THE UNIVERSITY *of* EDINBURGH
School of Physics & Astronomy

Observational Signatures of the Gas In and Around Galaxies

MPhys Project Report

Evan Jones

Submitted for the 40pt MPhys Project course PHYS11016

27th March 2023

Abstract

Observational censuses of galaxies and the universe do not find the amount of baryons predicted. One suggested reservoir is the circumgalactic medium, the diffuse gas bound to a galaxy that exists up to several hundred kiloparsecs from the centre. In this work we introduce a new 4D ionization table for use in post-processing simulations that accounts for self-shielding at non-solar metallicities, an important effect in the dense clouds of the circumgalactic medium. By studying low ions we determine that the assumption of solar metallicity has little effect of the total mass of gas, but can underestimate the number density by as much as 120%, especially of ions tracing hotter, less-enriched gas. Further work is necessary to quantify the effect on observations and estimates of mass in galaxies.

Supervisor: Dr Britton Smith

Personal Statement

For the first three weeks of the project I focused on taking raw ASCII ionization fraction data and converting them into one new 4D table, stored as an HDF5 file. Over the next two weeks I replaced any missing values by linearly interpolating over density. In these first five weeks I was reading up on the background, starting with some papers suggested by my supervisor.

The next two weeks consisted of updating the existing TRIDENT and YT libraries to include the ability to read in the 4D ionization table and interpolate over it. This included writing a 4D linear interpolation function in CYTHON and adding a new class to the YT library to interface between this and simulation data.

The final weeks of semester 1 I spent making sure my new code worked on cuillin, and started looking at data from a large low-resolution data set. For this I used the table I created and compared it to existing tables, which assume no self-shielding, and self-shielding with solar metallicity. This included directly comparing the tables, and comparing how properties of the simulated gas changed.

Over the Christmas break I noticed that there were issues with certain values being extrapolated to massively unphysical values, reaching infinity in some cases. I spent a few weeks, including the first two weeks of semester 2, identifying the cause of the problem and introducing a fix.

The majority of semester 2 was spent applying the tables to various high-resolution simulations of galaxies, calculating how physical properties and observable quantities change when accounting for self-shielding and metallicity, and also if it is possible to constrain these changes observationally.

The final two weeks were dedicated to preparing the report, presentation, and public summary. Unfortunately the Institute for Astronomy's computing cluster cuillin was shut down for several days before the deadline, so last minute tweaks were only possible with data that had been saved to my personal laptop, as opposed to the raw data. This limited my ability to make all the plots for this report to the standard I wanted.

Acknowledgements

I would like to thank my supervisor, Britton Smith, for his advice and help throughout the project. I would also like to thank my family and friends for supporting me throughout my time at University. This would not have been possible without them.

Contents

1 Introduction

The amount of visible baryonic matter in the local universe is different to the amount of baryonic matter in the early universe (e.g., Persic and Salucci 1992; Fukugita et al. 1998). This is called the *cosmic missing baryon problem*: when performing a census of the different phases of baryonic matter we find that it does not total to the expected value, by as much as $\sim 30\%$ (Shull et al. 2012). This observational mismatch is also true in galaxies, where the observed baryons in a galactic halo is a small percentage of what is expected when considering the cosmic ratio of baryons to dark matter—as little as 1% for satellite dwarf galaxies (McGaugh et al. 2010). This is known as the *galactic missing baryon problem*. Identifying the form this matter takes is important as it is a test of cosmological and galaxy evolution.

Over the last decade more research has been focused on the circumgalactic medium (CGM; e.g., Anderson and Bregman 2010; Tumlinson et al. 2011; Tumlinson et al. 2013)—the diffuse gas outside the visible extent of a galaxy—as a potential reservoir for these baryons. As observations are challenging, simulations provide valuable insight as to the structure of the CGM (e.g., Peeples et al. 2019; Appleby et al. 2023).

When increasing the resolution of simulations we see that the CGM consists of cooler and denser clouds than previously thought (Hummels et al. 2019; Peeples et al. 2019). These clouds are possibly no longer optically thin, as is typically assumed, but are instead self-shielded: the outside of the cloud stops radiation from penetrating to the inside. Self-shielding depends on the physical properties of the gas, such as the metallicity, which has been assumed to be solar when modelling self-shielding (e.g., in Zheng et al. 2020) but observationally spans several orders of magnitude (Prochaska et al. 2017).

The aim of this project was to investigate the effects of self-shielding and metallicity on the physical properties of the simulated circumgalactic medium, and if it is possible to differentiate between the different models observationally.

This was done by first assembling an ionization table that accounts for self-shielding and non-solar metallicities, and altering the existing YT[1] (Turk et al. 2011) and TRIDENT[2] (Hummels et al. 2017) libraries to use it to post-process simulations. We analysed the effect of including self-shielding with solar metallicity assumed, and self-shielding with varying metallicity on a large cosmological simulation of the warm-hot intergalactic medium at $z = 0$, and six simulated galaxies at both $z = 0$ and $z = 2$ to study the circumgalactic medium.

