces the observational data and describes our methods for continuum estimation and automatic absorption detection. We explore the resulting Mg II absorber catalogue and its statistical properties in Section

2.3. In Section 2.4 we analyse the galaxy-centric CGM properties of ELGs and LRGs. Finally we discuss the implications of our findings in Section 2.5 and summarize the results in Section 2.6.

2.2 Methods

2.2.1 Quasar catalogue

The latest Data Release 16 (DR16) quasar catalogue[1] compiled by Lyke et al. (2020) was released in late July 2020 as part of Value Added catalogue (VAC) of latest SDSS DR16 (Ahumada et al., 2020). Each SDSS data release is cumulative and includes all the objects observed in any previous release. The latest DR16 quasar catalogue contains 750,414 quasars. However, for our analysis we download the spectra of all the objects that are classified as QSOs in SDSS database[2]. It includes 983,317 objects identified as quasars (QSOs) with $0 < z < 7$ ($\sim$ 3,000 QSOs with $z > 4.8$).

In order to create robust continua of quasars we use the previously available quasar catalogue, DR14Q (Pâris et al., 2018; Abolfathi et al., 2018). It contains all quasars that were observed in SDSS-I/II/III (York et al., 2000; Schneider et al., 2010; Eisenstein et al., 2011; Dawson et al., 2013) and SDSS-IV/eBOSS (Dawson et al., 2016; Blanton et al., 2017) and classified as quasars with SDSS pipeline (Bolton et al., 2012). In order to compile a complete and pure catalogue, pipeline classified quasars were inspected visually to remove failed or uncertain classifications. The completeness and purity within a given target selection are as high as 99.5 percent in the catalogue. This high fidelity sample enables us to construct eigenspectra, as described below.

2.2.2 Continuum Estimation

The estimation of a robust continuum for a quasar is a crucial step in detecting absorbers in its spectrum. Among the empirical methods to model the quasar continuum the most common is principal component analysis (PCA). Though standard PCA is quite powerful, it does not use any information about the known uncertainties and missing data and assigns components to the variations that are purely due to errors. For more flexibility to work with known uncertainties and missing data, another powerful dimension reduction technique called Non-negative Matrix Factorization (NMF; Lee & Seung 1999) is used. We describe the details of NMF below.

2.2.2.1 Non-negative Matrix Factorization (NMF)

As the name implies, NMF factorizes a large non-negative matrix (X) into two much smaller non-negative matrices, minimizing the error ($||X - WH||_{\mathrm{F}}$, F denotes the Frobenius norm

[1] https://www.sdss.org/dr16/algorithms/qso_catalog/
[2] https://dr16.sdss.org/optical/spectrum/search

of the difference matrix), one of which can be thought as the basis (W) and other as the coefficient (H) (Zhu, 2016). Mathematically,

$$X \approx WH\,;\, X,\, H,\, W\, \geq 0 \tag{2.1}$$

In contrast to PCA, NMF naturally handles the data uncertainties because of its unique iterative update rules (Zhu, 2016)[3] that guarantee non-increasing approximate error. Even if the matrices are weighted by their data uncertainties (say the inverse variance matrix) it is guaranteed to converge to a local minimum (Blanton & Roweis, 2007; Zhu, 2016). Another advantage of NMF over traditional PCA is the ability to interpret NMF eigenpsectra in more physical manner, as it picks up the intrinsic features (prominent emission lines) of quasar spectra. Then each column of the original matrix can be approximately reconstructed as the linear combination of basis (or, eigenspectra, W), and corresponding coefficients (or, eigenvalues, H). We construct the NMF eigenspectra for DR14 quasars using the approach adopted by Zhu & Ménard (2013a), dividing the continuum estimation into four main steps.

2.2.2.2 Re-binning of Spectral Flux

First, the spectral flux must be re-binned on a common rest-frame wavelength array. Each SDSS spectrum is an observed frame spectrum and has a common grid i.e. $\log\lambda_{\rm i+1} - \log\lambda_{\rm i} = 0.0001$Å implying $\Delta v = 69\,{\rm km\,s^{-1}}$. We define a common rest-frame wavelength array that can incorporate all quasars in each redshift range defined in Table 2.1. We shift each quasar into the rest-frame and compute the total coadded fluxes (provided in each spectrum) using linear interpolation. While doing so we take care of the boundary values by restricting the interpolation to the edges of the redshift range appropriate for each quasar.

2.2.2.3 Flux Normalization

Second, we renormalize the flux for each quasar to a common scale at a fixed rest-frame wavelength, adopting the approach taken by Zhu & Ménard (2013a). As we move through different rest-frame wavelengths as a function of $z_{\rm QSO}$, a uniform normalization can not be applied due to the wavelength coverage of SDSS. Thus we select four different rest wavelength ranges and divide the observed spectra by the mean flux within this range (see column 1 of Table 2.1). The choices in the first and second columns of Table 2.1 are adapted from Zhu & Ménard (2013a) which are based on the median quasar flux distribution given in Vanden Berk et al. (2001) (see Figure A.1 for more details). The aim of selecting these four specific wavelength regions is to avoid any emission lines from the quasar itself entering the spectral regions that are searched for absorption systems. We want to normalize the

[3]The two update rules for H and W are: $H \leftarrow H \circ \frac{W^T X}{W^{\rm T} W H}$ and $W \leftarrow W \circ \frac{X H^T}{W H H^{\rm T}}$, where $\circ$ represents element-wise product and $\frac{()}{()}$ represents element-wise division. Even if X and WH are weighted by the inverse variance matrix (V): $X \rightarrow V \circ X$ and $WH \rightarrow V \circ (WH)$, the updates rule are guaranteed to approach local minima.

Table 2.1: Details of our normalization scheme used in the NMF fitting for DR16 QSOs, using eigenspectra constructed from DR14 QSOs.

Normalization Wavelength Range (QSO rest-frame)	Eigenspectra Construction Redshift Range ($z_{\rm QSO}$)	Continuum Fitting Range ($z_{\rm QSO}$)
4150-4250 Å	$z < 1.0$	$z < 0.97$
3020-3100 Å	$0.4 < z < 1.8$	$0.97 < z < 1.49$
2150-2250 Å	$0.8 < z < 2.8$	$1.49 < z < 2.10$
1420-1500 Å	$2.0 < z < 4.8$	$2.10 < z < 4.8$

quasar spectrum by the mean of flux in the region where the spectrum is as featureless as possible, masking missing or bad pixels as necessary.

2.2.2.4 NMF Eigenspectra Construction

Third, we run the NMF algorithm on our dataset. Note that, for the four redshift ranges, NMF can be run independently and the coefficient matrix (H) and eigenspectra (W) can be computed. The presence of spectroscopic artefacts can strongly affect the NMF eigenspectra which are tackled with an iterative approach. To avoid the peculiar objects we adopt the following approach: after constructing the set of eigenspectra we remove those quasars for which the coefficients differ by $\geq 5\sigma$ from the mean eigenvalues and repeat the NMF procedure until no outliers are found. The code converges in $\sim 15-20$ iterations. We always initialize the W and H matrices by the results obtained in the previous iteration to make convergence faster.

2.2.2.5 Continuum Fitting with NMF Eigenspectra

Finally, we use the DR14 quasar eigenspectra (an example eigenspectra set is shown in Figure 2.1) to fit the continuum of DR16 quasars, using the eigenspectra set derived from the redshift range whose median redshift is closest to quasar (see the third column of Table 2.1). This guarantees that the continuum of each quasar is modelled with a set of eigenspectra that is constructed from the maximal number of quasars covering the same wavelength range. Then we compute the residual, defined as the ratio of normalized flux to the NMF continuum. We smooth this residual with a median filter of kernel size = 141 pixels (~ 8 times the typical width of Mg II lines) to remove intermediate-scale fluctuations. We then derive a second NMF continuum as the product of the first NMF continuum and median filtered residual. We again compute the residual w.r.t this second NMF continuum and smooth the residual with the median filter of kernel size = 71 (~ 4 times the typical width of Mg II lines) to remove small-scale fluctuations. The final NMF continuum is then computed by multiplying the second NMF continuum with this median filtered residual.

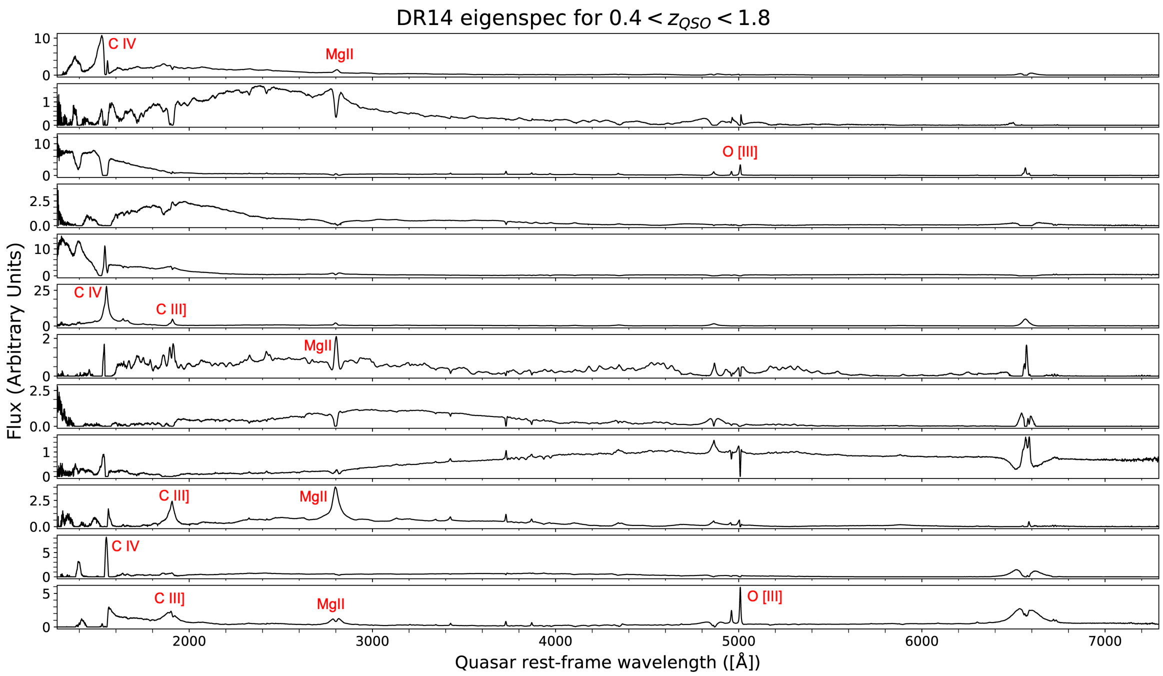

Figure 2.1: NMF eigenspectra for SDSS DR14 quasars with $0.4 < z_{\rm QSO} < 1.8$. We label the permitted metal emission lines and forbidden lines in some of the panels. NMF naturally filters out the intrinsic emission features from quasars.

2.2.3 Automatic Detection of Absorbers

The doublet nature of Mg II absorbers motivates us to develop an automated algorithm to detect its presence. Our approach is broken up into the following steps.

2.2.3.1 Wavelength Search Window

First, we define a wavelength search window using the quasar intrinsic emission lines. We start from $\Delta z = 0.018$ redshifted from the quasar's C IV line ($\lambda = 1549.48$ Å), or the blue end of SDSS (~ 3800 Å), and end at $\Delta z = 0.03$ blueshifted from quasar's Mg II emission line ($\lambda = 2799.117$ Å), or the red end of SDSS (~ 9200 Å). We select these Δz values to detect intervening absorbers rather than quasar associated absorbers. This also avoids misidentification due to the cases where the continuum is not well fitted to the quasar's intrinsic C IV or Mg II emission lines, which could produce spurious absorption dips. We also mask the possible Ca II $\lambda\lambda 3934, 3969$ lines (due to confusion with Mg II absorbers at $z \sim 0.4$) and O I lines ($\lambda 5577$ and $\lambda 6300$). In Figure 2.2 we show an example quasar spectrum in normalized flux (top) and residual (bottom), with the wavelength search window indicated.

Table 2.2: Parameters of threshold and weighting scheme. All parameters are set by atomic physics. Note that $\beta = \alpha/\delta$ and $\delta = f_{\rm strong}/f_{\rm weak}$ and $f_{\rm strong}$ and $f_{\rm weak}$ denote the oscillator strengths of doublet components. For selecting an optimized α, we show the detailed analysis in section 2.2.4.1.

Parameters	Mg II
α	2.50
$f_{\rm strong}$	0.6123
$f_{\rm weak}$	0.3054
strong	2796.35 Å
weak	2803.52 Å
$\lambda_{\rm prim}$	2799.935 Å
$\lambda_{\rm s1}$	2792.765 Å
$\lambda_{\rm s2}$	2807.405 Å
$\Delta\lambda_{\rm rest}$	7.17 Å
$\lambda_{\rm a}$	2795.65 Å
$\lambda_{\rm b}$	2797.05 Å

2.2.3.2 Absorber Candidate Selection

Next, we search for potential absorbers by defining a Gaussian kernel that mimics the Mg II $\lambda\lambda$ 2796, 2803 doublet. We convolve the residual with this kernel and apply a threshold on the convolved array ($\mathcal{C}_{\mathcal{R}}$). To define the threshold we use the local noise estimate $\sigma_{\mathcal{C}_{\mathcal{R}}}$. To estimate this local noise for the convolved array, for each pixel we take the noise as the standard deviation of the convolved array within ± 100 pixels around that pixel. This adaptive noise approach accounts for the noisy regions of the spectra, particularly at the edges. This convolution generically produces three peaks corresponding to the overlap of the kernel with the Mg II doublet. The primary peak ($\lambda_{\rm prim}$) corresponds to the case when two Gaussians of the kernel fully overlap with Mg II lines. The two secondary peaks ($\lambda_{\rm s1}$, $\lambda_{\rm s2}$) correspond to the case when kernel overlaps just one of the lines. We apply the following thresholds on the primary and secondary peaks to identify potential absorbers, selecting all the pixels that satisfy our threshold criterion. For the threshold parameters see Table 2.2.

Rule 1: sigma criteria on primary peak:

- select all pixels $\{1 + z_{\rm abs_i} = \frac{\lambda_{\rm abs_i}}{\lambda_{\rm prim}}\}$ where $\mathcal{C}_{\mathcal{R}} \leq {\rm median}(\mathcal{C}_{\mathcal{R}}) - \alpha \cdot \sigma_{\mathcal{C}_{\mathcal{R}}}$. This is a sigma criteria on the convolved residual and median($\mathcal{C}_{\mathcal{R}}$) implies global median of the convolved array.
- for each pixel with its associated redshift $z_{\rm abs}$ in $\{z_{\rm abs_i}\}$, define the neighbourhood of secondary peaks:

$$\lambda_{\rm sec_1} = \lambda_{\rm s1} \cdot (1 + z_{\rm abs}) \pm 1.3$$

$$\lambda_{\mathrm{sec}_2} = \lambda_{\mathrm{s2}} \cdot (1 + z_{\mathrm{abs}}) \pm 1.3$$

- measure the median residual amplitude for these pixels:

$$\mathcal{S}_{\mathcal{R}1} = \mathrm{median}(\mathcal{C}_{\mathcal{R},\lambda_{\mathrm{sec}_1}})$$
$$\mathcal{S}_{\mathcal{R}2} = \mathrm{median}(\mathcal{C}_{\mathcal{R},\lambda_{\mathrm{sec}_2}})$$
$$\mathcal{T} = \mathrm{median}(\mathcal{C}_{\mathcal{R}}) - \beta \cdot \sigma_{\mathcal{C}_{\mathcal{R}}} \text{ where } \beta = \frac{\alpha}{\delta}$$

Rule 2: sigma criteria on secondary peaks:

- if $\mathcal{S}_{\mathcal{R}1} \leq \mathcal{T}$ or $\mathcal{S}_{\mathcal{R}2} \leq \mathcal{T}$ then accept z_{abs}
- else, reject this z_{abs} as a detected absorption pixel

For our chosen values of α (primary threshold), β (secondary threshold), and $\delta = f_{\mathrm{strong}}/f_{\mathrm{weak}}$, where f_{strong} and f_{weak} are the oscillator strengths of Mg II λ2796 and λ2803 lines, see Table 2.2. Note that the threshold on the secondary peaks is based on the theoretical strength of lines given the oscillator strengths. The threshold $\mathcal{T}$ is defined for the two smaller peaks centered at λ_{s1} and λ_{s2}, which are weaker in the convolved residual.

Finally, note that we apply our thresholds to all individual pixels in the spectrum. The resulting list of accepted pixels will in general contain several contiguous ranges which need to be grouped and considered as one detected absorption feature. To do so we combine contiguous pixels absorbers with $\Delta z < 0.0026$, corresponding to $\Delta\lambda_{\mathrm{rest,\,MgII}} = 7.17$ Å. For each group we derive the absorber redshift as the mean of the pixel redshifts, weighting by the cube of the corresponding flux residual, such that the highest weight comes from the pixel closest to line center.

To provide the flexibility to detect absorbers of different strengths (rest equivalent widths) we separately run the kernel convolution varying the kernel FWHM from 3-8 pixels. We combine all the absorbers from each run, identifying and discarding duplicates by applying a similar weighted mean approach, weighting each redshift by the cube of the median of residual in the wavelength range $\lambda_{\mathrm{a}} \leq \frac{\lambda_{\mathrm{obs}}}{1+z_{\mathrm{abs}}} \leq \lambda_{\mathrm{b}}$ (see Table 2.2). Note that this grouping is applied only after running the detection pipeline for all widths to select the best location of the absorber.

Rule 3: S/N criteria for final candidate selection:

The matched kernel convolution gives a list of potential Mg II candidates. To select the final candidates we then use the S/N information of quasar spectra. Therefore, for each potential candidate from this list, we estimate the S/N (λ2796) and S/N (λ2803).[4] For the final candidate selection we apply the following criteria. We call this the *Mg II doublet*

[4] $\mathrm{S/N} = \sum_{\mathrm{i=p_1}}^{\mathrm{p_2}} \left(1 - \frac{\mathrm{F_i}}{\mathrm{C_i}}\right) / \left(\sum_{\mathrm{i=p_1}}^{\mathrm{p_2}} \sigma_{\mathrm{i}}^2\right)^{1/2}$; where F and C are flux and continuum respectively and σ is the corresponding error. p_1 and p_2 are the starting and ending pixels within ± 5 pixels from line centre.

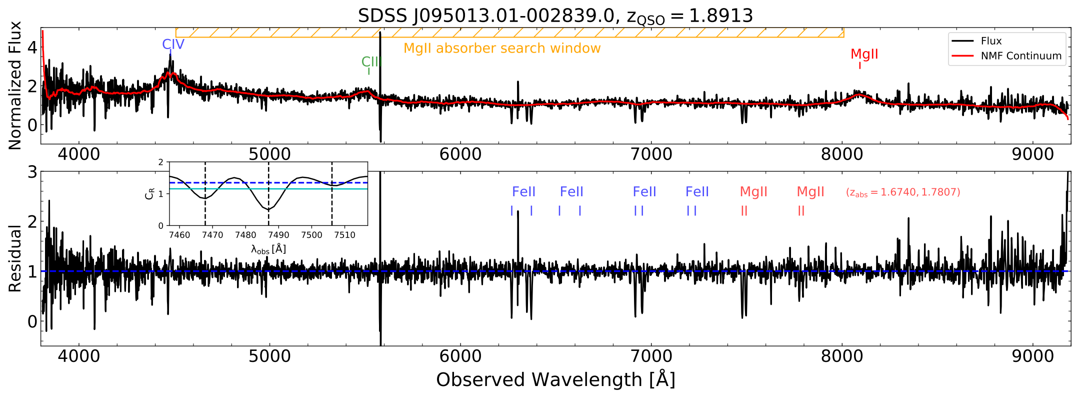

Figure 2.2: **Top:** the black line shows the normalized observed flux of quasar SDSS J095013.01-002839.0 (S/N = 7.46). The red line is the best-fitting NMF continuum after median filtering. The hatched orange rectangle shows the wavelength search window where the spectrum is searched for intervening Mg II absorbers. The intrinsic C IV, C III and Mg II emission lines of the quasar are shown in blue, green and red respectively. **Bottom:** the black line shows the residual spectrum defined as the ratio of the observed spectrum to the NMF continuum. The dashed blue line indicates unity. In this case the absorber detection pipeline identifies two intervening Mg II absorbers at $z_{\rm abs} = 1.6740, 1.7807$ (shown in red). The corresponding Fe II lines are shown in blue. Inset shows the convolved array with the thresholds (cyan: primary, blue: secondary) at the location of one of the absorbers. We clearly see three peaks as described in the text.

criteria.

$$\mathrm{S/N}(\lambda 2796) > 3 \text{ and } \mathrm{S/N}(\lambda 2803) > 2$$

2.2.3.3 Rejecting False Positives

Given the absorber candidates that passed the above criteria, we next reject likely false positives. We fit each absorption profile with a double gaussian function and discard candidates which have peculiar doublet separation. In fitting we allow all six parameters (two amplitudes, two-line centres, and two-line widths) to vary. We keep only those absorbers for which the line separation is within 1.2 Å of the fiducial value ($\Delta\lambda_{\rm rest} = 7.28$ Å, also see Table 2.2). At this stage we also estimate the mean redshift of the absorbers using the centroid of both lines. Finally, we check for cases where Fe II lines are incorrectly identified as Mg II lines. For spectra with more than one potential Mg II absorber, we check whether its observed wavelength corresponds to any Fe II $\lambda\lambda 2586, 2600$ lines, within ± 3 Å, corresponding to the redshifts of already identified Mg II lines. If this is the case, we reject the Mg II candidate having smaller S/N($\lambda 2796$) than any of the two Fe II lines as Mg II is a stronger transition than Fe II. This selection might also reject a few true cases in which the Fe II lines are also very strong, however, this is not very common. We show an

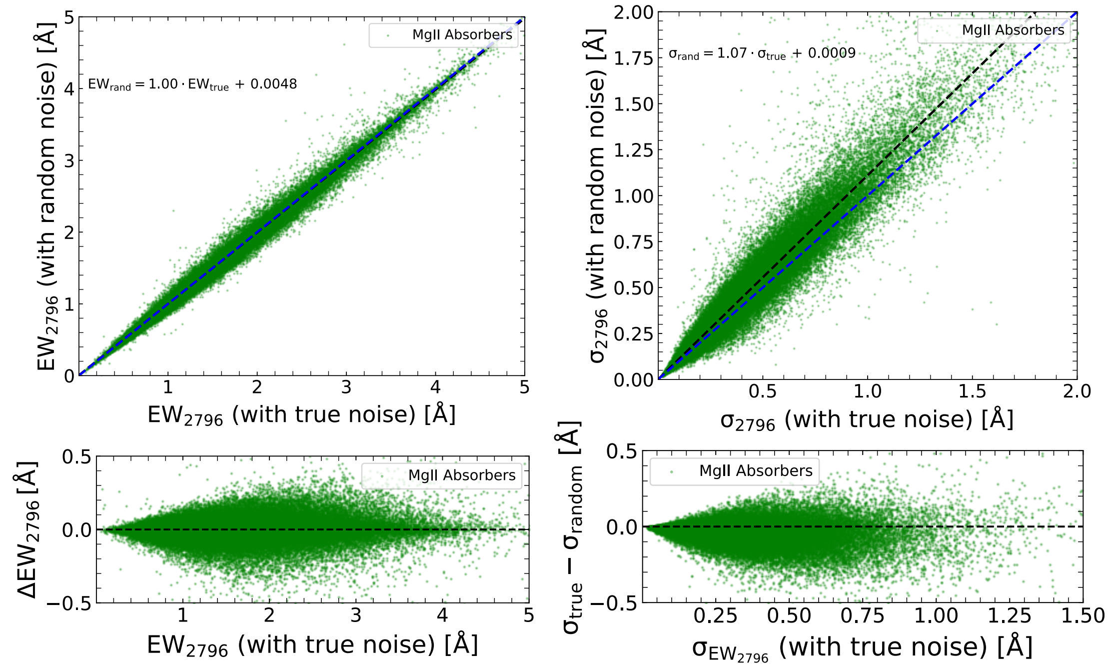

Figure 2.3: Comparison of rest EW and error measurements with random and true noise. **Top Left:** Measured rest EW_{2796} (by adding random noise) as a function of rest EW_{2796} (by adding true noise). **Top Right:** Corresponding EW errors measured using these approaches. They match very well though error from random noise approach is slightly higher (as visible in best-fit dashed line shown in black). **Bottom:** The corresponding smaller panel shows the difference between two as a function of rest EW_{2796} (by adding true noise). The agreement is very good and this further supports that measurements are robust.

example of a SDSS quasar spectrum with relatively high S/N ratio (= 7.5 pix^{-1}) in Figure 2.2, highlighting two detected Mg II absorbers and the locations of the corresponding Fe II lines.

2.2.3.4 Rest equivalent widths of Absorbers

With the detected absorbers in hand, we apply a Monte Carlo approach to estimate errors on the fit parameters and ultimately the measured rest equivalent widths. We fit each doublet with a standard Levenberg-Marquardt minimization, first by adding randomly generated noise, using a normal distribution centred at 0 and with standard deviation equal to the mean of the errors on the residual. We also try the alternative approach of adding the true noise of the spectra multiplied by a standard normal distribution. In both cases we repeat this process 200 times. In case of failures, we record all parameters and rest equivalent widths as zeros. This guarantees that errors will be large in cases with many failures, if there is a false positive all 200 runs fail and the rest equivalent width is consistent with zero. Finally, we take the median of all 200 runs as the estimate of the rest equivalent width of lines. For errors we take the 16th (p16) and 84th (p84) percentiles and compute sigma as $\sigma = (\mathrm{p84} - \mathrm{p16})/2$. Overall we find that these two error estimate methods agree well (shown in Figure 2.3), also indicating that the majority of absorbers are genuine. The uncertainties measured from adding random noise are slightly ($\sim 5\%$) higher than the errors measured by adding true noise from the spectra. For our purposes we use the uncertainties obtained from the true noise of the spectra. The typical error in the rest EW_{2796} ($\sigma_{EW_{2796}}$) is ~ 0.2 Å.

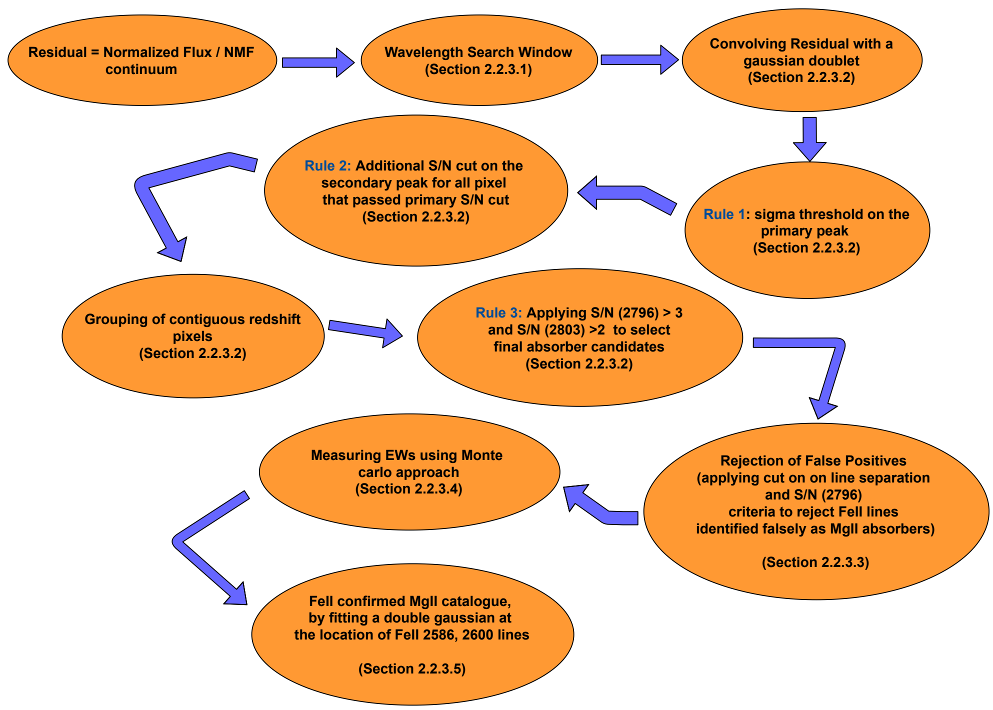

Figure 2.4: A short schematic of our automated detection pipeline.

2.2.3.5 Fe - confirmed Mg II catalogue

With our Mg II absorber catalogue we perform an additional step and attempt to confirm each Mg II absorber with Fe II $\lambda\lambda$ 2586, 2600 by fitting gaussian at the location of Fe II lines. If the fitting succeeds (parameters are finite and within the boundaries) and the line separation of Fe II lines is within 1.2 Å of the fiducial value ($\Delta\lambda_{\mathrm{rest}} = 13.48$ Å) we flag this Mg II absorber as 'confirmed' with Fe II. We also estimate the rest frame equivalent width and errors of the Fe II $\lambda\lambda$2586, 2600 lines using a similar Monte Carlo approach as above. Note that Fe II is a much weaker transition than Mg II and the minimum Fe II strength detectable with our algorithm is $\mathrm{EW}_{2600} \geq \mathrm{f}_{\mathrm{FeII},\,2600}/\mathrm{f}_{\mathrm{MgII},\,2796} \cdot \mathrm{EW}_{2796}$, assuming the same

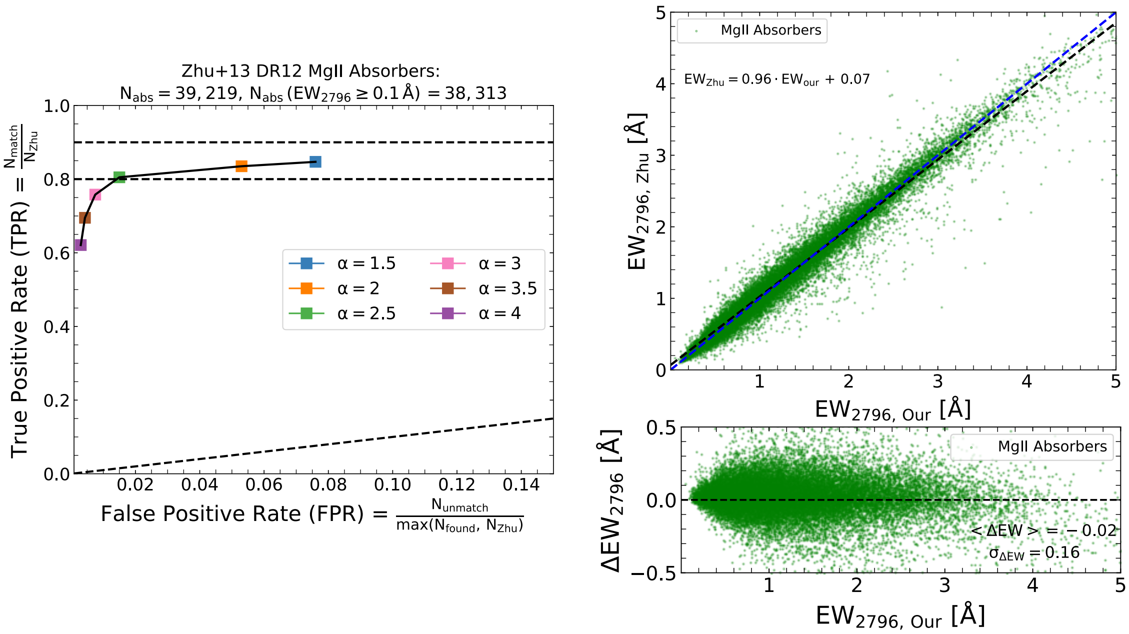

Figure 2.5: **Left:** ROC curve: True Positive Rate (TPR) as a function of False Positive Rate (FPR) for different α values using selection based on SNR condition. Higher α corresponds to lower TPR and FPR as expected. Overall the FPR is significantly lower in the adaptive S/N approach. Note that, x-axis is zoomed for clarity. The dashed horizontal lines show 0.8 and 0.9 on TPR axis and the dashed inclined line shows one-to-one line. **Top Right:** $EW_{2796,\,\mathrm{Zhu}}$ vs. $EW_{2796,\,\mathrm{our}}$, showing good agreement, though our EWs are slightly higher (as visible in best-fit dashed line shown in black). **Bottom Right:** Difference between these two values as a function of $EW_{2796,\,\mathrm{our}}$. The majority of the absorbers have $|\Delta\mathrm{EW}| \leq 0.2$ Å and the typical error between our values versus Zhu is ~ 0.16 Å.

abundance for both Mg and Fe. With the minimum detected $EW_{2796} \sim 0.2$ Å in our catalogue, and $f_{\mathrm{FeII},\,2600} \sim 0.3$, we obtain a theoretical minimum $EW_{2600} \sim 0.1$ Å. In total we have $\sim 70,000$ Fe II confirmed Mg II absorbers ($\sim 44\%$ of the sample). Finally, we show a schematic in Figure 2.4 that summarizes the steps of our automated detection pipeline.

2.2.4 Comparison with Previous Mg II Catalogue

2.2.4.1 Receiver Operating Characteristic (ROC) Analysis

To select an optimized value for the α parameter (see Section 2.2.3), we make use of the available DR12 Mg II catalogue compiled by Zhu & Ménard (2013a) to test our detection pipeline. The reference catalogue has 76,148 Mg II absorbers detected in 47,065 quasars. After applying all the masks as described in the section 2.2.3, we end up with 39,219 Mg II absorbers in 25,716 quasars. Then we run our complete detection pipeline on this set of quasars for several values of α and estimate the True Positive Rate (TPR) and False Positive Rate (FPR) to analyse the ROC curve to choose an optimized value for

α. To estimate the TPR and FPR we only use Mg II absorbers with $\mathrm{EW}_{2796} > 0.1$ Å ($\mathrm{N} = 38,313$). We define TPR and FPR as follows:

2.2.4.2 Adaptive SNR approach: Mg II doublet criteria

In this approach (using Rule 1, 2, and 3 described in section 2.2.3) we define TPR as the ratio of absorbers that passed our Mg II criteria to the total number of Zhu absorbers. FPR is defined as the ratio of all legitimate absorbers that passed our Mg II criteria, but did not match with Zhu absorbers, to the maximum of total Zhu absorbers and all absorbers detected with our pipeline.

We take the maximum of these two quantities to account for the case when our pipeline finds several absorbers that are not matched with the Zhu & Ménard (2013a) catalogue. Left panel of Figure 2.5 shows the ROC curve estimated for different α values in adaptive S/N approach. We clearly see that applying the adaptive S/N condition significantly reduces the false cases. For example at $\alpha = 2.5$ the corresponding TPR and FPR are ~ 0.8 and ~ 0.02, respectively. Therefore, we take advantage of the S/N of the spectra to reduce the incidence of false cases. As expected for larger values of α, the TPR and FPR are both low while, increasing for smaller α values.

Finally, for our current catalogue, we conservatively choose $\alpha = 2.5$ because our ROC analysis is based on quasars (from the DR12 absorber catalogue only) with possibly high S/N compared to the entire DR16 quasar set used in the current study. By choosing a slightly higher α value we select relatively strong potential absorbers at the convolution step thus reducing the pipeline runtime significantly. However, we miss a good fraction of absorbers as shown in the completeness analysis (see section 2.3.3, Figure 2.11). On the other hand, this guarantees a very high purity for our catalogue (see discussion in section 2.3.4 and Figure 2.11). We also show the comparison of our rest-frame EW_{2796} measurements with the values from Zhu & Ménard (2013a) catalogue in the right panel of Figure 2.5. We agree very well, though our EW_{2796} values are slightly higher (as visible in best-fit dashed line shown in black).

2.3 Metal absorber catalogue

2.3.1 Properties of Individual Absorbers

Running the detection pipeline on the DR16 quasars we compile our final Mg II absorber catalogue. In Figure 2.6 we show the redshift distributions of quasars (light gray shows all DR16 QSOs and dark gray shows QSOs with $\mathrm{S/N_{QSO}} > 2$) and detected absorbers (blue). We construct three absorber catalogues: (i) all Mg II detections (light blue), (ii) strong Mg II absorbers defined as having $EW_{2796} > 3\sigma_{2796}$ (medium blue), and (iii) 'Fe II confirmed' detections where we detect Fe II $\lambda\lambda 2586, 2600$ at the same redshift (dark blue). Within our wavelength search window we find 159,524 Mg II absorbers which satisfy our detection pipeline. Out of these 69,675 absorbers have also passed our Fe II confirmation test and 121,989 have $\mathrm{EW}_{2796} > 3\sigma_{2796}$ (i.e. $\mathrm{S/N}_{2796} > 3$).

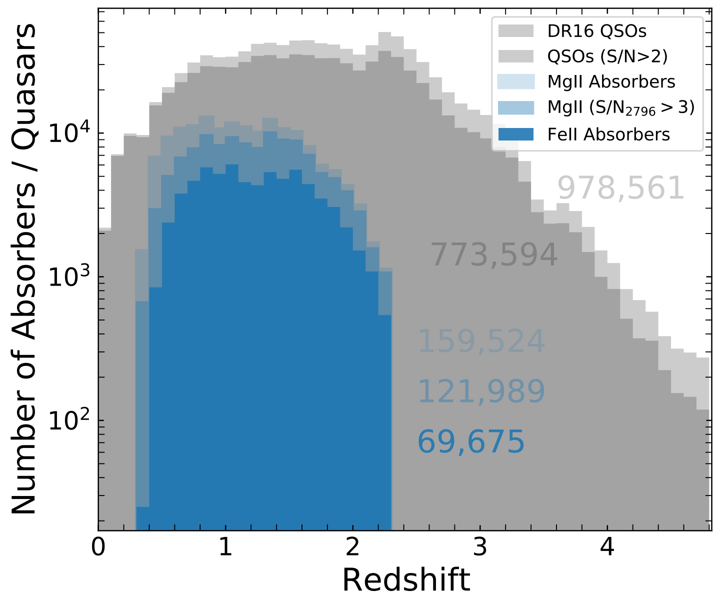

Figure 2.6: Redshift distribution of DR16 quasars (shown in light black, $\mathrm{N} = 978,561$, dark black shows quasars with $\mathrm{S/N_{QSO}} > 2$, $\mathrm{N} = 773,594$) and the Mg II absorbers detected by our pipeline (shown in light blue, N = 159,524). The medium blue region shows the Mg II absorbers with rest equivalent width greater than their three times the corresponding errors (N = 121,989). The dark blue region shows the Mg II absorbers which are 'confirmed' by Fe II lines (N = 69,675). The redshift range spans from $z = 0.36$ to $z = 2.28$ with bin size of $\Delta z = 0.1$.

In the four panels of Figure 2.7 we present the properties of individual Mg II absorbers. The top left panel shows the ratio of total rest equivalent width (EW) of Mg II to Fe II (using $\lambda\lambda2586, 2600$ only) as a function of the total rest equivalent width of Mg II absorption. This quantifies the relative strength of Fe II absorbers, which also trace $\sim 10^4$ K gas around galaxies. We see that most ($\gtrsim 80$ percent) of the Fe II confirmed Mg II absorbers lie within the theoretical limit (assuming abundances to be same for Fe and Mg) for the line strength ratio of 1 (for the saturated case) and $\sim$3 (for the completely unsaturated case). The majority of absorbers regardless of EW have a line ratio of $\gtrsim 1$, indicating they are intermediate between saturated and unsaturated. The median line ratio is ~ 2. We also note the large scatter and many outliers above ~ 3, likely indicating cases of non-solar abundance ratios.

The top right panel of Figure 2.7 shows the doublet ratio of EW_{2796} and EW_{2803} as a function of EW_{2796} for all Mg II absorbers. The doublet ratio indicates if the lines are saturated or unsaturated and is sensitive to the opacity of the medium. For Mg II $\lambda\lambda2796, 2803$ the theoretical value of the doublet ratio varies between 1 (fully saturated) and 2 (completely unsaturated). We see that most strong absorbers ($\mathrm{EW}_{2796} > 1$ Å) have a doublet ratio close to 1, indicating saturation. For weak absorbers ($\mathrm{EW}_{2796} < 1$ Å) we

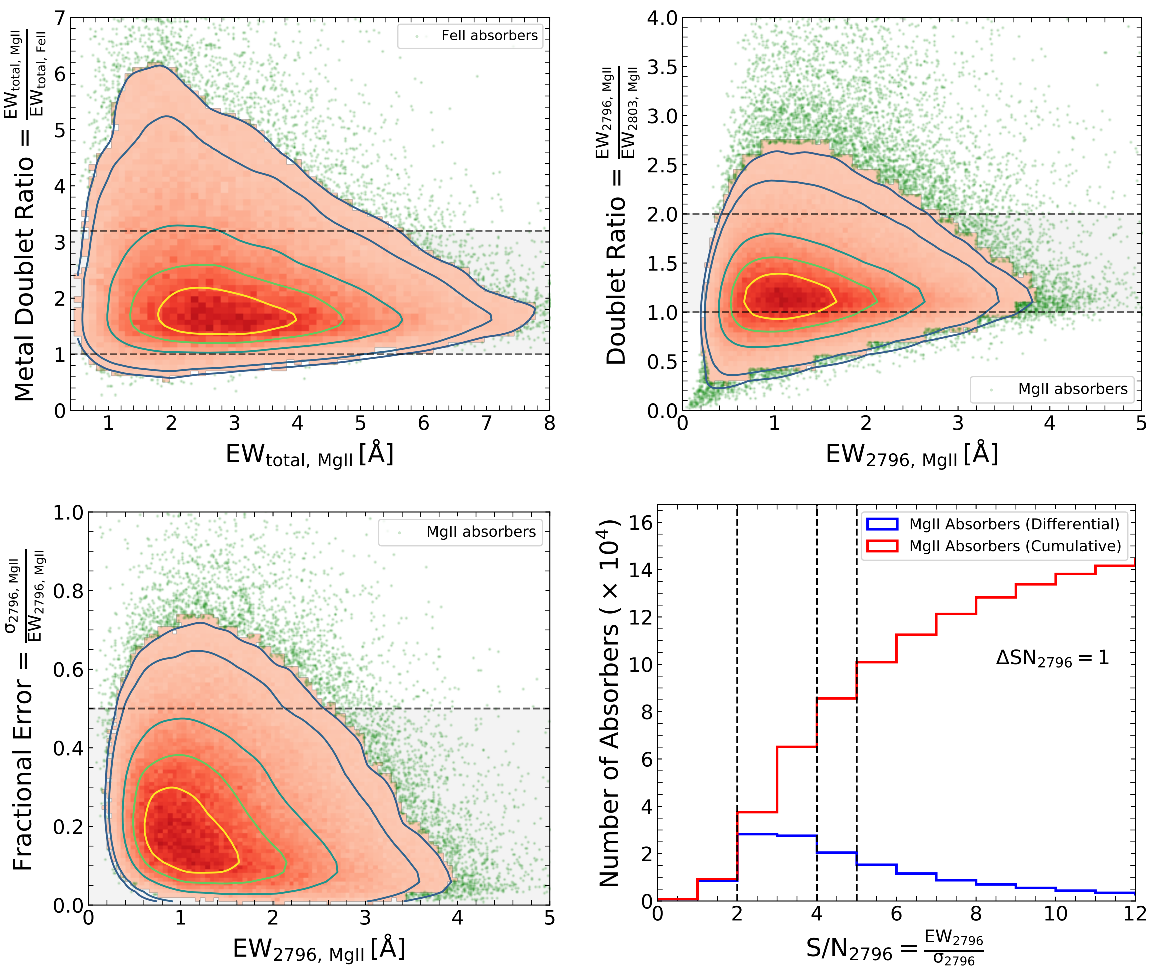

Figure 2.7: **Top Left:** Doublet ratio of total rest rest equivalent width of Mg II $\lambda\lambda$2796, 2803 and Fe II $\lambda\lambda$2586, 2600 as a function of total rest rest equivalent width of Mg II . The shaded rectangle highlights the region when all lines are completely saturated (doublet ratio=1) and when lines are unsaturated (doublet ratio$\sim$ 3). Note that this theoretical limit is estimated using oscillator strengths and assuming abundances of Fe and Mg be same. Most Mg II absorbers lie within this theoretical limit. **Top Right:** Mg II $\lambda\lambda$2796, 2803 doublet ratio as a function of EW_{2796}, where the shaded rectangle highlights the region when both lines are saturated (i.e. doublet ratio=1) and when neither are saturated (i.e. doublet ratio=2). Most Mg II absorbers lie within these two lines. **Bottom Left:** Fractional error in measurement of EW_{2796} as a function of EW. Errors are estimated with bootstrap approach. In all panels contours enclose 25, 50, 75, 95 and 97.5 percentiles of the sample. **Bottom Right:** Cumulative (red) and differential (blue) distribution of S/N_{2796} of all detected Mg II absorbers.