The report is laid out as follows. Section 2 provides the theoretical background to the project, introducing the missing baryon problem, the circumgalactic medium, and a brief overview of cosmological simulations. Section 3 introduces the creation of the new ionization table; how existing libraries were enhanced to include it; and a qualitative discussion on the effects of varying metallicity. Section 4 covers the application of the new table to simulation data

[1] https://yt-project.org/
[2] http://trident-project.org

of the warm-hot intergalactic medium and six simulated galaxies. Finally, we conclude in Section 5.

All simulations use flat ΛCDM cosmologies, but with different values of $\Omega_\Lambda = 1 - \Omega_m$ and H_0 that will be introduced alongside the simulations. Unless otherwise specified, all distances are comoving.

2 Background

2.1 The Missing Baryon Problem

Collapsed objects like stars, galaxies, and clusters only make up $\sim 10\%$ of the baryonic matter in the low-redshift universe (e.g. Persic and Salucci 1992; Fukugita and Peebles 2004), with the rest distributed in the various phases of the gas in-between and around galaxies. These are called the intergalactic medium (IGM) and circumgalactic medium (CGM) respectively. How the gas is distributed within these is uncertain, partly because of the lack of clear boundaries between phases and their definitions, and partly because observing the hottest and least dense of these phases is challenging (e.g., Sołtan 2006).

The IGM is typically divided into a warm ($T \lesssim 10^5$ K) and a warm-hot (10^5 K $\lesssim T \lesssim 10^7$ K) component. The warm IGM is detected by absorption in the Lyman-α forest, and makes up $\sim 30\%$ of baryons (Danforth and Shull 2008). The Warm-Hot Intergalactic Medium (WHIM) is predicted to make up an additional 25%, leaving $\sim 30\%$ of baryons 'missing'. The success of the ΛCDM cosmological model, agreement between different methods of predicting the baryon density (e.g., Pitrou et al. 2018; Planck Collaboration et al. 2020), cosmological simulations (e.g., Davé et al. 2001; Smith et al. 2011), and the Lyman-α forest at high ($z \gtrsim 3$) redshift (e.g., Fukugita et al. 1998) imply that the issue lies with low-redshift observational censuses being incomplete.

The WHIM is challenging to observe in emission because of the neutral hydrogen in the Milky Way efficiently absorbs these high-energy photons, and in absorption because it is primarily ionized hydrogen, so there are no line features to observe. However at these temperatures helium- and lithium-like metals (O VI, O VII, ...) have absorption lines in the UV, observable by the Hubble Space Telescope's Cosmic Origins Spectrograph, and X-ray, observable by XMM-Newton and Chandra. For the 10^5 K $< T < 10^6$ K WHIM, broad Lyman-α systems from the small amounts of neutral H I are also observable, but metal lines are the most important method of detection.

For $T > 10^6$ K, it is necessary to use X-ray observations, but looking for the absorption lines of small amounts of metals with our current instrumentation is difficult, and reported detections are often controversial. For example, Johnson et al. (2019) dispute the detections reported by Nicastro et al. (2018), that would have otherwise closed the budget.

New techniques designed to avoid these problems are also being proposed. Macquart et al. (2020) used Fast Radio Bursts to find a baryon density consistent with results from Big Bang

Nucleosynthesis (Pitrou et al. 2018) and the Cosmic Microwave Background (Planck Collaboration et al. 2020), but with substantial errors—ranging from 50–160% of these values.

There is a similar problem on the scale of galaxies. The *galactic missing baryon problem* refers to the fact that the observed ratio of baryonic to dark matter in galaxies is far below that of the cosmic ratio (McGaugh et al. 2010), with two suggested solutions: the baryons were ejected early on in galaxy formation, and hence have never been gravitationally bound to galaxies (Anderson and Bregman 2010); or they are just as of yet unobserved and are in a large halo around galaxies (e.g., Werk et al. 2014).

2.2 The Circumgalactic Medium

The CGM, or circumgalactic medium, is the diffuse gas that is gravitationally bound to a galaxy but lies outside the stellar extent. The existence of a hot, diffuse medium surrounding galaxies was first suggested by Spitzer (1956) and observed via absorption lines in quasar spectra in the 1960s (Bahcall and Spitzer 1969, and references therein), but it wasn't until the late 1990s that there was strong evidence of this being a generic feature of galaxies (Lanzetta et al. 1995). More recently, large spectroscopic surveys of galaxies such as COS-Halos (Tumlinson et al. 2013; Werk et al. 2013) and COS-Dwarfs (Bordoloi et al. 2014) find large quantities of cool, ionized, metal-rich gas that covers a range of densities extending out past $\sim 100\,\mathrm{kpc}$ from the centres of galaxies. The boundary of the circumgalactic medium is poorly defined, but is usually taken to be the virial radius, R_{200}, defined as the radius at which the mean density is 200 times the critical density (Shull 2014).

Simulations of the CGM have had a corresponding increase in attention, with recent research increasing the resolution of simulations to more accurately capture the physical properties of it (Hummels et al. 2019; Peeples et al. 2019). These studies show that it has traditionally been under-resolved, and has a much more varied structure of small cool and dense clouds in a halo of hot gas.

These dense clouds have typically been assumed to be optically thin radiation, but are dense enough that self-shielding begins to become important. Self-shielding is a phenomenon where a cloud of gas is optically dense enough that radiation cannot penetrate into the centre, only ionizing a layer on the outside—there is an outer shell shielding the rest of the gas from ionizing radiation.