Table 2.3: Sample statistics comparison versus SDSS data release (DR). References are: [a] Pâris et al. (2017); [b] Pâris et al. (2018); [c] Ahumada et al. (2020); [$\star$] Zhu & Ménard (2013a) (trimmed case with the same masks as ours); [†] **this work**. The increased quasar sample enables us to increase the statistics of Mg II absorption systems significantly (~ 4 times). Note that S/N_{2796} is defined as $EW_{2796}/\sigma_{EW_{2796}}$ and S/N_{QSO} implies signal-to-noise of quasar spectra.

Objects	DR12	DR14	DR16
QSOs	297,301[a]	526,356[b]	983,317[c]
QSOs ($S/N_{QSO} > 1$)	-	-	941,939
QSOs ($S/N_{QSO} > 2$)	-	-	773,594
LRGs	-	$\lesssim 1$ million	1,252,722[c]
ELGs	-	35,094[c]	269,889[c]
Mg II Absorbers	39,219[$\star$]	-	**159,524**[†]
Mg II ($S/N_{2796} > 1$)	38,327[$\star$]	-	**158,725**[†]
Mg II ($S/N_{2796} > 2$)	37,763[$\star$]	-	**150,236**[†]
Mg II ($S/N_{2796} > 4$)	33,376[$\star$]	-	**94,403**[†]
Fe II Absorbers	-	-	**69, 675**[†]
Mg II Absorbers (with $S/N_{QSO} > 2$)	-	-	**158, 494**[†]
Fe II Absorbers (with $S/N_{QSO} > 2$)	-	-	**69, 594**[†]

see doublet ratio ~ 2 as expected due to low saturation.

The bottom left panel of Figure 2.7 shows the fractional errors on the EW_{2796} measurement as a function of EW_{2796} for all deteced Mg II absorbers. We see that, most ($\sim$ 94 percent) of the absorbers have low errors (fractional error < 0.5). Weak absorbers tend to have large errors, making it more problematic to measure their properties. The median EW_{2796} is ~ 1.3 Å with typical $\sigma_{EW_{2796}} \sim 0.2$ Å for the sample. The bottom right panel of Figure 2.7 shows the cumulative (red) and differential (blue) distributions of S/N_{2796}. The majority of our absorbers have high S/N_{2796} (> 2) and median S/N_{2796} is ~ 4.7.

2.3.2 Properties of Absorbers in Stacked Spectra

We stack the residual spectra of quasars in the rest frame of detected Mg II absorbers to study features which become visible in these composite spectra. Figure 2.8 shows a composite median spectrum, stacking on all Mg II absorbers. We see several weak metal lines such as Si II, C IV, Al II, Al III, Fe II (several lines) and Mg I. The clear detection of these weak lines provides strong evidence that the majority of our absorbers are genuine. We have divided our sample into five EW_{2796} bins to understand the corresponding variation of the strength of metal absorbers, which increases with increasing EW_{2796}.

In spectral regions without absorption features the median flux residual is flat, indi-

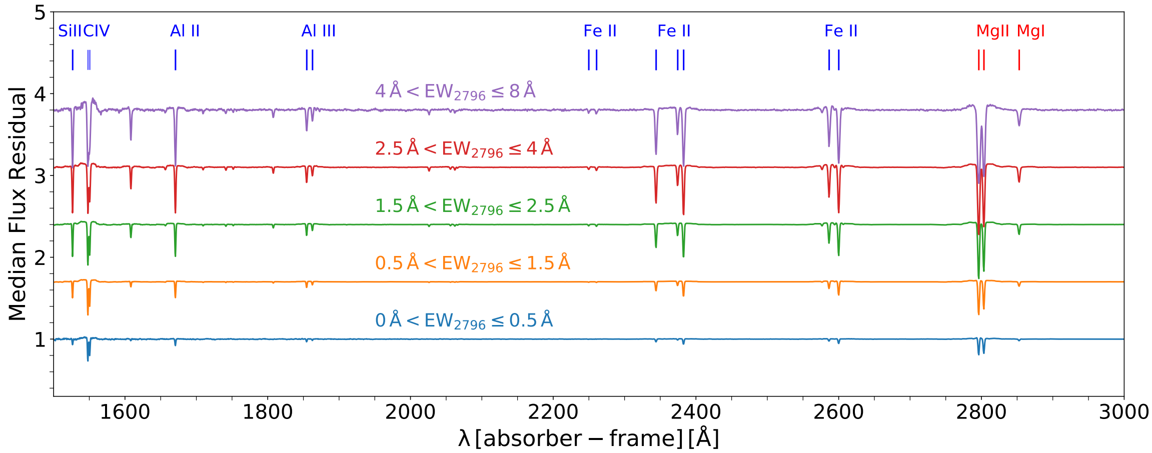

Figure 2.8: Median composite spectra of quasars, stacked in the rest-frame of all DR16 Mg II absorbers. We have divided the absorber sample into five sub-samples with $0\,\text{Å} < \text{EW}_{2796} \leq 0.5\,\text{Å}$ ($\text{N}_{\text{abs}} = 5718$), $0.5\,\text{Å} < \text{EW}_{2796} \leq 1.5\,\text{Å}$ ($\text{N}_{\text{abs}} = 90,317$), $1.5\,\text{Å} < \text{EW}_{2796} \leq 2.5\,\text{Å}$ ($\text{N}_{\text{abs}} = 50,083$), $2.5\,\text{Å} < \text{EW}_{2796} \leq 4\,\text{Å}$ ($\text{N}_{\text{abs}} = 12,869$), $4\,\text{Å} < \text{EW}_{2796} \leq 8\,\text{Å}$ ($\text{N}_{\text{abs}} = 537$). We have indicated the most prominent metal lines to guide the eye: in stacks we detect weaker transitions including Si II, C IV, Al II, Al III, Mg I and other weak transitions of Fe II.

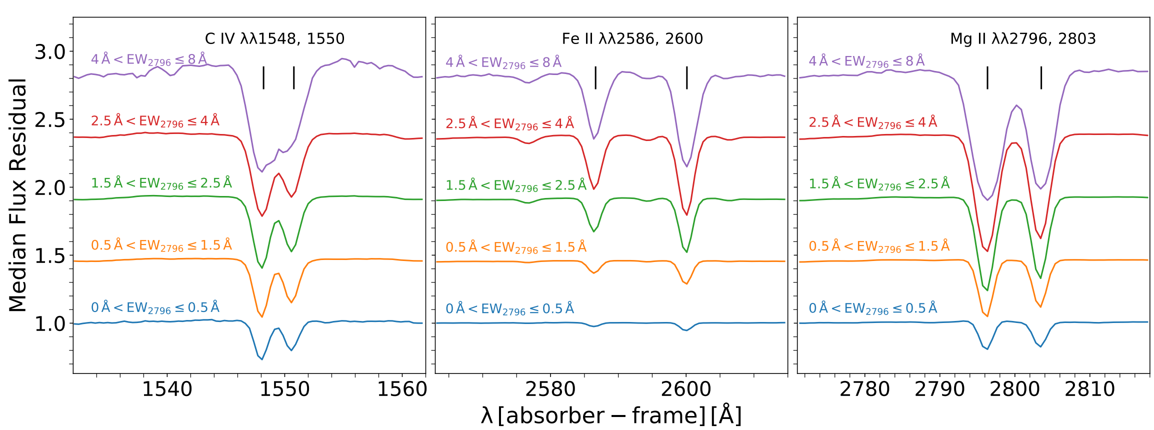

Figure 2.9: Zoom into the median composite spectra to show the structure of three of the most prominent metal lines. For display purposes we have shifted the lines of different EWs by a small amount in the vertical direction in all panels.

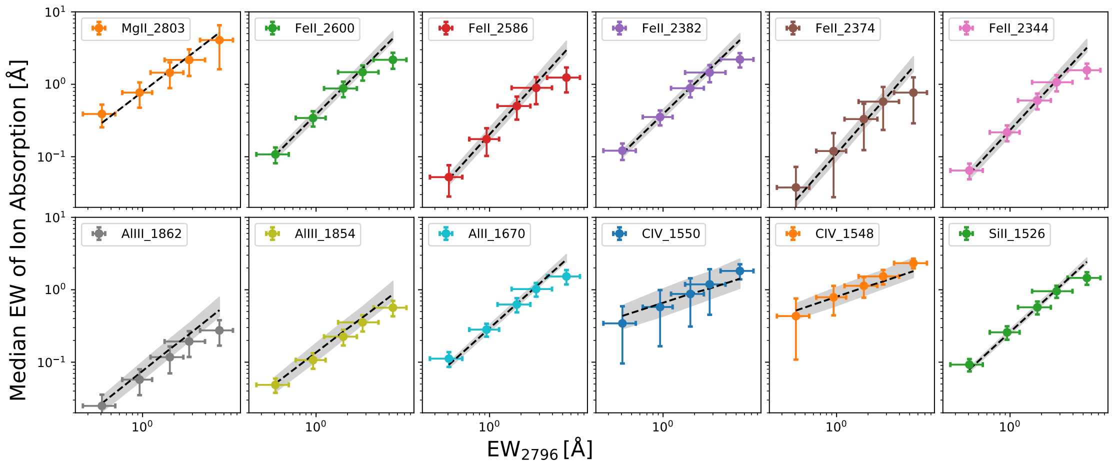

Figure 2.10: Rest equivalent width of different metal transitions as a function of rest equivalent width of Mg II 2796 line for the median composite spectra. The solid circles show the measured rest EWs (rEWs) and dashed solid lines show the empirical rEWs scaling as a function of $EW_{\mathrm{MgII},2796}$ from Lan & Fukugita (2017). To compare we use the median redshift of absorbers in each rEW (2796) bin. The shaded region shows the corresponding 5th and 95th percentiles of redshifts in each rEW (2796) bin. We have estimated the rest equivalent widths by fitting suitable gaussian profiles and errors using the bootstrap method.

cating that our continuum estimation pipeline works well. In Figure 2.9 we zoom into the profiles of three prominent metal lines, C IV $\lambda\lambda$1548,1550 (left), Fe II $\lambda\lambda$2586, 2600 (middle), Mg II $\lambda\lambda$2796, 2803 (right). The absorption features are well represented by Gaussian profiles, as expected in a stack due to the central limit theorem. As we move to strong absorber systems the lines become saturated (doublet ratio $\sim$1) and doublet profiles overlap, particularly for C IV. For the strongest EW_{2796} bin noise begins to be visible in the residual due to the low statistics. As a check we also stack the spectra of all quasars with Mg II absorbers without Fe II confirmation (not shown). In this case we also detect the same ensemble of weaker metal transitions.

In Figure 2.10 we show the rest equivalent widths of visible metal lines in the composite spectra as a function of the total rest equivalent width of Mg II . We see that Fe II and C IV lines are the second most prominent features though the absorption strength of Fe II is slightly higher compared to C IV lines. This is because Fe II is a stronger transition and also at longer wavelength, Fe II can be detected in SDSS quasar spectra with redshifts as low as $z \sim 0.5$, while C IV absorbers can only be detected in quasars with $z_{\mathrm{QSO}} \gtrsim 1.5$ due to the wavelength coverage of SDSS. We note that one should limit the stacks to similar redshift path-lengths in order to properly quantify the relative scaling between the lines.

In order to investigate whether our pipeline estimates the consistent rest equivalent widths, we also make a comparison of metal equivalent widths against the empirical scaling relations (as a function of $\mathrm{EW}_{\mathrm{MgII},2796}$ and redshift) derived in Lan & Fukugita (2017)[5]. We observe that our pipeline yields a consistent trend, though there is a slight discrepancy in the strongest $\mathrm{EW}_{\mathrm{MgII},2796}$ bin, possibly due to low statistics. We also note that a direct comparison is not straightforward as the scaling also depends on the redshift of the absorbers. For the current comparison we take the median redshift of absorbers in each $\mathrm{EW}_{\mathrm{MgII},2796}$ bin and this could also be a possible source of the slight discrepancy in some cases.

2.3.3 Completeness of Detection Algorithm

The search for absorption features in the relatively low S/N quasar spectra from SDSS is challenging, and an accurate characterization of our ability to recover absorbers of different strengths is required. To estimate the completeness of our detection pipeline we therefore implement a Monte Carlo simulation approach. We generate doublet profiles that mimic true absorber and insert these at a random location in a *real* residual chosen randomly from the set of DR16 QSOs. While inserting we take care of all the masks that we have defined in our detection algorithm. To generate doublets we uniformly sample EW_{2796} from $0 \leq \mathrm{EW}_{2796} \leq 8$ Å. We select both a doublet ratio ($0.25 <$ doublet ratio < 4.5) and width ($0.34 < \sigma < 3.5$ Å) from gaussian distributions. We then run our detection pipeline on the spectrum with a fake absorber and check for a successful detection. We define success as having satisfied the *Mg II doublet criteria* and recovered a sufficiently accurate rest equivalent width, $|\mathrm{EW}_{\mathrm{in}} - \mathrm{EW}_{\mathrm{out}}|/\mathrm{EW}_{\mathrm{in}} < 1$, where $\mathrm{EW}_{\mathrm{in}}$ is the injected EW_{2796} and $\mathrm{EW}_{\mathrm{out}}$ is the measured EW_{2796}. In total, we have simulated over ~ 33 million fake absorbers in ~ 1 million DR16 quasars. Since our completeness corrections are based on the final *Mg II doublet criteria*, this does not bias results in the galaxy section.

Finally, for absorbers in a given EW_{2796} and redshift bin, we estimate the completeness, $c(EW_{2796}, z)$ as the ratio of *detected* absorbers to the *injected* absorbers. Using this estimate we compute, for each absorber with given EW_{2796} and redshift, a completeness corrected effective 'number of absorbers' $N_{\mathrm{eff}} = 1/c(EW_{2796}, z)$, the count of absorbers had our detection method been perfect. By definition, $c \leq 1$ and so $N_{\mathrm{eff}} \geq 1$. Similarly we estimate the purity as the ratio of detected absorbers to legitimate absorbers (i.e. all the absorbers that pass the Mg II doublet criteria) found by our pipeline. We show the 2D distribution of completeness in $EW_{2796} - z$ space in the top left panel of Figure 2.11.

In the bottom left panel of Figure 2.11 we show the completeness (and purity) as a function of EW_{2796} averaged over all redshifts. The 2D distribution of purity is shown in Figure 2.11. As expected, the completeness is higher for strong absorbers and lower for weak absorbers. The bottom right panel shows the completeness (and purity) as a function of z averaged over all EW_{2796}, where dips at specific redshifts reflecting excluded wavelength search windows. Overall, the purity of our algorithm (shown in Figure 2.11) is extremely high, at the cost of moderate completeness, however, completeness goes up significantly

[5]The empirical scaling relation in Lan & Fukugita (2017) is: $EW_{\lambda} = C\,(EW_{2796})^{\alpha}\,(1+z)^{\beta}$. For the best-fitting values of C, α and β, see Table 1 of Lan & Fukugita (2017).

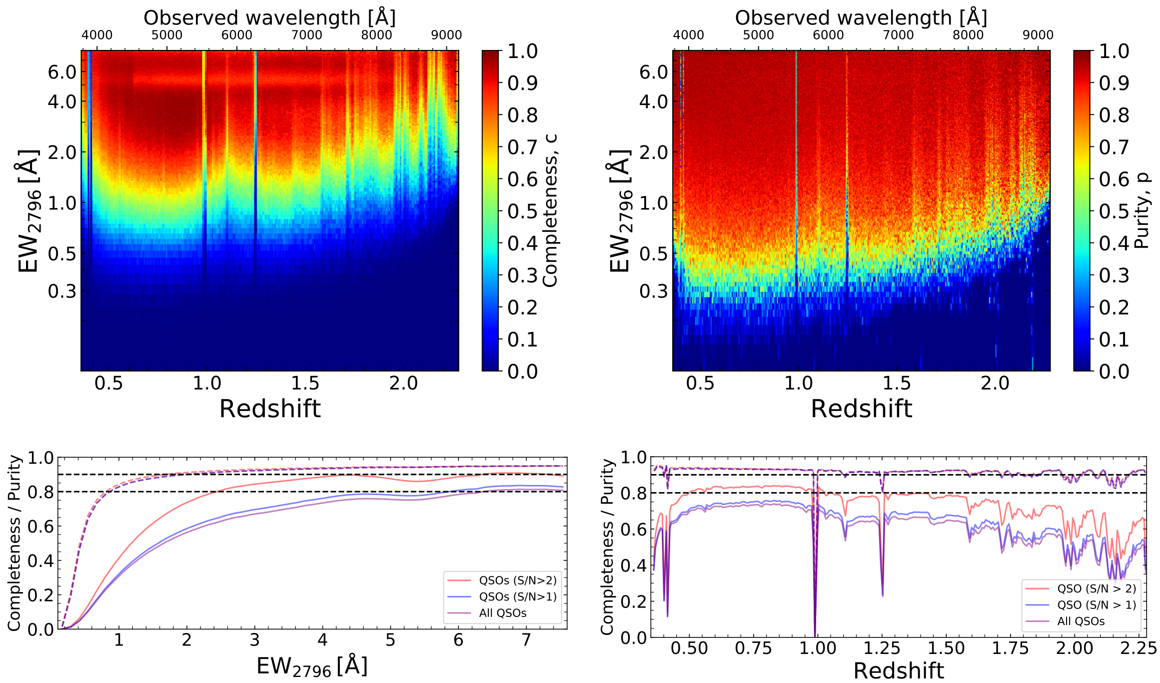

Figure 2.11: **Top left:** 2D Completeness function $c(EW_{2796}, z)$ as a function of rest equivalent width and redshift. We are showing the 2D completeness only for QSOs with $S/N_{QSO} > 2$. Completeness is low on both wavelength edges of SDSS spectra, as the edges are generally noisy. It also drops in regions with many sky lines such as OI λ 5577, OI λ 6300 and OH. The dip at $\sim$ 5900 Å is due to sodium dichoric, while CaII $\lambda\lambda$3934, 3969 is also present. **Top right:** Purity function $p(EW_{2796}, z)$ for QSOs with $S/N_{QSO} > 2$, as a function of rest equivalent width and redshift. The purity is very high for detected absorbers. For very weak absorbers ($EW_{2796} < 0.3$ Å) the purity is lower because the completeness is similarly low, and the pipeline is more likely to identify false cases due to noisy features. **Bottom left**: Completeness as a function of EW_{2796} averaged over all redshifts. **Bottom right:** Completeness as a function of redshift smoothed over all EW_{2796}. Weak and intermediate absorbers ($EW_{2796} \leq 1$ Å) have lower completeness, as expected. In both panels solid and dashed curves represent completeness and purity respectively. The dashed horizontal lines show 0.9 and 0.8 on vertical axis. The overall purity and completeness are very high for our detection pipeline for QSOs with $S/N_{QSO} > 2$.

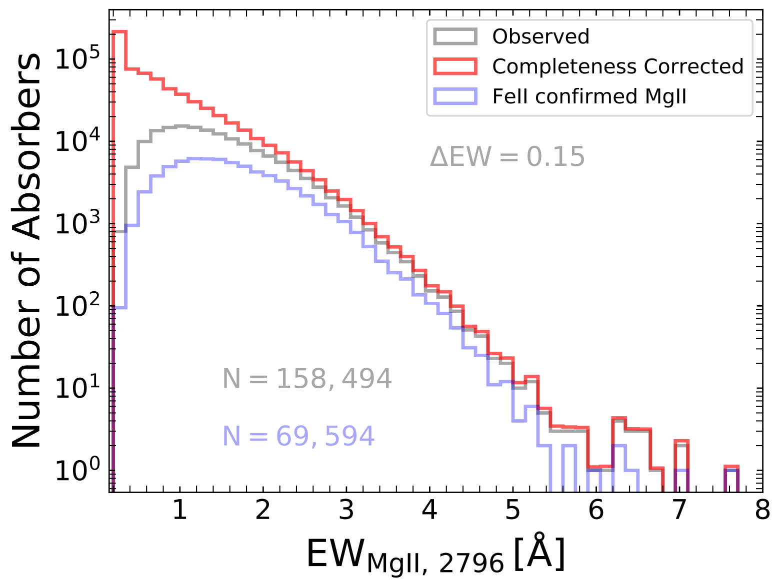

Figure 2.12: 1D histogram of intrinsic observed (gray) and completeness corrected (red) of EW_{2796} for all Mg II absorbers (associated with QSOs having $S/N_{QSO} > 2$), as well as the Fe II confirmed subset (blue). The completeness corrected distribution follows an exponential profile in rest equivalent width, such that the weakest absorbers are also the most frequent. Our Mg II catalogue spans 0.2 Å$\lesssim EW_{2796,MgII} \lesssim 8$ Å.

for QSOs with high S/N data as shown in the bottom left panel of Figure 2.11. Because the eBOSS spectra have relatively lower S/N (Dawson et al., 2016), we also include here completeness curves restricted to spectra above minimum S/N thresholds (i.e. $S/N > 1, 2$). As expected we find that it is difficult to detect absorbers primarily in noisy spectra, while our algorithm has excellent performance on high S/N data.

The observed and completeness corrected EW_{2796} distributions for all Mg II absorbers (that are associated with QSOs having $S/N_{QSO} > 2$, i.e. our "fiducial" sample, defined explicitly in section 2.4.1), as well as for the Fe II confirmed subset, are shown in Figure 2.12. The completeness corrected distribution follows an exponential profile, hence the mean value after completeness correction shifts to smaller values as weak absorbers dominate the distribution. The mean EW_{2796} for the observed distribution is $\sim$1.4 Å while for completeness corrected distribution it decreases to ~ 0.8 Å. We also see that the 'completeness corrected distribution' is very high in the lowest bin, which may be due to the small number statistics of absorbers and, such a peak is possibly not significant. We observe that the fraction of Fe II confirmed Mg II absorbers for high rest equivalent widths does not reach unity even for very strong absorbers. This is because Fe II is a lower wavelength transition than Mg II and does not fall within the SDSS wavelength coverage for Mg II absorbers at low redshifts.

2.3.4 Purity analysis of the Pipeline

In order to quantify the quality of our catalogue we also estimate the purity of simulated absorbers detected in our pipeline. As described in Section 2.3.3, we estimate the purity in each $EW - z$ bin as the ratio of detected absorbers to the total absorbers that pass our Mg II criteria.

We show the 2D purity distribution of simulated absorbers as a function of rest EW and redshift of the absorber in top right panel of Figure 2.11. We clearly see that the purity of detected absorbers is very high in almost every bin for $EW_{2796} > 0.4$ Å systems. For very weak absorbers ($EW_{2796} < 0.3$ Å) the purity is low because the corresponding completeness is also low, and possibly the pipeline finds false cases due to noisy features. The overall purity of our detection pipeline is $\sim 95\%$. Note that we only show the purity for absorbers detected in QSOs having $\mathrm{S/N_{QSO}} > 2$.

2.4 Connecting to Galaxies

With absorbers identified across a wide range of redshifts and over a significant fraction of the sky, we now need to connect them to nearby galaxies. Our main resource to obtain a large sample of galaxies with $0.4 < z < 1$ are the BOSS and eBOSS programs of SDSS. Namely, emission line galaxies (ELGs) and luminous red galaxies (LRGs), which we use to study the galaxy - Mg II absorber connection. When necessary, we use a Planck-consistent cosmology (Planck Collaboration et al., 2020) for our analysis, i.e. $\Omega_{\mathrm{m},0} = 0.307$, $H_0 = 67.7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2.4.1 Fiducial Quasar Sample

As discussed in Section 2.3.3, our detection pipeline has excellent performance on high S/N data. Therefore, for the galaxy-centric analysis we select only quasars having $\mathrm{S/N_{QSO}} > 2$ and Mg II absorbers associated with them. We make this cut to define the fiducial quasar sample and the corresponding Mg II absorber catalogue because the completeness and purity both are very high for our detection algorithm and we miss intermediate or strong absorbers rarely. For example, at $EW_{2796} \sim 2$ Å completeness and purity are $\gtrsim 75$ percent and $\gtrsim 90$ percent respectively as shown in bottom left panel of Figure 2.11. The high completeness catalogue also brings more confidence in the results as we are not applying big corrections. As presented in Table 2.3, out of $978,561$ quasars $773,594$ have $\mathrm{S/N_{QSO}} > 2$, which is $\sim 80\%$ of the total sample size. We show the histogram of redshifts of quasars with $\mathrm{S/N_{QSO}} > 2$ in Figure 2.6. We lose quasars uniformly in redshift space. The median $\mathrm{S/N_{QSO}}$ of QSOs is ~ 7.5. In our Mg II catalogue we have $\sim 1,030$ absorbers associated with quasars having $\mathrm{S/N_{QSO}} < 2$ and we remove them for consistency. From now onwards we perform all galaxy-centric analysis with this fiducial quasar sample and Mg II absorber catalogue.

2.4.2 Galaxy Samples

2.4.2.1 Emission Line Galaxies

The main selection criteria for ELGs in SDSS is a cut in the $g-r$ vs $r-z$ colour-colour diagram and g band magnitude (Raichoor et al., 2017). Their typical stellar mass is $\mathrm{M}_\star \sim 10^{10.5}\,\mathrm{M}_\odot$ (Raichoor et al., 2017) and they reside in dark matter haloes with $\mathrm{M}_{\mathrm{halo}} \sim 10^{12.2}\,\mathrm{M}_\odot$ (Favole et al., 2016).

The characteristic features of ELG spectra are strong gas emission lines such as [O II] $\lambda3727$ [O III] $\lambda5007$ and H β, due to high star formation activity. Based on [O II] $\lambda3727$ luminosity and models described in Kennicutt (1998) the star formation rate in ELGs is estimated to be from 1 to 20 $\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. The presence of hot and young stars make ELGs appear blue.

We select the latest DR16 ELG catalogue[6] compiled by Raichoor et al. (2021), which contains 269,178 objects. The catalogue includes stellar masses estimated by the FAST spectral fitting code (Kriek et al., 2009), with the Bruzual & Charlot (2003) stellar model and Chabrier (2003) initial mass function (IMF). We include only ELGs with reliable redshifts (selected by IMATCH==1 condition). We also apply cuts on stellar mass and redshift to connect with Mg II absorbers. The final sample contains 188,323 ELGs at $z > 0.4$ and with $9 < \log\mathrm{M}_\star\,[\mathrm{M}_\odot] < 12$. The typical measurement error in redshift is ~ 20 $\mathrm{km\,s^{-1}}$. The median stellar mass, halo mass and redshift of the ELGs in the final sample are $\mathrm{M}_\star \sim 10^{10.4}\,\mathrm{M}_\odot$, $\mathrm{M}_{\mathrm{halo}} \sim 10^{12.1}\,\mathrm{M}_\odot$ and $z \sim 0.84$, respectively. The ELG catalogue does not include a stellar mass uncertainty for each galaxy, however the typical uncertainty in stellar mass for ELGs is 0.05 dex and the majority are within 0.25 dex (Raichoor et al., 2017, see section 6.3)[7]. The stellar mass-redshift distribution is shown in the left panel of Figure 2.13. For this sample we estimate the star-formation rate (SFR) using the scaling relation from Kewley et al. (2004).

$$\mathrm{SFR}\,[\mathrm{M}_\odot\mathrm{yr}^{-1}] = \frac{\mathrm{L}_{[\mathrm{O\,II}]}\,[\mathrm{ergs\,s^{-1}}]}{1.52 \times 10^{-41}} \tag{2.2}$$

where $\mathrm{L}_{[\mathrm{O\,II}]}$ is luminosity measured as $\mathrm{L}_{[\mathrm{O\,II}]} = \mathrm{F} \cdot 4\pi \mathrm{D}_\mathrm{L}^2$; F is the measured [O II] $\lambda3727$ flux and D_L is the luminosity distance. We note that the measured [O II] $\lambda3727$ flux (F) is not corrected for reddening and the SFR may be underestimated by a factor of a few on average and up to factor of ~ 10 or more for the most massive, star-forming systems. Possible contamination by AGN is also discussed below. Our measured SFRs vary from 1 to 25 $\mathrm{M}_\odot\mathrm{yr}^{-1}$ with a median value of ~ 7 $\mathrm{M}_\odot\mathrm{yr}^{-1}$.

2.4.2.2 Luminous Red Galaxies

The luminous red galaxies in SDSS are observed as described in Padmanabhan et al. (2012). The selection of LRGs is based on cuts on SDSS g, r, and i band magnitudes. Due to low

[6] https://data.sdss.org/sas/dr16/eboss/lss/catalogs/DR16/eBOSS_ELG_full_ALLdata-vDR16.fits
[7] Hong Guo, private communication.

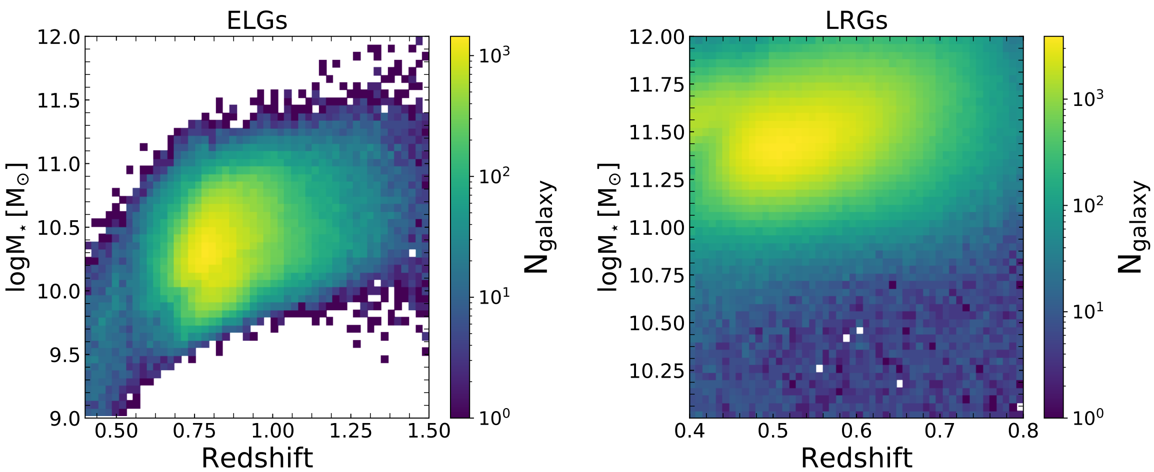

Figure 2.13: Joint distribution of stellar mass and redshift of all the galaxies used in this work. **Left:** Distribution of ELGs, where we have selected galaxies with $z > 0.4$ and $9 < \log M_\star\,[\mathrm{M}_\odot] < 12$. The mean redshift and stellar mass of ELGs are $\langle z \rangle \sim 0.86$ and $\langle M_\star \rangle \sim 10^{10.4}\,\mathrm{M}_\odot$ respectively. **Right:** Distribution of LRGs, where we have selected galaxies with $z > 0.4$ and $10 < \log M_\star\,[\mathrm{M}_\odot] < 12$. The mean redshift and stellar mass of LRGs are $\langle z \rangle \sim 0.54$ and $\langle M_\star \rangle \sim 10^{11.5}\,\mathrm{M}_\odot$ respectively.

star formation activity, a large fraction of these galaxies have no emission lines in their spectra. The presence of old stellar populations makes them look optically red. They are typically more massive than ELGs and have $M_\star \sim 10^{11.2}\,\mathrm{M}_\odot$ and reside in dark matter haloes with $M_{\mathrm{halo}} \sim 10^{13.5}\,\mathrm{M}_\odot$ on average (White et al., 2011).

For LRGs we take the Wisconsin PCA-based catalogue[8] which contains a total of $\sim 1,489,670$ objects. The stellar masses were estimated using a principal component analysis (PCA) method described in Chen et al. (2012a), assuming Kroupa (2001) IMF. The authors also added 0.057 dex to adjust to a Chabrier (2003) IMF (see Herrmann et al., 2016). The stellar libraries and single stellar population models were taken from Maraston & Strömbäck (2011). The typical measurement error in redshift is $\sim 30\ \mathrm{km\,s^{-1}}$. For our galaxy-absorber correlation analysis we select LRGs having $z > 0.4$ with $10 < \log M_\star\,[\mathrm{M}_\odot] < 12$. Our final LRG catalogue contains $\sim 1,081,329$ galaxies and the median stellar mass, median halo mass, and redshift $M_\star \sim 10^{11.5}\,\mathrm{M}_\odot$, $M_{\mathrm{halo}} \sim 10^{14}\,\mathrm{M}_\odot$ and $z \sim 0.54$, respectively. The typical uncertainty in stellar mass is 0.16 dex. The stellar mass-redshift distribution is shown in the right panel of Figure 2.13.

[8]https://www.sdss.org/dr16/spectro/galaxy_wisconsin/

2.4.3 Methods

2.4.3.1 Covering Fraction of Mg II absorbers

We define $N^{\mathrm{abs}}_{\mathrm{gal,j}}|^{\Delta z}_{\Delta \mathrm{R}}$ the number of detected absorbers (corrected for their completeness factor) around the j^{th} SDSS galaxy within an annulus ΔR satisfying a maximum Δz separation. Analogously, $N^{\mathrm{QSO}}_{\mathrm{gal,j}}|_{\Delta \mathrm{R}}$ is the corresponding number of quasars, i.e. the number of sightlines. The subscript ΔR denotes an annulus with projected inner and outer radii R_1 and R_2, and $|\Delta z| = |z_{\mathrm{galaxy}} - z_{\mathrm{abs}}|$ is the redshift separation; we adopt $|\Delta z| \leq 0.01$.

Given a galaxy with a position on the sky and redshift (α, δ, z) we derive the apparent covering fraction of absorbers, defined as the fraction of quasar sightlines, in a given radial bin ΔR, which have one or more absorbers satisfying a chosen EW_{2796} threshold

$$f'_{\mathrm{c}}|_{\Delta R} = \frac{N(\text{sightlines with absorbers})}{N(\text{QSO sightlines})} = \frac{\sum_j N^{\mathrm{abs}}_{\mathrm{gal,j}}|^{\Delta z}_{\Delta \mathrm{R}}}{\sum_j N^{\mathrm{QSO}}_{\mathrm{gal,j}}|_{\Delta \mathrm{R}}}. \tag{2.3}$$

where j varies over all galaxies in given galaxy sample. Using our estimated completeness as a function of absorber rest equivalent width and redshift we then derive the true covering fraction f_c by correcting the apparent covering fraction f'_c as $f_c = f'_c/c(\mathrm{EW}_{2796}, z)$. As $c(\mathrm{EW}_{2796}, z) \leq 1$, the corrected $f_c \geq f'_c$.

In addition to this differential covering fraction, we also estimate cumulative covering fractions. For a given projected distance $D < D_{\mathrm{proj}}$ we count all absorber-galaxy pairs and QSO-galaxy pairs up to that distance, applying the same definition as above.

We measure the differential (cumulative) mean absorption strength per absorber around a given galaxy sample, in (up to) a given radial bin, as the sum of rest equivalent widths of all absorbers (weighted by their corresponding completeness corrected number), normalized by the sum of Mg II absorbers (completeness corrected) in (up to) that bin. This quantity is the average total (both lines of the doublet) rest equivalent width per absorber.

2.4.3.2 Random Galaxy Samples

To better understand the bias and clustering of Mg II absorbers around galaxies, we compare to the expected average absorption signals for random sky sightlines. To do so, we define 40 random galaxy samples, each equal in size to our fiducial galaxy sample. Starting from true galaxies, we shuffle sky positions as well as redshifts. This shuffling de-correlates the positions and redshifts of the galaxies while preserving the original redshift distribution and sky coverage. We compute all observational measurements for both the true and random galaxy samples.

2.4.3.3 Bootstrap Error Estimates

To estimate errors we use the bootstrap approach, repeating the absorber/QSO-galaxy pair identification procedure 100 times. In each iteration we select galaxies from the original sample at random, with repetition, to account for Poisson statistics. We take the mean

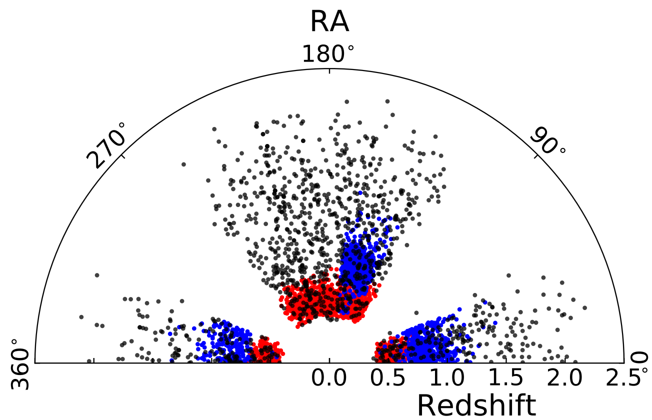

Figure 2.14: Wedge map of DR16 Mg II absorbers, and the two galaxy samples in RA-z space. A subsample of N = 1000 objects is selected randomly to show the distribution of galaxies and absorbers on the sky. The black circles indicate the Mg II absorbers, blue circles indicate the ELGs and red circles indicate the positions and redshifts of LRGs.

over these 100 iterations to estimate each quantity and take the error as the standard deviation of these samples.

2.4.3.4 Halo Mass and Virial Radius

We perform analysis requiring two unobservable quantities: halo mass and virial radius. To estimate halo mass for a given stellar mass and redshift we make use of the Stellar Mass-Halo Mass (SMHM) model developed in Behroozi et al. (2010). Given halo mass, the virial radius r_{200} of the galactic halo is found using the $\Delta_c \sim 200$ definition of Bryan & Norman (1998)

$$r_{200} = r_{vir} = 211.83\, M_{12}^{1/3}\, [\Omega_{m,0}(1+z)^3 \, + \, \Omega_{\Lambda,0}]^{-1/3}\, [\mathrm{kpc}] \tag{2.4}$$

where $M_{12} = M_{\rm halo}/10^{12}\,[M_\odot]$. Note that adopting other SMHM relations would lead to different halo mass and radii estimates.

2.4.4 Galaxy-Absorber Correlation

Figure 2.14 shows the positions and redshifts of 1000 objects chosen randomly from the galaxy samples and our DR16 Mg II absorber catalogue. The black dots indicate the positions of detected Mg II absorbers and blue and red dots denote ELGs and LRGs, respectively. ELGs clearly extend out to $z \sim 1$ compared to LRGs that extend to $z \sim 0.7$. The map also shows that the Mg II absorber distribution is quite uniform in RA-z space and that many absorbers fall within the volume of the galaxy samples.

We begin with the *excess mean number of Mg II absorbers per kpc*^{2} (hereafter, excess mean surface density), i.e. the surface density around true galaxies divided by the surface

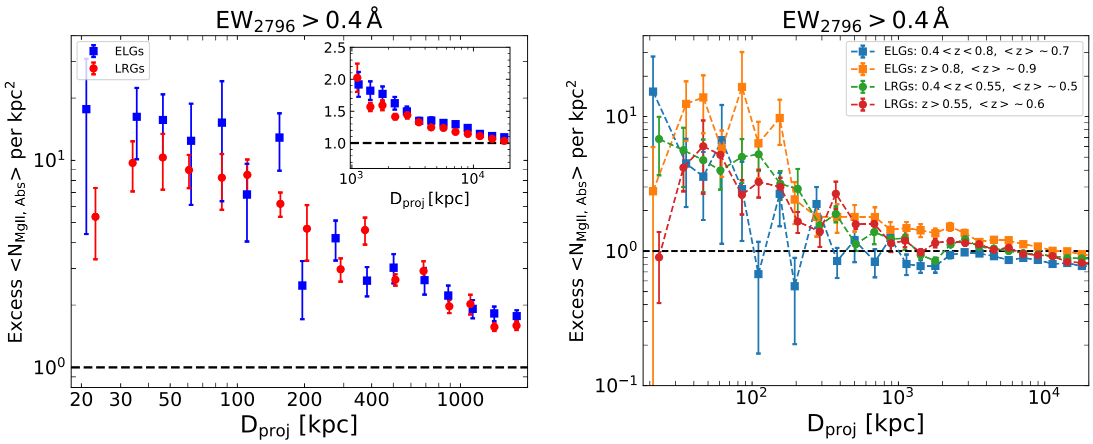

Figure 2.15: Excess mean surface density of absorbers ($d_1 < D_{\rm proj} < d_2$): the mean number of absorbers per galaxy per kpc^2 for Mg II absorbers with EW$_{2796}$ > 0.4 Å divided by the corresponding value around the random galaxy sample. **Left:** The blue and red squares denote ELGs and LRGs, respectively. The corresponding dashed black line denotes unity, i.e. the expectation for random sightlines. The inset shows that the values converge to their random expectation only at $D_{\rm proj} \gtrsim 10\,$Mpc. **Right:** Excess mean surface density of absorbers around ELGs and LRGs in two redshift bins. The surface density of absorbers increases towards higher redshift at larger distances.

density around the random galaxy sample, as a function of projected distance ($D_{\rm proj}$). Figure 2.15 shows the result for Mg II absorbers with EW$_{2796}$ > 0.4 Å. For other EW bins the trends are similar, with mean surface density decreasing for stronger absorbers, due to their relative scarcity.

For ELGs the excess surface density rises to a maximum at $\sim 50 - 80$ kpc and then declines with projected distance, however, for LRGs, the trend is consistent with decreasing as a function of distance. The decreasing trend of Mg II surface density with projected distance around LRGs was also identified with a different stacking based methodology (Zhu et al., 2014), as well as with individual galaxy-absorber pairs in the COS-LRG survey (Zahedy et al., 2019). We also see an enhancement of Mg II absorbers around ELGs relative to LRGs below $D_{\rm proj} \sim 100$ kpc, by a factor of $2 - 4$. For these two samples the mean surface density is consistent within the error bars beyond $D_{\rm proj} \sim 100$ kpc.

The excess mean surface density of absorbers converges to the random expectation (unity) only for $D_{\rm proj} \gtrsim 15\,$Mpc for both LRGs, and ELGs, as shown in the inset. To understand how excess surface density depends on redshift of galaxies we further divided ELGs and LRGs into redshift bins and estimate the excess surface density around each. As shown in the right panel of Figure 2.15 we find that within a given sample the excess surface density is slightly higher for low-z galaxies below $D_{\rm proj} \sim 100$ kpc. They all converge to random values at $D_{\rm proj} \gtrsim 10$ Mpc.

Next, we estimate the covering fraction, $f_c(d_1 < D_{\rm proj} < d_2)$, of Mg II absorbers around

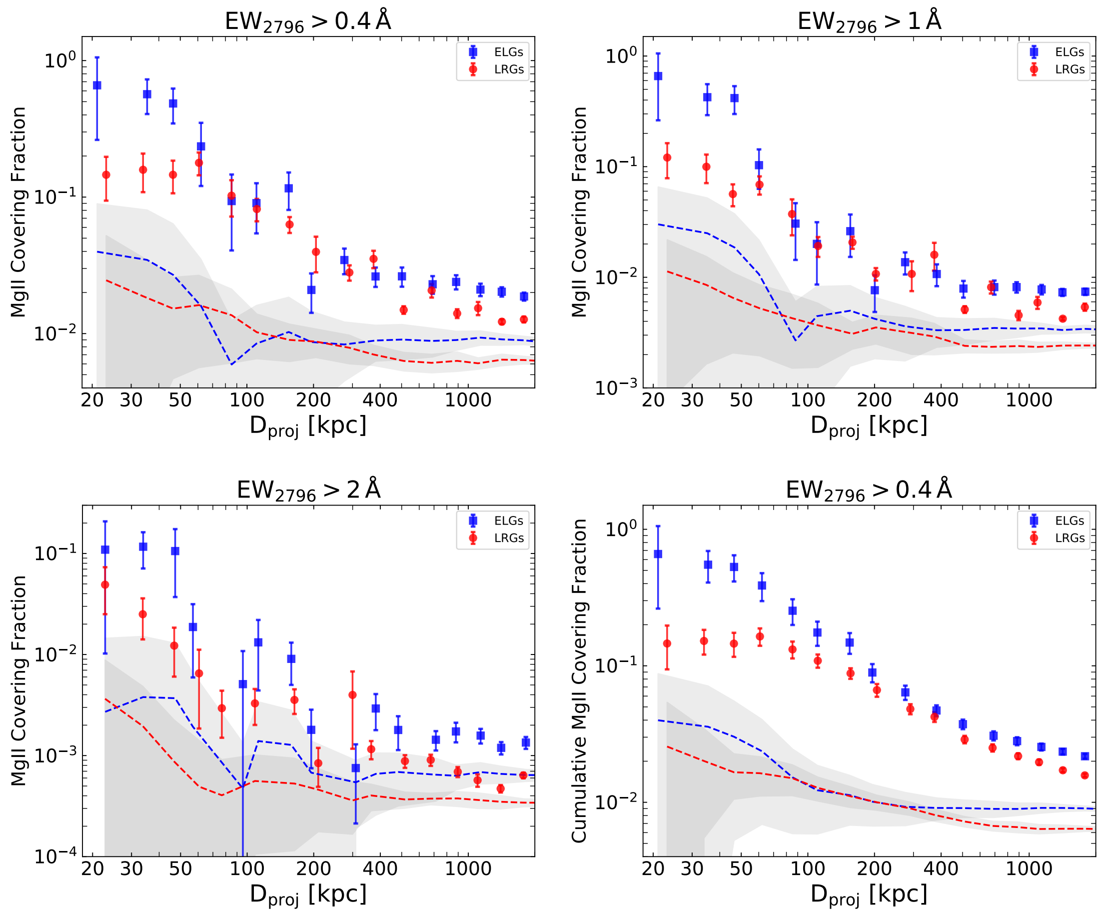

Figure 2.16: Differential covering fraction of Mg II absorbers in three EW_{2796} bins (shown at the top of each panel). The blue and red squares denote ELGs and LRGs, respectively. The corresponding dashed color lines show the covering fractions expected around the random samples. The ELG sample has a pronounced excess relative to the LRGs within $D_{\rm proj} \lesssim 50$ kpc. **Bottom right:** Cumulative covering fraction of Mg II absorbers ($d < D_{\rm proj}$) $\mathrm{EW}_{2796} > 0.4$ Å bin around star-forming ELGs (blue) and passive LRGs (red) as a function of projected distance. The corresponding dashed color lines show cumulative covering fractions estimated using the random samples. Excess absorption around ELGs is visible out to $D_{\rm proj} \sim 200$ kpc.

galaxies. We show the variation of f_c as a function of projected distance in Figure 2.16. The covering fraction decreases with projected distance and varies strongly with galaxy type. In each EW_{2796} bin ELGs have 2-5 times higher covering fraction than LRGs below $D_{\mathrm{proj}} < 50$ kpc. For example, in the $\mathrm{EW}_{2796} > 0.4$ Å bin, ELGs have a Mg II covering fraction of $f_c \gtrsim 50-70$ percent compared to $f_c \lesssim 15$ percent for LRGs. At large distances ($D_{\mathrm{proj}} > 100$ kpc) there is no significant variation with galaxy type. Several previous studies have reported similar trends using different samples and analysis (Lan et al., 2014; Nielsen et al., 2013; Lovegrove & Simcoe, 2011). The covering fraction converges to the random expectation for $D_{\mathrm{proj}} \gtrsim 10$ Mpc.