2.3 Cosmological Simulations

Given the vast range of scales and complexity of the processes involved in galaxy formation and evolution, it is impossible to write down a complete analytical description. Often the only way to study the effect of different processes is to simulate them. In theory, this is very simple; write down a list of equations to approximate numerically. This is what simulations aim to do, but unfortunately there is a major trade-off between accuracy and speed, so it is

necessary to make decisions about what is important to simulate and what can be ignored or approximated.

In this report we use snapshots from the Eulerian+N-body code Enzo[3] (Bryan et al. 2014; Brummel-Smith et al. 2019), where a snapshot is the state of the simulation at a specific time. 'Eulerian' refers to the fact that Enzo evolves the properties of gas occupying cells in a grid, as opposed to a Lagrangian approach of considering parcels of gas moving through space. 'N-body' refers to how it models dark matter and stars as an arbitrary number (N) of particles (bodies) in space. Note these are not individual dark matter particles and stars, but rather a representation of an amount of dark matter or a population of stars.

The equations of gravity and hydrodynamics are solved for the particles and grid cells, but the scales of other processes that need to be modelled, such as star formation and feedback, are often much smaller than a grid cell with sides several hundred to thousands of parsecs long. These processes are collectively referred to as *sub-grid physics*, and simulations rely on simple or empirical models to account for their effects. For example, a star particle represents a population of stars, and is formed when the gas in a cell obeys certain criteria. It then follows a simple model of converting gas mass into stellar mass and returning enriched gas and energy into the region surrounding the particle, as a group of stars would.

One of the main benefits of using grid-based codes is the ability to use Adaptive Mesh Refinement (AMR), where a grid cell can be divided into smaller cells (*refined*) to increase the resolution where it is needed for dense and/or cool gas, and reduced when lower resolution suffices. This allows the majority of resources to be applied to dense regions, while diffuse regions are given larger cells that do not over-resolve the scale of the system. This enables large simulations that capture both cosmological and galactic evolution to efficiently cover large regions of space.

However, as mentioned above, this can lead to under-resolving areas of potential interest such as the circumgalactic medium (Hummels et al. 2019; Peeples et al. 2019). Zoom simulations (also called zoom-in simulations or zoom re-simulations; e.g., Hahn and Abel 2011) are one method to capture the effects of large-scale structure on a galaxy, while also resolving the smaller scales of interest. First, a large-scale low–moderate resolution simulation is run, and potential galaxies (or any other structures) of interest (for example, halos resembling the Milky Way; e.g., Peeples et al. 2019) are identified. The simulation is then run for a second time, but with increased resolution in the region surrounding the galaxy. The galaxies used in this report were simulated using this method (Simons et al. 2020).

2.3.1 Post-processing

One way to balance detail with computational efficiency in simulations is *post-processing*, where quantities that would be too expensive to evolve in time are ignored or accounted for in some other way in the simulation, then are calculated from the simulation output.

[3] https://enzo-project.org

The most common example is simulating networks of different ions. Increasing the number of ions studied exponentially increases the number of equations that need to be solved as you consider how electrons and ions recombine or are ionized to higher states. While ions of interest can be explicitly tracked (e.g., Cen and Fang 2006) simulations tend to track only the primordial elements hydrogen and helium, the electron population, and the metallicity. Metallicity is defined as the ratio of mass in metals (any element heavier than helium) to total mass. The metal cooling rates are pre-computed for a range of density, temperature, and metallicity values and saved as a table (initially described in Smith et al. 2008; built upon in Smith et al. 2011; Smith et al. 2017) that is used by the simulation. This enables more accurate cooling rates, without the increased time cost otherwise needed, but means the final simulation product lacks the details of the state metals are in. This poses a problem, as metals are the main source of absorption lines considered observationally.

Quantities such as number density, density, and mass of metal ions can be calculated using codes like TRIDENT, a package for generating synthetic spectra. Synthetic spectra are spectra derived from simulated data, modelling the absorption and emission features of different elements in simulated gas, and while beyond the scope of this project are mentioned as they are an important tool for comparing simulations to observations of galaxies. Mass and mass density are simple to calculate from number density, which is more complex and requires assumptions about the abundances of elements and the physics involved.

Assuming solar abundances, the number density of an element X can be calculated from the hydrogen number density, n_{H}, and metallicity, Z, both tracked by the simulation, using the equation

$$n_X = n_{\mathrm{H}} Z \left(\frac{n_X}{n_{\mathrm{H}}} \right)_{\odot}, \tag{1}$$

where $(n_X/n_{\mathrm{H}})_{\odot}$ represent the solar abundance of element X (extracted from the CLOUDY[4] documentation: Ferland et al. 1998; Ferland et al. 2017). To convert this to the number density of a specific ion X_i, multiply by the *ionization fraction* f_{X_i}—the proportion of the element X that is in its i-th ionization state.

2.3.2 Ionization Fractions and Tables

The ionization fraction of an ion can be calculated as a function of density, temperature, and incident radiation under the assumption of ionization equilibrium. The ionizing radiation is assumed to be a uniform metagalactic ultraviolet background (UVB), like those detailed in Faucher-Giguère et al. (2009) or Haardt and Madau (2012). Generally, this background radiation comes from high-mass stars, so increases with star formation rate to a peak at redshift $z \sim 2$ (Madau and Dickinson 2014), where it begins to decline again. This means the intensity can be written as a function of redshift, which translates between time in cosmological simulations and the flux used to model ionization. As such, redshift is often used as a proxy for ionizing radiation, and referring to redshift is shorthand for saying this.