In addition to the trend with galaxy type, the covering fraction also depends strongly with the strength of absorbers (note the vertical limits vary with each panel). There is a clear anti correlation between covering fraction and absorber strength. This can be attributed to the rarity of strong absorbers around galaxies, such that absorption with $EW_{2796} > 2$ Å is roughly an order of magnitude less frequent in the inner halo. Particularly for LRGs, however, there are low number statistics and the measurement errors are high. Nonetheless, the strong enhancement of f_c seen within 50 kpc is clearly present for all rest equivalent widths. As we discuss below, this small distance component implies an association to the ongoing star formation activity of ELGs and the resultant galactic-scale outflows.

We also investigate the cumulative covering fraction $f_c(d < D_{\mathrm{proj}})$, and the result as a function of projected distance is shown in bottom right panel of Figure 2.16. The trend reflects what we have previously seen with the differential covering fraction: star-forming galaxies (ELGs) have 2-4 times higher cumulative values than passive galaxies (LRGs) up to $D_{\mathrm{proj}} < 100$ kpc. The measurement errors are also smaller as the cumulative estimation significantly enhances the statistics. We now clearly see a difference in f_c between the two samples which extends up to ~ 200 kpc. This partially reflects the different physical sizes of the gaseous (and dark) haloes hosting LRGs versus ELGs. For LRGs, 100 kpc corresponds to $\lesssim 0.2r_{\mathrm{vir}}$ (near the central galaxy), and gas accreted from halo or IGM may not reach into these central regions (Huang et al., 2016). However, for ELGs, this corresponds to $\sim 0.3r_{\mathrm{vir}}$ where metal rich gas ejected by powerful galactic outflows or winds can be deposited (Muratov et al., 2017; Nelson et al., 2019; Mitchell et al., 2020).

2.4.5 Dependence on Stellar and Halo Mass

To understand how the covering fraction evolves with the size of the dark matter halo we normalize the projected distances by the virial radii of galaxies. In Figure 2.17 we show the cumulative Mg II covering fraction in stellar mass bins (different for ELGs and LRGs). We see a weak anti-correlation between the stellar mass of ELGs and the covering fraction of Mg II absorbers when normalizing by r_{vir}. In contrast, LRGs show a strong stellar mass dependence in f_c. Higher mass passive galaxies have systematically lower cumulative Mg II covering fractions at all distances. For example, LRGs with $\mathrm{M}_\star < 10^{11}\,\mathrm{M}_\odot$ have cumulative Mg II covering fractions up to ~ 25 percent compared to $2-3$ percent around more massive LRGs ($\mathrm{M}_\star > 10^{11.5}\,\mathrm{M}_\odot$) below $D_{\mathrm{proj}} \sim 0.3r_{\mathrm{vir}}$.

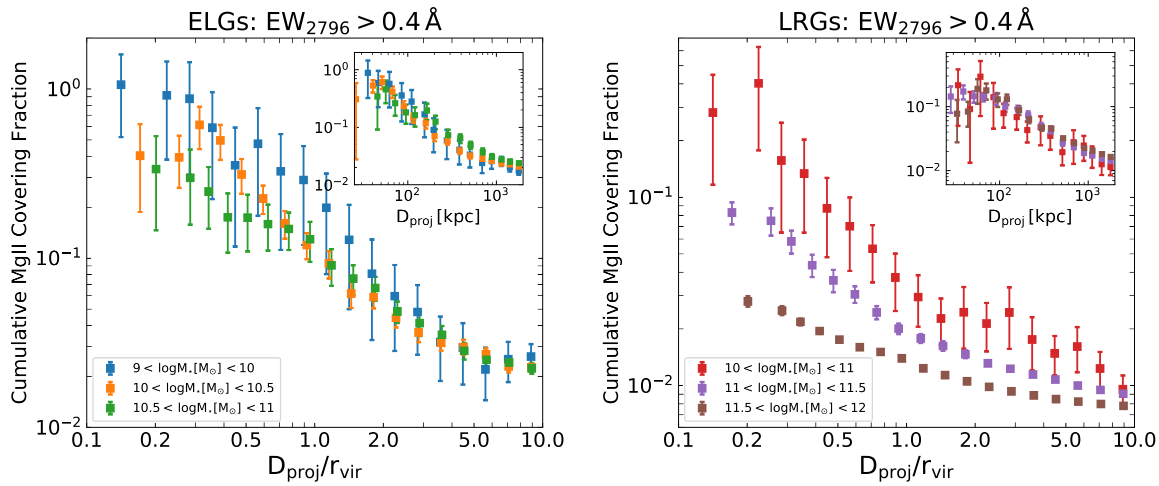

Figure 2.17: Dependence of the cumulative Mg II covering fraction on stellar mass, around ELGs (left panel) and LRGs (right panel), as a function of distance normalized by the virial radius. The insets show the covering fraction as a function of physical kpc. In each panel, the galaxies are divided into stellar mass bins as shown. We focus here on absorbers with $EW_{2796} > 0.4$ Å for clarity, since the other two EW bins show similar behavior. We see a very weak stellar mass trend for ELGs, however, there is a strong negative correlation for LRGs. Symbols are shifted horizontally for visual clarity.

The reason for this strong stellar mass dependence may be two-fold. First, the tight correlation between stellar and halo mass implies that the size, and total mass, of the circumgalactic medium increases rapidly for more massive galaxies. Second, gas is also thermalized to higher temperatures in more massive haloes, which would naturally inhibit the formation of cooler gas phases. In this regime, heating by the virial shock produces long-term 'hot-mode' growth, which is the dominant accretion mode for LRG-type galaxies (Birnboim & Dekel, 2003; Kereš et al., 2005; Nelson et al., 2013).

We show the stellar mass dependence of covering fractions versus physical kpc in the insets of Figure 2.17. Here we see only a weak stellar mass dependence of f_c on physical separation in kpc, and predominantly at large distances, where massive galaxies have higher covering fractions. The lack of a strong trend at fixed physical distance suggests that the signature above is largely driven by the $r_{\rm vir}$ normalization, i.e. the increasing size of more massive haloes.

To better explore these mass trends, we derive the cumulative Mg II covering fraction at $D_{\rm proj} \lesssim r_{\rm vir}$, a rough outer boundary of the CGM, as a function of stellar mass. The result is shown in the left panel of Figure 2.18, where red and blue markers represent LRGs and ELGs, respectively. For ELGs we see a decreasing trend of Mg II covering fraction from $\sim$ 11 percent for $M_\star \sim 10^{10}\,M_\odot$ to $\sim$ 9 percent for $M_\star \sim 10^{11}\,M_\odot$. For LRGs we see a similar decreasing trend of covering fraction with stellar mass from $\sim$ 3 percent for low mass galaxies to just $\sim$ 1 percent for massive galaxies. Qualitatively this shows a clear

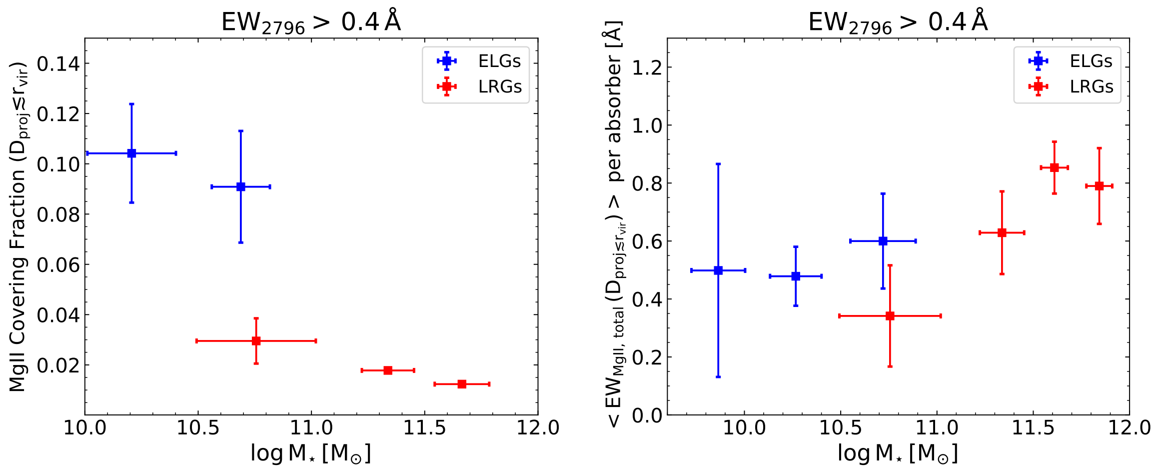

Figure 2.18: **Left:** Mg II covering fraction at $D_{\rm proj} \lesssim r_{\rm vir}$ around ELGs (blue) and LRGs (red) as function of stellar mass. Both samples show the same signal: more massive galaxies have lower covering fractions of cool gas in their CGM. **Right:** Total Mg II rest equivalent width per absorber within $D_{\rm proj} \lesssim r_{\rm vir}$ as a function of stellar mass. The absorption strength increases with the stellar mass of galaxies for both galaxy types. In both panels we show absorbers with $\mathrm{EW}_{2796} > 0.4$ Å only as the other two bins show qualitatively similar results.

picture: more massive galaxies host less cool gas, on average, in their circumgalactic media.

In the right panel of Figure 2.18 we show the cumulative total Mg II absorption EW per absorber within $D_{\rm proj} \lesssim r_{\rm vir}$ (see section 2.4.3.1) as a function of stellar mass. We see an increasing trend in absorption strength with stellar mass for ELGs, albeit with large error bars. The cumulative absorption EW varies from ~ 0.5 Å below $\mathrm{M}_\star \sim 10^{10}\,\mathrm{M}_\odot$ to ~ 0.6 Å around $\mathrm{M}_\star \sim 10^{11}\,\mathrm{M}_\odot$. For LRGs, the trend is much stronger and values vary from 0.3 Å to 0.8 Å from the least to most massive galaxies. Broadly, this result implies that either Mg II absorbers have larger EWs around more massive galaxies, or that sightlines intersect a larger number of individual absorbers in more massive haloes. A similar positive correlation between equivalent width of Mg II absorbers and stellar mass was observed in star-forming galaxies (Bradshaw et al., 2013).

In the left panel of Figure 2.19 we show this same mean total rest equivalent width per absorber, i.e. $\mathrm{EW}_{2796} + \mathrm{EW}_{2803}$, cumulative as a function of $d \lesssim D_{\rm proj}$. At small distances (less than 100 kpc) the absorption strength declines with projected distance. For LRGs at $D_{\rm proj} < 50$ kpc, we see an enhancement up to 3 Å, compared to $\lesssim 0.8$ Å, at larger radii. ELGs show a similar behavior, but have systematically lower rest equivalent widths, up to $D_{\rm proj} \lesssim 400$ kpc, remaining slightly above the LRGs at larger distances. This is principally due to the significant difference in the average redshift between the two galaxy samples and the reason could be possibly related to the different gas properties at different epochs. To understand the possible redshift contribution to this difference we divide the ELGs and LRGs into low and high-z bins (splitting at the medians). We clearly see a slightly higher

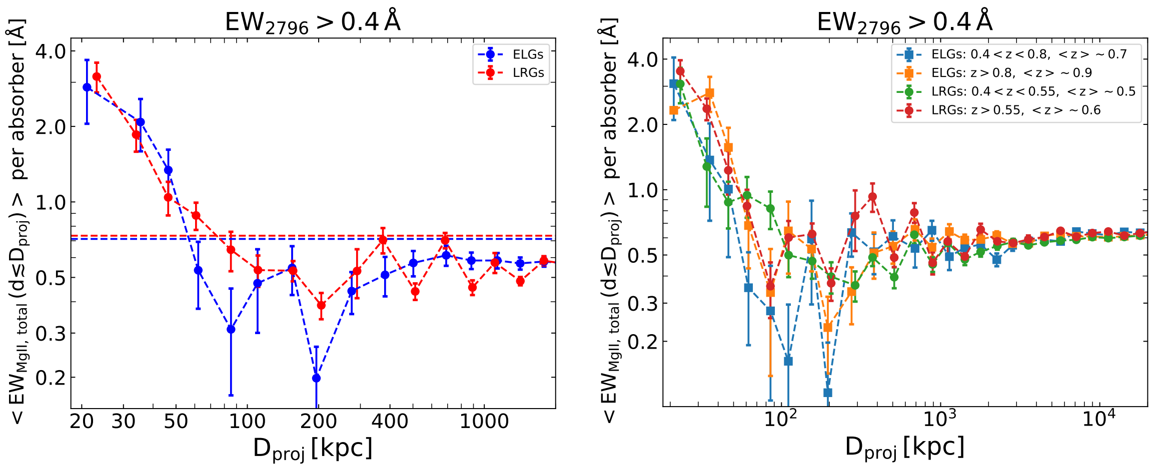

Figure 2.19: **Left:** Total Mg II rest equivalent width (EW_{2796}+EW_{2803}) per absorber within $d \lesssim D_{\rm proj}$ around ELGs (blue) and LRGs (red) as a function of projected distance. The dashed lines show the mean values. There is a clear decreasing trend below $D_{\rm proj} < 50$ kpc for both ELGs and LRGs. **Right:** Redshift evolution of total Mg II rest equivalent width around ELGs and LRGs. We see a weak positive correlation between average redshift and total EW.

EWs around high-z galaxies at distances $D_{\rm proj} \gtrsim 1$ Mpc. We show the redshift evolution of EW in the right panel of Figure 2.19 for both ELGs and LRGs. The weak positive correlation between total EW and redshift is clearly visible for a given galaxy sample.

Nevertheless, we can still see in each redshift interval the same characteristic scale dependence of the total rest equivalent width per absorber: a strong decrease to a tentative minimum value at projected radii of $\sim$ 150 kpc, followed by a slight rise to value that remains constant beyond 1 Mpc. This feature is more clear for ELGs than LRGs, and may be related to gas inflow processes onto dark matter haloes, which act to shock-heat infalling gas at the boundary of the halo.

2.4.6 Dependence on Star Formation Rate

A key physical property of galaxies that may affect the distribution of cold gas traced by Mg II absorbers is the star formation rate (SFR; Lan & Mo, 2018; Rubin et al., 2018). The ELG catalogue includes measurements for [O II] λ3727 flux which traces the star formation activity in ELGs for $z > 0.4$ (Kennicutt, 1998). Despite its simplicity and neglect of dust attenuation effects, this SFR value can still be used to perform qualitative studies.

In the left panel of Figure 2.20 we show the dependence of cumulative and differential Mg II covering fraction on star formation rate for ELGs. We divide the galaxies into two SFR bins: low (SFR $< 10\,{\rm M_\odot yr^{-1}}$) and high (SFR $> 10\,{\rm M_\odot yr^{-1}}$). We observe a strong variation of covering fraction with star-formation activity. Galaxies with high star-formation rate have 2-3 times higher cumulative covering fraction than their low SFR

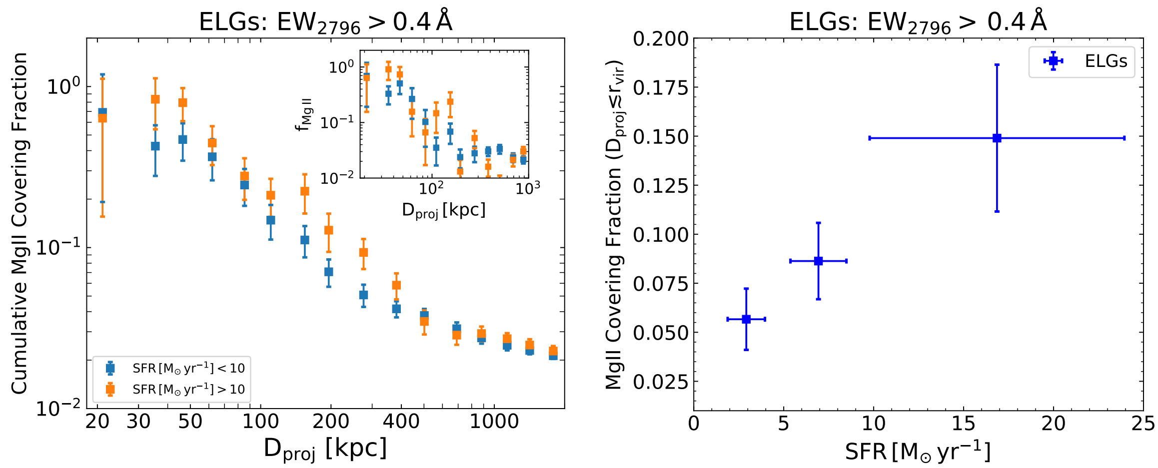

Figure 2.20: **Left:** Star formation rate dependence of cumulative Mg II covering fraction, $f_c(d < D_{\rm proj})$ around ELGs as a function of projected distance from galaxy. Inset shows the differential covering fraction as a function of projected distance from galaxy in two SFR bins. **Right:** Mg II covering fraction at $D_{\rm proj} \lesssim r_{\rm vir}$ as a function of SFR of ELGs. A strong positive correlation between covering fraction and SFR of ELGs is visible.

counterparts, all the way up to $D_{\rm proj} \sim 400$ kpc, close to the $r_{\rm vir}$ of ELGs. One the other hand the differential covering fraction (shown in the inset) is 2-5 times higher below ~ 50 kpc in galaxies with high star-formation rate.

In the right panel we show the cumulative covering fraction at $D_{\rm proj} \lesssim r_{\rm vir}$ as a function of SFR. We again see a strong positive correlation with SFR, though the highest SFR bin has large uncertainties. This increasing trend of covering fraction with SFR indicates that SF activity in ELGs plays an important role in enriching the cold gas in their CGM. We also point out that for massive star-forming galaxies a possible source of influence is the central active galactic nucleus (AGN). We have not tried to exclude AGN from our samples. The galactic outflows powered by AGNs can significantly enrich the metal abundance in the CGM of massive star-forming galaxies (Veilleux et al., 2005), and future work will focus on disentangling the roles of star formation versus AGN activity.

2.4.7 The Relative Kinematics of Galaxies and Absorbers

With robust spectroscopic information for both the absorbing gas clouds and parent galaxies we can study the line-of-sight velocities to constrain the relative motion between the two. For this purpose we estimate the line-of-sight velocity separation, $\Delta v = c\Delta z/(1+z)$, where c is the speed of light and z is the galaxy redshift.

In the left panel of Figure 2.21 we show the distribution of Δv for Mg II absorber-galaxy pairs, separating ELGs (blue) from LRGs (red). Here we show absorbers with $EW_{2796} > 0.4$ Å within $10\,{\rm kpc} < D_{\rm proj} < 400\,{\rm kpc}$ from the galaxy. The velocity distribution of Mg II absorbers around ELGs (shown in blue solid line) is well characterized by a single gaussian with mean $\langle \Delta v \rangle \sim -11(\pm 8)$ $\mathrm{km\,s^{-1}}$and dispersion, $\sigma_v \sim 135$ $\mathrm{km\,s^{-1}}$(shown in the dashed blue line). The velocity distribution for LRGs (shown in the red solid line) in the central region is best-fit by a single gaussian of mean $\langle \Delta v \rangle \sim -9(\pm 10)$ $\mathrm{km\,s^{-1}}$and dispersion, $\sigma_v \sim 200$ $\mathrm{km\,s^{-1}}$(shown in the dashed red line). The mean velocity difference is in agreement with previous studies such as Huang et al. (2021), even though the samples and methods are significantly different. This further suggests that our analysis yields consistent trends. The typical dark matter halo velocity dispersion (Δv_{200}) for ELGs is $\sim$ 140 $\mathrm{km\,s^{-1}}$ and ~ 350 $\mathrm{km\,s^{-1}}$ for LRGs (Elahi et al., 2018). This implies $\sigma_{\mathrm{ELG}} \sim \sigma_{\mathrm{ELG,\,halo}}$ while $\sigma_{\mathrm{LRG}} \sim 0.6\,\sigma_{\mathrm{LRG,\,halo}}$. While the cool CGM is consistent with virial motion around star-forming galaxies, it is sub-virial for massive quiescent systems.

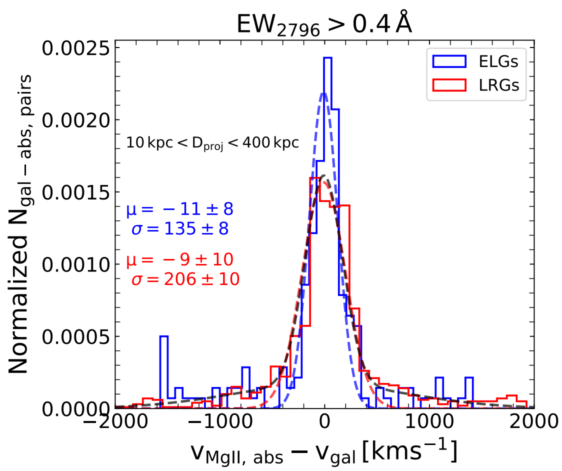

Figure 2.21: Line-of-sight (LOS) velocity difference between galaxies and detected cool gas absorption. **Left:** Distribution of relative velocities for galaxy-absorber pairs within $10\,\mathrm{kpc} < \mathrm{D_{proj}} < 400\,\mathrm{kpc}$, blue and red curves showing ELGs and LRGs, respectively (solid lines). The distributions are characterized by single gaussian profiles shown in dashed color lines. A two-component gaussian fit to the LRG distribution is also shown in the dashed black line, which better characterizes the high-velocity tails. Best fit parameters (only for single gaussian) are shown in the panel. **Right:** Line-of-sight velocity separation between galaxy-absorber pairs as a function of projected distance. The colored circles denote the median value in each radial bin and vertical bars show the 5th to 95th percentile ranges in each bin. The blue and red dashed line shows the escape velocity corresponding to the typical halo masses of ELGs and LRGs, respectively. The Hubble flow is indicated by the shaded grey region. Note that we have slightly offset horizontally the points and lines for visual clarity.

Visual inspection shows that a single gaussian is not a good representation of the tails of these two Δv distributions. A multi-component gaussian has been shown to better characterize the velocity distribution around LRGs (Huang et al., 2016; Chen, 2017). We therefore fit a double gaussian (shown with the dashed black line) to the LRG distribution. The best-fit means are consistent with zero, while the two velocity dispersions represent

Table 2.4: Line-of-sight velocity separation between galaxy-absorber pairs (within 10 kpc < D_{proj} < 400 kpc) as a function of galaxy stellar mass. $\langle \Delta v \rangle$ and σ are the mean and standard deviation of each distribution. Δv_{10}, Δv_{50}, Δv_{90} are 10th, 50th and 90th percentiles of the velocity separations.

	$\log M_\star$ [$M_\odot$]	$\langle \log M_\star \rangle$ [$M_\odot$]	$\sigma_{\log M_\star}$ [$M_\odot$]	$\langle \Delta v \rangle$ [$\mathrm{km\,s^{-1}}$]	σ [$\mathrm{km\,s^{-1}}$]	Δv_{10} [$\mathrm{km\,s^{-1}}$]	Δv_{50} [$\mathrm{km\,s^{-1}}$]	Δv_{90} [$\mathrm{km\,s^{-1}}$]
ELGs:	Full Sample	10.4	0.33	-55	543	-768	28	325
LRGs:	Full Sample	11.4	0.33	51	529	-462	33	584
	[9, 10]	9.8	0.2	151	341	-68	52	513
ELGs:	[10, 10.5]	10.3	0.14	-47	531	-760	39	360
	[10.5, 11]	10.7	0.15	-120	576	-1100	-17	278
	[10, 11]	10.5	0.24	122	394	-100	85	286
LRGs:	[11, 11.5]	11.3	0.12	28	547	-455	-34	607
	[11.5, 12]	11.7	0.12	66	528	-498	125	578

a narrow component with $\sigma_{\mathrm{narrow}} \sim 180\ \mathrm{km\,s^{-1}}$, describing the majority of pairs (≈ 85 percent, central distribution), together with a broad component with $\sigma_{\mathrm{broad}} \sim 1100\ \mathrm{km\,s^{-1}}$ describing the tails. We discuss possible interpretations in Section 2.5.

In the right panel of Figure 2.21 we show the relative velocity separation of galaxy-absorber pairs as a function of projected distance from the galaxy. The median values of the distribution in each radial bin are shown in colored circles for ELGs (blue) and LRGs (red). The vertical bars show the 5th to 95th percentiles of the distribution in each radial bin. The corresponding dashed color lines show the decreasing escape velocity for a dark matter halo mass, estimated using the average masses of the galaxies. The shaded gray region shows the increasing Hubble flow at the average redshift of the galaxy samples.

The median values are always within $\sim 100\ \mathrm{km\,s^{-1}}$and well below the escape velocity of the dark matter haloes. There is a small positive offset for ELGs, possibly due to gas moving away from ELGs towards the observer powered by strong galactic outflows, while gas behind the galaxy is detected less frequently due to obscuration effects. At small distances, $D_{\mathrm{proj}} < 50$ kpc, the distributions are narrow and the vast majority of galaxy-absorber pairs have low relative velocities. At larger distances $D_{\mathrm{proj}} > 200$ kpc the tails of the distributions imply a small fraction of absorbing clouds with velocities higher than the escape velocity of dark matter halo for both ELGs and LRGs.

To study the dependence of gas kinematics on the stellar mass of galaxies we divide the galaxy samples into stellar mass bins and estimate the mean $\langle \Delta v \rangle$ and standard deviation σ, as well as the 10th, 50th, and 90th percentiles of the velocity distributions within 10 kpc < D_{proj} < 400 kpc. The values are compiled in Table 2.4. For both ELGs and LRGs, we see an increasing trend of gas velocity dispersion σ with stellar mass. For ELGs, this can be attributed to the positive correlation between SFR and stellar mass, and consequently stronger galactic outflows. However, for LRGs, this is more likely related

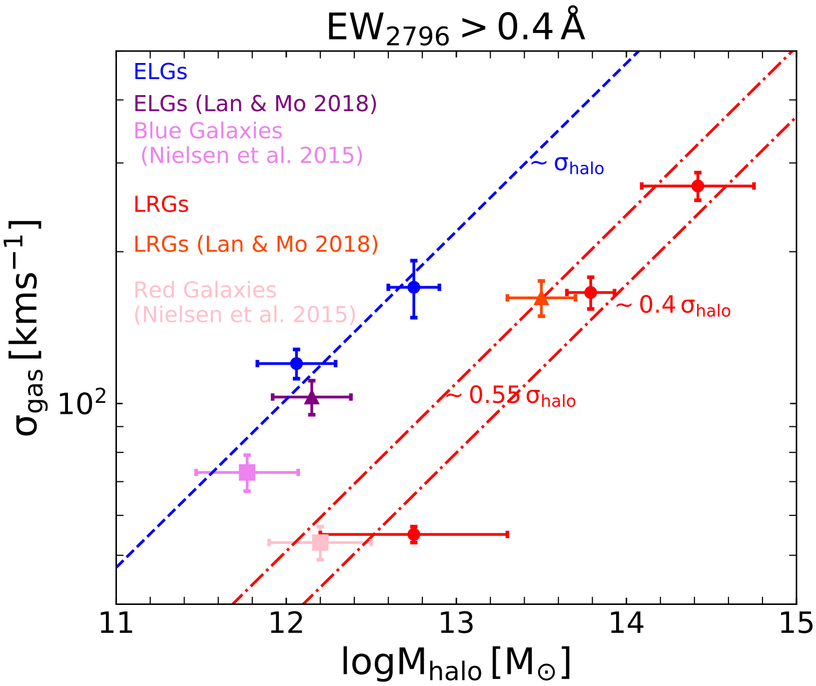

Figure 2.22: The line-of-sight velocity dispersion, $\sigma_{\rm gas}$, of galaxy-absorber pairs, within $10\,{\rm kpc} < {\rm D_{proj}} < 400\,{\rm kpc}$, as a function of halo mass and galaxy types. The dark blue and red dots represent ELGs and LRGs, respectively. The dashed blue line shows the expected velocity dispersion as a function of halo mass of ELGs. The two dash-dotted red lines show the 0.5 and 0.4 times the dark matter velocity dispersion of LRGs. We also show the comparison with previous measurements from Nielsen et al. (2015) in violet and pink squares, and Lan & Mo (2018) in purple and orange-red triangles. The motion of gas around LRGs is suppressed relative to the expected velocity dispersion given their dark matter halo masses.

to a trend of gas accretion rate with halo mass, either from the IGM or from satellite galaxies and their interaction with pre-existing halo gas.

Finally, we directly assess how our measured dispersion values depend on halo mass. The result is shown in Figure 2.22. For simplicity the velocity dispersion is estimated by fitting a single gaussian in each halo mass range for both ELGs and LRGs. We clearly see that the gas motion around LRGs is suppressed ($0.4 - 0.55\,\sigma_{\rm halo}$, as indicated by the dot-dashed red lines) over a large range of halo masses ($M_{\rm halo} \sim 10^{13} - 10^{14.5}\,M_\odot$). This is in contrast to ELGs (blue points), which exhibit dispersions similar to $\sigma_{\rm halo}$, as indicated by the dashed blue line. The trend for ELGs is also visible up to $M_{\rm halo} \sim 10^{13}\,M_\odot$. We contrast against measurements from previous studies (Nielsen et al., 2015; Lan & Mo, 2018), which we here extend to a larger mass range – this comparison is discussed in the next section.

2.5 Discussion: The Nature of Cold Gas around Galaxies at $0.5 < z < 1$

2.5.1 Comparison with Previous Studies

2.5.1.1 Mg II Covering Fraction

We now make a comparison between our new results and previous studies. First, we quantitatively compare our covering fractions with values obtained in Lan et al. (2014) for $\mathrm{EW}_{2796} > 1$ Å and Lan & Mo (2018) for $\mathrm{EW}_{2796} > 0.4$ Å, shown in Figure 2.23. Our analysis extends over larger spatial scales due to the increased statistics offered in DR16, but within the same scales probed, the results agree well within the error bars. We observe that for $\mathrm{EW}_{2796} > 0.4$ Å, f_c slightly lower than Lan & Mo (2018) below $D_{\mathrm{proj}} \sim 50$ kpc. However, a similar covering fraction dependence with galaxy type is present. The minor discrepancy is likely due to the non-trivial differences in methodologies and analysis choices.

For $\mathrm{EW}_{2796} > 0.4$ Å Lan & Mo (2018) estimated covering fraction using a stacking analysis. They assumed that the average Mg II absorption is dominated by absorbers with $\mathrm{EW}_{2796} > 0.4$ Å and can be expressed as the product of covering fraction and the average value of EW_{2796} estimated from the intrinsic distribution of individual absorbers toward random quasar sightlines. Lan et al. (2014) estimated the excess in the number of galaxies found within a given impact parameter w.r.t reference quasars as opposed to random positions. Then by counting galaxies around the selected quasars they estimated the denominator in the covering fraction expression. The error bars were estimated with Poisson statistics.

The uncertainties on Lan & Mo (2018) values are smaller than ours even though the absorber sample is smaller. This is due to the small errors on the average Mg II strength in stacked spectra. For $\mathrm{EW}_{2796} > 1$ Å absorbers we agree well with Lan et al. (2014) at $D_{\mathrm{proj}} <$ 100 kpc even though our methods and resulting absorber catalogues are substantially different. For instance, Lan et al. (2014) used the Zhu & Ménard (2013a) Mg II catalogue (based on SDSS DR7 quasars) which has different completeness characteristics than our catalogue. We also note that the Lan & Mo (2018) results show less clear evidence for a sudden decline in covering fraction at a projected radius of 50 kpc.

2.5.1.2 Trends with Galaxy Properties

A detailed quantitative comparison of the stellar mass dependence of the Mg II covering fraction with results from previous studies lies beyond the scope of this chapter. The study of Lan (2020) of the stellar mass dependence of covering fraction indicates that it is important to study these trends in relatively narrow redshift intervals, because there are relatively strong evolutionary trends in covering fraction with redshift at fixed mass. When this is accomplished, Lan (2020) find that after normalizing the projected distance by the virial radii the covering fraction of absorbers with rest equivalent widths (EW_{2796})

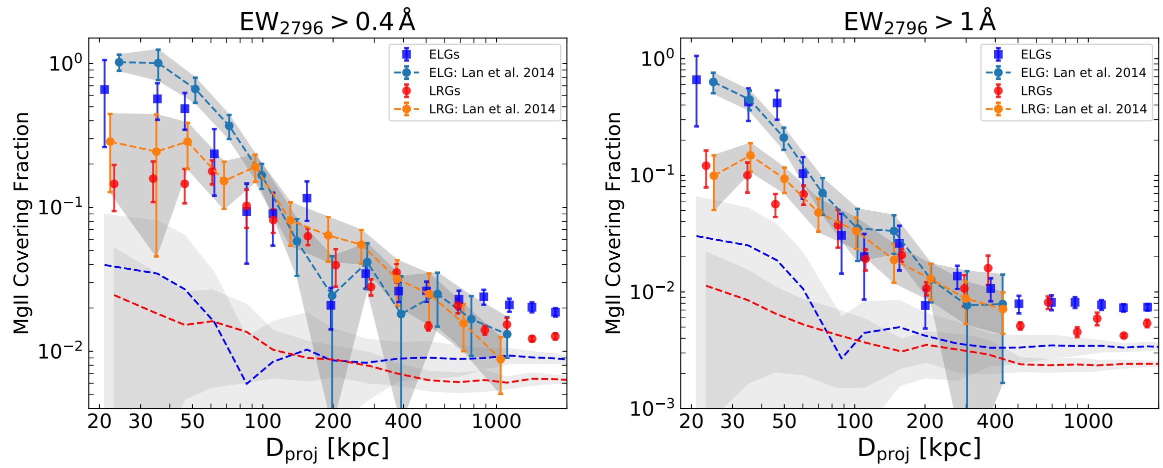

Figure 2.23: Comparison of the differential Mg II covering fraction with values obtained in Lan et al. (2014) and Lan & Mo (2018) around ELGs and LRGs. **Left:** Comparison for Mg II absorbers with $\mathrm{EW}_{2796} > 0.4$ Å. **Right:** Comparison for Mg II absorbers with $\mathrm{EW}_{2796} > 1$ Å. Our results agree well with these literature values, recovering the same qualitative and relative trends with respect to galaxy type, despite significant methodological and analysis differences.

greater than 1 Å shows very little dependence on stellar mass for star-forming galaxies, but decreases with stellar mass for the passive galaxy population. The weak dependence on mass that we find for our ELG sample is in agreement with these findings.

For LRGs we see a strong anti-correlation of covering fraction with stellar mass after normalizing projected distance by the virial radii, such that f_c is lower for more massive galaxies (bottom panel of Figure 2.17). These results are also in agreement with Lan (2020) for strong absorbers ($\mathrm{EW}_{2796} > 1$ Å), who found that the covering fraction (on projected scale) varies weakly with stellar mass as $f_\mathrm{c} \propto M_\star^{0.3}$, and varies from 2 percent from low mass to 3 percent for high mass LRGs at $D_\mathrm{proj} \gtrsim 100$ kpc. For similar stellar masses, we find that the covering fraction varies from 3 percent to 7 percent. Note that our LRGs are more massive than DESI passive galaxies.

2.5.1.3 The Kinematics of Galaxy-Absorber Pairs

By connecting galaxy-absorber pairs in velocity space we estimated the ensemble velocity distributions of gas clouds associated with galaxies. Using single Gaussian profiles, we have characterized the velocity dispersion for absorbers (Figure 2.22). Using the line width of the Mg II doublet in stacked spectra, similar velocity dispersion values were obtained in Lan & Mo (2018). That work also reported a large velocity dispersion comparable to the dark matter halo velocity dispersion for ELGs and sub-virial gas motion ($\sim 0.5\sigma_\mathrm{halo}$) around LRGs. The sub-virial velocity of cold gas was also observed by Huang et al. (2016) for both passive LRGs and [O II] emitting LRGs (see also Chen, 2017). Furthermore, the velocity dispersion scaling is in rough agreement with Nielsen et al. (2015, 2016) who characterized the galaxy-absorber relative velocity around blue and red galaxies using high-resolution spectra of quasars. The existence of broad components in the velocity distributions of Mg II absorbers around massive red galaxies is in agreement with the results of Kauffmann et al. (2017), who proposed that the large velocity separation Mg II absorbers trace gas that has been pushed out of the dark matter haloes by multiple episodes of AGN-driven mechanical feedback acting over long time-scales.

2.5.1.4 The main new results in this chapter

Thanks to the larger quasar and galaxy samples that are included in our study, we are able to characterize the scale dependence of Mg II absorber covering fraction and rest equivalent width with greater accuracy. We find that the Mg II covering fraction for both LRGs and ELGs declines very rapidly at a projected radius of 50 kpc. Our results are consistent with previous findings, but the greater S/N of our measurements allows us to better characterize the sharpness of the break, indicative of a rather sudden transition between the regime where physical properties of the CGM are regulated by galactic outflows and the regime where the CGM is in thermal balance set by the gravitational field of the dark matter halo.

On scales of a few hundred kpc, the average total Mg II EW per absorber dips below the characteristic field value. This is seen both for LRGs and ELGs, but is a stronger effect for the latter. We speculate that this dip is related to gas inflow processes onto dark matter haloes.

2.5.2 Clues to the origin of Mg II absorbers

The velocity separation between galaxy-absorber pairs provides clues to the kinematics of absorbing gas around galaxies. The gas around galaxies can acquire motion due to several physical processes. The parent dark matter halo can accrete pristine gas from the intergalactic medium which can be enriched by metals ejected via galactic outflows, stellar winds, or supernovae in star-forming galaxies. The inflow velocity of gas accreting from the IGM can be smaller or similar to the velocities expected from gravitational free-fall. In addition, cold gas can be in virial equilibrium with orbiting halo gas e.g. rotating with the halo orbital velocity. Alternatively, the metals forming inside stars can be thrown out of the galaxies by powerful supernovae with velocities up to $v \sim 500 - 1000\ \mathrm{km\,s^{-1}}$(Dalla Vecchia & Schaye, 2008; Sharma & Nath, 2013). The powerful outflows from supermassive black holes can likewise eject gas out to $\sim$ tens of kpc with velocities exceeding ~ 3000 $\mathrm{km\,s^{-1}}$(Circosta et al., 2018; Perrotta et al., 2019; Nelson et al., 2019).

The high velocity dispersion around ELGs suggests that the origin of cold gas traced by Mg II absorbers around ELGs is likely due to powerful galactic outflows. Such powerful outflows around star-forming galaxies have been observed (Steidel et al., 2010; Bordoloi et al., 2014; Rubin et al., 2014; Zhu et al., 2015). This conclusion is also supported by the particularly strong enhancement of the Mg II covering fraction within $D_{\mathrm{proj}} < 50$ kpc. Using higher fidelity imaging, Lan & Mo (2018) found an enhancement of Mg II absorption along the minor axes of ELGs, which also supports the outflow scenario. In contrast, accreted gas from the halo or satellite galaxies is predicted to align along the major axes of galaxies (Péroux et al., 2020b). Finally, we also see a strong positive correlation between the SFR of ELGs and covering fractions (differential values) (Figure 2.20), evidence that star driven outflows play a pivotal role in metal enrichment at $\lesssim 100$ kpc or $D_{\rm proj} \lesssim r_{\rm vir}$ in star-forming galaxies.

On the other hand, the suppression of gas motion around LRGs indicates that the absorbing gas associated with them is gravitationally bound and unlikely to have originated from galactic outflows. This is further supported by the low star formation activity in LRGs and negligible contribution of stellar activity to the abundance of metals around them (Afruni et al., 2019). A similar suppression of gas velocity has been found for Ly α, and O VI gas around COS-halos galaxies (Tumlinson et al., 2011, 2013). The interaction between cold and hot gas inside the halo also plays an important role in deciding the fate of cold gas around LRGs. As pointed in Huang et al. (2016), the drag exerted by the hot CGM can slow down the gas motion around LRGs and, cool clumps of gas can fall towards the central galaxy due to orbital decay. However, the evaporation time is significantly smaller than the infall time and most cool clumps would therefore evaporate before reaching the galaxy, or small distances (Zahedy et al., 2019). The observed existence of cold gas around LRGs implies that there is some form of balance between heating and cooling inside the halo such that the cold gas is routinely formed and destroyed in the halo (Sharma et al., 2012; Voit, 2018). The origin, and survival, of such clouds of cool gas remains an unsolved theoretical question (Schneider et al., 2018; Nelson et al., 2020; Das et al., 2021a).

Furthermore, our analysis also shows that a non-negligible fraction (≈ 15 percent) of LRGs have high-velocity gas clouds associated with them. As it is unlikely that galactic outflows powered by stellar activity could account for this population, one possible explanation could be AGN-driven outflows. The origin of cool gas around LRGs is undoubtedly therefore diverse and results from a combination of processes including accretion from the IGM, ram-stripping of the gas from satellite galaxies, mass loss from massive halo stars, and metal-enrichment by AGN driven outflows (Bordoloi et al., 2011; Huang et al., 2016; Lan & Mo, 2018).

Combining our results with previous measurements (Figure 2.22), we see that the different behaviour of the gas motion around star-forming versus quiescent galaxies supports the picture that the origin of cold gas is fundamentally different for star-forming versus passive galaxies.

2.6 Summary

In this work we have developed a fully automated quasar continuum estimator and absorption line detection pipeline. We have run the pipeline on the SDSS DR16 quasar sample and compiled the largest Mg II metal absorber catalogue to date. Our main findings are:

- Our Mg II catalogue contains 160,000 Mg II absorbers ($0.36 < z < 2.3$) based on the SDSS DR16 final quasar sample. The median redshift and Mg II absorption strength of the catalogue are $z \sim 1.14$ and $EW_{2796} \sim 1.3$ Å, respectively. Fe II lines are also commonly detected and have smaller rest equivalent widths than Mg II as they have weaker oscillator strengths.

- Stacking quasar spectra in the rest-frame of Mg II absorbers, we detect the presence of other weak metal lines including Si I , C IV, Al II, Al II, Mg I and Fe II. We study their properties as a function of Mg II line strength.
- The measured doublet ratio of Mg II absorbers shows that most lie within the theoretical limit of the Mg II doublet ratio, i.e. 1 (saturated case) and 2 (unsaturated case). Strong absorbers are almost always saturated.
- To investigate the completeness of our detection pipeline we simulate ~ 33 million fake absorbers. Completeness is a strong function of both rest equivalent width and redshift of absorbers. Our method is naturally less complete for weak absorbers ($EW_{2796} < 0.5$ Å) and reasonably complete for strong absorbers. Our detection pipeline performs excellent on quasars with relatively high S/N. On the other hand purity is also very high ($\sim 95\%$) for our catalogue.

In the second half of the chapter we connect Mg II absorbers to the latest SDSS DR16 catalogue of ~ 1.3 million galaxies, divided into star-forming, emission line galaxies (ELGs) and quiescent, luminous red galaxies (LRGs). We use these samples to study the incidence and properties of metal absorption in the circumgalactic medium and nearby intergalactic medium. We investigate the absorber-galaxy correlation based on several properties of galaxies to understand the physics which most affect the properties of gas around galaxies. Our main results are:

- The mean surface density of Mg II absorbers is larger than the expected random background for both ELGs and LRGs. At large distance, it reaches the expectation for random sightlines at a distance ($D_{\rm proj} \sim 15$ Mpc) for both ELGs and LRGs. The mean surface density is also a strong function of galaxy type within $D_{\rm proj} < 50$ kpc, with ELGs showing a strong enhancement in the inner halo.
- The covering fraction of Mg II absorbers varies strongly with galaxy type. ELGs have 2 - 4 times higher covering fractions than LRGs within $D_{\rm proj} < 50$ kpc, regardless of $\rm EW_{2796}$ strength. The covering fraction decreases with projected distance and decreases strongly with increasing rest equivalent width (EW). For $D_{\rm proj} < 400$ kpc, the covering fraction is larger than the expected random background and converges to this value only at $D_{\rm proj} \gtrsim 10$ Mpc. A similar trend is visible for the cumulative covering fraction, which is clearly enhanced in star-forming versus passive galaxies up to $D_{\rm proj} \lesssim 200$ kpc.
- The average rest equivalent width of Mg II absorption per absorber shows a clear enhancement close to galaxies, $D_{\rm proj} < 50$ kpc for both ELGs and LRGs. LRGs always have systematically higher values than ELGs. There is a weak evolution of

EW with galaxy redshift, such that higher z galaxies have slightly higher EWs at all $D_{\rm proj} \gtrsim 1$ Mpc. In all redshift bins, the Mg II average rest equivalent width declines to the field value beyond ~ 150 kpc.

- When normalizing projected distances by the virial radii of galaxies we find that the covering fraction varies strongly with galaxy type. There is a weak stellar mass trend for ELGs, but a strong stellar mass trend for LRGs, whereby the covering fraction decreases with increasing stellar mass, consistent with previous studies. The trend is visible even at $D_{\rm proj} \sim 2r_{\rm vir}$.

- Splitting ELGs by star-formation activity we find that Mg II covering fraction is positively correlated with SFR, supporting the idea of a galactic outflow origin for the cold gas around ELGs.

- We trace the kinematics of galaxy-absorber pairs with the distribution of line-of-sight relative velocity Δv. Around ELGs, this is well-characterized by a single gaussian profile, while LRGs require a second, broad component with $\sigma_v \sim 1100$ km s^{-1} to capture high-velocity tails, likely indicative of AGN-driven outflows. The velocity dispersion is higher around LRGs than ELGs. The majority of the absorbers close to galaxies ($D_{\rm proj} < 100$) kpc have low velocities and are likely gravitationally bound.

- The velocity dispersion of cool CGM gas increases with stellar (and halo) mass for both samples. The gas motions around ELGs are similar to the expected dark matter halo velocity dispersion, while gas motions around LRGs are suppressed and closer to $\sim 0.5\sigma_{\rm DM,halo}$. The different properties and, moreover, trends of Mg II absorption with galactic properties between ELGs and LRGs implies that cool circumgalactic gas around star-forming versus quiescent galaxies has a fundamentally different physical origin.

For me, the distant future and far-off galaxies is where it's at. That's where my imagination can really come out to play.

Alastair Reynolds

Chapter 3

Cool circumgalactic gas in galaxy clusters: The nature of Mg II absorbers in clusters

The contents of this chapter is based on Anand et al. (2022) published in Monthly Notices of Royal Astronomical Society, 513, 3210

3.1 Introduction

This chapter extends our galaxy-absorber cross-correlation analysis to galaxy clusters and characterises the spatial distribution and relative kinematics of Mg II absorbers in $z \sim 0.5$ clusters. In this high-mass regime we have the potential to shed light on the role of environment in shaping the CGM of galaxies. Cluster halo gas, known as the intracluster medium (ICM), is heated up to high temperatures $T \sim 10^7 - 10^8$ K due to gravitational collapse. In addition, outflows powered by supermassive black holes in the centre of clusters can also heat the gas (Zuhone & Markevitch, 2009; Yang & Reynolds, 2016). The hot ICM emits mostly at X-ray wavelengths due to radiation from thermal bremsstrahlung produced in the highly ionized gas (for a review, see Böhringer & Werner, 2010; Kravtsov & Borgani, 2012). Although the ICM is hot, cold/cool gas has sometimes been detected in and around clusters. The most frequently observed elements are hydrogen (Hα, Lyα) (Yoon & Putman, 2017; Muzahid et al., 2017; Pradeep et al., 2019) and metal absorption lines (Mg II , O VI) (Lopez et al., 2008; Burchett et al., 2018; Lee et al., 2021), which are detected in the spectra of background quasars.

Similarly, recent radio observations with the Very Large Array (VLA) have revealed a large amount of neutral hydrogen (H I) in the Virgo cluster, mostly stripped from cluster galaxies due to ram-pressure of hot ICM (Vollmer et al., 2012; Bahé et al., 2013). In addition to H I, observations have also revealed the presence of cool low-ionization gas (e.g., Mg II) in groups and clusters (Fossati et al., 2012; Dutta et al., 2020; Hamanowicz et al., 2020; Dutta et al., 2021b) and submillimeter observations have also revealed long extended filaments of cold molecular gas in cool-core clusters (Edge, 2001). In a recent work (Olivares et al., 2019), the authors analysed ALMA (CO lines) and MUSE (Hα) data of three clusters: Centaurus, Abell S1101, and RXJ1539.5 and detected long (3-25 kpc) extended molecular gas filaments in the inner part of cluster halo. By cross-correlating SDSS DR3 quasars and clusters from the Red-Sequence Cluster Survey (RCS), Lopez et al. (2008) and Padilla et al. (2009) found that strong absorbers ($EW_{2796} > 1$ Å) are more abundant within < 1 Mpc than < 2 Mpc from the cluster centre, while weak absorbers follow the expected distribution of Mg II absorbers in field. The authors argued that the difference could be explained if the strong absorbers trace the overdensity of galaxies in clusters. In a recent analysis, Lee et al. (2021) performed a similar Mg II - galaxy cluster cross-correlation with SDSS DR14 quasars and redMaPPer clusters and found ~ 3 times higher detection rate of Mg II absorbers in inner part ($D_{\rm proj} \lesssim r_{\rm vir}$) of clusters than at large distances ($D_{\rm proj} \gtrsim 3 - 4r_{\rm vir}$).

On the other hand, simulations have also predicted the formation and survival of cold gas in cluster environments. Recently, Qiu et al. (2020) reported a radiation-hydrodynamic simulation of AGN feedback in galaxy clusters (particularly for cool-core clusters) and proposed a scenario in which cold gas filaments ($T \sim 10^4$ K) can condense out of warm-hot (10^4 to 10^7 K) AGN outflows, as they move out of the central galaxy. Idealized simulations of gas in the ICM have shown that cold gas can form in-situ out of hot ICM due to local temperature fluctuations and thermal instabilities (TI), if the ratio of thermal instabilities time-scale to the free-fall time-scale is below some critical threshold ($t_{\rm TI}/t_{\rm ff} \lesssim 10$) (Sharma et al., 2012; McCourt et al., 2012; Voit & Donahue, 2015; Voit et al., 2017). Besides the in-situ formation of cold gas in clusters (Dutta et al., 2021a), other phenomena such as interactions between cluster galaxies (Wang, 1993) and ram pressure stripping (Gunn & Gott, 1972) of cold gas from satellite galaxies (Cortese et al., 2007; McCarthy et al., 2008; Yun et al., 2019) can also give rise to cold gas in clusters. Therefore, it is important to explore the nature of cool low-ionization gas in cluster environments and compare its incidence and physical properties with cool gas surrounding isolated galaxies in order to understand its origin and impact on galaxy evolution.

In this chapter, our goal is to use the latest datasets of Mg II absorbers and galaxy clusters to perform a high S/N absorber-cluster cross-correlation study to shed light on the nature and origin of cool metal gas ($T \sim 10^4$ K, traced by Mg II absorbers) in cluster environments. We aim to detect Mg II absorption and understand how it traces the underlying halo, as well as to connect cool gas with the physical properties of clusters and their member galaxies. We also explore the nature of Mg II absorbers as a function of equivalent width, and study their properties in the inner and outer parts of clusters to shed light on physical processes that might impact the observed properties of cool gas in dense environments. As far as we are aware, this is the largest statistical study of Mg II absorbers in galaxy clusters to date, over a broad range of redshift, $0.4 < z < 1$ and halo mass, $10^{13.8}\,{\rm M}_\odot < {\rm M}_{500} < 10^{14.8}\,{\rm M}_\odot$.

The outline of the chapter is as follows: Section 3.2 introduces the samples of Mg II absorbers, galaxy clusters and galaxies. We describe the methods in Section 3.3 and present our main results in Section 3.4. Finally, we discuss the possible interpretations of our findings and conclusions in Section 3.5, followed by a detailed summary of our results in Section 3.6. We adopt a Planck-consistent cosmology (Planck Collaboration et al., 2020) for our analysis, i.e. $\Omega_{\mathrm{m},0} = 0.307$, $H_0 = 67.7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\Lambda,0} = 1 - \Omega_{\mathrm{m},0}$ throughout the chapter.

3.2 Data

3.2.1 Mg II absorber catalogue

We use our recent Mg II /Fe II absorber catalogue[1] (MPA-SDSS catalogue, Anand et al. (2021)) based on ~ 1 million quasars from SDSS data release (DR16) (Ahumada et al., 2020; Lyke et al., 2020). The catalogue includes $\sim 160,000$ Mg II systems and $\sim 70,000$ Fe II confirmed systems (see Table 3.1). The completeness and purity both are very high for the catalogue as shown in Anand et al. (2021), where the pipeline to model quasar continuum and detect absorbers is described in detail.

To summarize, the detection of absorption lines uses a dimensionality reduction technique, called Non-negative Matrix Factorization (NMF, see also Lee & Seung, 1999; Zhu & Ménard, 2013a), to estimate the quasar continuum, followed by a matched kernel convolution technique and adaptive S/N criteria to automatically detect the Mg II /Fe II doublets in spectra. In this chapter, we focus our analysis on systems with $EW_{2796} > 0.4\,\text{Å}$ (denotes the equivalent width of Mg II $\lambda2796$ line) and redshifts $z > 0.4$, which constitutes $\sim 95\%$ of the absorbers in the parent catalogue for which the completeness estimates are determined to be high ($\gtrsim 40 - 50\,\%$, and goes up to $\gtrsim 90\,\%$ for strong systems). The sample with $EW_{2796} > 0.4\,\text{Å}$ is fairly complete ($\gtrsim 30 - 40\,\%$) and results obtained for this sample should be considered as most robust to uncertainties resulting from observational selection effects.

In the current analysis, we also divide absorbers into two EW_{2796} bins, namely $0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$ (*weak* absorbers) and $EW_{2796} > 1\,\text{Å}$ (*strong* absorbers) to characterize the absorber properties in cluster environments as a function of absorber strength. In addition, we also show some results where we combine these two bins (denoted by $EW_{2796} > 0.4\,\text{Å}$) to investigate a few global trends and compare with previous studies. Note that we take only those absorbers that are associated with quasars having $\mathrm{S/N_{QSO}} > 2$, as both the completeness and the detection purity of absorbers are high for these quasars (see section 3.3 in Paper I, for more discussion).

3.2.2 Galaxy Clusters

We use the galaxy clusters identified in Data Release 8 (DR8) of the legacy imaging survey (Dey et al., 2019) of the Dark Energy Spectroscopic Instrument (DESI). The full galaxy cluster catalogue was compiled by Zou et al. (2021)[2]. It includes 540,432 clusters with

[1] publicly available at www.mpa-garching.mpg.de/SDSS/MgII/

[2] publicly available at https://cdsarc.cds.unistra.fr/viz-bin/cat/J/ApJS/253/56

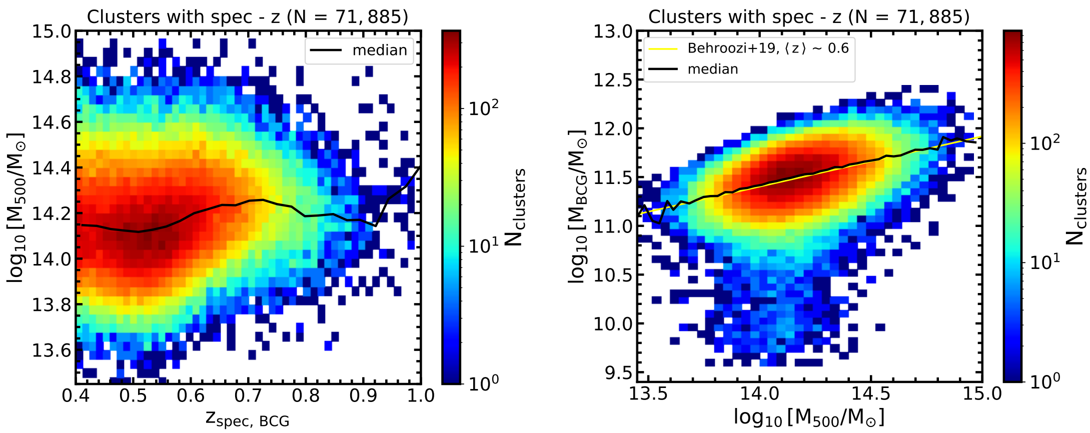

Figure 3.1: **Left:** 2D histogram of DESI galaxy cluster total mass (M_{500}) and spectroscopic redshifts of BCGs. **Right:** 2D distribution of DESI galaxy cluster's BCG stellar mass and total halo mass (M_{500}) of clusters. The yellow solid line show the analytical stellar mass-halo mass (SMHM) model at $z \sim 0.6$ from Behroozi et al. (2019). In both panels, the black solid lines show the median values in square bins.

photometric redshifts, $z_{\rm photo} \lesssim 1$, for which $122,390$ systems also have the spectroscopic redshifts[3] for their brightest cluster galaxies (BCG) (given as 'SZ_BCG' in the catalogue). In this chapter, we use the clusters with spectroscopic redshifts for their BCGs in order to perform a robust cross-correlation study in both projected and redshift space. Since only $z > 0.4$ Mg II absorbers can be detected in SDSS spectra due to wavelength coverage, we apply the same redshift cut $z_{\rm BCG} > 0.4$ for selecting the final cluster sample with which we will perform the absorber-cluster correlation study. After applying this redshift cut, the final sample includes $N = 71,885$ galaxy clusters. Zou et al. (2021) do not provide the completeness limits for the stellar mass of BCGs and cluster halo mass as a function of redshift. However, as described in Zou et al. (2021) the completeness of photo $-z$ galaxy catalogue (Zou et al., 2019) is higher than $\gtrsim 90\%$ for galaxies with $r < 23$ ($r-$band magnitude), which is the magnitude cut for the photo $-z$ galaxies used for cluster identification.

We summarize the sample sizes with different cuts in Table 3.1. The mean redshift of BCGs in the sample and their corresponding errors are $\langle\, z_{\rm BCG}\,\rangle \approx 0.55$ and $\sigma_{z_{\rm BCG}} \lesssim 60$ km s^{-1}, respectively. The mean cluster mass M_{500}, defined as the total mass within the radius r_{500}, where r_{500} is the radius at which the total mass density is estimated to be 500 times the critical density of the universe[4] is $\langle\, M_{500}\,\rangle \sim 10^{14.2}\,{\rm M}_\odot$. The cluster halo masses M_{500} are derived with an empirical scaling relation between luminosity ($L_{1\,\rm Mpc}$) and mass (see also Gao et al., 2020). The empirical relation was derived based on a calibrated sample of galaxy clusters detected in X-ray and Sunyaev–Zeldovich (SZ) surveys (Wen & Han, 2015) that have reliable estimates of M_{500} and R_{500}. The typical r_{500} of clusters in our

[3]Mostly taken from SDSS, in private communication with Hu Zou.
[4]We derive r_{200} from r_{500} as: $r_{200} \approx 1.55\, r_{500}$ (Reiprich et al., 2013).

sample is ~ 700 kpc. The mean stellar mass of the cluster BCGs is $\langle M_{\rm BCG} \rangle \sim 10^{11.5}\,{\rm M}_\odot$. We show the distributions of M_{500} vs redshift in the left panel of Figure 3.1 and M_{500} vs $M_{\rm BCG}$ in the right panel of Figure 3.1. We also plot the median values (solid black line) in each bin in both panels. In the $M_{\rm BCG}$ vs M_{500} panel, we contrast the median values with the analytical stellar mass-halo mass (SMHM) relation from Behroozi et al. (2019) (solid yellow line) and they are consistent. There is considerable scatter around this median line. Part of this scatter arises because clusters are distributed over a broad range of redshifts, $0.4 < z_{\rm BCG} < 0.8$, while the rest is intrinsic scatter.

To better understand the statistical significance of our cluster-absorber correlation, we also estimate the expected average absorption signals for random sky sightlines. For this purpose, following the recipe in Anand et al. (2021), we define 100 *random galaxy cluster samples*, each equal in size to our parent cluster sample. To do so, we shuffle both the true sky positions and redshifts. The shuffling de-correlates the positions and redshifts of the clusters while preserving the original sky coverage and redshift distribution of the DESI legacy imaging surveys. We compute all observational measurements for both true and random samples and show the results.

3.2.3 Photometric Galaxy Sample

To investigate the association of Mg II absorbers in clusters with their galaxy members, we use the photometric sample of galaxies detected in same legacy imaging surveys of DESI. Zou et al. (2019) constructed the photo$-z$ galaxy catalogue which is publicly available[5]. It contains more than ~ 300 million galaxies with $z_{\rm photo} < 1$ out of which ~ 244 million galaxies have photometric redshifts $z_{\rm photo} > 0.4$ and stellar masses in the range $8.5 < \log\,[{\rm M}_\star/{\rm M}_\odot] < 11.8$ (see Table 3.1). The stellar mass cut is applied to retain high completeness independent of galaxy redshift. As described in the previous section, the overall completeness of galaxies used in the current analysis is $\gtrsim 90\%$ as all galaxies have $r < 23$ in the final sample. The photometric z is estimated with a linear regression method developed in Beck et al. (2016). The stellar mass and other galaxy parameters such as star formation rate (SFR) were derived using the Le Phare code (Ilbert et al., 2009) which adopts Bruzual & Charlot (2003) stellar population models and the Chabrier (2003) initial mass function (IMF).

The median photo$-z$ redshift accuracy of the galaxies in our sample is $z_{\rm photo} \approx 0.66$ ($\sigma_{\Delta z_{\rm norm}} \approx 0.017$).[6] The median stellar mass of the galaxies in the sample is $M_\star \sim 10^{10.3}\,{\rm M}_\odot$ (with 0.1 dex uncertainty). Similarly the median SFR and specific SFR (sSFR) are $\approx 2\,{\rm M}_\odot\,{\rm yr}^{-1}$ and $\log[{\rm sSFR/yr}^{-1}] \approx -9.92$, respectively. We show the SFR distribution of the DESI legacy photo$-z$ galaxies (blue) in Figure 3.2.

[5] http://cdsarc.u-strasbg.fr/viz-bin/cat/J/ApJS/242/8

[6] Defined as the standard deviation of $\Delta z_{\rm norm} = \frac{z_{\rm phot} - z_{\rm spec}}{1 + z_{\rm spec}}$ (Zou et al., 2019).

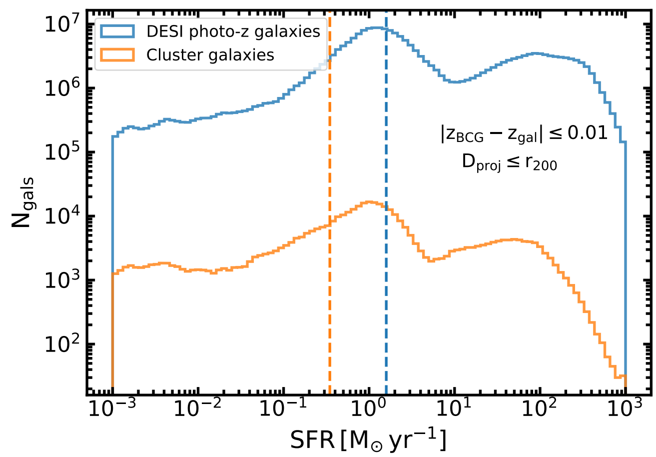

Figure 3.2: 1D distribution of the star formation rate (SFR) of DESI photo$-z$ main galaxy sample (blue) and galaxies that lie within r_{200} of clusters (orange) as defined in section 3.2.4. We see that typical SFR (shown in dashed vertical orange line) of cluster galaxies is $\sim 6-7$ times lower than main galaxy sample (shown in dashed vertical blue line).

3.2.4 Connecting Galaxies with Clusters

In order to find the members of a given cluster, we apply cuts on projected distance and redshift separation between a cluster BCG and nearby galaxies. To this end, we take all the galaxies (g_i) within the projected distance, $D_{\rm proj} \lesssim 2.5$ Mpc[7] from the BCG of the cluster with redshift separation, $|\Delta z| = |z_{\rm BCG} - z_{\rm g_i}| \leq 0.01$[8]. Note that we are using photo$-z$ for our calculations which have typically large uncertainties, and hence we are applying a conservative cut. This means we will intentionally exclude some of the cluster members, which should be kept in mind when interpreting the results. In practice, we experiment with a number of cuts on $|\Delta z|$ to make sure that our main conclusions are not sensitive to this choice.

We show the SFR distribution (in orange) for the cluster galaxies that are found within the r_{200} of the clusters in Figure 3.2. The typical r_{200} is ~ 1 Mpc for our clusters. Note that the typical SFR of cluster galaxies is $\approx 0.3\,{\rm M_\odot\,yr^{-1}}$ which is $\sim 6-7$ times lower than the SFR of galaxies in the full photo$-z$ sample (see section 3.2.3, blue curve in Figure 3.2). On the other hand, the typical stellar mass of the full DESI legacy photo-z galaxy sample is roughly the same as for cluster member galaxies. This implies that galaxies that lie in cluster environments have lower SFR at a given stellar mass, as is expected in cluster environments because most of the cluster galaxies have lost their cool gas due to environmental effects (see also, Cen et al., 2014).

Moreover, we also want to explore the connection between Mg II absorption in clusters

[7]This corresponds to $3-4\,r_{500}$ of our clusters.

[8]The $|\Delta z|$ condition corresponds to $|\Delta v| = \left|\frac{c\Delta z}{1+z_{\rm BCG}}\right| \lesssim 1900\ \ {\rm km\,s^{-1}} \lesssim 2\,v_{500}$, (for $M_{500} \sim 10^{14.2}\,{\rm M_\odot}$) where $c = 3\times10^5\ \ {\rm km\,s^{-1}}$ is the speed of light.

Table 3.1: Sample statistics of Mg II absorbers, clusters and galaxies in this chapter. References are: [a] Anand et al. (2021); [b] Zou et al. (2021); [c] Zou et al. (2019)

Objects	N
SDSS DR16 Mg II Absorbers (S/$N_{QSO} > 2$)	158,508[a]
Mg II Absorbers ($EW_{2796} > 0.4$ Å and $z > 0.4$)	155,141[a]
DESI DR8 photo−z Clusters	540,432[b]
Clusters (with $z_{\rm spec}$)	122,490[b]
Clusters ($z_{\rm photo} > 0.4$)	383,085[b]
Clusters ($z_{\rm spec} > 0.4$)	71,885[b]
DESI Galaxies (with photo−z)	303,379,640[c]
DESI Galaxies ($z_{\rm photo} > 0.4$)	244,428,600[c]

and the stellar activity of its member galaxies. To this end, we divide our clusters into two subsamples: a) clusters with predominantly star-forming galaxies, i.e. the star-forming fraction is larger than half ($f_{\rm star} > 0.5$), or b) clusters with more passive galaxies. In both cases, we consider galaxies within r_{500}, (or r_{200}) of the BCG. For this purpose, we use the specific star-formation rate (sSFR) of the member galaxies from our photo−z galaxy catalogue to estimate the fraction of star-forming and passive galaxies in cluster halo. A value of log(sSFR) $= -11$ often defines the boundary (at $z = 0$) below which galaxies are classified as passive, while galaxies above this limit are considered star-forming. A similar cut was applied by Werk et al. (2012) in COS-Halos survey, though other studies may use slightly different cuts for this purpose (Donnari et al., 2019), e.g, a higher cut, log(sSFR) $= -10$ was applied by Tchernyshyov et al. (2022) in CGM^2 survey galaxies.

We adopt the log(sSFR) $= -11$ threshold, we find that this procedure divides the sample roughly equally. The average fraction of star-forming galaxies within r_{500} is $f_{\rm star} \approx 0.35$ in our clusters. As we move outwards from the cluster centre, the fraction of star-forming galaxies increases and galaxies in the outskirts of cluster are more star-forming (Cen et al., 2014; Butsky et al., 2019).

Previous studies of Mg II absorption aiming specifically to characterize the abundance of cool gas in clusters have had significantly smaller sample sizes ($\approx 3-8$) than the current study – a comparison is shown in Table 3.2. Recently, Mishra & Muzahid (2022) performed a stacking analysis of Mg II absorption around Wen & Han (2015) SDSS clusters with SDSS DR16 quasars (Lyke et al., 2020) and found a significant reservoir of metal-rich gas in the outskirts of clusters.

Table 3.2: Sample statistics comparison versus previous samples. Notes: [1] Number of Mg II absorbers - cluster pairs ($N_{\rm pair}$) within $D_{\rm proj} < 2h^{-1}$ Mpc from the cluster centre; [2] Number of Mg II absorbers found in 82,000 quasar sightlines that are within $D_{\rm proj}/r_{\rm vir} < 5$ from the cluster centre; [3] $N_{\rm pair}$ within $D_{\rm proj} < 2h^{-1}$Mpc from the BCG in our sample; [4] $N_{\rm pair}$ within $D_{\rm proj}/r_{200} < 5$ from the BCG in our study. Our sample sizes are larger by $5-8$ times. [$\star$] S/N$_{\rm QSO} > 2$, i.e. signal-to-noise of quasar spectra. Refs: [a] Schneider et al. (2005); [b] Gladders & Yee (2005); [c] Pâris et al. (2018); [d] Rykoff et al. (2014, 2016); [e] Anand et al. (2021); [f] Ahumada et al. (2020)

Study	EW_{2796} limit	QSOs	Clusters	Mg II Pairs
Lopez et al. (2008)	> 0.3 Å	$46,420^{\rm a}$	$\lesssim 500^{\rm b}$	421^1
Lee et al. (2021)	> 0.3 Å	$526,356^{\rm c}$	$26,000^{\rm d}$	$\sim 200^2$
Current study	> 0.4 Å	$773,594^{\rm e,f,\star}$	$71,885$	1337^3
Current study	> 0.4 Å	$773,594$	$71,885$	2768^4

3.3 Methods

3.3.1 Mg II Covering Fraction

A key value that quantifies the incidence rate of Mg II absorbers in a given environment is the covering fraction (f_c). Similar to our previous analysis (Paper I), we define the covering fraction as the fraction of quasar sightlines in a given projected distance bin ΔR (optionally normalized by r_{500} of the cluster), having one or more absorbers with EW_{2796} larger than a given EW_{2796} threshold.

More precisely, we define the following quantities in the same manner as in Anand et al. (2021). First, $N^{\rm abs}_{\rm BCG,j}|^{\Delta z}_{\Delta \rm R}$ is the number of Mg II absorbers (corrected for their completeness, see below) detected around the $j^{\rm th}$ cluster (defined by its BCG) within an annulus ΔR satisfying a maximum Δz separation. We also correct the absorbers by their completeness factor, $c(\mathrm{EW}_{2796}, z)$, that we estimated in Anand et al. (2021) and the 'effective number' is given as $1/c(\mathrm{EW}_{2796}, z)$. As the completeness factor, $c(\mathrm{EW}_{2796}, z) \leq 1$, the corrected number of absorbers for a given EW_{2796}, z is, $N_{\rm abs} \geq 1$.

Analogously, $N^{\rm QSO}_{\rm BCG,j}|_{\Delta R}$ is the corresponding number of quasars, i.e. the number of sightlines in that annular bin. The subscript ΔR denotes an annulus with projected inner and outer radii R$_1$ and R$_2$, and $|\Delta z| = |z_{\rm BCG} - z_{\rm Mg\,II}|$ is the redshift separation between BCG and the Mg II ; we adopt $|\Delta z| \leq 0.01$[9] in our analysis same as Anand et al. (2021). Moreover, we also show a comparison for the $|\Delta z| \leq 0.003$[10] case in Appendix B.1, to understand if our conclusions change significantly with this choice. We find that decreasing $|\Delta z|$ to such small values does not change our overall conclusions, particularly at small distances from the BCG (see Figures B.1 and B.2).

[9] The $|\Delta z|$ condition corresponds to $|\Delta v| = \left|\frac{c\Delta z}{1+z_{\rm BCG}}\right| \lesssim 1900\ \ {\rm km\,s^{-1}} \lesssim 2\,v_{500}$, (for $M_{500} \sim 10^{14.2}\,{\rm M_\odot}$) where $c = 3\times10^5\ \ {\rm km\,s^{-1}}$ is the speed of light.

[10] This corresponds to $|\Delta v| \leq 600\ \ {\rm km\,s^{-1}}$ which is smaller than the typical v_{500} of cluster halo.

Given a cluster BCG with angular positions and redshift (α, δ, z), we derive the covering fraction (defined above) of absorbers around clusters.

$$f_{\rm c}|^{\Delta R} = \frac{\rm N(sightlines\,with\,absorbers)}{\rm N(QSO\,sightlines)} = \frac{\sum_{\rm j} {\rm N}^{\rm abs}_{\rm BCG,\,j}|^{\Delta z}_{\Delta R}}{\sum_{\rm j} {\rm N}^{\rm QSO}_{\rm BCG,\,j}|_{\Delta R}}. \tag{3.1}$$

where j varies over all clusters in given galaxy cluster sample. If multiple absorbers satisfy the given EW_{2796} threshold, they are counted as multiple sightlines as we are counting absorbers over all clusters (the sum is over all j) in a given radial bin. As described in Anand et al. (2021) and Zhu & Ménard (2013a), the typical width (FWHM) of a single Mg II absorption system can vary from $4-8$ pixels that correspond to $280-500$ $\mathrm{km\,s^{-1}}$in SDSS quasar spectra[11].

We estimate errors by bootstrapping the Mg II - cluster cross-correlation 100 times, as described in Anand et al. (2021).

3.3.2 Mg II Absorption and Surface Mass Density

Another physical quantity that characterizes the strength and overall distribution of Mg II absorbers in a given space around cluster or galaxies, is the mean Mg II absorption strength. To this end, we first estimate the total $\langle\, EW_{2796}\,\rangle$ by summing EW_{2796} , contributed by all absorbers with $EW_{2796}\ >\ 0.4$ Å in a given radial bin ΔR around clusters. Then we divide this total EW_{2796} by the number of absorbers in that radial bin. Next, to account for the incidence rate of absorbers, we also multiply this statistical mean by the covering fraction, to measure the mean $\langle\, EW_{2796}\,\rangle$ of Mg II absorbers per sightline. We combine both weak and strong absorbers here because we want to measure the total Mg II absorption. Mathematically,

$$\langle EW_{2796}\rangle^{\Delta R} = \frac{\sum_i EW^{\Delta R}_{2796,\,\rm i}}{N_{\rm abs}}\, f_c^{\Delta R} \tag{3.2}$$

where i varies over all the absorbers in a given radial bin, $N_{\rm abs}$ is the number of absorbers (i.e. the sum over all i) and $f_c^{\Delta R}$ is the covering fraction in that annular bin (ΔR). Note that these mean values should be close to the mean absorption that we would have obtained by stacking the quasar spectra directly, i.e. accounting for rare (weak on average) absorption in the spectra. We test this hypothesis for SDSS LRGs from our previous analysis (Paper I) and compare the results with Zhu et al. (2014); Lan & Mo (2018)[12], the values are consistent within error bars (see Figure 3.11).

Once we have the mean equivalent widths ($\langle\, EW_{2796}\,\rangle^{\Delta R}$) we use the curve of growth for Mg II assuming we are in the linear regime (Eqn. 3.3), following the derivation in Draine (2011), to estimate the column density (see also Zhu et al., 2014, Eqn. 11). Assuming a thermal broadening factor of $b \sim 4$ $\mathrm{km\,s^{-1}} \implies T \sim 25{,}000\,\mathrm{K}$, the stronger Mg II line (i.e. $\lambda\,2796$) starts saturating at $EW_{2796} \gtrsim 0.15-0.2$ Å (Churchill et al., 2000; Zhu et al., 2014).

[11] Each wavelength pixel in SDSS spectra corresponds to 70 $\mathrm{km\,s^{-1}}$in SDSS spectra.

[12] For more details see also section 2.2 of Lan & Mo (2018).

As a result, only a lower limit on $N_{Mg\,II}$ can be estimated for the strongest absorbers, if we use the linear part of curve of growth. For very weak absorbers in optically thin ($\tau \ll 1$, i.e. unsaturated), regions $\langle N_{Mg\,II} \rangle^{\Delta R}$ can be estimated as

$$\langle N_{\mathrm{Mg\,II}} \rangle^{\Delta R} = 1.13 \times 10^{20} \frac{\langle EW_{2796} \rangle^{\Delta R}}{f_{2796} \lambda_{2796}^2} \mathrm{~cm}^{-2} \tag{3.3}$$

where, EW_{2796} and λ_{2796} are in Å and f_{2796} is the oscillator strength[13] of Mg II $\lambda\,2796$ line. Note that this will only give lower limits on column densities for saturated systems, and we must use the full curve of growth in such cases. Next we estimate the Mg II surface mass density, which is defined as the mass of Mg II absorbers per unit surface area, by multiplying the column densities by the atomic mass of Mg II [14]. We show the results in units of $M_\odot\,kpc^{-2}$.

$$\sum\nolimits_{\mathrm{Mg\,II}}^{\Delta R} = \langle \mathrm{N_{Mg\,II}} \rangle^{\Delta \mathrm{R}} \, \mathrm{m_{Mg\,II}} \tag{3.4}$$

In the section 3.4.1.2, we will further illustrate that most of the absorbers have small $\langle EW_{2796} \rangle$ in clusters and therefore our assumption that they are unsaturated and lie on the linear part of the curve of growth is fairly well justified.

Next, we also estimate the total mass of Mg II gas in clusters and SDSS LRGs within a given radius, using our surface mass density estimates. We describe the methods in Section 3.4.2.2.

3.3.3 Connecting Mg II Absorbers with Cluster Members

To further investigate the connection between Mg II absorbers (detected in clusters) and cluster members (described in section 3.2.3), we also connect Mg II absorbers with cluster galaxies in projected and redshift space. In particular, for each cluster we first find all the absorbers within a given D_{proj} (e.g., $\lesssim r_{200}$). Then we find the nearest member galaxy to each absorber – the results are discussed in Section 3.4.3.2.

Furthermore, we also want to test the hypothesis that absorbers are not related to the locations of cluster member galaxies. To do so we construct a pseudo-random Mg II - galaxy cross-correlation sample. First, for each cluster we take all the absorbers within a maximum aperture ($D_{\mathrm{proj}} \lesssim r_{200}$) and distribute them uniformly in angle, keeping their distance to the BCG fixed.[15] While distributing absorbers randomly we keep the number of absorbers fixed to preserve their number density profile. As above, we then find the nearest member galaxy to each random absorber – the results of this exercise are discussed in Section 3.4.3.2.

[13] $f_{2796} = 0.608$, $\lambda_{2796} = 2796.35$ Å.

[14] $\mathrm{m_{Mg\,II}} = 24.305$ amu.

[15] To distribute absorbers randomly in angle (both polar and azimuthal) on a spherical shell (centred at BCG with coordinates x_o, y_o, z_o) with radii R, where R is the true distance between BCG and the absorber, we use the following recipe: Generate azimuthal and polar angles with uniform distributions, $\phi = \mathrm{U}\,[0, 2\pi]$, $\cos\theta = \mathrm{U}\,[-1, 1]$, respectively. The coordinates in 3D space for new absorber (x_a, y_a, z_a) would then be: $x_a = x_o + \mathrm{R}\cos\phi\sin\theta$, $y_a = y_o + \mathrm{R}\sin\phi\sin\theta$, $z_a = z_o + \mathrm{R}\cos\theta$.

Table 3.3: Summary of galaxy properties of ELGs, LRGs and BCGs.

Objects	$\langle z \rangle$	$\langle M_\star \rangle$ [$\mathrm{M}_\odot$]	$\langle M_{500} \rangle$ [$\mathrm{M}_\odot$]	$\langle R_{500} \rangle$ [kpc]	$\langle R_{200} \rangle$ [kpc]
ELGs	0.84	$10^{10.5}$	$10^{12.2}$	120	200
LRGs	0.54	$10^{11.3}$	$10^{13.5}$	350	550
BCGs	0.55	$10^{11.5}$	$10^{14.2}$	700	1100

3.3.4 Galaxy Samples for a Comparative Study

The measurements from previous studies of Mg II absorption around star-forming and passive galaxies (described in section 3.1) allow us to compare the properties of Mg II absorption in clusters with measurements around field galaxies. To perform such a comparative study, we contrast our Mg II measurements in clusters with cool gas properties around emission-line galaxies (ELGs, Raichoor et al. (2021)) and luminous red galaxies (LRGs, Chen et al. (2012b)) that we measured in Anand et al. (2021). As ELGs, LRGs and BCGs trace a wide range of halo masses from around $10^{12}\,\mathrm{M}_\odot$ to $10^{15}\,\mathrm{M}_\odot$, this comparison will help understand how Mg II absorber properties vary as a function of halo mass. We summarize the basic properties (redshift, halo mass, stellar mass and virial radii) of these objects in Table 3.3 for a detailed comparison.

3.4 Results

3.4.1 Galaxy Cluster - Mg II Correlation

3.4.1.1 Mg II Covering Fractions in Galaxy Clusters

We start by cross-correlating clusters with Mg II absorbers in redshift and projected distance space. First we derive the Mg II covering fraction (f_c), described in Section 3.3.1, as a function of projected distance ($D_{\rm proj}$) from the BCG of the clusters for both *weak* *(*$0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$*)* and *strong (*$EW_{2796} > 1\,\text{Å}$*)* absorbers. We show the covering fractions ($d_1 < D_{\rm proj} < d_2$) for weak absorbers in top left panel of Figure 3.3 and for strong absorbers in top right panel of Figure 3.3. The corresponding dashed lines show the measurements for the random mocks.

We see that on smaller scales ($D_{\rm proj} \lesssim 1$ Mpc), the true measurements are $3-5$ times higher than random expectations, indicating the statistical significance of our measurements. The covering fraction decreases with the projected distance. We see that for the weak absorbers, the f_c varies from $\sim 2-3$ per cent (similar for $|\Delta z| \leq 0.003$, upper left panel of Figure B.1) within $D_{\rm proj} \lesssim 200$ kpc to $\lesssim 1$ per cent ($\lesssim 0.5$ per cent for $|\Delta z| \leq 0.003$) at $D_{\rm proj} \sim 2$ Mpc. For $EW_{2796} > 1\,\text{Å}$ absorbers the Mg II covering fraction varies from $\sim 2-3$ per cent ($\sim 1-2$ per cent for $|\Delta z| \leq 0.003$, upper right panel of Figure B.1) within $D_{\rm proj} \lesssim 200$ kpc to $\lesssim 0.5$ per cent ($\lesssim 0.3$ per cent for $|\Delta z| \leq 0.003$) at $D_{\rm proj} \sim 2$ Mpc. This implies that the covering fraction is also a function of absorber strength, particularly at larger distances. The covering fractions converge to the random expectation at $D \gtrsim 10$ Mpc in each EW bin.

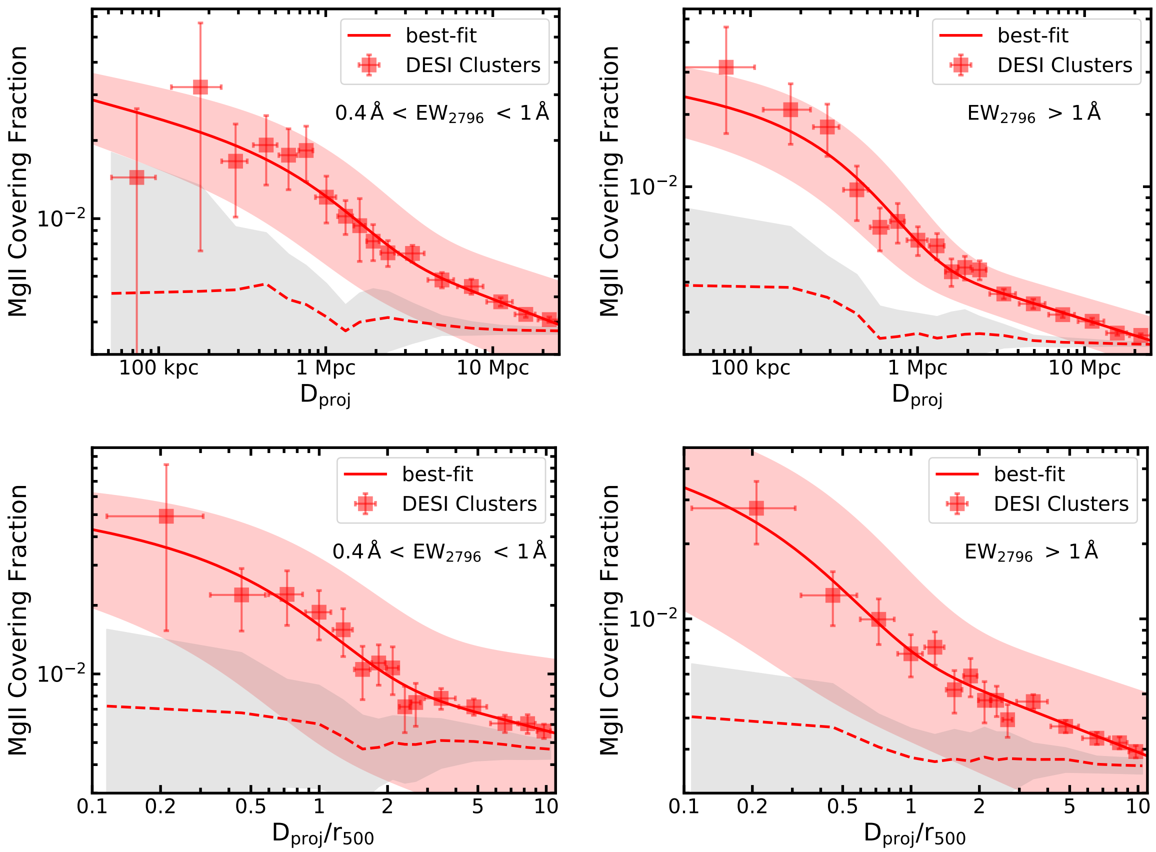

Figure 3.3: **Top:** Differential covering fraction of weak and strong Mg II absorbers around DESI galaxy clusters as a function of projected distance from the BCG. **Bottom:** Mg II covering fractions of weak and strong systems as a function of projected distance normalized by r_{500} of cluster. The corresponding dashed lines show the values measured for the random clusters. We clearly see a decreasing trend with both distance and absorber strength. The solid lines show the best-fitting profiles described in Section 3.4.2, and the shaded regions show the corresponding 1σ uncertainty intervals.

Next, we normalize the projected distance by dividing by the r_{500} of the clusters. Results are shown in the bottom panels of Figure 3.3. The covering fraction of weak absorbers is $\sim 2-5$ per cent ($1-3$ per cent for $|\Delta z| \leq 0.003$, lower left panel of Figure B.1) at $D_{\rm proj} \lesssim 0.5 r_{500}$, while it is $\lesssim 3$ per cent ($\lesssim 2$ per cent for $|\Delta z| \leq 0.003$, lower right panel of Figure B.1) for strong absorbers at similar scales. In the outer parts of cluster halo $D_{\rm proj} \gtrsim r_{500}$, it is about ~ 1 per cent (similar for $|\Delta z| \leq 0.003$) for weak absorbers and $\lesssim 0.5$ per cent ($\lesssim 0.3-0.4$ per cent for $|\Delta z| \leq 0.003$) for strong absorbers, similar to projected distance trends. The values converge to the random expectations at large distances.