[4] https://nublado.org

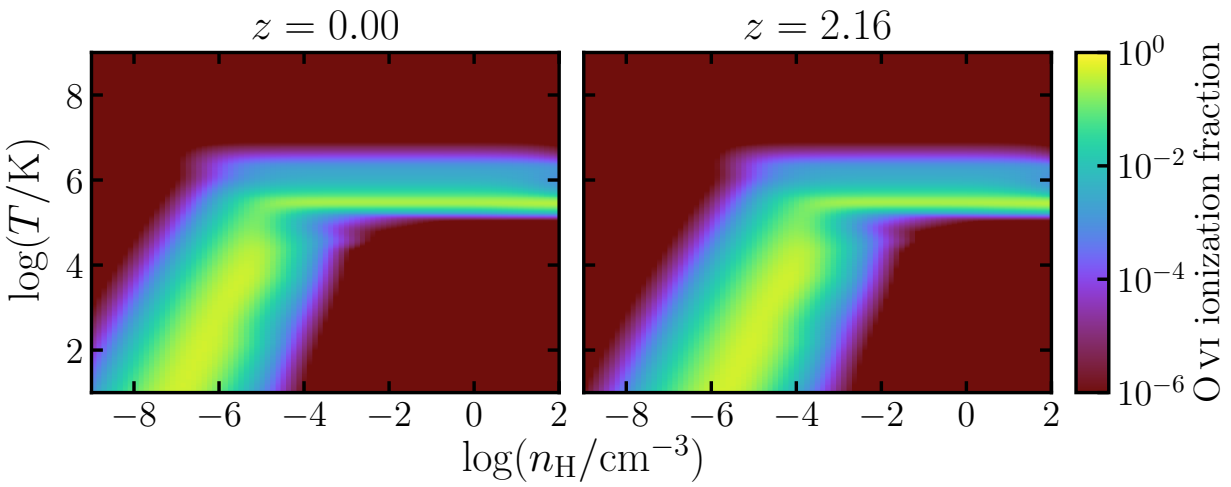

Figure 1. Ionization fractions at solar metallicity for redshifts $z = 0$ and $z \sim 2$. Redshift 2 is when star formation rate, and hence the ultraviolet background, peaks. For a given temperature and density, increasing the ionizing radiation will promote ions into a higher state, which has the effect of increasing the density that a given ion peaks at, moving the 'tail' to the right. Alternatively, it can be described as increasing radiation shifts the point where collisional ionization dominates over photoionization (the knee/kink in the plot) to higher densities.

Figure 1 plots the ionization fraction of an ion, in this case O VI, as a function of hydrogen number density, n_H, and temperature, T, of a cloud of gas as it is exposed to radiation. This is a slice through an ionization table at a given redshift and metallicity. At higher densities the gas is in collisional excitation equilibrium because it is self-shielded from the incident radiation, and ionization fraction is solely a function of temperature. As density decreases the effects of photoionization become more important and what would otherwise be O VI is ionized to higher states (O VII, O VIII, O IX). The position of this photoionization 'tail' is therefore a function of incident radiation as it changes the balance between photo- and collisional ionization. This is demonstrated in the change from the left plot to the right plot in Figure 1, corresponding to increasing the intensity of the metagalactic ultraviolet background (UVB), which decreases with redshift from its peak at $z \sim 2$. Increasing the UVB means that for a given temperature and density there are more ionizing photons, so ions are excited into higher states.

The ionization fractions are modelled using CLOUDY for a range of range of density, temperature, and redshift values, and compiled into a three-dimensional look-up table. The ionization fraction for any grid cell in the simulation output can be calculated by linearly interpolating between the points of the table (see Figure 2 for a diagram of interpolation in 1–3 dimensions). This is the final step to calculating the number density of any ion, from which mass density and mass follow.

It is common to assume that gas in a simulation is optically thin to radiation, and this was the case for the initial TRIDENT release, but full radiative transfer simulations show that the photoionization rate of a gas changes as a function of density and the strength of UVB in a way that allows self-shielding to be modelled without performing full radiative

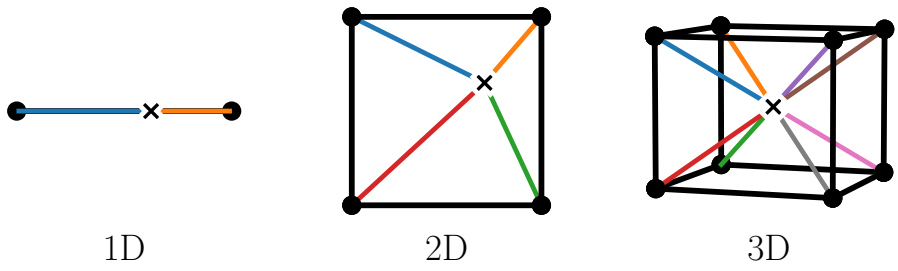

Figure 2. From left to right, diagrams of 1-, 2-, and 3-dimensional interpolation. Black points represent values where $\mathbf{x} = (x, y, \dots)$ and $f(\mathbf{x})$ are known. Using these known values and the vectors (coloured lines) from the known points to an arbitrary point (black cross), $f(\mathbf{x})$ can be approximated.