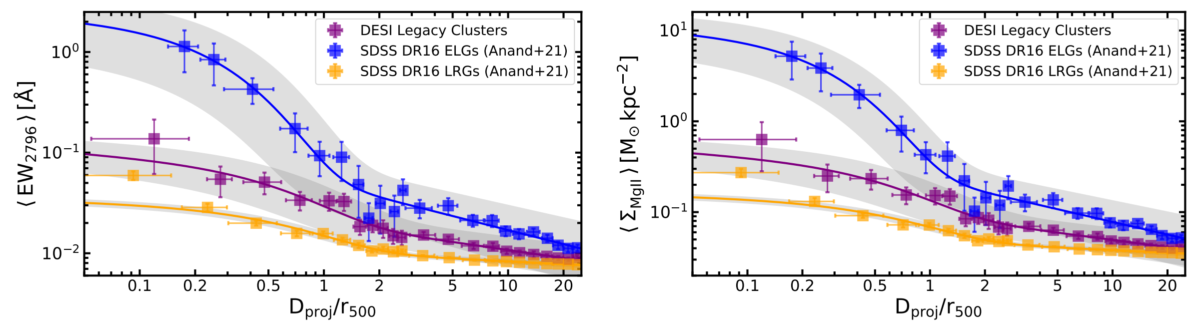

Figure 3.4: **Left:** The mean $\langle\, EW_{2796}\, \rangle$ in and around DESI legacy galaxy clusters (purple), SDSS DR16 ELGs (blue) and SDSS DR16 LRGs (orange) as a function of projected distance normalized by r_{500} of the corresponding haloes. **Right:** Differential average surface mass density of Mg II absorbers in and around DESI legacy galaxy clusters (purple), SDSS DR16 ELGs (blue) and SDSS LRGs (orange) as a function of normalized distances. The solid lines are the best-fitting curves described in section 3.4.2 and the best-fitting parameters are summarized in Table 3.4. The shaded regions show the corresponding 1σ uncertainty intervals. For all three samples: SDSS ELGs, SDSS LRGs and DESI legacy clusters, the values converge at large distances.

In comparison to SDSS LRGs (Lan, 2020; Anand et al., 2021) the covering fractions are slightly larger for DESI legacy clusters, likely due to the bigger halo masses and denser environments. Recently, Nielsen et al. (2018) performed a similar cross-correlation study of Mg II absorption in groups based on a heterogeneous literature sample and Dutta et al. (2021b) did another study based on MUSE and HST observations and found similar higher covering fractions of Mg II absorbers in groups relative to isolated environments.

3.4.1.2 Mean Equivalent Width of Mg II Absorbers

We next estimate the mean equivalent width $\langle\, EW_{2796}\, \rangle$ in galaxy clusters as described in Eqn 3.2. We show the results in the left panel of Figure 3.4, where $\langle\, EW_{2796}\, \rangle$ is shown in purple squares as a function of projected distance ($D_{\rm proj}$) from the BCG, normalized by r_{500}. We also contrast our results with SDSS DR16 LRGs (orange squares) SDSS DR16 ELGs (blue) (Anand et al., 2021) where the projected distance is likewise normalized by r_{500}.

We find that the average Mg II absorption strength is systematically higher in clusters than SDSS LRGs at a given normalized distance from the central galaxy, up to $D_{\rm proj} \lesssim 10\, r_{500}$, however, the mean $\langle\, EW_{2796}\, \rangle$ around ELGs is $5-7$ times higher than both clusters and LRGs. It is not surprising, as previous studies (Lan et al., 2014; Lan & Mo, 2018; Anand et al., 2021) have shown that high SFR of ELGs contributes (via stellar outflows) significantly to the overall Mg II content in the inner part of halo ($\lesssim$ 50 kpc). All three measurements converge at large distances ($D_{\rm proj} \gtrsim 15\, r_{500}$, $\langle\, EW_{2796}\, \rangle < 0.01$ Å). The mean $\langle\, EW_{2796}\, \rangle$ varies from $\approx 0.15 - 0.05$Å (similar values for $|\Delta z| \leq 0.003$ case, left panel of

Figure B.2) within $D_{\rm proj} \lesssim r_{500}$ in DESI legacy clusters while on similar scales it varies from $\approx 0.05 - 0.02$ Å for LRGs and from $\approx 1 - 0.03$ Å for ELGs, respectively. Note that $D_{\rm proj} \approx 10\, r_{500}$ will correspond to $2 - 10$ Mpc on projected scales for clusters, ELGs and LRGs because of their different halo sizes.

3.4.1.3 Mg II Surface Mass Density

We proceed to estimate the average column density, $\langle N_{\rm Mg\,II} \rangle$ and surface mass density, $\langle \sum_{\rm Mg\,II} \rangle$ (see Eqn 3.4) of Mg II absorbers as a function of projected distance ($D_{\rm proj}$) normalized by r_{500} of the clusters. We use the linear curve of growth (Eqn 3.3), as justified by the relatively weak equivalent widths measured, particularly for clusters and SDSS LRGs. However, we point out that our measurements are only a lower limit on the total column densities and surface mass densities (particularly for SDSS ELGs, as most of the Mg II absorbers are saturated ($EW_{2796} \gtrsim 0.2$ Å) around them), as our method is based on the detection of individual absorbers and thus insensitive to very weak Mg II sightlines. A more complete estimate would be possible by measuring both $\langle EW_{2796} \rangle$ and $\langle \sum_{\rm Mg\,II} \rangle$ in stacked spectra in order to account for the contribution from the weakest systems (Zhu et al., 2014; Pérez-Ràfols et al., 2015; Zu, 2021).

We show our results in Figure 3.4 (right panel) as a function of $D_{\rm proj}/r_{500}$. The average surface mass density of Mg II absorbers in galaxy clusters (purple squares) decreases with cluster-centric distance. It varies from $\langle \sum_{\rm Mg\,II} \rangle \approx 0.6\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim 0.2\,r_{500}$ to $\langle \sum_{\rm Mg\,II} \rangle \lesssim 0.2\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim r_{500}$. At large distances, $D_{\rm proj} \gtrsim 10\,r_{500}$ the mean $\langle \sum_{\rm Mg\,II} \rangle \lesssim 0.05\,{\rm M_\odot\,kpc^{-2}}$, ≈ 5 times lower than inner halo. We also note that distribution becomes flat, possibly indicating that we have started tracing the IGM. We also performed the analysis for $|\Delta z| \leq 0.003$ case (right panel of Figure B.2) and we do not find any difference in the trend and measurements.

We contrast the cluster measurements with the surface mass density of Mg II absorbers around SDSS LRGs (orange squares) SDSS ELGs (blue squares) (Anand et al., 2021). For SDSS LRGs, the measurement varies from $\lesssim 0.2 - 0.3\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim 0.2\,r_{500}$ to $\lesssim 0.1\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim r_{500}$, while for SDSS ELGs it varies from $\lesssim 5 - 6\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim 0.2\,r_{500}$ to $\lesssim 0.6\,{\rm M_\odot\,kpc^{-2}}$ at $D_{\rm proj} \lesssim r_{500}$ (note that these values are just the lower limits as described above). This indicates that in the inner part ($D_{\rm proj} \lesssim r_{500}$) of cluster haloes, the surface mass density is $\sim 3-4$ times higher than LRGs, but ~ 10 times lower than ELGs, while at large distances ($D_{\rm proj} \gtrsim 2 - 3\,r_{500}$) the corresponding ratio is $\sim 2 - 3$, indicating a slower decline of Mg II surface mass density in the CGM of LRGs. On the other hand, at such distances the corresponding ratio between ELGs and cluster is $\lesssim 2 - 3$, indicating a much faster decline of Mg II surface mass density in the CGM of ELGs.

In summary, our analysis shows that the DESI clusters have higher (lower) Mg II surface mass densities out to $D_{\rm proj} \lesssim 10\,r_{500}$ relative to SDSS LRGs (ELGs), while all three converge at large distances ($D_{\rm proj} \gtrsim 15\,r_{500}$, $\langle \sum_{\rm Mg\,II} \rangle \lesssim 0.01 - 0.02\,{\rm M_\odot\,kpc^{-2}}$). It is worth pointing out that r_{500} of clusters are $2 - 3$ $(5 - 6)$ times larger than LRGs (ELGs).

Table 3.4: Best-fitting parameters for weak and strong Mg II absorber covering fractions (f_c), mean equivalent widths of Mg II absorbers ($\langle EW_{2796} \rangle$) and Mg II surface mass density ($\langle \sum_{\mathrm{Mg\,II}} \rangle$) around DESI legacy clusters, SDSS DR16 ELGs and SDSS DR16 LRGs (Anand et al., 2021). Note that, we show the best-fitting parameters for both $|\Delta z = z_{\mathrm{BCG}} - z_{\mathrm{Mg\,II}}| \leq 0.01, 0.003$ choices for DESI legacy clusters.

Mg II absorbers	$\|\Delta z\|$	Quantity	x	f_a	x_o	f_b	α
Weak: $0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$	0.01	f_c	D_{proj}	0.0087 ± 0.0017	980 ± 421 kpc	0.0099 ± 0.0049	-0.25 ± 0.04
			$D_{\mathrm{proj}}/r_{500}$	0.01 ± 0.004	0.67 ± 0.34	0.027 ± 0.015	-0.26 ± 0.13
Strong: $EW_{2796} > 1\,\text{Å}$	0.01	f_c	D_{proj}	0.0057 ± 0.0006	395 ± 112 kpc	0.0157 ± 0.0067	-0.22 ± 0.018
			$D_{\mathrm{proj}}/r_{500}$	0.01 ± 0.003	0.293 ± 0.154	0.025 ± 0.019	-0.35 ± 0.05
Weak: $0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$	0.003	f_c	D_{proj}	0.006 ± 0.0026	1165 ± 820 kpc	0.0028 ± 0.023	-0.54 ± 0.048
			$D_{\mathrm{proj}}/r_{500}$	0.007 ± 0.003	1.21 ± 0.78	0.0062 ± 0.004	-0.523 ± 0.08
Strong: $EW_{2796} > 1\,\text{Å}$	0.003	f_c	D_{proj}	0.003 ± 0.0007	765 ± 270 kpc	0.0046 ± 0.002	-0.42 ± 0.037
			$D_{\mathrm{proj}}/r_{500}$	0.007 ± 0.003	0.393 ± 0.257	0.0094 ± 0.009	-0.54 ± 0.04
Mg II absorbers (DESI):	0.01	$\langle EW_{2796} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.023 ± 0.0036	0.55 ± 0.18	0.06 ± 0.026	-0.264 ± 0.03
Mg II absorbers (DESI):	0.003	$\langle EW_{2796} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.023 ± 0.002	0.49 ± 0.12	0.075 ± 0.028	-0.262 ± 0.022
Mg II absorbers (LRGs):	0.01	$\langle EW_{2796} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.010 ± 0.0002	0.69 ± 0.06	0.019 ± 0.002	-0.09 ± 0.005
Mg II absorbers (ELGs):	0.01	$\langle EW_{2796} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.097 ± 0.0287	0.23 ± 0.070	2.12 ± 1.32	-0.46 ± 0.053
Mg II absorbers (DESI):	0.01	$\langle \sum_{\mathrm{Mg\,II}} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.106 ± 0.017	0.55 ± 0.18	0.26 ± 0.12	-0.264 ± 0.03
Mg II absorbers (DESI):	0.003	$\langle \sum_{\mathrm{Mg\,II}} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.107 ± 0.013	0.49 ± 0.13	0.34 ± 0.13	-0.262 ± 0.022
Mg II absorbers (LRGs):	0.01	$\langle \sum_{\mathrm{Mg\,II}} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.048 ± 0.001	0.69 ± 0.06	0.09 ± 0.01	-0.09 ± 0.005
Mg II absorbers (ELGs):	0.01	$\langle \sum_{\mathrm{Mg\,II}} \rangle$	$D_{\mathrm{proj}}/r_{500}$	0.44 ± 0.12	0.23 ± 0.070	9.84 ± 4.71	-0.46 ± 0.053

3.4.2 Characterizing Scale Dependence of Mg II Properties

3.4.2.1 Covering Fractions

As visible by eye, the scale-dependence of the Mg II covering fraction, $\langle EW_{2796} \rangle$ and the surface mass density cannot be described by a simple power-law function. We therefore fit the observed distributions with a superposition of power-law and exponential functions to account for the steep radial dependence of Mg II absorption in the inner regions and the flat behaviour at large distances. Specifically, we fit the following function:

$$f(x) = f_{\rm a} \left(\frac{x}{x_{\rm o}} \right)^{\alpha} + f_{\rm b} \, e^{-(x/x_{\rm o})} \tag{3.5}$$

where $x = D_{\rm proj}/r_{500}$ or $D_{\rm proj}$. The function has four free parameters namely $f_{\rm a}$, $f_{\rm b}$, $x_{\rm o}$ and α. The $x_{\rm o}$ characterizes the scale where the slope (defined by α) changes and both functions are comparable and contribute significantly to the observed distributions. Note that similar functional form (including redshift dependence, with 5-6 parameters) has been used to characterize Mg II and O VI absorber properties around passive and star-forming galaxies from the DESI legacy imaging survey (Lan, 2020) and COS-Halos and COS-Dwarfs surveys (Tchernyshyov et al., 2022).

We fit the stacked profile of each sample with a standard Levenberg-Marquardt minimization, and show the best fitting parameters and their uncertainties (measured with bootstrap approach) in Table 3.4. The best-fitting curves are shown by solid lines in corresponding figures (see Figures 3.3, 3.4).

We observe that the characteristic scale is smaller, $x_{\rm o} \approx 400$ kpc and $x_{\rm o} \approx 0.3$ (normalized) for strong absorbers compared to $x_{\rm o} \approx 1$ Mpc and $x_{\rm o} \approx 0.7$ (normalized) for weak absorbers. Furthermore, we note that the covering fraction declines faster (characterized by slope of power law, α) on normalized scales for strong absorbers ($\alpha \approx -0.35$) than weak systems ($\alpha \approx -0.26$). This implies that strong absorbers are destroyed rapidly in the inner part of the halo, while weak absorbers survive up to large distances.

We also compile the best-fitting parameters of Mg II covering fractions corresponding to $|\Delta z| \leq 0.003$ choice in Table 3.4. There is one trend that is of particular interest: the characteristic scales ($x_{\rm o}$) are slightly larger (albeit large error bars) for both weak and strong absorbers on both projected and normalized scales. On the other hand, slopes (α) are significantly larger (~ 2 times). It implies that covering fraction declines more rapidly than $|\Delta z| \leq 0.01$ for both strong and weak absorbers, and the Mg II scales shift to larger values. It is reassuring that the overall trend and conclusions (as described above) remain the same.

3.4.2.2 Mg II Mass in Clusters and Galaxies

We also fit the mean $\langle EW_{2796} \rangle$ and Mg II surface mass density profiles (solid lines in Figure 3.4) in clusters, SDSS ELGs and SDSS LRGs with the same fitting function (Eqn. 3.5). The best-fitting parameters with their errors are also reported in Table 3.4. We see that radial profiles of both quantities decline more steeply in DESI clusters (slope $\alpha \approx -0.26$, same for $|\Delta z| \leq 0.003$ case, see Table 3.4), than SDSS LRGs ($\alpha \approx -0.1$), while the decline is much faster in SDSS ELGs ($\alpha \approx -0.45$). Mg II absorbers in cluster environments seem to be concentrated in the inner parts of halo ($D_{\rm proj} \lesssim r_{500}$). At large distances from the cluster centre ($D_{\rm proj} \gtrsim 2-3r_{500}$), the probability of detecting absorbers is much lower. The EW_{2796} in the inner haloes of clusters is also higher than in SDSS LRGs, but much lower than SDSS ELGs.

To estimate the total mass of Mg II gas within r_{500} of the cluster halo, SDSS ELGs and SDSS LRGs, we integrate the surface mass density profile up to $D_{\rm proj} \lesssim r_{500}$ in all haloes.

$$M_{\rm Mg\,II}\,(x<1) = 2\pi r_{500}^2 \int_0^1 \left\{ f_{\rm a}\left(\frac{x}{x_{\rm o}}\right)^{\alpha} + f_{\rm b}e^{-x/x_{\rm o}} \right\} x\,dx \tag{3.6}$$

where $x = D_{\rm proj}/r_{500}$ and r_{500} is in kpc. The best-fitting values for $f_{\rm a}$, $f_{\rm b}$, $x_{\rm o}$ and α are compiled in Table 3.4.

Using this approach, we find that the total Mg II mass within r_{500} (see Table 3.3) of the clusters is $\rm M_{Mg\,II,\,cluster}(D_{proj} < r_{500}) \approx (3.26 \pm 1.58) \times 10^5\,M_\odot$ ($(3.43 \pm 1.44) \times 10^5\,\rm M_\odot$, for $|\Delta z| \leq 0.003$ case), while for LRGs within their r_{500} it is $\rm M_{Mg\,II,\,LRG}(D_{proj} < r_{500}) \approx (3.68 \pm 0.34) \times 10^4\,M_\odot$. This implies that $\rm M_{Mg\,II,\,LRG}/M_{Mg\,II,\,cluster} \approx 0.11 \pm 0.05$ such that there is more Mg II gas (≈ 10 times) in cluster haloes. Similarly, the total Mg II mass within r_{500} (see Table 3.3) of ELGs is $\rm M_{Mg\,II,\,ELG}(D_{proj} < r_{500}) \approx (6.1 \pm 4.5) \times 10^4\,M_\odot$, implying that $\rm M_{Mg\,II,\,ELG}/M_{Mg\,II,\,cluster} \approx 0.19 \pm 0.16$ (albeit large error bar), i.e. clusters have ~ 5 times more Mg II gas than ELGs, even though both mean $\langle EW_{2796} \rangle$ and surface mass density are higher in ELGs. However, as noted above, our column densities are only the lower limits for ELGs, given that majority of Mg II absorbers in their halo are saturated. On the other hand, this also implies that $\rm M_{Mg\,II,\,ELG}/M_{Mg\,II,\,LRG} \approx 1.7 \pm 1.2$ such that both have quite comparable (error bar is large) Mg II mass at virial scales. Similar trends were found for H I mass (traced by Mg II absorbers in the stacked spectra of quasars) around ELGs and LRGs (Lan & Mo, 2018). It is clear from above analysis that clusters host more Mg II gas in their haloes compared to LRGs, in part because they reside in larger haloes (see Table 3.3) and they trace overdense regions.

To test the robustness of the measurement, we also estimated the Mg II mass by simply summing the surface mass densities out to $D_{\rm proj} \lesssim r_{500}$ and the corresponding masses are $\rm M_{Mg\,II,\,cluster}(D_{proj} < r_{500}) \approx (3.54 \pm 0.41) \times 10^5\,M_\odot$ ($(3.68 \pm 0.41) \times 10^5\,\rm M_\odot$, for $|\Delta z| \leq 0.003$ case) and $\rm M_{Mg\,II,\,LRG}(D_{proj} < r_{500}) \approx (4.01 \pm 0.11) \times 10^4\,M_\odot$ and $\rm M_{Mg\,II,\,ELG}(D_{proj} < r_{500}) \approx (6.2 \pm 1.1) \times 10^4\,M_\odot$ for clusters, LRGs and ELGs, respectively, consistent with the previous values.

Finally, we estimate the Mg II mass within 200 kpc of the central galaxy: for clusters we find that $\rm M_{Mg\,II,\,cluster}(D_{proj} < 200\,kpc) \approx (4.23 \pm 1.70) \times 10^4\,M_\odot$ and for SDSS LRGs the $\rm M_{Mg\,II,\,LRG}(D_{proj} < 200\,kpc) \approx (1.32 \pm 0.12) \times 10^4\,M_\odot$. This gives $\rm M_{Mg\,II,\,LRG}/M_{Mg\,II,\,cluster} \approx 0.31 \pm 0.12$. Clearly, even on similar scales there is ≈ 3 times more cool gas in clusters than SDSS LRGs. On the other hand, for ELGs, the corresponding mass is $\rm M_{Mg\,II,\,ERG}(D_{proj} < 200\,kpc) \approx (8.3 \pm 1.4) \times 10^4\,M_\odot$, larger than $\rm M_{Mg\,II,\,ELG}(D_{proj} < r_{500})$, because $D_{\rm proj} = 200$ kpc is larger than the r_{500} of ELGs (see Table 3.3), i.e. we are tracing larger area.

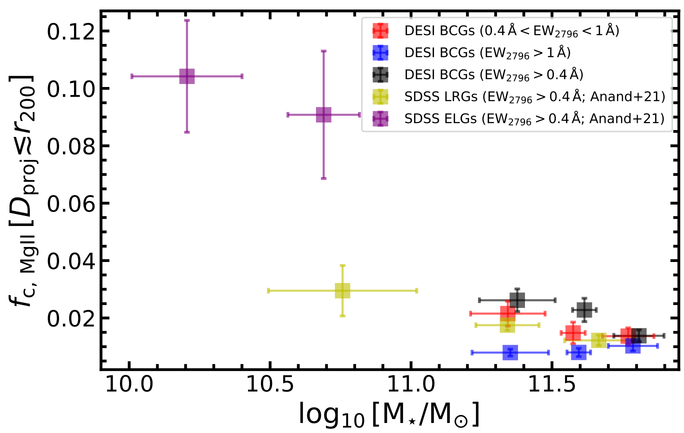

Figure 3.5: Total covering fraction of weak (red squares) and strong (blue squares) absorbers Mg II absorbers within the r_{200} of cluster haloes, as a function of stellar mass of the BCG. We also show the comparison of DESI BCGs (black squares) with SDSS LRGs (yellow squares) and SDSS ELGs (purple square) from Anand et al. (2021). We clearly see that massive haloes have lower Mg II covering faction for both clusters and galaxies (ELGs and LRGs). Note that we have slightly offset horizontally the points for visual clarity.

3.4.3 Dependence on Cluster Properties

3.4.3.1 Stellar Mass of BCGs

We now study the connection between cluster properties and Mg II properties. We begin with total covering fraction within the cluster halo, $f_c\,(D_{\mathrm{proj}} \lesssim r_{200})$, as a function of the stellar mass of the cluster BCG. We show the correlations in Figure 3.5, where we see that weak absorbers (red squares) have higher covering fractions than strong absorbers (blue squares) at a given BCG stellar mass, implying that weak absorbers are more ubiquitous than strong ones.

At lower stellar mass ($\mathrm{M}_\star \approx 10^{11.3}\,\mathrm{M}_\odot$) the covering fraction of weak absorbers (shown in red) within r_{200} is ≈ 2.5 per cent compared to ≈ 1 per cent for strong absorbers (shown in blue). However, for more massive BCGs ($\mathrm{M}_\star \approx 10^{11.8}\,\mathrm{M}_\odot$) the covering fraction for both weak and strong systems are consistent within the error bars. That is, the decreasing trend of Mg II covering fraction as a function of BCG mass is clearly visible for weak systems, but not for strong systems. This suggests that weak and strong systems may have a different physical origin.

We also contrast our cluster measurements ($EW_{2796} > 0.4\,\text{Å}$, black squares) with SDSS LRGs ($EW_{2796} > 0.4\,\text{Å}$, yellow squares) and SDSS ELGs ($EW_{2796} > 0.4\,\text{Å}$, purple squares) from Anand et al. (2021). We combine weak and strong absorbers for a consistent comparison. The boost in covering fraction for BCGs in clusters is approximately the same for the two stellar mass bins. SDSS ELGs trace comparatively smaller haloes and have $3-5$ times higher covering fraction than LRGs and clusters, though BCGs have $5-10$ times larger stellar mass than ELGs.

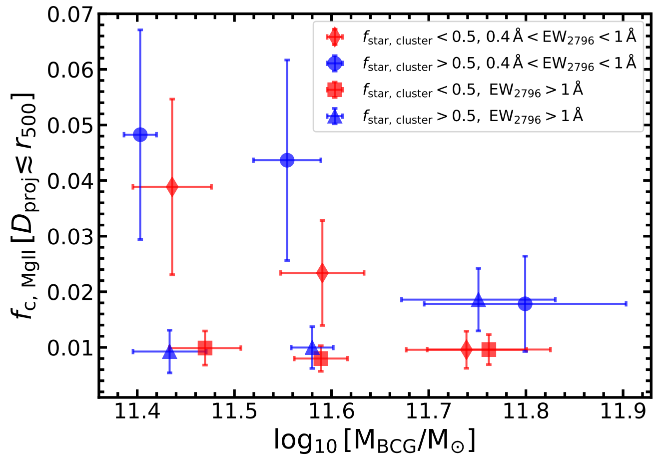

Figure 3.6: Mg II covering fraction of weak (blue circles and red diamond) and strong (blue triangles and red squares) as a function of fraction of star-forming galaxies in the cluster halo at a given stellar mass of the BCG. Blue shows the total covering fraction of Mg II absorbers in clusters that have more star-forming galaxies within their r_{500}, i.e.$f_{\rm star} > 0.5$, while red shows the measurements for clusters having more passive galaxies inside their halo ($f_{\rm star} < 0.5$). For weak absorbers, clusters with more blue galaxies in their haloes have slightly higher covering fraction at a given stellar mass of BCG. Note that we have slightly offset horizontally the points for visual clarity.

3.4.3.2 Mg II absorbers connection with cluster galaxies?

Next, as described in Section 3.3.3, we divide our clusters into two subsamples based on the fraction of star-forming galaxies $f_{\rm star}$ (> 0.5 and < 0.5) within r_{500}. For each subsample, we estimate the total covering fraction within $D_{\rm proj} \lesssim r_{500}$ as a function of the stellar mass of the BCG.

We show the results in Figure 3.6, for clusters with more star-forming galaxies (in blue circles for weak absorbers, blue triangles for strong absorbers) and clusters with more passive galaxies (in red diamonds for weak absorbers and red squares for strong absorbers). The overall anti-correlation of covering fraction with the stellar mass of the BCG is once again visible (Figure 3.5). The covering fraction of weak absorbers is higher in clusters with more blue galaxies (compare blue circles and red diamonds), although the offset is small. This hints that some of the Mg II absorbers in the cluster environment could be connected to member galaxies with ongoing star formation. However, the same trend is not apparent for strong absorbers (compare blue triangles and red squares).

To understand the connection to satellite galaxies better, we locate the nearest member galaxy (in projected distance) to each absorber within $D_{\rm proj} \lesssim r_{200}$ of the cluster halo. If the Mg II absorbers are physically associated with the star-forming ISM or extended gas surrounding the member galaxies, the connection should be visible in the distribution of distances between Mg II - galaxy pairs. In this case, a significant fraction of absorbers would be at small distances from their host galaxies.

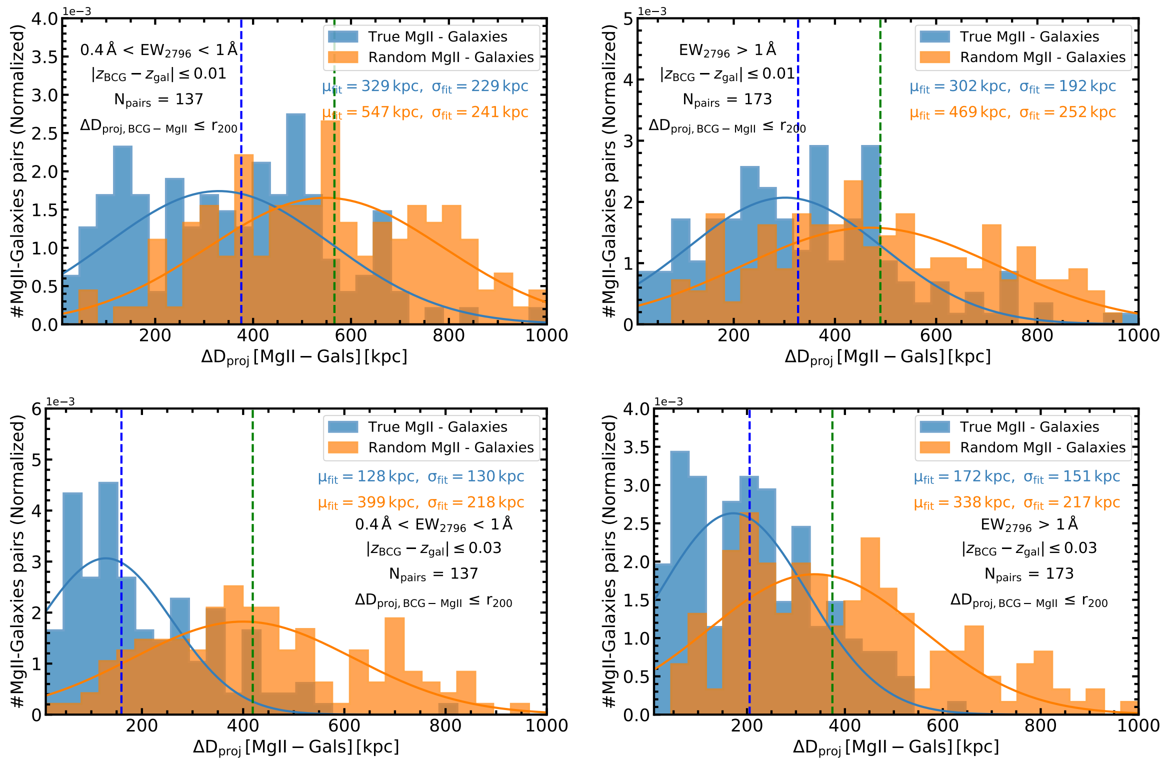

Figure 3.7: 1D distribution of projected separation between Mg II absorbers and its nearest member galaxy in clusters. We only take those absorbers that lie within $D_{\rm proj} \lesssim r_{200}$ of cluster halo. The blue shows the true distribution of absorber-galaxy pair separation and orange shows the distribution when absorbers are distributed randomly in the halo, only in angles, keeping their distances to the BCG fixed. We show the results for two $|\Delta z| = |z_{\rm BCG} - z_{\rm gal}|$ values, namely, a strict cut $|\Delta z| \leq 0.01$ (top row) as well as a more relaxed criterion $|\Delta z| \leq 0.03$ (bottom row). In each row: **Left:** Distribution of projected distance between weak absorbers and their nearest cluster galaxies. **Right:** Same as left panel, for strong absorbers. The corresponding solid blue and green lines represent the best fitting gaussians to the distributions. The dashed vertical lines show the corresponding median values of the distributions.

We show the histogram of absorber-nearest galaxy distance in Figure 3.7 for both weak (blue, left panels) and strong (blue, right panels) absorbers, for two $|\Delta z| = |z_{\rm BCG} - z_{\rm gal}|$ values namely, $|\Delta z| \leq 0.01$ (top row) and a less stringent $|\Delta z| \leq 0.03$ (bottom row). The distributions are roughly gaussian, and the resulting fit for the weak absorbers (shown in blue in the top left panel) is an average of $\mu \approx 330\,\rm kpc$ with $\sigma \approx 230\,\rm kpc$. The median separation is about $\approx$ 350 kpc. Similarly, for strong absorbers, the distribution (shown in blue in the top right panel) can be described by a gaussian with $\mu \approx 300\,\rm kpc$ and $\sigma \approx 200\,\rm kpc$. The median separation is also about $\approx$ 325 kpc.

When we increase the redshift separation ($|z_{\rm BCG} - z_{\rm gal}|$) to $|\Delta z| \leq 0.03$, the mean and median separations further decrease (see bottom row). This is of order, or slightly larger than, the uncertainty on the photo$-z$ estimates. Even for this less stringent $|\Delta z| \leq 0.03$, the separations between absorbers and the nearest galaxies are larger than the expected size of the star-forming disc and more comparable to the typical halo size ($r_{\rm vir} \approx 200$ kpc) of these galaxies (Kravtsov, 2013).

This suggests that the detected Mg II absorbers are not strongly associated with the ISM of the cluster galaxies. We note that faint, low-mass galaxies below the sensitivity of DESI would be present but not currently detected. It is likely that some Mg II absorbers are associated with the CGM of these fainter galaxies. The completeness of the DESI photo$-z$ galaxy sample is a strong function of redshift and stellar mass. For e.g. at redshift, $z \sim 0.6$, galaxies below $\mathrm{M}_\star \sim 10^{10.5}\mathrm{M}_\odot$ have low completeness in the catalogue (Zou et al., 2019).

To understand the significance of a possible connection between absorbers and satellites we turn to our random catalogues, as described in Section 3.3.3. We recall that the randoms place absorbers with the same radial distribution as real absorbers, but redistributed randomly in angle.

The corresponding distributions for randomly distributed absorbers are shown in Figure 3.7 in orange, in all panels. A single gaussian function can also describe the random distributions, and in all cases the mean and widths of these distributions are larger than for the true catalogs. In the case of $|\Delta z| \leq 0.01$ (top panel of Figure 3.7, orange), for weak absorbers, the best-fitting values, $\mu \approx 550$ kpc, $\sigma \approx 240$ kpc, are ≈ 1.5 higher than true distribution ($\mu \approx 330$ kpc,). The result is qualitatively similar for the strong absorber case, as well as for larger redshift intervals – in all cases a statistically significant difference is present between the random mocks and the true data.

To summarize, we find that in cluster environments the Mg II absorbers are preferentially found near cluster galaxies, as opposed to being randomly distributed throughout the halo. Galaxies and absorbers are clearly clustered together to some degree. Because the typical separations are larger than the sizes of the actual galaxy discs, we conclude that the ISM of cluster galaxies (above the detected mass threshold) is not directly responsible for the observed Mg II absorption in clusters. In contrast, the typical separations suggest that Mg II absorption arises at least in part from cool gas that has been removed from the ISM due to e.g. ram-pressure stripping. Our findings are consistent with a picture where the gas is no longer within the disc, but may be still gravitationally bound to, or recently stripped from, the dark matter subhaloes hosting member galaxies.

3.4.3.3 Mg II Equivalent Width vs Cluster Galaxy Properties

We now study the dependence of Mg II absorption (within r_{200} of the cluster) on properties of the nearest galaxy, such as its SFR and stellar mass. In Figure 3.8 we show EW_{2796} as a function of the projected distance between the absorber and the nearest cluster galaxy. We do not see a clear correlation between Mg II absorption and distance between the absorber-galaxy pair. The equivalent widths have significant scatter at fixed distance, and the median line (red) is largely consistent with being flat. This is true regardless of

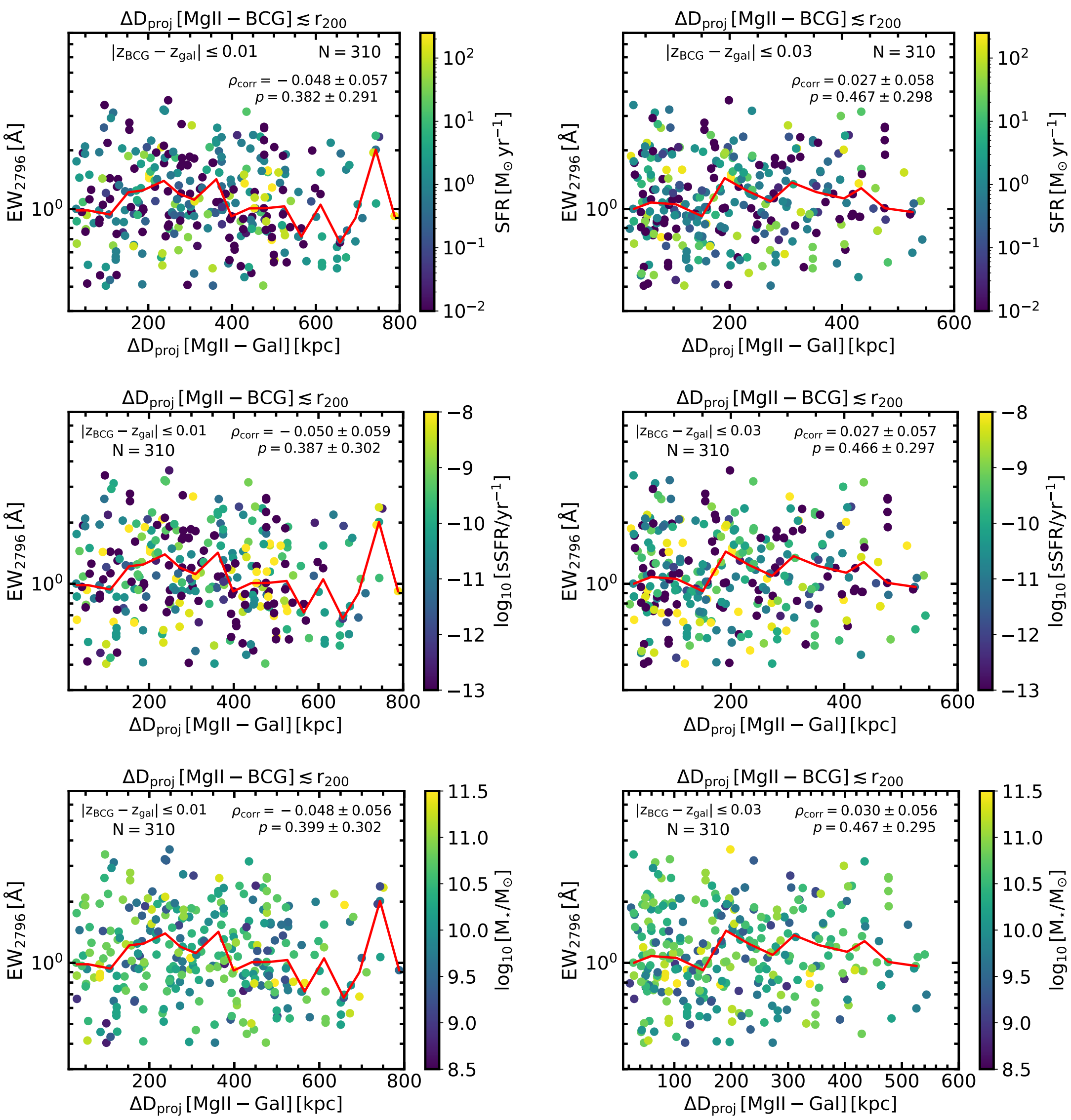

Figure 3.8: EW_{2796} of Mg II absorbers (within r_{200} of the cluster) as a function of impact parameter ($D_{\rm proj}$ from the nearest cluster member). We show the results for a strict $|\Delta z| = |z_{\rm BCG} - z_{\rm gal}|$ value, namely, $|\Delta z| \leq 0.01$ (left panels), as well as for a more relaxed $|\Delta z| \leq 0.03$ (right panels). In each row: **Top:** Distribution is colored by the SFR of the nearest galaxy. **Middle:** Distribution colored by the specific SFR (sSFR) of the nearest galaxy. **Bottom:** Distribution colored by the stellar mass of the nearest galaxy to the absorbers. Note that here, we have taken all the absorbers ($EW_{2796} > 0.4$ Å). No significant correlation is visible between Mg II absorption and properties of its nearest galaxy. We also show the Pearson correlation coefficients and the corresponding p-values in each panel. The low correlation coefficient and high $p-$ values indicate no significant correlation between quantities. The solid red line in each panel shows the median values. Note that, we do not show the error bars on EW_{2796} for visual clarity.

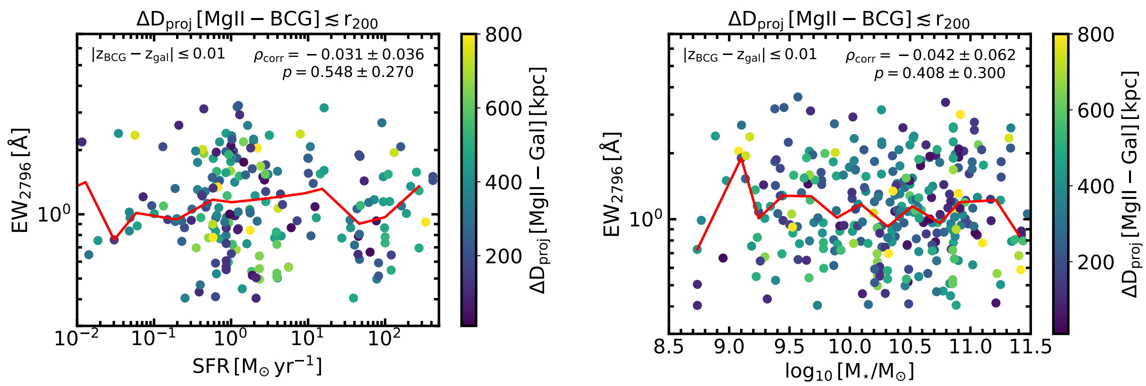

Figure 3.9: EW_{2796} of Mg II absorbers (within r_{200} of the cluster) as a function of SFR (left) and stellar mass (right) of cluster galaxies. We show the results for $|\Delta z| = |z_{\rm BCG} - z_{\rm gal}| \leq 0.01$. Points are colored by the distance between absorber and the nearest cluster galaxy. Note that here, we have taken all the absorbers ($EW_{2796} > 0.4$ Å). No significant correlation is visible between Mg II absorption and properties of its nearest galaxy – Pearson correlation coefficients and the corresponding p-values are given in each panel. The solid red line in each panel shows the median values.

whether the delta redshift criterion for association is strict (left panels) or more relaxed (right panels).

Previous studies of field galaxies (Nielsen et al., 2013; Chen, 2017; Lan & Mo, 2018; Anand et al., 2021) found that Mg II equivalent width decreases with the impact parameter, though the scatter is large (see Dutta et al., 2021b). We do not see such an anti-correlation here, in fact the Pearson correlation coefficient is consistent with zero ($\rho_{\rm corr} = -0.048 \pm 0.057$) and the p-value is large ($p = 0.4$), clearly indicating that there is no significant correlation in EW_{2796} and distance between absorber-galaxy pairs in clusters. Even after increasing the Δz cuts, we do not see any significant correlation (e.g. for a three times larger $|\Delta z| \leq 0.03$). This suggests that the connection discussed previously is, at best, weak. We also looked at $EW_{2796} < 0.4$ Å(not shown here) absorbers, but the statistics are poor, with which no clear additional trends were seen and nothing robust can be said. We discuss more about the limitations and caveats to this analysis in section 3.5.2.

We further color the points in each panel according to a given property of the nearest galaxy: star formation rate (top panels), specific star formation rate (middle panels), and stellar mass (bottom panels). We do not observe any strong dependence of EW_{2796} on these galactic properties (top: SFR, middle: sSFR and bottom: $M_\star$), even for the closest pairs ($D_{\rm proj} \lesssim 200$ kpc).

To investigate further, we show EW_{2796} as a function of several properties of the nearest galaxy in Figure 3.9, for $|\Delta z| \leq 0.01$, as the results are similar for other choices. We show EW_{2796} as a function of SFR (left panel) and stellar mass (right panel) of galaxies. We do not observe any strong correlations, even for closest pairs in (see the colour bar, showing distance between absorber-galaxy pairs). The distribution is consistent with random and,

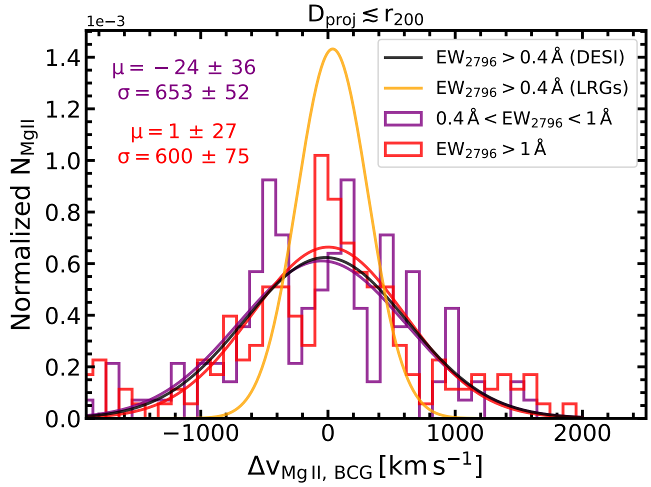

Figure 3.10: Distribution of LOS velocity difference between cluster BCG and Mg II absorbers pairs within $D_{\mathrm{proj}} < r_{200}$ of the clusters. The purple and red curves show distributions for weak and strong absorbers, respectively. The black curve shows the distribution for all absorbers (weak and strong). We also show the Mg II LOS velocity around LRGs in orange from Anand et al. (2021), which is much narrower. The distributions are characterized by single gaussian profiles shown in solid color lines. Best fit parameters for the weak and strong cases are shown in each panel (see text).

the median line is roughly flat, disfavoring any strong connection between Mg II absorption and the SFR or stellar mass of the nearest cluster galaxies.

Previous studies of field galaxies (Bordoloi et al., 2011; Rubin et al., 2014; Lan & Mo, 2018; Anand et al., 2021) have shown that Mg II equivalent widths and covering fractions correlate positively with the SFR of the host galaxies, particularly for emission line or star-forming galaxies. This correlation is attributed to the strong outflows due to the stellar activity of the galaxy like stellar winds or supernovae driven winds (Bordoloi et al., 2011; Rubin et al., 2014; Tumlinson et al., 2017). In the case of quiescent or passive luminous red galaxies, Mg II properties are found to correlate with stellar and halo masses (Nielsen et al., 2015; Anand et al., 2021). As a result, it appears that Mg II absorbers in clusters are different to Mg II absorbers in the field, in that there is little connection between EW_{2796} and the SFR or mass of the nearest galaxy. If the Mg II originated from satellite ISM gas, then it is likely that there has been enough time for the properties of the Mg II clouds to have evolved so that the patterns observed in field galaxies are washed out.