transfer calculations bu reducing the strength of incident radiation (Rahmati et al. 2013). Accounting for self-shielding using this prescription is becoming more common in codes that compute heating/cooling rates (e.g., Smith et al. 2017), and in simulations ranging from dwarf galaxies to regions 100 Mpc wide (Davé et al. 2019; Emerick et al. 2019). TRIDENT now includes a table that accounts for self shielding at solar metallicity, and this project extends the three-dimensional tables currently available to a four-dimensional table, with an extra axis for varying metallicity. This is important as metal line cooling increases the efficiency of self-shielding.

3 Creating a New Ionization Table

This section details the creation of a new ionization table for use in post-processing simulations, that includes the effects of self-shielding with varying metallicity. This is to more accurately model the circumgalactic medium, as current tables either do not account for self-shielding, or account for self-shielding but assume solar metallicity. We describe assembling the table, expanding existing libraries to use the table when post-processing simulations, and end with a qualitative analysis of the effect of varying metallicity.

3.1 Assembling the New Table

The first task in this project was to aggregate data stored in raw ASCII files into a multi-dimensional ionization table saved in the HDF5 format. These data were generated by my supervisor, by running CLOUDY radiative transfer models over the range of parameter values presented in Table 1. I was also provided two ionization tables that assumed optically thin gas, and accounted for self-shielding but assuming solar metallicity.

Parameter	Start	Stop (inclusive)	Step	Number of points
$\log(n_{\rm H}/{\rm cm}^{-3})$	-9	2	0.125	89
$\log(1+z)$	0	1.2	0.05	25
$\log(Z/Z_\odot)$	-2	1	0.2	16
$\log(T/{\rm K})$	1	9	0.025	321

Table 1. Parameter ranges used to create the new ionization table. Points are spaced uniformly in log space. The parameters are the number density of hydrogen, redshift (as a proxy for the intensity of the ultraviolet background), metallicity, and temperature.

Not all of the CLOUDY runs converged to a solution, but these were all at sufficiently high redshifts and surrounded by physically correct values that linearly interpolating through density was deemed an appropriate choice for calculating these values. Points were determined as missing if: (i) CLOUDY did not converge on a solution (no output files were generated) or (ii) the ionization fractions across all states, for a given element, were all 10^{-30}, which is unphysical as ionization fractions should sum to one. The value of 10^{-30} is used in the output as the zero-point for ionization fractions because the logarithm of 0 diverges and 10^{-30} is sufficiently small. See Figure 3 for the distribution of interpolated points, demonstrating that the majority of interpolated points are at redshifts not normally considered, and Figure 4 for an example comparison of the raw and final ionization fractions.

3.2 Interfacing the Table with Simulated Data

The YT and TRIDENT libraries were used to post-process simulations in this project. YT is an analysis and visualisation suite, and TRIDENT is a package for post-processing simulations to add information about ions not initially simulated.

For the existing three-dimensional tables the software stack for adding additional information is:

- TRIDENT takes the data read in by YT, and an ionization table.
- For an ion it iterates over each cell in the simulation grid and finds the ionization fraction as a function of density, temperature, and redshift, by linearly interpolating between values in the ionization table.
- TRIDENT calculates the number density of an ion from the density, metallicity, and (assumed solar) abundances using Equation (1).
- Ion density is found by multiplying the number density by the mass of the ion, and the total mass is the product of the density and volume of the cell.

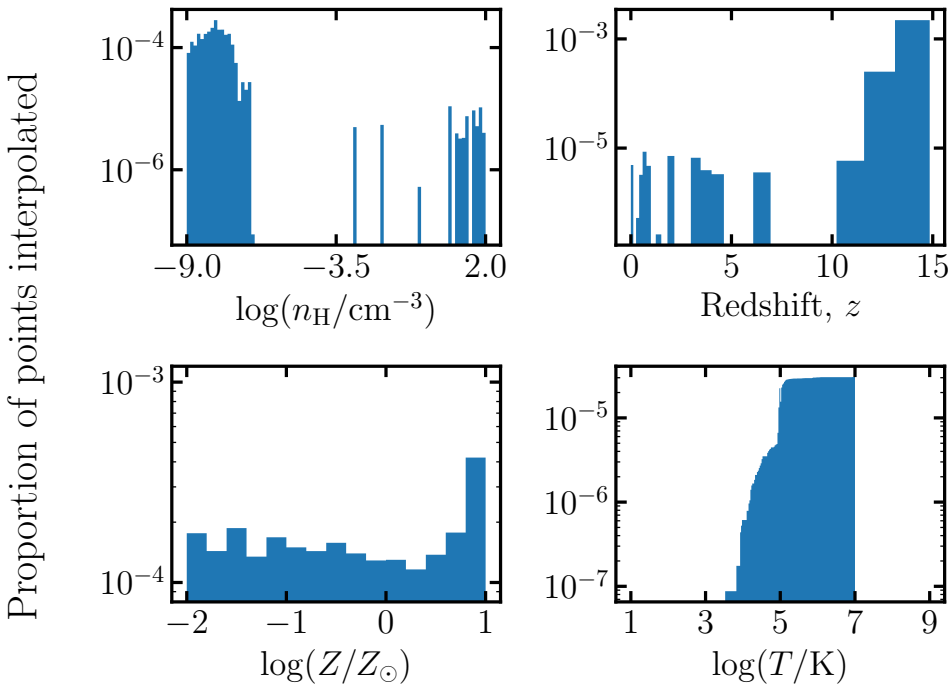

Figure 3. Histograms showing the location in phase space of interpolated points. Very few points are at low redshift, where this project is focused. From top left to bottom right the x-axis is hydrogen number density, redshift, metallicity, and temperature.

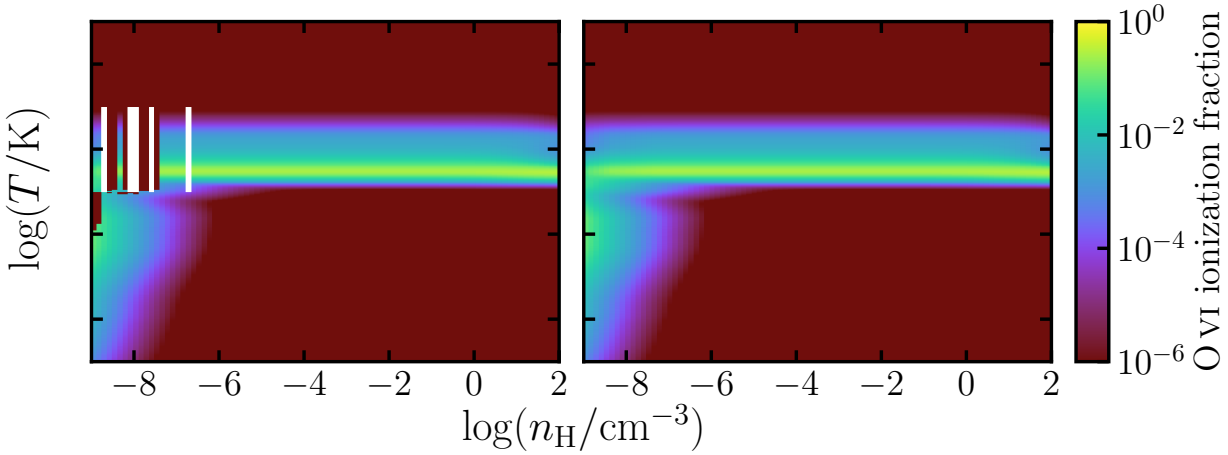

Figure 4. Plots of a slice through the new ionization table before (left) and after (right) interpolation was performed, at the highest redshift ($z = 14.849$) and solar metallicity. The white regions are missing points where CLOUDY did not converge, and the rectangles of 10^{-30} persist across all ionization states so are also replaced by interpolated results.

The interpolation functionality is part of YT, which has a series of `FieldInterpolator` classes to interface between a table of named fields, which is how simulation data is stored by YT, and generic interpolator functions, written in compiled CYTHON for efficiency.

To include the ability to use 4D ionization tables I added a new 4D `QuadrilinearFieldInterpolator` class to YT, alongside its corresponding generic 4D interpolator CYTHON function. I expanded TRIDENT to be able to read in the 4D ionization table and calculate the desired quantities by interpolating over density, redshift, temperature, and the new metallicity axis.

3.3 The Effect of Changing Metallicity

We now qualitatively examine the effect that this additional metallicity axis has on ionization fraction, using oxygen as an example. We describe how metallicity changes the state of oxygen, by considering O VI as a function of density, temperature, and metallicity at $z = 0$, before averaging over density and temperature to consider how the balance between all ionization states of oxygen changes with metallicity and redshift.

The effect of increasing metallicity on the ionization fraction is shown by plotting the difference between the fraction of oxygen in O VI at $Z = 0.01Z_\odot$ and $Z = 10Z_\odot$ in Figure 5. Higher numbers (shown in red) mean that at higher metallicities more oxygen is found in its fifth ionization state, O VI[5], at a given density and temperature. There are two regions where change occurs: a high-density one, which corresponds to collisional ionization equilibrium; and a low-density one, where photoionization dominates.

The change in the collisional ionization regime indicates that lower temperatures are needed to achieve the same ionization fraction of O VI. This could be explained by the fact that increasing the proportion of metals will introduce more heavy particles, changing the dynamics of the gas. The change in the photoionization regime has the expected effect of moving the photoionzation tail to the left, as increasing the number of metals will increase the efficiency of self-shielding. Metals increase self-shielding as they contribute more line cooling than primordial elements, allowing the outside of a cloud to sustain a higher rate of incident radiation before the inside is ionized, decreasing the amount of ionizing radiation the whole cloud experiences. This is in fact how it is modelled by Rahmati et al. (2013).

As a way of reducing the dimensionality of the table we considered the mean change in ionization fraction inside the grey dashed outline in Figure 5. This is roughly the region that a galaxy occupies in phase space. By looking at the mean change in ionization fraction, we can more easily examine the effect metallicity has across all the ions of an element at different redshifts. Figure 6 shows this for oxygen, but the general trend is true for other elements.