3.4.4 Relative Kinematics of Mg II absorbers in galaxy clusters

We now turn to the motion of Mg II absorbing gas. Using the BCGs which have robust spectroscopic redshifts, we estimate the line-of-sight (LOS) velocity difference between Mg II absorbers and cluster BCG as, $\Delta v = c\Delta z/(1 + z)$, where z is the redshift of the cluster

BCG. The distribution of LOS velocity separations allows us to constrain the motion of absorbing clouds in cluster environments.

Figure 3.10 shows results for both weak (blue) and strong (red) absorbers. We see that for weak absorbers, the distribution can be characterized by a single gaussian with mean $\langle \Delta v \rangle \approx -24(\pm 36)$ $\mathrm{km\,s^{-1}}$and dispersion, $\sigma_v \approx 650$ $\mathrm{km\,s^{-1}}$(shown in the solid purple line). There is weak evidence for an additional peak around -500 $\mathrm{km\,s^{-1}}$, which may signpost an inflow scenario where cool gas is flowing toward the centre of the cluster. Similarly, for strong absorbers the best fit gaussian parameters are mean $\langle \Delta v \rangle \approx 1(\pm 26)$ $\mathrm{km\,s^{-1}}$and dispersion, $\sigma_v \approx 600$ $\mathrm{km\,s^{-1}}$(shown in the solid red line). The mean velocity difference (consistent with zero) is in agreement with results from Mg II -galaxy cross-correlation studies (Huang et al., 2016; Chen, 2017; Huang et al., 2021). We also show the gaussian fit for all absorbers ($EW_{2796} > 0.4\,$Å, black line) and it is also consistent with the other two distributions.

The typical dark matter halo velocity dispersion within r_{200} (Δv_{200}) for the clusters in our sample is expected to be ≈ 900 $\mathrm{km\,s^{-1}}$. This implies $\sigma_{\mathrm{weak,\,strong\,Mg\,II}} \approx 0.6 - 0.7\sigma_{\mathrm{cluster,\,200}}$, i.e. suppressed by $30 - 40\%$ relative to the velocity dispersion of the cluster dark matter halo. The motion of both weak and strong absorbers in cluster halo is sub-virial. This is very different from the motion of cool CGM around star-forming galaxies (ELGs), where absorbing clouds move with velocities as high as the velocity dispersion of their dark matter halo (Nielsen et al., 2015; Lan & Mo, 2018; Anand et al., 2021).

Results for SDSS LRGs are shown for comparison as an orange curve in the figure – the Gaussian width is ~ 200 $\mathrm{km\,s^{-1}}$, i.e. a factor of 3 smaller than for the clusters. Clearly the motion of Mg II absorbers in higher mass halos is significantly larger than in lower mass halos, which is likely tracing the larger gravitationally induced motions in these halos.

3.5 Discussion

3.5.1 Comparison with Previous Studies

We now compare our results with previous studies that performed similar analysis of cool gas in galaxies and clusters. We contrast our mean $\langle\, EW_{2796}\,\rangle$ measurements with results from Zhu et al. (2014) and Lan & Mo (2018), where authors estimate the median EW_{2796} by stacking thousands of quasar spectra from SDSS DR11/DR14 in the rest-frame of LRGs from the same data release. The comparison is shown in Figure 3.11 (left panel), where orange squares show the values for SDSS LRGs using our approach (see section 3.3.2), green shows the results from Lan & Mo (2018) and magenta represents the measurements from Zhu et al. (2014).

Our measurements are in good agreement with previous studies at small distances ($D_{\mathrm{proj}} \lesssim 1\,\mathrm{Mpc}$), even though the methods and samples are significantly different. We see that at large distances ($D_{\mathrm{proj}} \gtrsim 5\,\mathrm{Mpc}$) our EW_{2796} is larger than previous measurements, and this discrepancy is expected given the different samples and methodology. In previous studies, the EW_{2796} is the mean equivalent width of the Mg II absorbers in the stacked quasar spectra. When we go far from galaxies and start tracing the IGM, the absorption is dominated by weak absorbers. In contrast, we measure the mean equivalent width of absorbers (Section 3.3.2) based on individual detections, which is limited by the sensitivity of the detection pipeline. Furthermore, the detection limit is a strong function of absorber strength, redshift and S/N of the spectra (see completeness in Anand et al., 2021) so it is difficult to detect weak systems in the individual spectra. By definition, our method additionally estimates the average equivalent width of absorbers above a threshold, and weak systems are much more ubiquitous than strong systems as we see from a completeness analysis of our (Anand et al., 2021) and JHU-SDSS (Zhu & Ménard, 2013a) Mg II catalogues.

We also compare our mean Mg II surface mass density (Section 3.3.2) in and around SDSS LRGs (orange) with the measurements from Zhu et al. (2014) (magenta) analysis and this comparison is shown in Figure 3.11 (right panel). As above, the agreement is excellent at small distances, but begins to diverge at large scales due to the methodological differences.

Our measurements exhibit a change of slope at $D_{\mathrm{proj}} \approx 1\,\mathrm{Mpc}$ (see also Section 3.4.2 for a detailed discussion on Mg II scale), which is consistent with previous studies (Zhu et al., 2014; Pérez-Ràfols et al., 2015; Zu, 2021). As discussed in Zhu et al. (2014) and Zu (2021), this slope change is indicative of a transition from scales where the measurements are dominated by gas within a single dark matter halo to a regime where the gas properties are determined by contributions from gas in multiple haloes – the 2-halo term.

In Section 3.4.1.2, we found that Mg II equivalent widths and surface mass densities are higher in clusters than SDSS LRGs, at any given impact parameter ($D_{\mathrm{proj}}/r_{500}$). Similar enhancements in Mg II covering fractions (Nielsen et al., 2018; Dutta et al., 2020) in group environments have been noted. On the other hand, Bordoloi et al. (2011) found extended radial distribution of Mg II equivalent widths in group galaxies in a stacking analysis with zCOSMOS galaxy survey. In this case, it has been previously assumed that the sum of EW_{2796} in isolated galaxies can account for total EW_{2796} in groups. We note that our results on the independence of Mg II equivalent widths on galaxy properties show that this cannot be true in detail.

3.5.2 Possible Origin of Mg II absorbers in clusters

As summarized in Section 3.4.1, our analysis points towards a scenario where Mg II absorption in clusters is higher and more extended than for less massive galaxies. The exact processes that give rise to the observed Mg II absorption in clusters remain unclear, because of the complex nature of galaxy dynamics and gas properties in clusters. A number of different mechanisms have been proposed in the literature to explain the observed properties of Mg II absorbers in overdense regions: (i) the contribution of multiple individual galaxy haloes within the clusters or group (Padilla et al., 2009; Bordoloi et al., 2011; Nielsen et al., 2018; Dutta et al., 2020); (ii) strong galactic outflows from a single massive galaxy within the cluster; (iii) stripping of cold gas by hot ICM as galaxies move through the ambient hot medium or by tidal interactions between galaxies (Fumagalli et al., 2014; Pearson et al.,

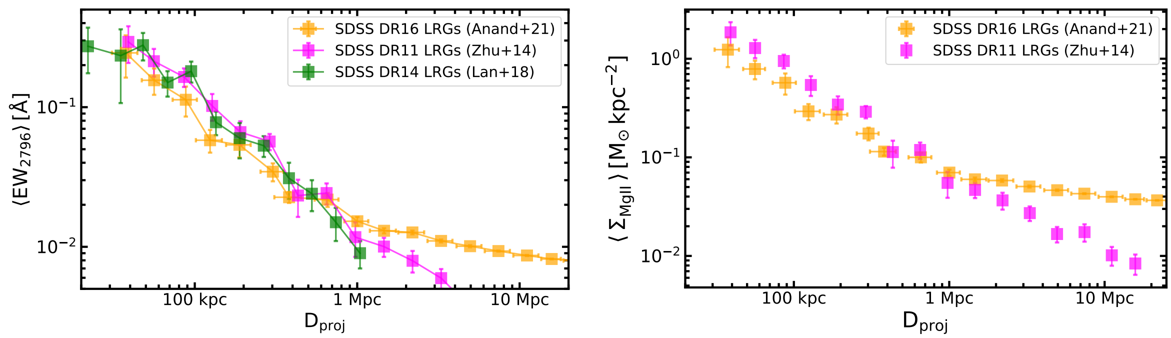

Figure 3.11: Mean $\langle\, EW_{2796} \,\rangle$ (left) and Mg II surface density (right) as a function of projected distance from the central galaxy. In the first case, we contrast the values around SDSS LRGs from Anand et al. (2021) (estimated using Eqn 3.2, orange) with measurements from Zhu et al. (2014) (magenta) Lan & Mo (2018) (green) where authors measured the mean absorption by stacking SDSS quasar spectra in the rest-frame of SDSS LRGs. For surface densities, we contrast our measurements around SDSS LRGs from Anand et al. (2021) (estimated using Eqn 3.4, orange) with Zhu et al. (2014) (magenta) measurements which was measured by stacking SDSS DR11 quasar spectra in the rest-frame of SDSS LRGs. In both cases the agreement is good at small distances, but methodological differences emerge at large scales (see text).

2016; Johnson et al., 2018; Fossati et al., 2019); and (iv) in-situ formed cool gas clumps floating in the intragroup medium or ICM (Sharma et al., 2012; Voit & Donahue, 2015; Qiu et al., 2020). A combination of satellite interactions/stripping and in-situ thermal instability driven cooling has also been explored (Nelson et al., 2020; Dutta et al., 2021a).

In our analysis, we find that the median separation between absorbers and the closest cluster galaxy is roughly a few hundred kpc on average. This places a large fraction of absorbers outside the expected radius of the subhaloes that host these systems. The possibility that absorbers may be associated with low-mass galaxies remains open. Tchernyshyov et al. (2022) recently reported an association of O VI absorbers with low mass star-forming galaxies in the CGM^2 survey. However, Dutta et al. (2020) did a comparative study of EW_{2796} measurements in groups with predictions from simple superposition models and concluded that such models could not explain the enhanced Mg II absorption in groups. We note that in Section 3.4.2.2 we find that within r_{500}, the cool gas mass in clusters is $\sim$ 10 times higher than SDSS LRGs, while the Mg II mass is $\sim$ 3 times larger in clusters than LRGs within $D_{\rm proj} \lesssim 200\,$kpc from the central galaxy. This indicates that the Mg II enhancement in clusters is scale dependent, which is not consistent with a simple superposition model.

The observation that Mg II absorbers are spatially correlated with cluster member galaxies (see Figure 3.7) compared to the random samples (though the average spatial separation is larger than the size of galaxy discs) points toward the scenario where these absorbers are more likely to be associated with the CGM of cluster galaxies than their ISM. Furthermore, the fact that there is no visible trend (characterized by small Pearson coefficients) between Mg II absorption and stellar activity (SFR) of the nearest galaxy (see Figures 3.8, 3.9) leads us to speculate that current stellar activity could not be the dominant mechanism for the origin of Mg II absorbers in clusters. Increasing the redshift separation between absorbers and cluster galaxies does not change this conclusion (see bottom panels of Figures 3.8).

Our analysis has a number of notable limitations. First, the large errors in the photo$-z$ of the galaxies can introduce large uncertainties in cluster member and Mg II absorbers correlation. Some of the pairs may be spurious and far in redshift space. Hence, the current discrepancy can also arise because previous studies have used spectroscopic samples that allow us to connect galaxies and absorbers more robustly in velocity space. Given these limitations, a direct comparison between previous results and current photo $-z$ galaxies may not be accurate. Second, our analysis is based on a particular EW_{2796} threshold ($EW_{2796} > 0.4$ Å, to retain high completeness) and does not take into account the non-detections or weak systems. It is possible that including such weak systems can show a significant anti-correlation between EW_{2796} and the impact parameter. We included the absorbers with $EW_{2796} < 0.4$ Å (not shown here), but the statistics are poor, and we can not test this hypothesis. Nonetheless, we do not observe any correlation between EW_{2796} and galactic properties for $EW_{2796} > 0.4$ Å absorbers. Hence, the presence of these absorbers perhaps can not be fully explained by the ISM or stellar activity of cluster galaxies. Moreover, while comparing EW_{2796} or Mg II surface mass density with previous studies, it is important to note that previous studies are based on stacking methods, which also account for non-detections and weak systems. Given our EW_{2796} threshold, we do not account for such non-detections in the current study. Hence, we conclude that Mg II absorbers ($EW_{2796} > 0.4$ Å) in clusters is likely regulated by the interaction between galaxies and the ICM or intragroup medium besides any contribution from stellar outflows of individual galaxies. However, the caveats and limitations should be kept in mind while interpreting the results.

Several studies have shown that galaxies are impacted by environmental effects before they enter the inner regions ($D_{\rm proj} \lesssim r_{500}$) of clusters (Cen et al., 2014; Marasco et al., 2016; Jung et al., 2018; Ayromlou et al., 2021). One important mechanism that can remove the gas from the CGM of an infalling galaxy is ram-pressure stripping, where the hot ambient ICM or intragroup medium strips cold gas from the outskirts of the galactic halo. Tidal disruption of gas from a small satellite galaxy by a massive halo can also play role in removing the gas in dense environments. Both observations and simulations have revealed stripping of gas in groups and clusters (Mihos et al., 2012; Marasco et al., 2016; O'Sullivan et al., 2018; For et al., 2019; Yun et al., 2019). It is plausible that some fraction of the stripped gas may extend to large distances from the host galaxy. In the recent cosmological magnetohydrodynamical simulation TNG50 (Pillepich et al., 2019; Nelson et al., 2019) it was found that cold gas can form due to local density fluctuations driven by ram-pressure stripped gas in massive LRG-like haloes (Nelson et al., 2020).

As shown in Section 3.4.4, the motion of absorbing clouds in the cluster is suppressed by $30 - 40$ per cent compared to the expected dark matter velocity dispersion of the cluster halo. If the absorbing clouds are gravitationally bound to satellite galaxies, the absorbers would follow the dark matter kinematics (More et al., 2011). As pointed in several studies (Huang et al., 2016; Lan & Mo, 2018; Anand et al., 2021), even though the stripped/accreted gas has initial velocities comparable to cluster velocity dispersion, the drag force exerted by the hot intragroup medium or ICM can decelerate the gas motion. Therefore gas clumps falling towards cluster centre will slow down, as slow-moving clouds will survive longer due to inverse relation between evaporation time scale and cloud velocity (Zahedy et al., 2019).

Our finding that there is little correlation between absorbers and the properties of the nearest cluster galaxy also supports the gas-stripping scenario. In clusters, the observed abundance and properties of cool, low-ionization gas is undoubtedly a result of the complex interplay between satellite stripping, cloud destruction, cloud formation, and galactic outflows.

3.6 Summary

In this work, we combine ∼160,000 Mg II absorbers from SDSS DR16 Mg II catalogue with $72,000$ galaxy clusters from the DESI legacy imaging surveys to characterize the nature and origin of the cool gas in galaxy clusters. Our main findings are:

- There is a significant covering fraction of both weak ($0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$) and strong ($EW_{2796} > 1\,\text{Å}$) Mg II absorbers in clusters ($D_{\text{proj}} \lesssim r_{500}$). Compared to the random expectation the measured values are $\approx 2{-}5$ times higher within $D_{\text{proj}} \lesssim 2\,r_{500}$ (slightly depending on our $|\Delta z|$ choice). This implies that the cool gas is readily detected in galaxy clusters, despite the hot ICM.

- Characterizing the scale dependence (radial profile) of Mg II absorption, we find that the Mg II covering fraction declines faster for strong absorbers than weak systems, indicating that strong absorbers are preferentially found at small distances.

- The mean EW_{2796} decreases with the cluster-centric distance and is higher in clusters compared to SDSS LRGs but lower than SDSS ELGs.

- Clusters contain ~ 10 times more total Mg II gas mass than SDSS LRGs within r_{500}. Within $D_{\text{proj}} \lesssim 200$ kpc from the central galaxy, there is ~ 3 times more Mg II mass in clusters than SDSS LRGs. Comparison to SDSS ELGs is complicated by the fact that the majority of associated Mg II absorbers are saturated, such that only lower limits on column densities are possible.

- The integrated covering fraction of Mg II within r_{500} of clusters anti-correlates with the stellar mass of the brightest cluster galaxy. This trend is more prominent for weak absorbers.

We then investigate the correlation between Mg II absorbers and cluster member galaxies, including their physical properties: stellar mass, SFR, and sSFR. Our main results and conclusions are:

- Clusters with more star-forming galaxies in their haloes have slightly higher Mg II covering fractions, indicating some (weak) level of connection between stellar activity and Mg II absorption.
- However, there is no significant correlation between Mg II equivalent width (absorption strength) and the stellar activity or stellar mass of the nearest (detected) cluster neighbour.
- Mg II absorbers are not randomly distributed within the cluster volume: they are located preferentially closer to satellite galaxies. However, the majority of Mg II absorbers reside at distances greater than 100 kpc from the nearest cluster neighbour.
- These typical separations suggest that Mg II absorption arises at least in part from cool gas that has been removed from the ISM due to e.g. ram-pressure stripping.

Our findings are consistent with the picture whereby the cool Mg II traced gas in clusters is not found within the interstellar medium of satellite galaxies, but may be still gravitationally bound to, or recently stripped from, the dark matter subhaloes hosting member galaxies.

Surrounded by darkness yet enfolded in light.

Alan Brennert, Molokai

Chapter 4

Multiphase circumgalactic gas in quasars at $z \sim 2$

The contents of this chapter is based on Anand et al., in preparation

4.1 Introduction

The two most studied metal species are Mg II and C IV , which trace different gas phases in the CGM or the IGM. The Mg II $\lambda\lambda\,2796, 2803$ resonant fine-structure doublet traces low-ionized cool gas ($T \sim 10^{4-4.5}$ K), in the CGM or intracluster medium (ICM) (Zhu & Ménard, 2013a; Huang et al., 2021; Anand et al., 2021, 2022) while the C IV $\lambda1548$, 1550 doublet traces warm ionized phase ($T \sim 10^{5-5.5}$ K)[1] of the CGM (Prochaska et al., 2014; Landoni et al., 2016; Schroetter et al., 2021).

Thanks to the sensitive and powerful telescopes, our understanding of the cosmic enrichment cycle has deepened by constraining the physical properties of C IV absorbers at these epochs. As discussed above, carbon is formed inside the stars, so by constraining the C IV absorbers in and around galaxies or IGM, we can learn about the amount of carbon that all previous generation stars have produced over time (Cooksey et al., 2013; Hasan et al., 2022). On the other hand, the amount of C IV in the galactic halo can also shed light on the nature and availability of total ionizing radiation needed to maintain the C IV systems at those epochs Prochaska et al. (2014). It has been found that the doublet ratio (EW_{1548}/EW_{1550}) of C IV absorbers decreases with time, indicating an overall increase in mean C IV density over time (Péroux et al., 2004; Songaila, 2005). In addition to the photoionization, the studies have also explored the role of collisional ionization in shocks in C IV production (e.g., Charlton et al., 2003; Oppenheimer & Davé, 2006; Oppenheimer et al., 2016). The cosmological simulations (e.g., EAGLE, Oppenheimer et al. (2016)) have shown that the fraction of collisonally ionized metals increase at lower redshifts as

[1]The ionization potential of C III and C IV , are ≈ 50 eV, and ≈ 64 eV, respectively.

the haloes become more massive ($M_{\rm h} \sim 10^{12}\,M_{\odot}$) and have higher virial temperatures ($T \gtrsim 10^5\,$K).

C IV absorption in quasar spectra also provides us with an opportunity to study the metal content of high$-z$ quasars[2] (Péroux et al., 2004; Becker et al., 2011; Prochaska et al., 2014; Landoni et al., 2016; Finlator et al., 2020). The cross-correlation studies of C IV absorbers with $z \sim 2$ quasars have revealed a large reservoir of metal gas in their CGM, where the covering factor is as high as $60 - 80\,\%$ within the $r_{\rm vir}$ of quasar halo (Prochaska et al., 2014; Landoni et al., 2016). Based on quasar - C IV two-point cross-correlation analysis and bias estimation, it has been predicted that C IV absorbers are predominantly tracing the massive $M_{\rm halo} = 10^{12-12.5}\,{\rm M}_{\odot}$ haloes at $z \sim 2$ (Vikas et al., 2013; Prochaska et al., 2014; Gontcho A Gontcho et al., 2018).

In this chapter we perform a very preliminary analysis of statistical properties of C IV absorbers in the universe between $1.5 < z < 2.5$. We will also study the correlation between ($T \sim 10^{4-4.5}\,$K , traced by Mg II absorbers) and warm ($T \sim 10^{5-5.5}\,$K , traced by C IV absorbers) metal gas in the universe at $z \sim 2$ when the cosmic star formation density was at its peak. Moreover, our goal is to cross-correlate C IV and Mg II absorbers to quasars at $1.5 < z < 2.3$ and perform an extensive metal absorption study in the halo of quasars. This study will shed light on the nature of cool and warm metal gas in the CGM of quasars at $z \sim 2$. Our study is one of the most extensive C IV and Mg II cross-correlation studies with quasars. The analysis will be an essential addition to understanding gaseous haloes of galaxies, clusters, and quasars over a broad range of redshifts and galactic properties.

We organize this chapter into five sections. Section 4.2 presents the samples of metal absorbers (C IV and Mg II) and the method to detect them in quasar spectra. We analyse the statistical results of individual absorbers in Section 4.3. In section 4.4, we then cross-correlate our metal absorbers with quasars and study metal gas properties and its kinematics in their haloes. Finally, we summarize our results in Section 4.5. We assume a Planck-consistent cosmology (Planck Collaboration et al., 2020) in this chapter, i.e. $\Omega_{\rm m,0} = 0.307$, $H_0 = 67.7\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{\Lambda,0} = 1 - \Omega_{\rm m,0}$.

4.2 Data

4.2.1 Mg II selected C IV absorber catalogue

To search for C IV doublets ($\lambda\lambda$1548, 1550) in the SDSS DR16 quasars, we make use of our Mg II absorber catalogue. In this work, our goal is to compile a Mg II - selected C IV absorber catalogue. However, this is a very first step, and we are also interested in performing a blind search in future for C IV absorbers in the SDSS DR16 quasar spectra using the detection pipeline that we developed in Anand et al. (2021). As C IV is a shorter wavelength than Mg II absorbers, there is a lower redshift limit, below which we can not detect C IV absorbers in the SDSS spectra due to the wavelength coverage. For our purpose, we select only those Mg II absorbers that have $z_{\rm Mg\,II} > 1.5$ (corresponds to $\sim 3870\,$Å for

[2] On the other hand, we are limited in high$-z$ ($z > 1.5$) galaxy samples.

Table 4.1: Sample statistics for searching Mg II selected C IV absorbers. References are: [a] Anand et al. (2021); [$\star$] **Current study** ($1.5 < z < 2.3$); [b] C-13: Cooksey et al. (2013); [c] H-20: Hasan et al. (2020)

Objects	N
SDSS DR16 Mg II Absorbers	159,524[a]
Mg II Absorbers ($z > 1.5$)	40,016[a]
C IV Absorbers (detected)	21,466[$\star$]
C IV Absorbers (0.1 Å $<$ EW_{1548} $<$ 2.5 Å)	21,271[$\star$]
C-13 C IV Absorbers	16,710[b]
C-13 C IV (0.1 Å $<$ EW_{1548} $<$ 2.5 Å, $1.5 < z < 2.3$)	11,732[c]
H-20 C IV Absorbers	1268[c]
H-20 C IV (0.1 Å $<$ EW_{1548} $<$ 2.5 Å, $1.5 < z < 2.3$)	400[c]

C IV 1548 line) such that our pipeline can detect them in the spectra. We compile these selections and corresponding sample sizes in Table 4.1. In addition, we also mask a few regions in the neighbourhood of Ca II $\lambda\lambda$ 3934, 3969 to avoid any confusion. We also mask the regions near the observed wavelengths corresponding to O I λ1302 and Si II λ1304 lines of the parent quasar to avoid any confusion. After having compiled the final set of spectra where Mg II -selected C IV absorber detection can be applied, we implement the following detection algorithm to locate C IV systems for a given Mg II absorber.

The method is as follows: For each Mg II absorber in our Mg II catalogue (where we can detect C IV absorbers), we locate its corresponding C IV absorber wavelength and fit a double gaussian to the C IV doublet. If the fitting process succeeds, i.e. parameters are finite and within the bounds, and the line separation of C IV doublet is within 1.2 Å of the true value ($\Delta\lambda_{\rm rest} = 2.58$ Å), we proceed to the next step. We measure the S/N of the absorbers and doublet properties such as line centres and the line widths of the corresponding C IV absorbers. Additionally, using observed line centres, we also estimate the true redshift of the C IV absorbers ($z_{\rm C\,IV}$) and their uncertainties. One example spectrum with Mg II and C IV absorbers is shown in Figure 4.1. We present the sample statistics of the Mg II selected C IV absorbers in Table 4.1. In the following sections, we discuss the properties of our detected C IV systems and their comparison with corresponding Mg II absorbers to shed some light on their correlation.

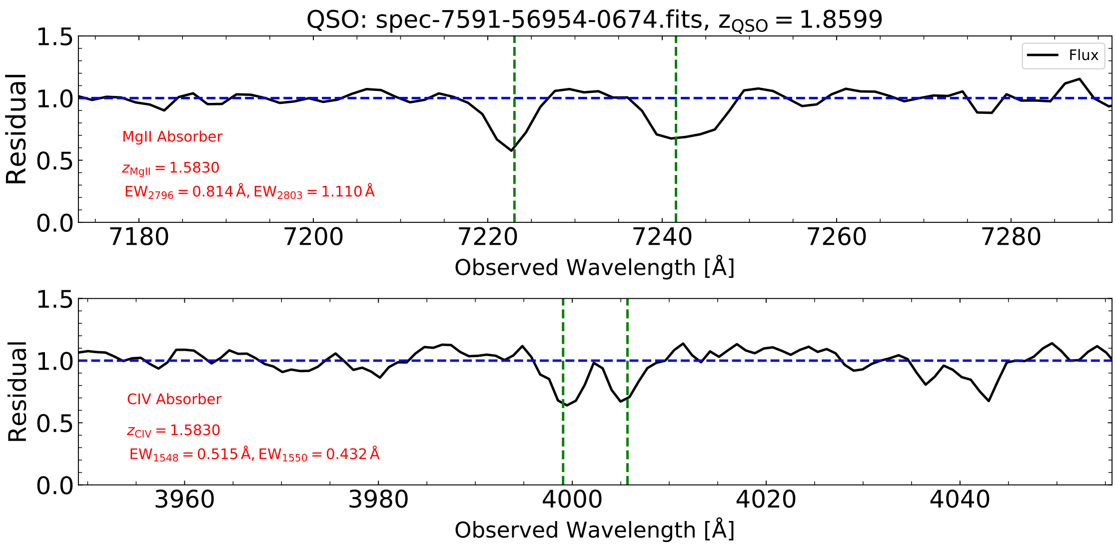

Figure 4.1: Quasar residual spectrum with Mg II and C IV absorbers. Residual is defined as the ratio of observed flux and continuum. **Top:** Mg II absorber and its properties. **Bottom:** C IV absorber and its properties. Both absorbers are in the same spectrum, and the observed wavelength is in Å. The redshifted line centres are shown in green dashed lines.

4.3 Results

4.3.1 C IV Absorber Properties

4.3.1.1 Redshifts of C IV absorbers

We show the redshift distribution in the left panel of Figure 4.2. We see that number of detected C IV systems declines with redshift. It is not surprising as the number of Mg II absorbers is also lower at higher redshifts (see also Figure 3 of Anand et al. (2021)). The median redshift of C IV absorbers in our sample is $z \sim 1.75$, and the typical error on the redshift is $\sigma_z \sim 15 - 20\ \mathrm{km\,s^{-1}}$.

4.3.1.2 Rest Equivalent Widths of C IV Absorbers

After having compiled our Mg II selected C IV absorbers, we estimate the rest equivalent widths and uncertainties on them. For this purpose, we follow the approach as described in Anand et al. (2021). In summary, we apply a Monte Carlo approach where we fit each C IV doublet with a double gaussian by adding the product of spectral noise of the residual and a standard normal distribution. We repeat the fitting process for 200 times and finally take the mean and standard deviation of 200 runs to estimate the rest equivalent widths and their errors. The EW_{1548} for our C IV systems varies between $0.1\,\text{Å} < EW_{1548} < 2.5\,\text{Å}$. The mean EW_{1548} is $\sim 0.9\,\text{Å}$ and the typical error is $\sim 0.15\,\text{Å}$. We show the observed

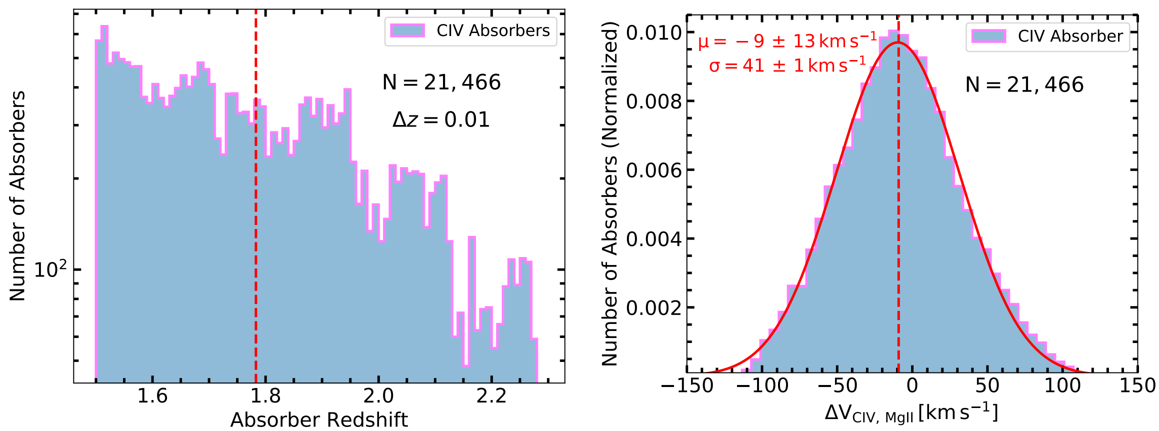

Figure 4.2: **Left:** Redshift distribution of Mg II - selected C IV absorbers. **Right:** 1D distribution (normalized) of the redshift difference (in velocity space) between C IV and Mg II absorbers. In both panels, the dashed vertical lines show the median values. The solid red curve in the right panel shows the best-fitting Gaussian to the data.

distribution (in purple) of EW_{1548} in Figure 4.5 and discuss more about it in section 4.3.4. on the other hand, the velocity dispersion($\sigma_{\rm abs}$, the line width) varies from $30-250$ km s^{-1} for our sample, much higher ($10-20$ km s^{-1}) than the C IV systems detected in high-resolution spectra (Hasan et al., 2020, 2022).

4.3.2 C IV versus Mg II Absorber Properties

4.3.2.1 Redshift Difference

C IV absorber may be either moving away from or towards the Mg II absorbers, which would randomly shift the line-centres of C IV systems. Therefore, we expect the mean difference between the two redshifts to be close to zero, with some dispersion. We show the 1D histogram of velocity difference[3] corresponding to their redshifts in Figure 4.2. The solid red line shows the best-fit single Gaussian profile that fully describes the histogram. The corresponding mean (μ) and dispersion (σ) are $-9\,(\pm 13)$ and $41\,(\pm 1)$ km s^{-1}, respectively. The mean difference is consistent with zero though there is a significant dispersion.

We now demonstrate that one way to interpret the dispersion physically is in terms of the virial velocities of C IV and Mg II clouds, assuming they are in virial equilibrium (may not be true, see Bouché et al. (2006)) and can be modelled by truncated isothermal spheres (Mo et al., 2010). Now if we assume that the virial temperature is close to the temperature that is traced by these systems, i.e. $T_{\rm C\,IV} \sim 10^5$ K and $T_{\rm Mg\,II} \sim 10^4$ K (Tumlinson et al., 2017), we can calculate the virial velocities[4] of these systems.

[3] $\Delta v = \frac{z_{\rm C\,IV} - z_{\rm Mg\,II}}{1+z_{\rm Mg\,II}} \cdot c$, where $c = 3\times10^5$ km s^{-1}, is the speed of light.

[4] $T_{\rm vir} \simeq 3.6\times10^5\,{\rm K}\left(\frac{\rm V_{vir}}{100\,{\rm km\,s^{-1}}}\right)^2$ (Mo et al., 2010).

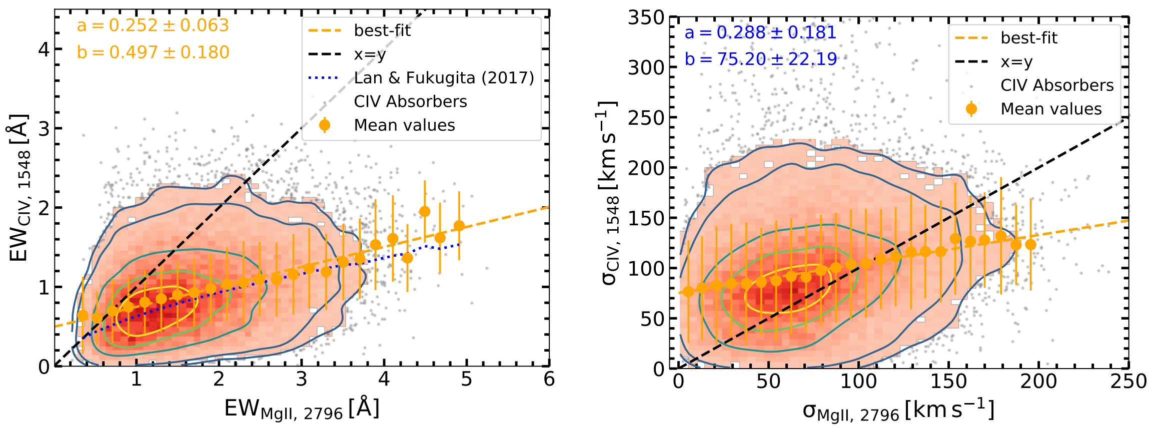

Figure 4.3: **Left:** Rest EW_{1548} vs EW_{2796} for our Mg II - selected C IV catalogue. **Right:** Line widths of C IV $\lambda1548$ ($\sigma_{\mathrm{C\,IV},1548}$) as a function of Mg II $\lambda2796$ ($\sigma_{\mathrm{Mg\,II},2796}$). In both panels, the orange circles denote the mean values in each EW_{2796} (or $\sigma_{\mathrm{Mg\,II},2796}$) bin and the error bars show 1σ standard deviation. The orange dashed line shows the best-fitting (parameters shown in upper left) line that describes the mean correlation. In left panel we also contrast our mean values with empirical relation (shown in dotted blue) derived in Lan & Fukugita (2017). They agree very well. Furthermore, in both panels, dashed black lines show the $y = x$ relation. The contours enclose 25, 50, 75, 95 and 97.5 percentiles of the sample.

For C IV absorbers with $T_{\mathrm{vir,C\,IV}} \sim 10^5\,\mathrm{K}$, we get $V_{\mathrm{vir,C\,IV}} \sim 52\,\mathrm{km\,s^{-1}}$, while for Mg II absorbers with $T_{\mathrm{vir,Mg\,II}} \sim 10^4\,\mathrm{K}$, we get $V_{\mathrm{vir,Mg\,II}} \sim 17\,\mathrm{km\,s^{-1}}$. Interestingly the difference between these two virial velocities is $\sim 36\ \mathrm{km\,s^{-1}}$, which is close to the velocity dispersion ($\sim 40\ \mathrm{km\,s^{-1}}$) that we observed Figure 4.2 (right panel). This is consistent with that both C IV and Mg II systems are possibly in virial equilibrium[5] or close to it at a given epoch, though the temperature of C IV and Mg II clouds may have a radial profile rather than one single temperature.

4.3.2.2 Equivalent Widths

As described in section 4.3.1.2, we estimate the rest equivalent width of C IV absorbers using gaussian fits to the doublet. We now compare them with the Mg II equivalent widths (Anand et al., 2021), to understand how these quantities are correlated. We show the comparison in left panel of Figure 4.3. We see that EW_{1548} is smaller than EW_{2796} (see dashed black line) and this is expected as C IV is a weaker transition than Mg II [6], in terms of oscillator strength. We also find that both quantities are positively correlated[7] and

[5]Though Bouché et al. (2006) concluded that Mg II absorbers in galactic haloes (particularly LRGs) may not be in virial equilibrium.

[6]$f_{1548,\,\mathrm{osc}} = 0.2$, $f_{2796,\,\mathrm{osc}} = 0.6$

[7]Spearman correlation is 0.34, with p value close to zero.

EW_{1548} increases with EW_{2796} . To understand this correlation better in average sense, we also show the mean values (in orange) of both quantities in each EW_{2796} bin, along with their $\pm 1\sigma$ standard deviation. The positive correlation is clearly visible now. We fit this mean correlation with a line[8] and the best-fitting line is shown in dashed orange line. The best-fitting parameters ($a = 0.25 \pm 0.06$, $b = 0.5 \pm 0.18$) are also shown in the panel. To compare with literature we also contrast these mean values with empirical relation[9] derived in Lan & Fukugita (2017) using stacking approach. The relation is shown in dotted blue line and they agree very well, even though our approaches and samples differ significantly. This further demonstrates that our equivalent width measurements are robust. We also show 25, 50, 75, 95 and 97.5 percentiles of the sample.

4.3.2.3 Line Width of Absorbers

In the right panel of Figure 4.3, we now compare the velocity dispersion of C IV λ1548 line ($\sigma_{\mathrm{C\,IV},1548}$) with Mg II λ2796 ($\sigma_{\mathrm{Mg\,II},2796}$). We calculate the velocity dispersion from the σ_{fit} parameter in Gaussian fitting. We also correct these for SDSS velocity resolution (σ_{SDSS}) with Gaussian aperture method[10].

We find that the velocity dispersion of C IV and Mg II systems are positively correlated (though the correlation seems weak). Here also, we fit the mean values with a line[11] (shown in orange) and the best-fitting parameters ($a = 0.29 \pm 0.18$, $b = 75.2 \pm 22.2$) are shown in blue. We also see that C IV systems have larger velocity dispersion than Mg II absorbers for $\sigma_{\mathrm{Mg\,II}} < 100$ km s^{-1}, while the trend changes for strong Mg II systems. Here also, we show 25, 50, 75, 95 and 97.5 percentiles of the sample.

4.3.3 Completeness Correction

To quantify the ability of our algorithm to detect C IV absorbers of different strengths in the spectra, we also estimate the completeness of our method as a function EW_{1548} and z. For this purpose, we implement a Monte Carlo simulation scheme. We generate fake C IV doublet with different EW_{1548} and doublet ratios and insert them in the quasar residual. Then, we also take care of the masks and redshift boundaries as described in section 4.2.1 to avoid any bias in our completeness estimation.

To select absorbers with different strengths, we draw EW_{1548} , line-widths and doublet ratios (DR) from a normal distribution with mean and standard deviations equal to the observed distributions of these quantities. After inserting the fake absorbers at some location in a randomly selected quasar from our parent quasar sample, we assume that

[8] $y = a\,x + b$.

[9] The empirical scaling relation in Lan & Fukugita (2017) is: $EW_\lambda = C\,(EW_{2796})^\alpha\,(1+z)^\beta$. For the best-fitting values of C, α and β, see Table 1 of Lan & Fukugita (2017). In case of C IV λ1548, the best-fit values are: $C = 1.524$, $\alpha = 0.58$, $\beta = -0.87$. For z in the equation we use the mean redshift in each EW_{2796} bin.

[10] $\sigma_{\mathrm{corr}} = \sqrt{\sigma_{\mathrm{fit}}^2 - \sigma_{\mathrm{SDSS}}^2}$, where $\sigma_{\mathrm{SDSS}} = 69\,\mathrm{km\,s^{-1}}$

[11] See footnote 8.

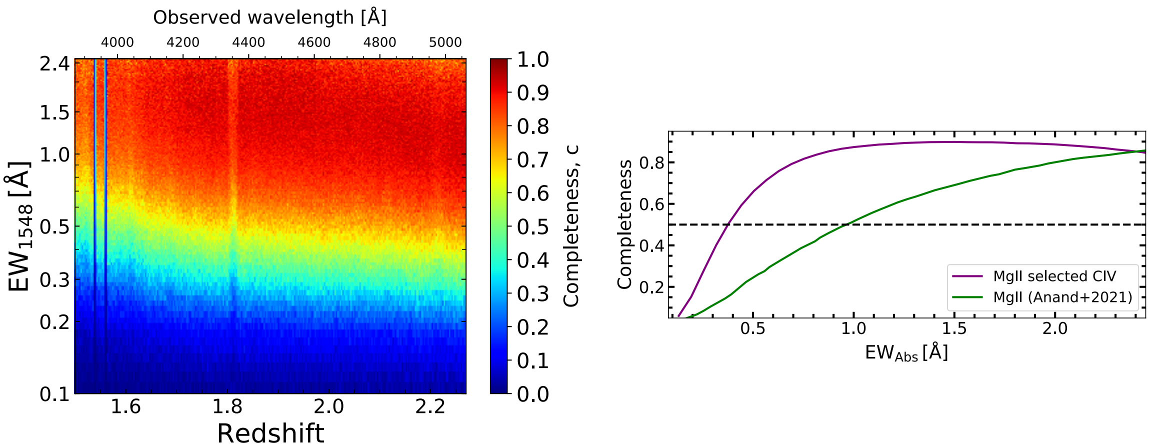

Figure 4.4: **Left:** 2D completeness as a function of EW and z of C IV absorbers. The low completeness stripes are visible at $\lambda 3934$ and $\lambda 3969$, corresponding to Ca II lines. **Right:** Completeness as a function of equivalent width smoothed over all z of detected absorbers for both C IV (purple) and Mg II (green) systems (Anand et al., 2021).

redshift is known. Since we are working with Mg II selected C IV absorbers, we already know the corresponding location of the C IV doublet in the residual. Now we run our C IV detection module (described in section 4.2.1) on fake absorbers. If the fitting succeeds (i.e. parameters are within the bounds) and line separation of C IV system is within ± 0.6 Å of the true separation, we estimate the $\mathrm{EW}_{1548,\mathrm{out}}$. If the fractional error, i.e. $|\mathrm{EW}_{1548,\,\mathrm{in}} - \mathrm{EW}_{1548,\,\mathrm{out}}|/\mathrm{EW}_{1548,\,\mathrm{in}} < 1$, we record it as detected system. In total we simulated over ~ 30 million fake C IV systems in over $\sim 20,000$ quasars.

Finally, for a fake absorber in a narrow EW_{1548} and redshift bin, we estimate completeness, $c(EW_{1548},\, z)$ as the fraction of *detected* C IV absorbers to the total *injected* systems. Using our completeness estimates we also estimate the corrected 'effective' number of asborbers, $N_{\mathrm{eff}} = 1/c(EW_{1548},\, z)$, the number had our C IV module been perfect. By definition $c \leq 1$, that implies $N_{\mathrm{eff}} \geq 1$. We show the 2D distribution of our completeness function, $c(EW_{1548},\, z)$ in $EW_{1548} - z$ plane in the left panel of Figure 4.4. We also show the completeness as a function of equivalent width of C IV absorbers averaged over redshifts (in purple) in the right panel of Figure 4.4. For contrast, we also show the completeness of our Mg II absorber catalogue (Anand et al., 2021) (in green), which clearly shows that completeness of C IV absorbers is higher than Mg II absorbers (at EW $= 1$ Å, $c_{\mathrm{Mg\,II}} = 0.5$ and $c_{\mathrm{C\,IV}} = 0.9$), as expected, because we already know their locations in the spectra. Furthermore, it is also evident that completeness is higher for strong absorbers, implying that it is easier to detect strong absorbers in the spectra compared to their weaker counterparts.

4.3.4 Rest Equivalent Width Distribution

After having corrected the C IV absorbers for completeness of our detection algorithm, we show both corrected (red square) and observed (purple square) equivalent distribution

(EW_{1548}) in Figure 4.5. To estimate the distribution, we bin the EW_{1548} measurements and count the number of absorbers (weighted by their completeness) in each bin. The errors on EW_{1548} show the $\pm 1\sigma$ standard deviation in each bin. To estimate error on counts, we take the $\sqrt{N}$ as the errors, assuming Poisson statistics; therefore, they are small for smaller EW_{1548} , where counts are large.

We see that the distribution is more or less complete (compare purple and red squares) above $EW_{1548} > 1$ Å, while it is incomplete at $EW_{1548} < 0.3$ Å. This implies that the turnover at $EW_{1548} \sim 0.8$ Å (see purple square) is artificial and due to detection incompleteness. After correcting the absorbers for their completeness (red squares), we do not observe such a turnover, and the number of absorbers is continuously increasing for small EW_{1548} . This implies that weak absorbers are much more ubiquitous than strong absorbers in the universe, as also pointed out by several studies (Cooksey et al., 2013; Burchett et al., 2015; Hasan et al., 2020) in the literature.