The specific details are sensitive to which ion is studied and at what redshift (as a proxy for UVB intensity), but the general trend is the same as above: as the effect of self-shielding

[5]We use the notation X I, X II, ..., to refer to the neutral, first ionized state, ..., for an element X.

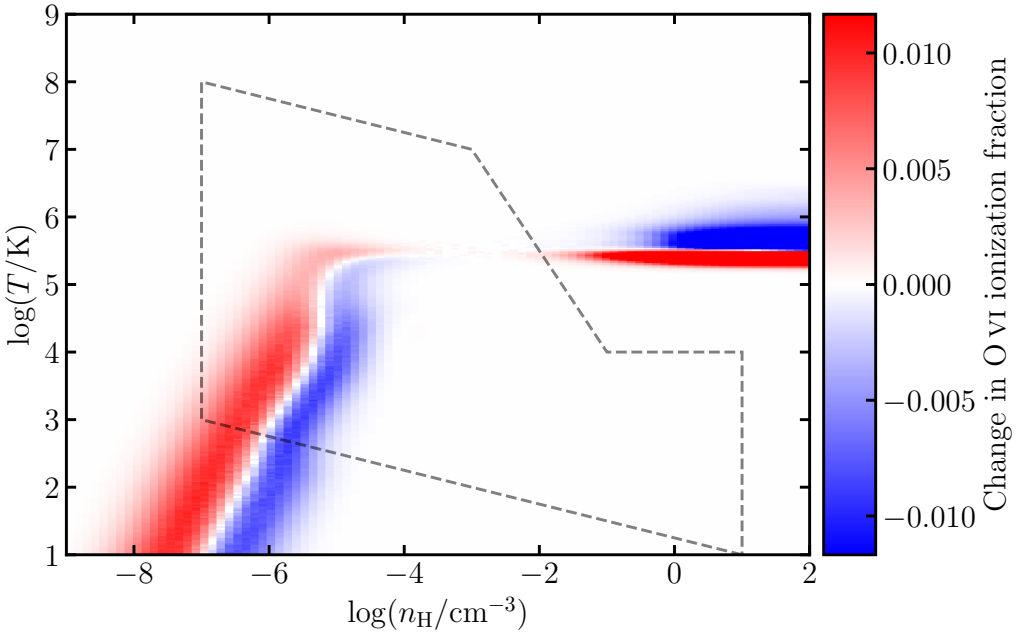

Figure 5. A plot of the change in ionization fraction from $0.01Z_\odot$ to $10Z_\odot$ at $z = 0$. The 'tail' of the fraction, where photoionization is important, moves to the left as incoming radiation is diminished as the effects of self-shielding increase with metallicity. This effectively increases the range of densities for which collisional ionization dominates. The grey dashed outline represents the typical area a galaxy might occupy, used to average over in Figure 6.

is increased with metallicity, the fraction of oxygen in its higher states decreases, and the fraction in lower states increases. The change is largest at redshift values ~ 2 because this is when the UVB peaks, so photoionization and processes affecting it are more dominant. At higher redshifts the UVB is still building up as more stars are forming, and at lower redshift the star formation rate is declining and the massive stars that provide the bulk of ionizing radiation have evolved off the main sequence.

4 Application to Simulation Data

We now turn our analysis to the effects of including the new metallicity-dependant self-shielded ionization table in post-processing simulation data. We focus on a subset of observationally motivated ions: Mg II, C II, C III, C IV, Si II, Si III, Si IV, and O VI. The CGM[6] is often studied by looking at these ions (e.g., Werk et al. 2013) as they trace the different phases of gas present.

We post-processed the simulations (as described in Section 2.3.1), adding masses and number densities for the ions using three ionization tables: one that assumes optically thin gas,

[6]With more time we would use two sets of ions, the one mentioned and a set better suited to studying the WHIM: e.g. C IV, N V, O VI, and O VII.