Next, we want to characterize EW_{1548} distribution with simple functions. In the past, studies have used different functions to characterize the observed and completeness corrected EW_{1548} distributions[12]. Studies have shown that surveys sensitive to low equivalent widths use power-law distributions while less-sensitive large surveys describe the distribution by an exponential function. We also fit our observed and completeness corrected EW_{1548} distributions with three different functions: (a) Schechter function (Kacprzak & Churchill, 2011; Hasan et al., 2020) (b) Power-law (Songaila, 2005; Boksenberg & Sargent, 2015) (c) Exponential function (Cooksey et al., 2010, 2013). The details and best-fitting parameters are compiled in Table 4.2. We discuss the best-fit parameters in section 4.3.5. Below are the details of these three functions.

4.3.4.1 Schechter Function

The Schechter function (Schechter, 1976) is a combination of power-law and exponential[13] and defined as:

$$f(W_{\rm r}) = N_{\rm o} \left(\frac{\rm W_r}{\rm W_o}\right)^{\alpha} \exp\Big(-\frac{\rm W_r}{\rm W_o}\Big) \tag{4.1}$$

where $W_{\rm r}$ is rest equivalent width of C IV $\lambda 1548$ line, α characterizes the slope of power-law, $W_{\rm o}$ is the characteristics EW_{1548} , where the behaviour changes from power-law to exponential decline and $N_{\rm o}$ defines the normalization.

4.3.4.2 Broken power-law

As visible by eyes, in Figure 4.5, a single power-law can not describe the behaviour of observed and completeness corrected EW_{1548} . A slope change is visible at EW_{1548} =

[12]Similar functions have been used to characterize the differential rest equivalent width distribution of Mg II absorbers detected in SDSS spectra (Nestor et al., 2005; Zhu & Ménard, 2013a).
[13]In statistics, Gamma distribution has the similar functional form.

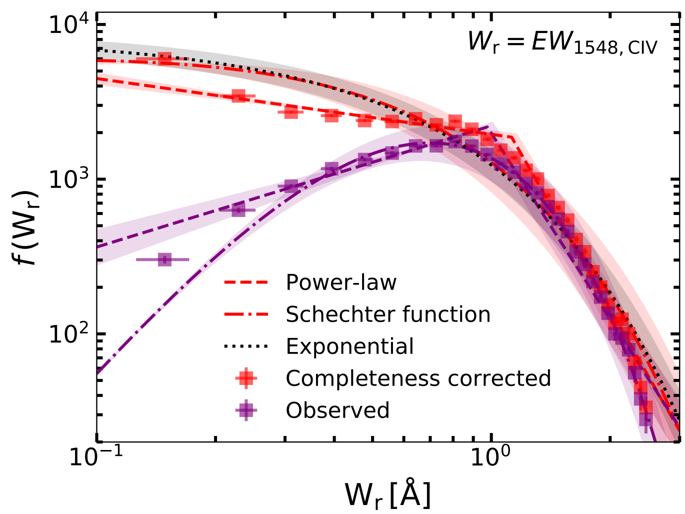

Figure 4.5: Observed (purple) and completeness corrected (red) EW_{1548} distribution of our Mg II - selected C IV catalogue. We also show the best-fitting broken power law (dashed lines), Schechter function (dotted dashed lines) and exponential function (dotted line) to the distribution as described in section 4.3.4. The best-fitting parameters and their corresponding errors are compiled in Table 4.2. The shaded region shows the $\pm 1\sigma$ spread of the best-fitting curves.

$0.8 - 0.9$ Å hence we use a broken power-law to characterize the complete behaviour of EW_{1548} distribution for our sample. We define a broken power-law as:

$$f(W_r) = \begin{cases} N_o \left(\frac{\mathrm{W_r}}{\mathrm{W_o}}\right)^{\alpha} & \text{if } \mathrm{W_r} < \mathrm{W_o} \\ N_o \left(\frac{\mathrm{W_r}}{\mathrm{W_o}}\right)^{\beta} & \text{if } \mathrm{W_r} \geq \mathrm{W_o} \end{cases} \tag{4.2}$$

where α and β are the slopes of two power-laws and EW_o characterizes the scale where the transition occurs. N_o is the normalization constant, such that the function is continuous at the characteristic scale.

4.3.4.3 Exponential Function

The exponential function is defined as:

$$f(W_r) = N_o \exp\left(\alpha \mathrm{W_r}\right) \tag{4.3}$$

where α characterizes the slope of exponential function and N_o defines the normalization.

We compile the best-fitting parameters in Table 4.2. To fit the distributions with the above-mentioned functions and obtain the best-fitting parameters and their errors, we make use of the least-square orthogonal distance regression (ODR) method[14].

[14]See https://docs.scipy.org/doc/scipy/reference/odr.html

Table 4.2: Best-fitting parameters for C IV equivalent width distribution shown in Figure 4.5.

Function	Quantity (EW_{1548}) *Obs:* Observed *Comp:* Completeness corrected	N_o [10^3 Å^{-1}]	W_o [Å]	α	β
Schechter:	Obs (dotted-dashed purple line)	(1.07 ± 0.13)	0.22 ± 0.005	3.1 ± 0.11	–
	Comp (dotted-dashed red line)	(9.2 ± 1.3)	0.48 ± 0.043	0.15 ± 0.11	–
Power law:	Obs (dashed purple line)	(2.18 ± 0.15)	0.98 ± 0.03	0.78 ± 0.14	-3.98 ± 0.21
	Comp (dashed red line)	(1.87 ± 0.08)	1.12 ± 0.03	-0.36 ± 0.06	-4.46 ± 0.25
Exponential:	Comp (dotted black)	(8.25 ± 0.99)	–	-1.9 ± 0.1	–

4.3.5 Best-fitting Parameters

In Figure 4.5, we see that for strong systems ($EW_{1548} > 0.8$ Å), all three functions fit equally well (within error bars) to both observed and completeness corrected EW_{1548} distributions. However, for weak systems ($EW_{1548} < 0.7$ Å), the broken power-law (dashed red line) fits the completeness corrected EW_{1548} distribution better than both Schechter (dotted dashed red line) and exponential (dotted black line) functions. The characteristics EW_{1548} where the slope changes in the power-law is $EW_{1548} = 1.1$ Å, also visible by eye. The slope describing the stronger EW_{1548} systems is much steeper ($\beta = -4.5$) than the weaker counterparts ($\alpha = -0.36$). It is also worth noting that both slopes are very tightly constrained (error bars are small) for our sample. Similarly, for observed distribution (dashed purple line) the two slopes are $\alpha = 0.78$ and $\beta = -3.98$ and the characteristics $EW_{1548} = 1$ Å, similar to the previous case. We also see that slopes describing the stronger systems (β) are approximately same (within 1σ) for observed and completeness corrected distributions, implying that the sample is complete for strong absorbers as expected.

On the other hand, the Schechter function does not fit well to the completeness corrected distribution (dashed red line) for intermediate systems ($EW_{1548} < 0.7$ Å) and over predicts the number, while it fits very well to the strong absorbers. The characteristics $EW_{1548} = 0.5$ Å is smaller than power-law case, while the slope $\alpha = 0.15$ (albeit large error bars). As the slope in the Schechter function characterizes the weaker tail of the distribution, a small slope implies that exponential decline mostly dominates the behaviour, which is the case for strong systems where function fits fairly well. In the case of observed distribution at the weaker end ($EW_{1548} < 0.3$ Å), the power-law (dashed purple line) slightly over predicts the numbers while the Schechter function (dotted dashed purple line) underestimates the distribution. The characteristics EW_{1548} for power-law is 1.1 Å, larger than 0.22 Å for Schechter function. Similar behaviour is seen for the slope α.

Finally, we show the best-fitting exponential function to the completeness corrected EW_{1548} (dotted black line). The slope is $\alpha = -1.9$, and the behaviour is very similar to the Schechter function. As discussed above, it is expected because the slope of the power-law component in the Schechter function is small (consistent with zero), and the exponential part dominates the distribution.

Our analysis shows that the frequency distribution for strong absorbers ($EW_{1548} > 1\,\text{Å}$) can be described by all three functions equally well. In contrast, the power-law slopes are very different, implying a dichotomy in the absorber population. The strong absorbers are rare, while the weak absorbers are more ubiquitous. In literature, this dichotomy has been discussed in the context of the physical processes responsible for the origin of these absorbers. Studies have shown that a large fraction (~ 40 per cent) of the strong C IV absorbers ($EW_{1548} > 0.5\,\text{Å}$) detected in background quasars have large velocities ($-1000\,\mathrm{km\,s^{-1}} < \Delta \mathrm{v} < 10000\,\mathrm{km\,s^{-1}}$) (Nestor et al., 2008). They are associated with the galactic outflows powered by stellar activity or the central AGN (Nestor et al., 2008; Bowler et al., 2014; Stone & Richards, 2019). On the other hand, the weak absorbers are mostly originating in the intervening material (e.g., IGM, low metallicity medium) (Nestor et al., 2008). Similar conclusions have been drawn for Mg II absorbers (see also Bouché et al., 2006; Tremonti et al., 2007).

4.4 Connecting C IV and Mg II with Quasars

4.4.1 Quasars

Quasars are the most luminous and powerful active galactic nuclei (AGN) in the universe (luminosity, $L > 10^{44}\,\mathrm{erg\,s^{-1}}$); therefore, they can be observed up to very far distances. The accretion disc fueled by continuous feeding of matter to the powerful central black hole gives rise to several emission features characteristic of the quasar (see Inayoshi et al., 2020, for a recent review). The supermassive black hole (SMBH) at the centre of a quasar can accrete matter from its surroundings triggering its AGN activity (Kauffmann & Haehnelt, 2000; Di Matteo et al., 2005). In this scenario, the powerful AGN outflows can expel large amounts of metals, previously formed in stars, to considerable distances in the halo; therefore large reservoir of metal gas is present in their haloes (Prochaska et al., 2011). On the other hand, the AGN outflow can ionize the gas in the halo up to large distances (Hennawi & Prochaska, 2007; Farina et al., 2013). These motivate us to look for these signatures using low and moderately ionized metal absorbers in background quasars which are useful tracers of cool and warm gas in the CGM.

With Mg II selected C IV systems detected over a broad range of redshifts, it is possible to connect these absorbers in both angular and redshift space to quasars. It will help shed some light on the nature of cool and warm gas in the CGM of quasar host galaxies. To connect C IV absorbers with quasars in redshift range $1.5 < z < 2.3$, we make use of the SDSS DR16 QSO sample (QSOs from SDSS, BOSS and eBOSS programs) (Ahumada et al., 2020; Lyke et al., 2020).

4.4.2 Connecting Quasars With Metal Absorbers

This chapter aims to understand the nature and properties of C IV and Mg II absorbers in quasar haloes. To connect C IV absorbers with quasar halo, we cross-correlate our C IV absorbers with SDSS DR16 quasars in projected and redshift space. To connect absorbers and quasars in redshift space we apply a cut on $|\Delta z = z_{\rm QSO} - z_{\rm abs}| \leq 0.01$[15]. In the next sections, we show the properties of C IV and Mg II absorbers in projected distance and velocity space around quasars.

With projected velocities, we can study the relative kinematic properties of absorbers in quasar haloes. It will help entangle the absorber kinematic profile in the quasar halo and its comparison with the galactic halo.

4.4.3 Metal Gas in Quasar haloes

As discussed in section 4.1, Mg II and C IV absorbers trace very different phases of metal gas in the galactic haloes. Hence, it is interesting to see how the properties of the cool ($T \sim 10^{4-4.5}$ K traced by Mg II absorbers) and warm ($T \sim 10^{5-5.5}$ K traced by C IV absorbers) phase of CGM correlate in the quasar haloes at $z \sim 2$ as a function of projected distance from the quasar. It will provide us with direct constraints on the spatial distribution of metals in the CGM of quasars.

4.4.3.1 Detection Fraction of Metal Absorbers

We start with the fraction of C IV and Mg II systems above certain rest equivalent width detected in quasar haloes relative to the total quasar sightlines within a given radial bin[16]. We divide the number of Mg II and C IV absorbers by the total quasar sightlines in that radial bin to estimate this fraction. The four projected distance ($D_{\rm proj}$) bins are: [10, 100], [100, 200], [200, 500], [500, 1000] kpc. We are taking larger bins to include more pairs and enhance the S/N of the measurements.

Figure 4.6 shows the detection fraction of C IV (magenta squares) and Mg II (purple squares) absorbers as a function of projected distance from the quasar. We do not apply the completeness correction here because we are interested in the observed fraction of C IV and Mg II systems around quasars as a function of projected distance. The error bars are estimated using error propagation of Poisson statistics on the counts in each bin. The detection fraction anti-correlates with distance (Spearman $\rho_{\rm corr} = -1$, $p \approx 0$, significant drop at ~ 400 kpc), though the error bars are large. The C IV fraction varies from $2-3$ per cent at $D_{\rm proj} < 100-200$ kpc to $\lesssim 0.1$ per cent at $D_{\rm proj} \sim 800$ kpc. Similarly, the Mg II detection fraction changes from $5-6$ per cent at $D_{\rm proj} < 100-200$ kpc (~ 2 times higher

[15] Converting this to $\Delta v = c\,\Delta z/(1+z)$, for $\langle z \rangle = 1.8$, we get $\Delta v \sim 1100$ km s^{-1}. The Δv is $2-3$ times larger than the typical redshift error of SDSS DR16 quasars, which is $300-500$ km s^{-1} (Lyke et al., 2020).

[16] Mathematically, $f = \frac{N_{\rm abs}(W_{\rm r} > W_{\rm o})}{N_{\rm QSO}}$, where $N_{\rm abs}(W_{\rm r} > W_{\rm o})$ is the number of absorbers with rest equivalent width larger than certain threshold and $N_{\rm QSO}$ is the number of quasar sightlines in that radial bin.

than C IV) to $\lesssim 0.2$ per cent at $D_{\rm proj} \sim 800$ kpc (0.07 per cent for C IV). It indicates that quasar halo ($r_{\rm vir} \approx 200$ kpc) is highly metal-enriched, and both cool and warm gas is readily detected, and cool gas has a higher incidence than warm gas. In contrast, as we move away from the quasar, the cool gas dominates the distribution, and warm gas becomes less frequent.

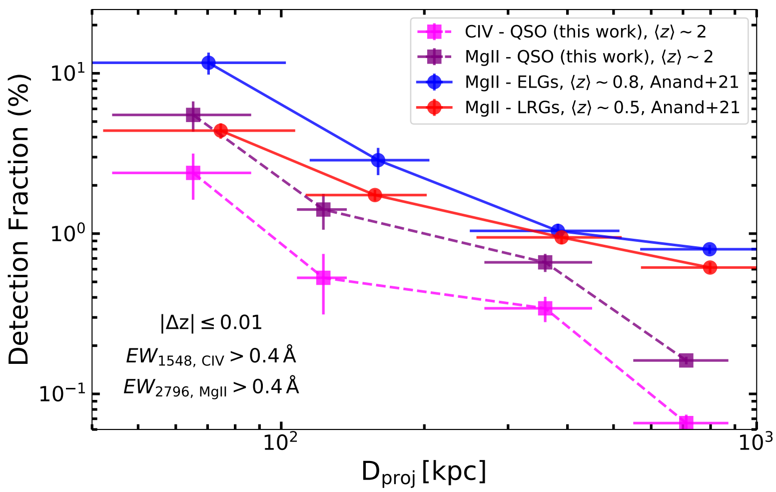

Figure 4.6: Fraction of detected C IV (magenta squares) and Mg II absorbers (purple squares) relative to the total number of quasar sightlines as a function of projected distance from the quasar. Error bars ($\pm 1\sigma$) are estimated using Poisson noise on the counts of Mg II and C IV systems in each radial bin. We also contrast the Mg II detection fraction around SDSS emission-line galaxies (ELGs) (blue circles) and luminous red galaxies (LRGs) (red circles) in the same radial bins based on Anand et al. (2021) analysis. We do not apply the completeness correction here, as we are interested in the trend as a function of projected distance. The equivalent cut is same for all absorbers $W_{\rm r} > 0.4$ Å. We see that the detection fraction declines with the projected distance. Mg II around ELGs has the highest detection fraction in each radial bin.

In addition, we also compare our results with Mg II measurements around SDSS emission-line galaxies (ELGs, $\langle z \rangle \sim 0.8$, blue circles) and luminous red galaxies (LRGs, $\langle z \rangle \sim 0.5$, red circles) based on analysis of Anand et al. (2021). We do not apply the completeness corrections and restrict to the same equivalent cut for a consistent comparison. We find that the Mg II detection fraction around ELGs is higher than both Mg II and C IV detection rate around quasars at all projected distances. For similar distances $D_{\rm proj} < 100 - 200$ kpc, Mg II detection around ELGs is $\gtrsim 10$ per cent compared to $\lesssim 5$ percent around LRGs (red circles) and $\lesssim 6$ per cent around quasars (purple squares). At very large distances ($D_{\rm proj} \sim 800$ kpc), Mg II detection fraction is much lower ($\lesssim 0.2$ per cent) around quasars than SDSS galaxies ($\sim 0.7 - 0.8$ per cent). The discrepancies could arise partly due to different redshifts of the samples, and galaxies and quasars reside in haloes with different masses, therefore, have different halo properties. ELGs typically reside in haloes with

$M_h \sim 10^{12}\,M_\odot$, while LRGs ($M_h \sim 10^{13}\,M_\odot$) and quasars ($M_h \sim 10^{12.5}\,M_\odot$) reside in more massive haloes. As shown in previous studies, massive haloes have less Mg II covering factors in their haloes, partially because gas is thermalized at high temperatures in massive haloes, which would inhibit the formation of cool gas (Lan, 2020; Anand et al., 2021). On the other hand, for ELGs, the higher Mg II absorption is also attributed to the strong stellar outflows powered by their stellar activity(Bordoloi et al., 2014; Rubin et al., 2014; Lan & Mo, 2018; Anand et al., 2021). In general, we find that C IV absorbers have a lower detection fraction than Mg II absorbers around galaxies and quasars, partly because we are only counting Mg II selected C IV absorbers. We plan to perform a blind search of C IV absorbers in the coming future for a detailed comparison.

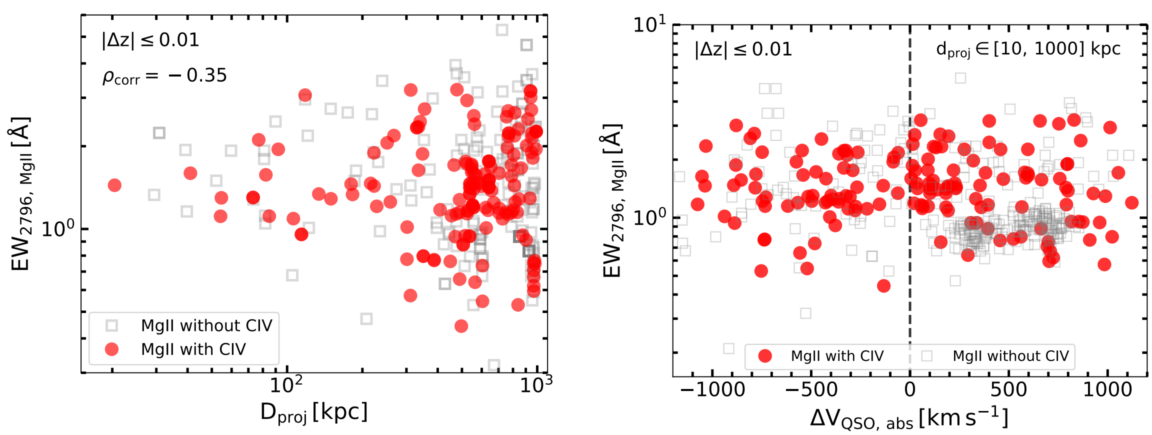

Figure 4.7: **Left:** Rest equivalent width of Mg II (EW_{2796}) as a function of projected distance (D_{proj}) from the quasar. **Right:** Rest equivalent width of Mg II (EW_{2796}) as a function of projected velocity ($\Delta V_{QSO,\,abs}$) for quasar-absorber pairs within $D_{proj} < 1$ Mpc. In both panels the red circles are the Mg II absorbers with confirmed C IV and the gray open squares are Mg II absorbers with no C IV detection. We see higher absorption (EW_{2796} $> 1 - 2$ Å) at smaller distances ($D_{proj} < 100$ kpc), though there is a large scatter at large distances. Mg II absorbers show weak anti-correlation with projected distance (Spearman correlation coefficient $= -0.35$ with $p-$ value close to zero). In projected velocity space, we see a symmetric distribution with no correlation. In both panels, results are shown for $|\Delta z = z_{QSO} - z_{abs}| \leq 0.01$ case. For visual clarity, we do not show the error bars on EW.

4.4.3.2 Rest Equivalent Width vs. Projected Distance

We now move to the equivalent width distributions of Mg II and C IV absorbers around quasars. In the left panel of Figure 4.7, we show the EW_{2796} as a function of projected distance (D_{proj}) from the quasar. The red circles show the Mg II systems for which we also could detect C IV absorbers, while open gray squares show the systems with no C IV detections. We find that the absorber strength is large ($EW_{2796} > 1 - 2$ Å, red circles) at small impact parameters ($D_{proj} < 100 - 200$ kpc). However, there is a significant scatter (EW_{2796} varying from $0.5 - 3\,\text{Å}$) at large distances ($D_{\rm proj} > 500$ kpc). Spearman correlation analysis shows a weak anti-correlation ($\rho_{\rm corr} = -0.35$, $p \approx 0$) with distance. Previous studies of Mg II absorbers in galaxies ($z \sim 0.5 - 1$) have found that EW_{2796} is anti-correlated with the impact parameter (Nielsen et al., 2013; Chen, 2017; Lan & Mo, 2018; Huang et al., 2021), though the scatter is also large (see Huang et al., 2021; Dutta et al., 2021b). In the right panel of Figure 4.7, we show the equivalent distribution in projected velocity space. The absorbers are distributed symmetrically with no correlation, though a few absorbers have large velocities ($> 500\ \rm km\,s^{-1}$). In summary, this analysis clearly shows that haloes of quasars at $z \sim 2$ are metal-enriched and host larger reservoirs of cool metal gas around them.

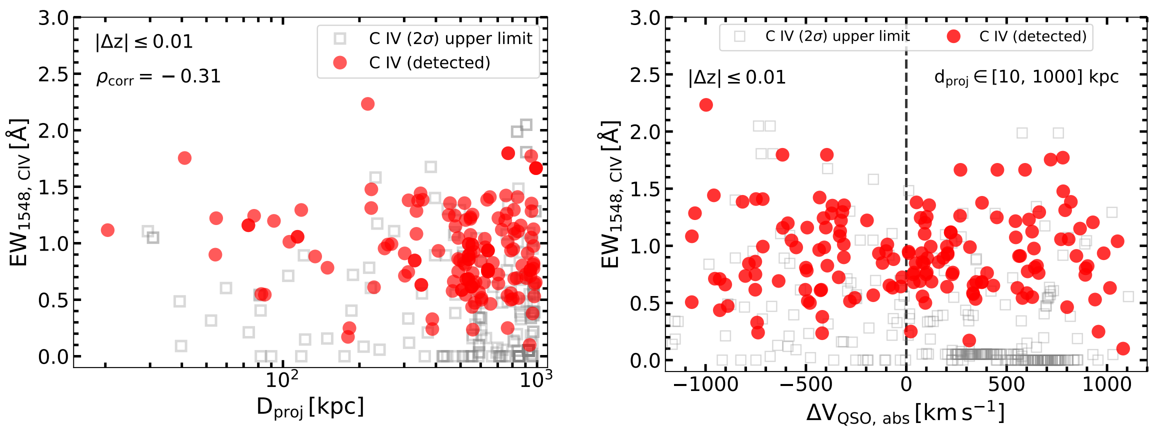

Figure 4.8: **Left:** Rest equivalent width of C IV (EW_{1548}) as a function of projected distance ($D_{\rm proj}$). **Right:** EW_{1548} as a function of projected velocity ($\Delta V_{\rm QSO,\,abs}$) for quasar- absorber pairs within $D_{\rm proj} < 1\,$Mpc. In both panels the red circles show the measured EW_{1548} for detected C IV systems while the gray open squares are 2σ upper limits for non-detections. We see higher absorption (EW $> 0.5 - 1\,\text{Å}$) at smaller distances ($D_{\rm proj} < 100\,$kpc). C IV absorbers show weak anti-correlation (Spearman correlation coefficient $= -0.31$ with $p-$ value close to zero). The absorbers are distributed symmetrically in projected velocity space with no correlation. In both panels, results are shown for $|\Delta z = z_{\rm QSO} - z_{\rm abs}| \leq 0.01$ case. For visual clarity, error bars are omitted in both panels.

The left panel of Figure 4.8 shows the equivalent width of C IV absorbers (EW_{1548}) as a function of projected distance around quasars. We find similar behaviour, there is strong C IV absorption ($EW_{1548} \approx 0.5 - 1\,\text{Å}$) at small distances ($D_{\rm proj} < 200$ kpc), while C IV absorption varies from $0.1 - 1.5\,\text{Å}$ at large distances ($D_{\rm proj} > 500$ kpc). The red circles show the EW_{1548} measurements (for detections), while open gray circles present the 2σ upper limits on EW_{1548} (for non-detections). In some cases, the upper limits are negative (possibly indicating emission like features or errors are large); therefore, we set them as zeros. Overall, we see a moderate anti-correlation (Spearman correlation coefficient $= -0.31$ with $p \approx 0$) between EW_{1548} and projected distance ($D_{\rm proj}$). Similar weak anti-correlation was found in $z \sim 2$ quasars by Prochaska et al. (2014) at $> 99.99\,\%$ confidence, using Keck data. On the other hand, in projected velocity space (right panel), the absorber distribution is very symmetric and with no visible correlation, though a few C IV absorbers also have large velocities (> 500 km s^{-1}).

Furthermore, we also observe that C IV absorption is systematically lower ($\sim 2-3$ times) than Mg II absorption (note the vertical scale in both panels). One possible reason could be the smaller oscillator strength of C IV compared to Mg II . Though, it may also be related to the relative abundances of these metals in quasar haloes at these redshifts. Indeed, the extensive surveys have revealed an increasing trend with redshift for the number of Mg II absorbers per unit redshift path (dN/dz) and the differential number peaks around $z \sim 2$ (Nestor et al., 2005; Zhu & Ménard, 2013a). In the literature, it has been interpreted as Mg II absorbers tracing cool gas associated with star-formation, which peaks around $z \sim 1.7 - 2$ (Madau & Dickinson, 2014). On the other hand, the differential C IV numbers per unit comoving path (dN/dX) increases smoothly from $z = 4 \rightarrow 0$ (Cooksey et al., 2010, 2013; Hasan et al., 2020). The physical interpretation has been that the comoving number density and absorber cross-section of C IV clouds have increased steadily for the last 10 billion years.

Finally, combining our data (~ 3 times larger) with Prochaska et al. (2014), we conclude that both ionized warm and cool gas is frequently present in quasar haloes, and cool gas absorption is higher compared to warm gas absorption. The two-point cross-correlation studies have shown that quasars typically reside in dark matter haloes with $M_{\rm halo} = 10^{12-12.5}\,{\rm M}_\odot$ (Vikas et al., 2013; Prochaska et al., 2014; Gontcho A Gontcho et al., 2018), which corresponds to $r_{\rm vir} = 160 - 200$ kpc. As discussed above, this indicates that in the inner regions of quasar haloes ($D_{\rm proj} \lesssim r_{\rm vir}$) majority of the C IV and Mg II absorbers have significantly higher absorption. In addition to metal-rich inner regions, the quasar environment is also significantly metal-enriched even at far distances ($D_{\rm proj} \gtrsim 3 - 5 r_{\rm vir}$, $r_{\rm vir} \sim 200$ kpc). We interpret this as a signpost that previous generation stars can enrich the haloes (and possibly IGM) up to considerably large distances.

4.4.4 Line-of-sight Relative Kinematics of metal absorbers in QSOs

Next, we study the line-of-sight kinematics of C IV and Mg II absorbers relative to the quasar to shed some light on the motion of absorbing clouds in the CGM of quasars at $z \sim 2$. With the robust spectroscopic redshifts of quasars and metal absorbers, we estimate Δv[17] within some projected distance from the quasar. In Figure 4.9, we show the Δv distribution of Mg II (green) and C IV (purple) absorbers within $D_{\rm proj} < 400$ kpc, $\lesssim 2{\rm r}_{\rm vir}$ from the quasar. We find that both distributions are very flat with mean $\langle \Delta v \rangle \lesssim -25$ km s^{-1}. There is a weak indication of two distinct peaks around -1000 km s^{-1}and $300 - 400$ km s^{-1}for both absorbers. The peak on the negative side (~ 13 % of the sample) is a possible signpost of accretion, while the positive peak (similar fraction) implies evidence for outflows. We speculate that our analysis reveals that the motion of C IV and Mg II absorbers in quasars

[17] $\Delta v = c \cdot \frac{z_{\rm abs} - z_{\rm QSO}}{(1+z_{\rm QSO})}$, where abs = Mg II, C IV.

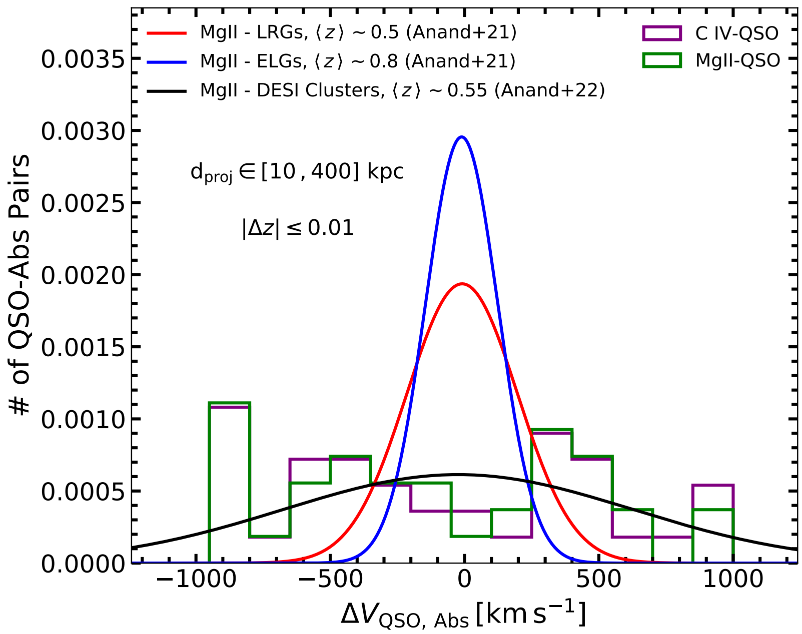

Figure 4.9: Relative line-of-sight velocity between metal absorbers and quasars within $D_{\mathrm{proj}} < 400\,\mathrm{kpc}$. The C IV and Mg II absorbers are shown in purple and green, respectively. We also contrast results for Mg II -LRGs (red) and Mg II -ELGs (blue) (Anand et al., 2021) and Mg II -DESI clusters (black) (Anand et al., 2022) for a comparison.

is a mixture of both accretion and galactic outflows. It will be interesting to explore this in more detail with high-resolution quasar spectra in future.

Furthermore, we also contrast our measurements with Mg II - galaxies and clusters cross-correlation measurements from previous studies (Anand et al., 2021, 2022). We compare the relative kinematics of Mg II absorbers in the halo of luminous red galaxies (LRGs, shown in red). In addition, we also show the measurements for Mg II - emission-line galaxies (ELGs, in blue) from (Anand et al., 2021). On the massive end, we also compare our measurements with Mg II kinematics in DESI legacy[18] clusters (black) from Anand et al. (2022). We find the Gaussian fit to LRGs and ELGs are much narrower ($\sigma_{\mathrm{LRG}} \sim 200$ $\mathrm{km\,s^{-1}}$and $\sigma_{\mathrm{ELG}} \sim 130\ \mathrm{km\,s^{-1}}$) than the observed distribution of C IV and Mg II absorbers in quasars (compare red and blue Gaussian to green and purple histogram). On the other hand, the distribution of C IV and Mg II motion in quasars are very similar to the motion of Mg II absorbers in DESI legacy clusters (black Gaussian profile, $\sigma_{\mathrm{DESI}} \sim 650\ \mathrm{km\,s^{-1}}$). Note that clusters reside in much more massive haloes ($M_{\mathrm{halo}} \gtrsim 10^{14}\,\mathrm{M_{\odot}} \implies \mathrm{v_{vir}} = 900\ \mathrm{km\,s^{-1}}$, see Kravtsov & Borgani (2012) for a review) than quasars ($M_{\mathrm{halo}} \sim 10^{12-12.5}\,\mathrm{M_{\odot}} \implies \mathrm{v_{vir}} = 150\ \mathrm{km\,s^{-1}}$). In clusters, the motion of absorbing clouds is suppressed relative to virial motion, while in quasar halo, the absorber motion is larger than typical halo virial velocities. Powerful quasar outflows may have enhanced the absorber motion in its CGM.

[18]Dark Energy Spectroscopic Instrument Legacy Imaging surveys https://www.legacysurvey.org/

Both the recent simulations (Nelson et al., 2019) and observations (Circosta et al., 2018; Perrotta et al., 2019) the central AGN of a quasar (powered by black hole accretion) can expel metals out to tens of kpc with velocities up to $\sim 3000\ \mathrm{km\,s^{-1}}$.

4.5 Summary

This chapter presents a large absorber-galaxy cross-correlation study for cool and warm gas in the CGM of quasars at $1.5 < z < 2.3$. The work is ongoing and the main results are:

- We constructed a large C IV absorber ($N \sim 22,000$) catalogue based on SDSS DR16 Mg II catalogue.
- The completeness corrected equivalent width distribution can be best described by a broken power-law, indicating different populations of C IV absorbers in the IGM at $z \sim 2$.
- By cross-correlating Mg II and C IV absorbers with quasars, we have studied the cool and warm gas distribution of the CGM at $z \sim 2$.
- Our study reveals a large reservoir of metals in the quasar halo, implying that quasar haloes are highly metal-enriched.
- The observed detection fraction of C IV and Mg II absorbers around quasars is $2 - 3$ times lower than SDSS ELGs and LRGs.
- The relative kinematics of absorbers around quasars point towards a scenario where a significant fraction of absorbers may be associated with quasar outflows.

Look up at the stars and not down at your feet. Try to make sense of what you see, and wonder about what makes the universe exist. Be curious.

Stephen Hawking

Chapter 5

Summary, Conclusion and Future Outlook

In this chapter, I summarize findings of this thesis and their implications in the context of the circumgalactic medium (CGM) and its connection to galaxy formation evolution. Moreover, I also illustrate future applications of some of the methods and techniques developed in this thesis. Finally, I also discuss briefly the next steps in the field where the next generation of upcoming telescopes can provide a holistic picture of this gas and its connection to galaxy formation and evolution.

5.1 Metal absorbers in Quasar spectra

In chapter 2 (Anand et al., 2021), I design a fully automated quasar continuum estimator and absorption line detection pipeline. The pipeline was run on ~ 1 million quasars from the SDSS DR16 and resulted in the most extensive Mg II absorber catalogue (MPA-SDSS catalogue) to date. Full catalogue is publicly available at https://www.mpa-garching.mpg.de/SDSS/MgII/. The main features of our catalogue are:

- The catalogue contains 160,000 Mg II absorbers ($0.36 < z < 2.3$) and 72,000 Fe II systems based on the SDSS DR16 final quasar sample.

- After stacking quasar spectra in the rest-frame of Mg II absorbers, several other weak metal lines including Si I , C IV, Al II, Al II, Mg I and Fe II were detected.

- By simulating millions of fake absorbers, I also characterize the completeness and purity of the detection pipeline. Our pipeline is less complete for weak absorbers ($EW_{2796} < 0.5$ Å) and fairly complete for strong absorbers. On the other hand, purity is also very high ($\sim 95\%$) for our catalogue.

5.2 Mg II absorbers in SDSS galaxies

In the second half of chapter 2, I cross-correlate these Mg II absorbers to ~ 1.3 million galaxies from the SDSS DR16, divided into star-forming, emission-line galaxies (ELGs) and quiescent, luminous red galaxies (LRGs). I use these samples to study the properties of metal absorption in the circumgalactic medium and intergalactic medium at $z \sim 0.5-0.8$. I also investigate the absorber-galaxy correlation based on several properties of galaxies to understand the impact of galactic processes on its CGM. The key results are:

- Thanks to the larger quasar and galaxy samples that our study uses, we could characterize the scale dependence of Mg II absorber covering fraction, surface number density and rest equivalent width with greater accuracy.

- This analysis reveals that Mg II covering fraction for both LRGs and ELGs declines rapidly at a projected radius of 50 kpc. The greater S/N of our measurements allows us to characterize the sharpness of the break better. The sharp break at ~ 50 kpc indicates the separation between two regimes. In the inner region, the physical properties of the CGM are regulated by galactic outflows and the outer CGM is in thermal equilibrium with the dark matter halo.

- The analysis reveals a weak stellar-mass dependence of Mg II covering fractions for ELGs but a strong stellar-mass trend for LRGs. The cumulative covering fraction is anti-correlated with stellar mass of LRGs on normalized distance scales (projected distance divided by $r_{\rm vir}$).

- In ELGs, SFR seems to have a strong impact on the Mg II properties as we find a strong positive correlation between Mg II covering fraction and SFR on virial scales.

- The line-of-sight relative kinematics of absorber-galaxy pairs around ELGs is well-characterized by a single gaussian profile. At the same time, LRGs require a second, broad component with $\sigma_v \sim 1100\ {\rm km\,s^{-1}}$ to capture high-velocity tails, likely indicative of AGN-driven outflows.

- The gas motions around ELGs are similar to the expected dark matter halo velocity dispersion. In contrast, gas motions around LRGs are suppressed and closer to $\sim 0.4 - 0.5\sigma_{\rm DM,\,halo}$. The suppressed gas motion around LRGs implies that gas is gravitationally bound to the dark matter halo and unlikely to have originated in outflows. Though the cool gas can be accreted from the neighbouring haloes or the IGM.

- The different properties and trends of Mg II absorption with galactic properties between ELGs and LRGs point towards a scenario where cool CGM of star-forming and quiescent galaxies has a fundamentally different physical origin.

5.3 The cool metal gas in galaxy clusters

In Chapter 3 (Anand et al., 2022), I extended the absorber cross-correlation study to the galaxy clusters to understand the nature and origin of the cool gas in galaxy clusters. I also investigated the difference between the cool CGM properties of isolated and group galaxies. Our study is the largest to date in this context. The key results and conclusions of our work are:

- The cross-correlation reveals a significant covering fraction of both weak ($0.4\,\text{Å} < EW_{2796} < 1\,\text{Å}$) and strong ($EW_{2796} > 1\,\text{Å}$) Mg II absorbers in cluster halo ($D_{\rm proj} \lesssim r_{500}$), implying that the cool gas has significantly higher incidence rate in galaxy clusters, despite the hot ICM.

- Our high S/N samples also allowed us to characterize the radial scales of Mg II absorption in more detail. We find that the Mg II incidence declines faster for strong absorbers than weak systems, indicating that strong absorbers are preferentially found at small distances.

- Clusters contain $\sim 3 - 10$ times more total Mg II gas mass than SDSS red galaxies (LRGs) within r_{500}, partly due to their larger halo sizes.

Additionally, I also investigate the correlation between Mg II absorbers and cluster member galaxies, including their physical properties: stellar mass, SFR, and sSFR. Our main findings are:

- Clusters with more star-forming galaxies in their haloes have slightly higher Mg II covering fractions, indicating some (weak) level of connection between stellar activity and Mg II absorption.

- Mg II absorbers are not randomly distributed within the cluster volume: they are located preferentially closer to cluster member galaxies.

- The typical separations between absorber and galaxies (~ 100 , kpc) suggest that Mg II absorption arises at least in part from cool gas that has been removed from their CGM due to ram-pressure or tidal stripping.

5.4 Quasar halo connection with metal absorbers

Finally, in chapter 4 (Anand et al., in prep.), I attempt to characterize the cool ($T \sim 10^{4-4.5}$ K) and warm ($T \sim 10^{5-5.5}$ K) CGM of quasars at $z \sim 2$ using a quasar-metal cross-correlation study with SDSS DR16 quasars. This is one of the first studies of simultaneous detection of quasar CGM's cool and warm phase at these redshifts. Our main findings and conclusions are:

- I compile a Mg II selected C IV absorber (a doublet with $\lambda\lambda$ 1548, 1550) catalogue. It includes $\sim 22,000$ C IV systems with $1.5 < z < 2.3$ and $EW_{1548} < 2.5$ Å.
- C IV properties and Mg II properties are positively correlated, though C IV rest equivalent widths are smaller than Mg II , partly due to their weak oscillator strength.
- By cross-correlating these C IV absorbers and their parent Mg II systems with quasars at $1.5 < z < 2.3$, I characterize the metal distribution around quasars.
- Quasar halo is very metal-enriched, and both C IV and Mg II absorbers have higher equivalent widths in the inner part of the quasar halo.
- The relative line-of-sight analysis reveals that absorbing clouds in quasar halo have $2-3$ times higher velocity dispersion than ELGs, which also reside in similar dark matter haloes ($M_{\mathrm{h}} \sim 10^{12-12.5}\, M_{\odot}$). This possibly indicates that quasar outflow may be responsible for boosting the motion of metal absorbers in their haloes.

5.5 Future Outlook

As the absorber detection pipeline developed in this thesis is very generic in detecting doublet lines, it can easily be extended to search for other doublet lines in the quasar spectra such as Si II, Ca II, Al III and Fe II . Using these metal lines, we can estimate the metallicity of the CGM at different epochs, given that we have some measurements on hydrogen content. On the other hand, we can also model the galaxy continuum with the exact NMF based continuum modelling. Zhu et al. (2015) implemented Nonnegative matrix factorization (NMF) to model the spectra of ELGs and measure the emission properties.

In this thesis, we have benefited from the cross-correlation of an extensive Mg II absorber catalogue with similarly extensive spectroscopic and photometric galaxy samples. However, in understanding the galaxy-CGM connection, we were limited in the physical properties of galaxies available. The ground-based wavelength limitations also restricted the available redshift range to connect galaxies and absorbers. Finally, available SDSS imaging does not generally allow fine-grained galactic morphologies to be inferred. However, such a morphology-based analysis can be performed with the recent data from the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Survey (Dey et al., 2019).

In the future, other large imaging surveys, such as the Large Synoptic Survey Telescope (LSST) at the Rubin Observatory (The LSST Collaboration et al., 2018), will provide enormous datasets of galaxies, up to higher redshifts, and with high-quality imaging. Together with upcoming large spectroscopic galaxy surveys such as the Prime Focus Spectrograph (PFS) on the Subaru telescope (Tamura et al., 2016), statistical analyses of the circumgalactic medium will be an even more powerful tool to understand the formation and evolution of galaxies across cosmic time.

On the other hand, in clusters, we found little direct connection with the stellar activity or the CGM of member galaxies. Therefore, a more detailed analysis is needed, given the uncertainties in current photo$-z$ measurements, not accounting for non-detection or weak Mg II absorber systems and the stellar mass limits of the DESI legacy survey. The currently ongoing DESI spectroscopic survey will provide accurate redshifts (Besuner et al., 2021), detailed relative velocity analysis and definitive associations will be possible. An 'Early Data Release' (EDR) and the first 'Data Release (DR) are expected by the end of the year.

Moreover, with deeper surveys such as the LSST at the Rubin Observatory and PFS on the Subaru telescope, enormous datasets of low mass galaxies at high redshifts will become available. This will allow us to test our hypotheses in more detail, including the possible association between Mg II absorbers and faint galaxies in groups and clusters. Finally, in addition to optical data, the ongoing *eROSITA* all-sky survey promises direct characterization of the hot gas components in galaxy clusters (Bulbul et al., 2021; Liu et al., 2021). In the future, it will be possible to compare the gas contents of clusters as a function of temperature and phase (see Truong et al., 2020; Oppenheimer et al., 2020; Truong et al., 2021; Das et al., 2021b) through multi-wavelength and multi-dataset analyses.

Similarly, the metal-quasar cross-correlation study can be extended to include quasar properties such as its black hole mass and SFR to explore how they impact the nature of their halo gas. Finally, a comparative study of cool and warm halo gas in galaxies, clusters and quasars at different epochs would be desirable. Ultimately, connecting these observations with theoretical and numerical models of this multi-phase gas would be a key goal in the near future.

Appendix A

The cool circumgalactic medium in SDSS galaxies

A.1 Composite Median SDSS Quasar Spectra

As discussed in section 2.2.2.3, our choice for flux normalization is based on median composite spectra constructed in Vanden Berk et al. (2001) using 2200 quasars from SDSS commissioning phase. In Figure A.1 (taken from Vanden Berk et al. (2001)), we show the median composite spectra of SDSS quasars. To summarize the method, the quasar spectra were shifted in their rest-frames and re-binned on a common wavelength scale. Then the composite flux is estimated as the median flux value (scaled in an arbitrary manner, see section 3.1 in Vanden Berk et al. (2001)) in each wavelength bin. The red dashed arrows show the wavelength regions of interest for normalization purposes. As one can see, the quasar spectrum is mostly featureless in those regions; therefore, mean flux values in those regions are not affected by strong emission lines.

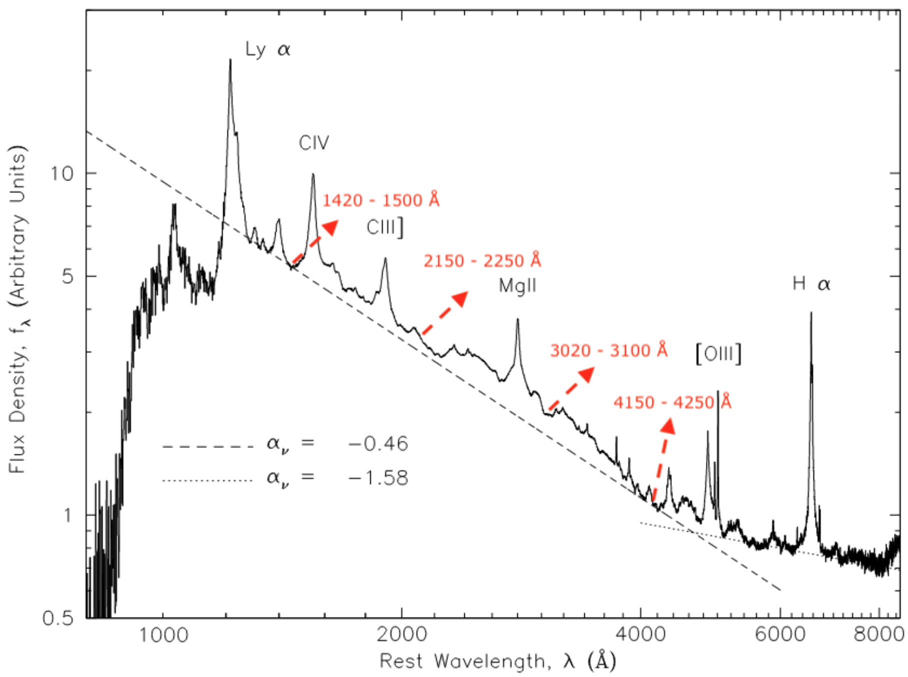

Figure A.1: Composite spectra of 2200 quasars, constructed using median flux values in each wavelength pixel. The Figure is taken from Vanden Berk et al. (2001) and based on quasar data from SDSS commissioning phase. The spectrum has a resolution of $R \approx 1800$. The intrinsic quasar emission features such as Lyα, C IV , Mg II , [O III] are clearly visible. The red dashed arrows show the wavelength range used for normalizing the quasar flux for NMF eigenspectra construction (see section 2.2.2.3). The spectrum is featureless in these wavelength ranges. Also shown is the power-law describing the slopes of the quasar continuum.

Appendix B

Cool circumgalactic gas in galaxy clusters

B.1 Effect of Redshift Difference between BCG and Mg II Absorbers

B.1.1 Mg II Covering Fraction

To investigate the effect of $|\Delta z| = |z_{\mathrm{BCG}} - z_{\mathrm{Mg\,II}}|$ choice on the measurements (particularly the Mg II covering fractions such that there are not many spurious absorber-cluster pairs), we also estimate the Mg II covering fraction for absorbers that reside within $|\Delta z| \leq 0.003$ ($|\Delta v| \leq 600$ km s^{-1}). As we understand, this is smaller than the typical $v_{500} \sim 900$ km s^{-1}, of the clusters hence, all the absorbers can be assumed to be gravitationally bound to the cluster halo. In Figure B.1, we compare the Mg II covering fraction (for both weak and strong absorbers) as a function of $D_{\mathrm{proj}}/r_{500}$ and D_{proj} for both $|\Delta z| \leq 0.003$ (blue squares) and $|\Delta z| \leq 0.01$ (purple squares) choices. The best-fitting parameters (solid lines) are compiled in Table 3.4. We find that even after decreasing the $|\Delta z|$, Mg II covering fractions (compare blue squares with purple squares) do not differ significantly, particularly at small distances on both projected and normalized scales though the difference is significant at large distances. At large scales, a large $|\Delta z|$ choice will include many absorber-BCG pairs, implying a higher covering fraction in that case. The analysis confirms that a significant fraction of the absorbers at small distances indeed have small Δz, relative to the cluster BCG. Hence, our results do not vary significantly at small distances even after decreasing the Δz to small values. Besides this, we also see that the slope (parametrized by α) is also larger, implying a faster decline in covering fractions on larger scales. On the other hand, the characteristic scales (parametrized by x_{o}) of Mg II covering fraction are slightly larger (see Table 3.4), than the $|\Delta z| \leq 0.01$, though the error bars are also larger as S/N of the measurements also goes down.

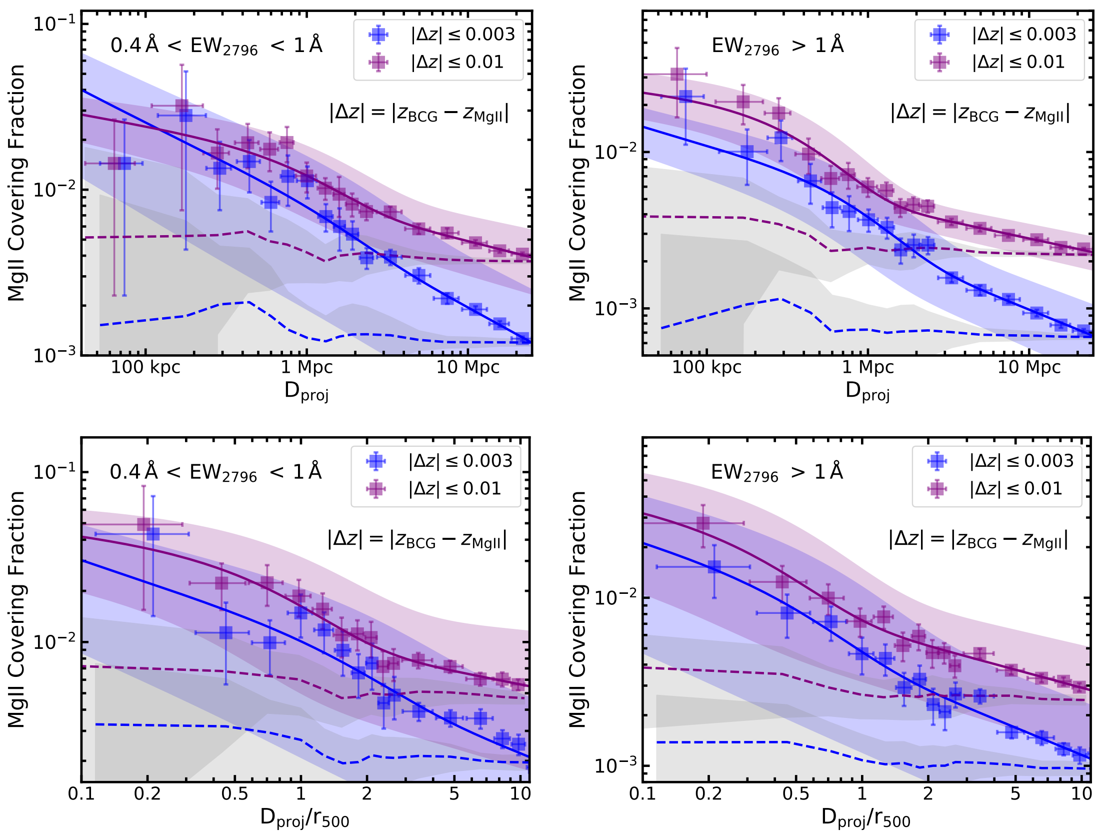

Figure B.1: Effect of changing redshift interval on Mg II covering fraction. Comparing results for $|\Delta z| \leq 0.01$ (purple squares) and $|\Delta z| \leq 0.003$ (blue squares). **Top:** Differential covering fraction of weak and strong Mg II absorbers around DESI galaxy clusters as a function of projected distance from the BCG. **Bottom:** Mg II covering fractions of weak and strong systems as a function of projected distance normalized by r_{500} of cluster. In all panels the red squares show the Mg II covering fraction for absorber-BCG pairs with $|\Delta z| \leq 0.003$. The corresponding dashed lines show the measurements around random quasar sightlines. We compile the fitting parameters for these measurements in Table 3.4 and the shaded regions show the corresponding 1σ uncertainty intervals.

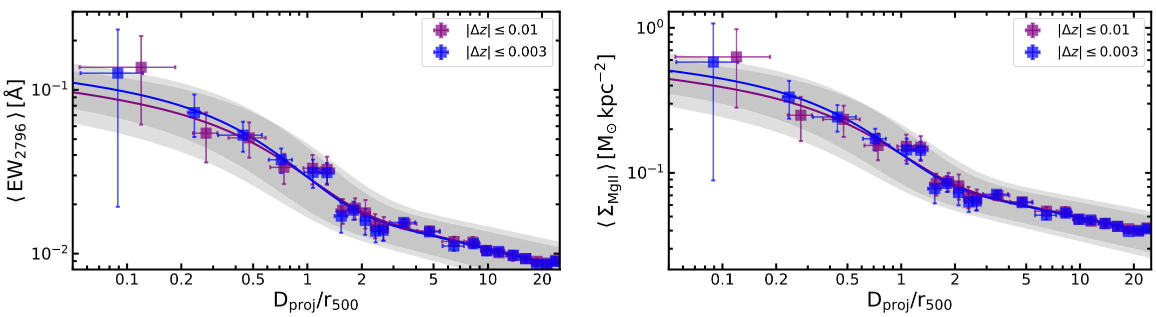

Figure B.2: **Left:** The mean $\langle EW_{2796} \rangle$ in and around DESI legacy galaxy clusters (purple squares for $|\Delta z| \leq 0.01$ and blue squares for $|\Delta z| \leq 0.003$) as a function of projected distance normalized by r_{500} of the corresponding haloes. **Right:** Differential average surface mass density of Mg II absorbers in and around DESI legacy galaxy for both $|\Delta z|$ cuts. The solid lines are the best-fitting curves described in section 3.4.2 and the best-fitting parameters are summarized in Table 3.4.

B.1.2 Mg II Mean Equivalent Width and Surface Mass Density

Similarly, we do not see any significant variation in mean equivalent widths and surface mass densities of Mg II absorbers in clusters with $|\Delta z|$ choice, particularly at small scales. This analysis shows that our results and conclusions are not dependent on this choice. We show the measurements for $|\Delta z| \leq 0.003$ case in Figure B.2. The best-fitting parameters for these measurements are also shown in Table 3.4.

Bibliography

Abolfathi B., et al., 2018, ApJS, 235, 42

Afruni A., Fraternali F., Pezzulli G., 2019, A&A, 625, A11

Afshordi N., Lin Y.-T., Nagai D., Sanderson A. J. R., 2007, MNRAS, 378, 293

Ahumada R., et al., 2020, ApJS, 249, 3

Anand A., Nelson D., Kauffmann G., 2021, MNRAS, 504, 65

Anand A., Kauffmann G., Nelson D., 2022, MNRAS, 513, 3210

Anderson M. E., Churazov E., Bregman J. N., 2016, MNRAS, 455, 227

Arrigoni Battaia F., Hennawi J. F., Prochaska J. X., Cantalupo S., 2015, ApJ, 809, 163

Arrigoni Battaia F., Hennawi J. F., Prochaska J. X., Oñorbe J., Farina E. P., Cantalupo S., Lusso E., 2019, MNRAS, 482, 3162

Ayromlou M., Kauffmann G., Yates R. M., Nelson D., White S. D. M., 2021, MNRAS, 505, 492

Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017

Beck R., Dobos L., Budavári T., Szalay A. S., Csabai I., 2016, MNRAS, 460, 1371

Becker G. D., Sargent W. L. W., Rauch M., Calverley A. P., 2011, ApJ, 735, 93

Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379

Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143

Bergeron J., 1986, A&A, 155, L8

Bertone S., Schaye J., Dolag K., 2008, Space Science Reviews, 134, 295

Besuner R., et al., 2021, arXiv e-prints, p. arXiv:2101.11794

Birnboim Y., Dekel A., 2003, MNRAS, 345, 349

Blanton M. R., Roweis S., 2007, AJ, 133, 734

Blanton M. R., et al., 2017, AJ, 154, 28

Bluck A. F. L., et al., 2020, MNRAS, 499, 230

Blumenthal G. R., Pagels H., Primack J. R., 1982, Nature, 299, 37

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Böhringer H., Werner N., 2010, A&A, 18, 127

Boksenberg A., Sargent W. L. W., 2015, ApJS, 218, 7

Bolton A. S., et al., 2012, AJ, 144, 144

Bond H. E., Nelan E. P., VandenBerg D. A., Schaefer G. H., Harmer D., 2013, ApJL, 765, L12

Bordoloi R., et al., 2011, ApJ, 743, 10

Bordoloi R., et al., 2014, ApJ, 794, 130

Bordoloi R., Prochaska J. X., Tumlinson J., Werk J. K., Tripp T. M., Burchett J. N., 2018, ApJ, 864, 132

Borthakur S., et al., 2015, ApJ, 813, 46

Borthakur S., et al., 2016, ApJ, 833, 259

Bouché N., Murphy M. T., Péroux C., Csabai I., Wild V., 2006, MNRAS, 371, 495

Bowler R. A. A., Hewett P. C., Allen J. T., Ferland G. J., 2014, MNRAS, 445, 359

Bradshaw E. J., et al., 2013, MNRAS, 433, 194

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

Bulbul E., et al., 2021, arXiv e-prints, p. arXiv:2110.09544

Burchett J. N., et al., 2015, ApJ, 815, 91

Burchett J. N., et al., 2016, ApJ, 832, 124

Burchett J. N., Tripp T. M., Wang Q. D., Willmer C. N. A., Bowen D. V., Jenkins E. B., 2018, MNRAS, 475, 2067

Burchett J. N., et al., 2019, ApJL, 877, L20

Burchett J. N., Rubin K. H. R., Prochaska J. X., Coil A. L., Vaught R. R., Hennawi J. F., 2021, ApJ, 909, 151

Butsky I. S., Burchett J. N., Nagai D., Tremmel M., Quinn T. R., Werk J. K., 2019, MNRAS, 490, 4292

Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, Nature, 506, 63

Caulet A., 1989, ApJ, 340, 90

Cen R., Ostriker J. P., 1999, ApJ, 514, 1

Cen R., Roxana Pop A., Bahcall N. A., 2014, Proceedings of the National Academy of Science, 111, 7914

Chabrier G., 2003, PASP, 115, 763

Charlton J. C., Ding J., Zonak S. G., Churchill C. W., Bond N. A., Rigby J. R., 2003, ApJ, 589, 111

Chen H.-W., 2017, The Circumgalactic Medium in Massive Halos. Springer International Publishing, p. 167, doi:10.1007/978-3-319-52512-9˙8

Chen H.-W., Lanzetta K. M., Webb J. K., 2001, ApJ, 556, 158

Chen H.-W., Helsby J. E., Gauthier J.-R., Shectman S. A., Thompson I. B., Tinker J. L., 2010, ApJ, 714, 1521

Chen Y.-M., et al., 2012a, MNRAS, 421, 314

Chen Y.-M., et al., 2012b, MNRAS, 421, 314

Chen Z.-F., Gu Q.-S., Zhou L., Chen Y.-M., 2016, MNRAS, 462, 2980

Choi E., Ostriker J. P., Naab T., Oser L., Moster B. P., 2015, MNRAS, 449, 4105

Churchill C. W., Rigby J. R., Charlton J. C., Vogt S. S., 1999, ApJS, 120, 51

Churchill C. W., Mellon R. R., Charlton J. C., Jannuzi B. T., Kirhakos S., Steidel C. C., Schneider D. P., 2000, ApJS, 130, 91

Churchill C. W., Kacprzak G. G., Steidel C. C., 2005, in Williams P., Shu C.-G., Menard B., eds, IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines. pp 24–41 (`arXiv:astro-ph/0504392`), doi:10.1017/S1743921305002401

Churchill C. W., Trujillo-Gomez S., Nielsen N. M., Kacprzak G. G., 2013, ApJ, 779, 87

Churchill C. W., Evans J. L., Stemock B., Nielsen N. M., Kacprzak G. G., Murphy M. T., 2020, ApJ, 904, 28

Circosta C., et al., 2018, A&A, 620, A82

Codoreanu A., Ryan-Weber E. V., García L. Á., Crighton N. H. M., Becker G., Pettini M., Madau P., Venemans B., 2018, MNRAS, 481, 4940

Cooksey K. L., Thom C., Prochaska J. X., Chen H.-W., 2010, ApJ, 708, 868

Cooksey K. L., Kao M. M., Simcoe R. A., O'Meara J. M., Prochaska J. X., 2013, ApJ, 763, 37

Cooray A., Milosavljević M., 2005, ApJL, 627, L89

Cortese L., et al., 2007, MNRAS, 376, 157

D'Odorico V., Calura F., Cristiani S., Viel M., 2010, MNRAS, 401, 2715

Dalla Vecchia C., Schaye J., 2008, MNRAS, 387, 1431

Das H. K., Choudhury P. P., Sharma P., 2021a, MNRAS, 502, 4935

Das S., Mathur S., Gupta A., Krongold Y., 2021b, ApJ, 918, 83

Davé R., et al., 2001, ApJ, 552, 473

Davé R., Finlator K., Oppenheimer B. D., 2011, MNRAS, 416, 1354

Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, MNRAS, 434, 2645

Dawson K. S., et al., 2013, AJ, 145, 10

Dawson K. S., et al., 2016, AJ, 151, 44

Dekel A., et al., 2009, Nature, 457, 451

Dey A., et al., 2019, AJ, 157, 168

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Donahue M., Voit G. M., 2022, arXiv e-prints, p. arXiv:2204.08099

Donnari M., et al., 2019, MNRAS, 485, 4817

Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium. Princeton University Press

Dutta R., et al., 2020, MNRAS, 499, 5022

Dutta A., Sharma P., Nelson D., 2021a, MNRAS,

Dutta R., et al., 2021b, MNRAS, 508, 4573

Edge A. C., 2001, MNRAS, 328, 762

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Einasto J., Kaasik A., Saar E., 1974, Nature, 250, 309

Eisenstein D. J., et al., 2011, AJ, 142, 72

Elahi P. J., Power C., Lagos C. d. P., Poulton R., Robotham A. S. G., 2018, MNRAS, 477, 616

Fang T., Buote D., Bullock J., Ma R., 2015, ApJS, 217, 21

Farina E. P., Falomo R., Decarli R., Treves A., Kotilainen J. K., 2013, MNRAS, 429, 1267

Farina E. P., et al., 2017, ApJ, 848, 78

Favole G., et al., 2016, MNRAS, 461, 3421

Fielding D. B., et al., 2020, ApJ, 903, 32

Finlator K., Doughty C., Cai Z., Díaz G., 2020, MNRAS, 493, 3223

For B. Q., et al., 2019, MNRAS, 489, 5723

Fossati M., Gavazzi G., Boselli A., Fumagalli M., 2012, A&A, 544, A128

Fossati M., et al., 2019, MNRAS, 490, 1451

Fumagalli M., Fossati M., Hau G. K. T., Gavazzi G., Bower R., Sun M., Boselli A., 2014, MNRAS, 445, 4335

Gabor J. M., Davé R., Finlator K., Oppenheimer B. D., 2010, MNRAS, 407, 749

Gamow G., Teller E., 1939, Physical Review, 55, 654

Gao J., Zou H., Zhou X., Kong X., 2020, PASP, 132, 024101

Geha M., Blanton M. R., Yan R., Tinker J. L., 2012, ApJ, 757, 85

Genel S., et al., 2014, MNRAS, 445, 175

Gladders M. D., Yee H. K. C., 2005, ApJS, 157, 1

Gontcho A Gontcho S., Miralda-Escudé J., Font-Ribera A., Blomqvist M., Busca N. G., Rich J., 2018, MNRAS, 480, 610

Gunn J. E., Gott J. Richard I., 1972, ApJ, 176, 1

Guth A. H., Pi S. Y., 1982, PRL, 49, 1110

Hamanowicz A., et al., 2020, MNRAS, 492, 2347

Hasan F., et al., 2020, ApJ, 904, 44

Hasan F., Churchill C. W., Stemock B., Nielsen N. M., Kacprzak G. G., Croom M., Murphy M. T., 2022, ApJ, 924, 12

Hasselquist S., et al., 2021, ApJ, 923, 172

Hennawi J. F., Prochaska J. X., 2007, ApJ, 655, 735

Henning T., Salama F., 1998, Science, 282, 2204

Herrmann K. A., Hunter D. A., Zhang H.-X., Elmegreen B. G., 2016, AJ, 152, 177

Huang Y.-H., Chen H.-W., Johnson S. D., Weiner B. J., 2016, MNRAS, 455, 1713

Huang Y.-H., Chen H.-W., Shectman S. A., Johnson S. D., Zahedy F. S., Helsby J. E., Gauthier J.-R., Thompson I. B., 2021, MNRAS, 502, 4743

Hubble E. P., 1926, ApJ, 64, 321

Ilbert O., et al., 2009, ApJ, 690, 1236

Inayoshi K., Visbal E., Haiman Z., 2020, ARA&A, 58, 27

Ishigaki M. N., Chiba M., Aoki W., 2012, ApJ, 753, 64

Johnson S. D., et al., 2018, ApJL, 869, L1

Jung S. L., Choi H., Wong O. I., Kimm T., Chung A., Yi S. K., 2018, ApJ, 865, 156

Kaaret P., et al., 2020, Nature Astronomy, 4, 1072

Kacprzak G. G., Churchill C. W., 2011, ApJl, 743, L34

Kalberla P. M. W., Kerp J., 2009, ARA&A, 47, 27

Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576

Kauffmann G., Nelson D., Ménard B., Zhu G., 2017, MNRAS, 468, 3737

Keeney B. A., et al., 2017, ApJS, 230, 6

Kelleher D. E., Podobedova L. I., 2008, Journal of Physical and Chemical Reference Data, 37, 709

Kennicutt Robert C. J., 1998, ARA&A, 36, 189

Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2

Kewley L. J., Geller M. J., Jansen R. A., 2004, AJ, 127, 2002

Komatsu E., et al., 2011, ApJS, 192, 18

Kramida A., Yu. Ralchenko Reader J., and NIST ASD Team 2021, NIST Atomic Spectra Database (ver. 5.9), [Online]. Available: `https://physics.nist.gov/asd` [2021, March 1]. National Institute of Standards and Technology, Gaithersburg, MD.

Kravtsov A. V., 2013, ApJL, 764, L31

Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353

Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221

Kroupa P., 2001, MNRAS, 322, 231

Lan T.-W., 2020, ApJ, 897, 97

Lan T.-W., Fukugita M., 2017, ApJ, 850, 156

Lan T.-W., Mo H., 2018, ApJ, 866, 36

Lan T.-W., Ménard B., Zhu G., 2014, ApJ, 795, 31

Landoni M., Falomo R., Treves A., Scarpa R., Farina E. P., 2016, MNRAS, 457, 267

Larson R. B., 1976, MNRAS, 176, 31

Lee D. D., Seung H. S., 1999, Nature, 401, 788

Lee J. C., Hwang H. S., Song H., 2021, MNRAS, 503, 4309

Lehner N., O'Meara J. M., Fox A. J., Howk J. C., Prochaska J. X., Burns V., Armstrong A. A., 2014, ApJ, 788, 119

Lehner N., Howk J. C., Wakker B. P., 2015, ApJ, 804, 79

Lifshitz E. M., 1946, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki, 16, 587

Liu A., et al., 2021, arXiv e-prints, p. arXiv:2106.14518

Lochhaas C., Tumlinson J., O'Shea B. W., Peeples M. S., Smith B. D., Werk J. K., Augustin R., Simons R. C., 2021, ApJ, 922, 121

Lopez S., et al., 2008, ApJ, 679, 1144

Lopez S., et al., 2018, Nature, 554, 493

Lovegrove E., Simcoe R. A., 2011, ApJ, 740, 30

Lundgren B. F., et al., 2009, ApJ, 698, 819

Lyke B. W., et al., 2020, ApJS, 250, 8

Madau P., Dickinson M., 2014, ARA&A, 52, 415

Marasco A., Crain R. A., Schaye J., Bahé Y. M., van der Hulst T., Theuns T., Bower R. G., 2016, MNRAS, 461, 2630

Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785

Matejek M. S., Simcoe R. A., 2012, ApJ, 761, 112

McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593

McCourt M., Sharma P., Quataert E., Parrish I. J., 2012, MNRAS, 419, 3319

McGaugh S. S., 2008, in Davies J. I., Disney M. J., eds, Proceedings of the International Astronomical Union, IAU Symposium Vol. 244, Dark Galaxies and Lost Baryons. pp 136–145 (arXiv:0707.3795), doi:10.1017/S1743921307013920

Ménard B., Wild V., Nestor D., Quider A., Zibetti S., Rao S., Turnshek D., 2011, MNRAS, 417, 801

Mihos J. C., Keating K. M., Holley-Bockelmann K., Pisano D. J., Kassim N. E., 2012, ApJ, 761, 186

Mishra S., Muzahid S., 2022, arXiv e-prints, p. arXiv:2201.08545

Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020, MNRAS, 494, 3971

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press

More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

Moster B. P., Naab T., White S. D. M., 2018, MNRAS, 477, 1822

Muratov A. L., et al., 2017, MNRAS, 468, 4170

Muzahid S., Charlton J., Nagai D., Schaye J., Srianand R., 2017, ApJL, 846, L8

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59

Narayanan A., Misawa T., Charlton J. C., Kim T.-S., 2007, ApJ, 660, 1093

Navarro J. F., White S. D. M., 1994, MNRAS, 267, 401

Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353

Nelson D., et al., 2019, MNRAS, 490, 3234

Nelson D., et al., 2020, MNRAS, 498, 2391

Nelson D., Byrohl C., Peroux C., Rubin K. H. R., Burchett J. N., 2021, MNRAS, 507, 4445

Nestor D. B., Turnshek D. A., Rao S. M., 2005, ApJ, 628, 637

Nestor D., Hamann F., Rodriguez Hidalgo P., 2008, MNRAS, 386, 2055

Nicastro F., et al., 2018, Nature, 558, 406

Nielsen N. M., Churchill C. W., Kacprzak G. G., Murphy M. T., 2013, ApJ, 776, 114

Nielsen N. M., Churchill C. W., Kacprzak G. G., Murphy M. T., Evans J. L., 2015, ApJ, 812, 83

Nielsen N. M., Churchill C. W., Kacprzak G. G., Murphy M. T., Evans J. L., 2016, ApJ, 818, 171

Nielsen N. M., Kacprzak G. G., Pointon S. K., Churchill C. W., Murphy M. T., 2018, ApJ, 869, 153

Nielsen N. M., Kacprzak G. G., Pointon S. K., Murphy M. T., Churchill C. W., Davé R., 2020, ApJ, 904, 164

O'Sullivan E., Kolokythas K., Kantharia N. G., Raychaudhury S., David L. P., Vrtilek J. M., 2018, MNRAS, 473, 5248

Olivares V., et al., 2019, A&A, 631, A22

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

Oppenheimer B. D., et al., 2016, MNRAS, 460, 2157

Oppenheimer B. D., et al., 2020, ApJL, 893, L24

Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJL, 193, L1

Padilla N., Lacerna I., Lopez S., Barrientos L. F., Lira P., Andrews H., Tejos N., 2009, MNRAS, 395, 1135

Padmanabhan N., Xu X., Eisenstein D. J., Scalzo R., Cuesta A. J., Mehta K. T., Kazin E., 2012, MNRAS, 427, 2132

Pâris I., et al., 2017, A&A, 597, A79

Pâris I., et al., 2018, A&A, 613, A51

Pearson S., et al., 2016, MNRAS, 459, 1827

Peebles P. J. E., 1965, ApJ, 142, 1317

Peebles P. J. E., 1982, ApJL, 263, L1

Peebles P. J., Ratra B., 2003, Reviews of Modern Physics, 75, 559

Peebles P. J. E., Yu J. T., 1970, ApJ, 162, 815

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

Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, ApJ, 786, 54

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Pérez-Ràfols I., Miralda-Escudé J., Lundgren B., Ge J., Petitjean P., Schneider D. P., York D. G., Weaver B. A., 2015, MNRAS, 447, 2784

Perlmutter S., et al., 1999, ApJ, 517, 565

Péroux C., Howk J. C., 2020, ARA&A, 58, 363

Péroux C., Petitjean P., Aracil B., Irwin M., McMahon R. G., 2004, A&A, 417, 443

Péroux C., Nelson D., van de Voort F., Pillepich A., Marinacci F., Vogelsberger M., Hernquist L., 2020a, MNRAS, 499, 2462

Péroux C., Nelson D., van de Voort F., Pillepich A., Marinacci F., Vogelsberger M., Hernquist L., 2020b, MNRAS, 499, 2462

Perrotta S., Hamann F., Zakamska N. L., Alexand roff R. M., Rupke D., Wylezalek D., 2019, MNRAS, 488, 4126

Pieri M. M., et al., 2014, MNRAS, 441, 1718

Pillepich A., et al., 2018, MNRAS, 473, 4077

Pillepich A., et al., 2019, MNRAS, 490, 3196

Planck Collaboration et al., 2020, A&A, 641, A6

Pointon S. K., Nielsen N. M., Kacprzak G. G., Muzahid S., Churchill C. W., Charlton J. C., 2017, ApJ, 844, 23

Pradeep J., Narayanan A., Muzahid S., Nagai D., Charlton J. C., Srianand R., 2019, MNRAS, 488, 5327

Press W. H., Schechter P., 1974, ApJ, 187, 425

Prochaska J. X., Kasen D., Rubin K., 2011, ApJ, 734, 24

Prochaska J. X., Lau M. W., Hennawi J. F., 2014, ApJ, 796, 140

Prochter G. E., Prochaska J. X., Burles S. M., 2006, ApJ, 639, 766

Puech M., et al., 2018, in Evans C. J., Simard L., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 10702, Ground-based and Airborne Instrumentation for Astronomy VII. p. 107028R (`arXiv:1806.03296`), doi:10.1117/12.2311292

Qiu Y., Bogdanović T., Li Y., McDonald M., McNamara B. R., 2020, Nature Astronomy, 4, 900

Quider A. M., Nestor D. B., Turnshek D. A., Rao S. M., Monier E. M., Weyant A. N., Busche J. R., 2011, AJ, 141, 137

Raichoor A., et al., 2017, MNRAS, 471, 3955

Raichoor A., et al., 2021, MNRAS, 500, 3254

Reimers D., 2002, Space Science Reviews, 100, 89

Reiprich T. H., Basu K., Ettori S., Israel H., Lovisari L., Molendi S., Pointecouteau E., Roncarelli M., 2013, Space Science Reviews, 177, 195

Riess A. G., et al., 1998, AJ, 116, 1009

Rubin V. C., Ford W. Kent J., 1970, ApJ, 159, 379

Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, ApJ, 794, 156

Rubin K. H. R., Diamond-Stanic A. M., Coil A. L., Crighton N. H. M., Moustakas J., 2018, ApJ, 853, 95

Rudie G. C., et al., 2012, ApJ, 750, 67

Rykoff E. S., et al., 2014, ApJ, 785, 104

Rykoff E. S., et al., 2016, ApJS, 224, 1

Sargent W. L. W., Steidel C. C., Boksenberg A., 1988, ApJ, 334, 22

Schaye J., et al., 2015, MNRAS, 446, 521

Schechter P., 1976, ApJ, 203, 297

Schiminovich D., et al., 2010, MNRAS, 408, 919

Schneider D. P., et al., 2005, AJ, 130, 367

Schneider D. P., et al., 2010, AJ, 139, 2360

Schneider E. E., Robertson B. E., Thompson T. A., 2018, ApJ, 862, 56

Schroetter I., et al., 2021, MNRAS, 506, 1355

Sharma M., Nath B. B., 2013, ApJ, 763, 17

Sharma P., McCourt M., Quataert E., Parrish I. J., 2012, MNRAS, 420, 3174

Silk J., 1968, ApJ, 151, 459

Simcoe R. A., et al., 2011, ApJ, 743, 21

Smoot G. F., et al., 1992, ApJL, 396, L1

Somerville R. S., Davé R., 2015, ARA&A, 53, 51

Songaila A., 2001, ApJl, 561, L153

Songaila A., 2005, AJ, 130, 1996

Springel V., et al., 2005, Nature, 435, 629

Starobinsky A. A., 1982, Physics Letters B, 117, 175

Steidel C. C., Sargent W. L. W., 1992, ApJS, 80, 1

Steidel C. C., Dickinson M., Persson S. E., 1994, ApJL, 437, L75

Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289

Stone R. B., Richards G. T., 2019, MNRAS, 488, 5916

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Suresh J., Rubin K. H. R., Kannan R., Werk J. K., Hernquist L., Vogelsberger M., 2017, MNRAS, 465, 2966

Tamura N., et al., 2016, in Ground-based and Airborne Instrumentation for Astronomy VI. p. 99081M (arXiv:1608.01075), doi:10.1117/12.2232103

Tanimura H., et al., 2019, MNRAS, 483, 223

Tchernyshyov K., et al., 2022, ApJ, 927, 147

The LSST Collaboration et al., 2018, arXiv e-prints, p. arXiv:1809.01669

Toomre A., Toomre J., 1972, ApJ, 178, 623

Tremonti C. A., Moustakas J., Diamond-Stanic A. a. M., 2007, ApJL, 663, L77

Truong N., et al., 2020, MNRAS, 494, 549

Truong N., Pillepich A., Nelson D., Werner N., Hernquist L., 2021, MNRAS, 508, 1563

Tumlinson J., et al., 2011, Science, 334, 948

Tumlinson J., et al., 2013, ApJ, 777, 59

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

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

Tytler D., Boksenberg A., Sargent W. L. W., Young P., Kunth D., 1987, ApJS, 64, 667

Vanden Berk D. E., et al., 2001, AJ, 122, 549

Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769

Vikas S., et al., 2013, ApJ, 768, 38

Vogelsberger M., et al., 2014, Nature, 509, 177

Voit G. M., 2018, ApJ, 868, 102

Voit G. M., Donahue M., 2015, ApJL, 799, L1

Voit G. M., Meece G., Li Y., O'Shea B. W., Bryan G. L., Donahue M., 2017, ApJ, 845, 80

Vollmer B., et al., 2012, A&A, 537, A143

Walters D., Woo J., Ellison S. L., 2022, MNRAS, 511, 6126

Wang B., 1993, ApJ, 415, 174

Wen Z. L., Han J. L., 2015, ApJ, 807, 178

Werk J. K., Prochaska J. X., Thom C., Tumlinson J., Tripp T. M., O'Meara J. M., Meiring J. D., 2012, ApJS, 198, 3

Weymann R. J., Williams R. E., Peterson B. M., Turnshek D. A., 1979, ApJ, 234, 33

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJL, 754, L29

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

White M., et al., 2011, ApJ, 728, 126

Wisotzki L., et al., 2018, Nature, 562, 229

Yang H. Y. K., Reynolds C. S., 2016, ApJ, 829, 90

Yoon J. H., Putman M. E., 2017, ApJ, 839, 117

York D. G., et al., 2000, AJ, 120, 1579

York D. G., et al., 2006, MNRAS, 367, 945

Yun K., et al., 2019, MNRAS, 483, 1042

Zahedy F. S., Chen H.-W., Johnson S. D., Pierce R. M., Rauch M., Huang Y.-H., Weiner B. J., Gauthier J.-R., 2019, MNRAS, 484, 2257

Zel'dovich Y. B., 1970, A&A, 5, 84

Zheng Y., Peek J. E. G., Putman M. E., Werk J. K., 2019, ApJ, 871, 35

Zhu G., 2016, arXiv e-prints, p. arXiv:1612.06037

Zhu G., Ménard B., 2013a, ApJ, 770, 130

Zhu G., Ménard B., 2013b, ApJ, 773, 16

Zhu G., et al., 2014, MNRAS, 439, 3139

Zhu G. B., et al., 2015, ApJ, 815, 48

Zibetti S., Ménard B., Nestor D., Turnshek D., 2005, ApJL, 631, L105

Zibetti S., Ménard B., Nestor D. B., Quider A. M., Rao S. M., Turnshek D. A., 2007, ApJ, 658, 161

Zou H., Gao J., Zhou X., Kong X., 2019, ApJS, 242, 8

Zou H., et al., 2021, ApJS, 253, 56

Zu Y., 2021, MNRAS, 506, 115

Zuhone J., Markevitch M., 2009, in Heinz S., Wilcots E., eds, American Institute of Physics Conference Series Vol. 1201, The Monster's Fiery Breath: Feedback in Galaxies, Groups, and Clusters. pp 383–386 (arXiv:0909.0560), doi:10.1063/1.3293082

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2019, A&A, 624, A48

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

Acknowledgements

Finally, after a tedious write-up of thesis chapters, it's the time to extend my gratitude and thanks to all those who have held my hand and didn't let me fall or sink. Of course, I'll miss many, but everyone has played their role. I don't know where to start, so I'll go ahead with my supervisors.

First and foremost, I would thank Guinevere, who provided me with the opportunity to join her group. Throughout my four years at MPA, she has mentored and supported me in every possible way. Starting from the first project, she has supported my ideas and plans and helped me whenever I needed it. She has been a constant support during the whole terrible pandemic phase and sought over the progress. She always encouraged me to participate in conferences, workshops and schools, which helped me meet new people. So a huge thanks to her!

Dylan, I am not sure if I would be able to thank you enough. I could not have expected more from a co-mentor. This thesis would not have been possible without your constant support, thanks to you. Starting from my arrival at MPA, you taught me so many things and provided invaluable suggestions that have shaped me. Whether it's some code in Python, creating nice plots or understanding some basic concepts, you were always there. Even after you moved to Heidelberg, you continued guiding me. You are an inspiring advisor. I wish you a healthy and very successful life.

Next, many thanks to the secretaries, Gabi, Maria, Cornelia, Sonja, Sonia, and Annette, for taking care of all the paperwork needed to let things flow smoothly during PhD years. In particular, I would like to extend a special thanks to Gabi. She went beyond her duties to help me settle down in a very different country with all the proper documents. She is always ready to help with visa issues or other logistic or administrative issues. She is one of the most helpful persons I've ever met. I wish her a healthy and happy life. Another person is Maria, who has helped almost every time with translating and filling the business and vacation forms and answering any doubt related to the administration. I wish her a very healthy life too.

Now, I would like to thank the members of my thesis committee at MPA (apart from my supervisors), Eiichiro, Benedetta and Volker. They have supported my extension of contract due to the pandemic and gave some valuable suggestions during the meetings. Many thanks to Klaus Dolag, Ariel Sánchez, Jan von Delft and Jörg Schreiber at LMU for agreeing to be on my defense committee. A huge thanks to Simon, a fantastic scientist and a wonderful person. I should also thank Andi and Goran for helping me with computers.

Thanks to Fabrizio for organizing group meetings and encouraging us to talk, discuss and learn many exciting things. Many thanks to Max and Celine for insightful discussion over coffee. Thanks to the people who organized weekly galaxy group meeting. It's an excellent place to present exciting research and get feedback from people. Many thanks to Anna and Anja for organizing German classes, though I still get scared of 'die, der, das', but it was fun learning a new language. I'm still a beginner after all these years :)

It's time to thank my colleagues at MPA. First is Reza, my senior in the group. He has been a great friend and always helped with my questions or doubts (sometimes even silly ones), either personal or professional. I have learned so much from him and wish him a great future. Periklis, the best office mate in world. He was always there, whether it was about exploring the Bavarian countryside, movie nights in Munich, exploring new restaurants, or discussing astrophysics. Next is Ilham; he is like my brother in Munich. I can not forget talking to him for hours in my apartment, going to McDonald's on Fridays night and listening to his music. I don't know how many times I forgot my apartment keys and went to him for extra ones, even in weird hours :) I learned a beautiful lesson from him; to stay calm and relaxed even in the worst times. Thanks, Ilham! Next is Spandan; you're truly a great friend; thanks for all the memories together. We four (Periklis, Ilkham, Spandan and Me) call ourselves, *the fantastic four*, and it suits us well. I hope we will explore more of Europe that we missed because of covid. Finally, I would like to thank my other colleagues at MPA: Jessie, Aniket, Jere, Ildar, Pavan, Wolfgang, Abinaya, Daniela, Nahir, Vale, Hitesh, Maja, Constantina, Laura, Deepika, Anshuman and many others who have always been nice to me.

A special thanks to Naveen. He's like my elder brother. One of the wisest men I have ever met. He is full of knowledge, calm and always willing to help. I can not forget going to his apartment, having some delicious food (he's damn good at cooking) and talking on almost all kinds of topics. I have learned so many things from him. I wish him a healthy and successful life. Another person is Avinash at ESO, who has been so nice to me. I wish him great success in life. A special thanks to Laura, who helped me with the translation of the abstract to German.

I am also very grateful to some of my teachers and professors. First is KN Pandey sir, at St. Xavier's Patna, who encouraged me to explore this beautiful subject. I am so lucky that I became part of that prestigious school. I am grateful to all my teachers, Tavishi Ma'am, Jyoti Ma'am, Zahidi Sir, Sudhir Sir, Gracy and Helen Ma'am. Next is Chandrashekhar Sharma at Resonance, Kota, who further cultivated my interest. I am also very thankful to Mukul Sir and Parmar Sir in Patna. Many thanks to my friends from school: Anand, Ashok, Ved, Abhijay, Aditya, Shwetank. Special thanks to my friends in Kota: Amit (Choubey and Ojha), Jyoti, and Akash.

Next, at IISc, I have had the chance to meet some great professors who taught me many things and cultivated my interest in pursuing a research career. A few of them are: Arnab Rai Choudhuri, Late V. Venkatraman, Subroto Mukerjee, Prateek Sharma, Nirupam Roy, Chandan Dasgupta, HR Krishnamurthi, Diptiman Sen, Manjunath Krishnapur, Raghavendra Gadagkar and Kaushal Verma. A special thanks to Nirupam Roy, who gave me an opportunity to explore radio astronomy during my MS project and helped me publish my first research paper. Thanks to Saurabh Sharma at ARIES, Nainital, who introduced me to observational astronomy through a summer project. I also thank Gajendra Pandey at IIA Bangalore for allowing me to work on some exciting topics in stellar physics. I am also very thankful to Prof. Janaki Balakrishnan, Sudharshana ma'am, Sourina, Avinash, Nasima, Anjani, Sriram and Aparupa at NIAS for making my time full of good memories.

There are a few more people who have a special place in my heart: my friends from IISc, Abinash, Ankit, Bhavna, Samprita, Ojesh, Shayan, Hamilton, Shubham, Abhishek, Ankush, Abhinay, Saksham, Japsal, Shikhar, Gagan, Arvind, Aravind, Sabareesh, Tapan, Avinash, Harshit, Siddharth, Kishalay, Rahul, Irfan, Sudeb, Atanu, Tushar, Prasanjit and Kaushik. It was one of the most memorable times of my life. Thank you all. A special thanks to Abinash and Shayan, whom I have been in constant touch with throughout my PhD. It was relaxing to chat with you guys over the phone for hours and discuss everything about life and the world. I wish you all great success and health.

Finally, my deepest gratitude and thanks to my family, who has been my strength throughout my life. I have dedicated this thesis to my grandparents, who have always showered their blessings on me. I will begin with my mother, Bandita. Maa, you taught me not to lose patience and keep working for the best. Next is my father (Ajit); thank you so much, Papa, for always encouraging me to do what I liked and never letting me worried about anything else. My elder brother, Premjeet, bhaiya, you are my constant support. Just because of you, I am here writing my thesis. Whenever I need you, you are always there. Thank you so much! My elder sister, Richa, Didi, I don't know how to thank you, you're like my guardian, always with me. Thank you so much! I am so lucky to have two new members in my family now, Niraj (my sister's husband) and Richa (my brother's wife). You two have showered all the love and brought happiness to our lives. My first ever visit to California became memorable because of my sister and her husband and their friends (Dheeraj, Rasha (in fact, these two wanted to know what I was doing :)) and Gaurav). Many thanks to Purnima (mother's sister), Sanjib (her husband), and my two little sisters, Sakshi and Mahi, for their love and support. Thank you so much, everyone, and I wish you all a pleasant and healthy life!