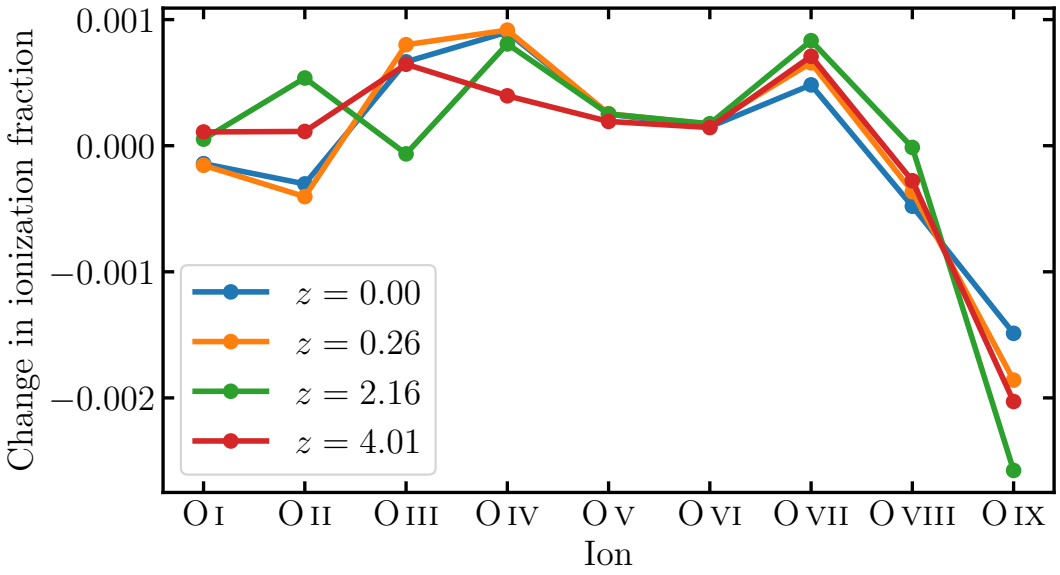

Figure 6. Plot of the average change of ionization fraction in a galaxy as metallicity is increased from $0.01 Z_\odot$ to $10 Z_\odot$. Changing metallicity, and thus self-shielding, has the largest effect at $z \sim 2$, where cosmic star formation rate and hence the ionizing background peaks.

one that takes self-shielding into account but assumes solar metallicity, and the new table described in Section 3.1 that accounts for self-shielding and non-solar metallicities. The masses and number densities of each ion are totalled and averaged[7] respectively, over the region of interest. For the WHIM this is a box covering a subset of the simulation, and for the CGM this is a sphere of radius 100 kpc around the galaxy. The centre of the galaxy was found by iteratively creating a $(300\,h^{-1}\,\mathrm{kpc})^3$ box around a point, finding the centre-of-mass, and re-centering the region around that point. The initial guesses were informed by the densest point of the simulation, and the centre-of-mass of the gas and star particles of an arbitrarily large region around the centre of the simulation box, and confirmed by visual inspection. The process was iterated over until the centre stopped changing at a scale below ~ 100 pc.

The total mass and average number density can be compared for the three possible pairings of the ionization tables by considering the difference between:

1. The optically thin table and the table with self-shielding, to quantify the effect of self-shielding with assumed solar metallicity.

2. The self-shielding table that assumes solar metallicity and the metallicity-dependant self-shielding table, which quantifies the effect that just allowing metallicity to vary has.

3. The optically thin table and the metallicity-dependant self-shielded table.

[7] Averages are specifically weighted averages, $\sum_i w_i f_i / \sum_i w_i$, weighted by the mass of a cell.

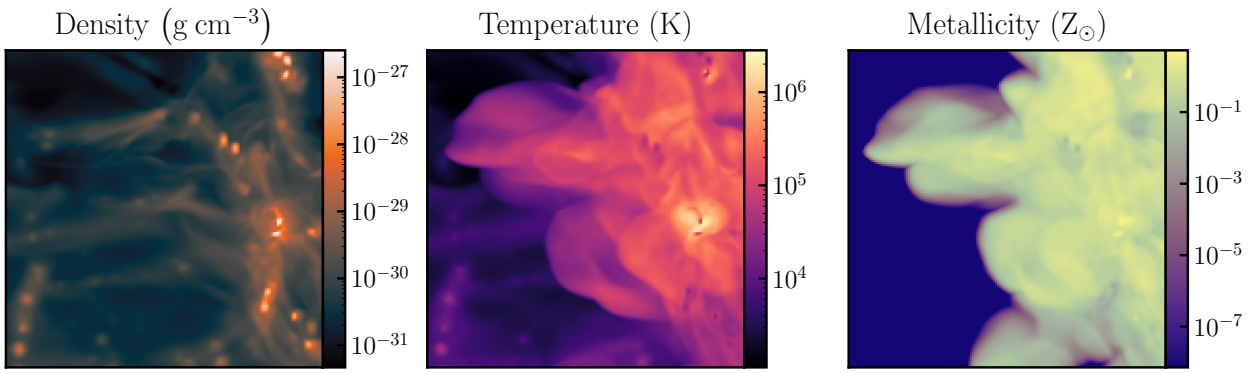

Figure 7. Left to right: Average density, temperature, and metallicity in the region of the WHIM studied. Each box has a side length and depth of $5\,h^{-1}\,\mathrm{Mpc}$.

4.1 Physical Properties of the WHIM

To study the WHIM we used a higher resolution run of the simulation described in Smith et al. (2011). It is a cosmological hydrodynamical+N-body Enzo simulation of a $(50\,h^{-1}\,\mathrm{Mpc})^3$ cube, including star formation, radiative heating and cooling, and chemical evolution prescriptions. The simulation box consists of 1536^3 cells, each with side length $32.6\,h^{-1}\,\mathrm{kpc}$. The simulation assumed a flat ΛCDM cosmology with $\Omega_\Lambda = 0.727$, $\Omega_\mathrm{m} = 0.273$, and $H_0 = 70.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. For analysis we took a cube of side length $5\,h^{-1}\,\mathrm{Mpc}$, which is large enough to capture a variety of phases of gas and structures. Figure 7 shows a projection of average density, temperature, and metallicity in this box.

In Figure 8 the differences in total mass and average number density for the three ionization table pairings are plotted for the diagnostic ions of interest. As a reminder, the three pairings are: optically thin versus self-shielded with solar metallicity; self-shielded with solar metallicity versus self-shielded with varying metallicity; and optically thin versus self-shielded with varying metallicity. Generally, the mass and number density have small changes; typically $\lesssim 1\%$ with a maximum change of 6.9% in O VI number density as self-shielding is included.

This lack of change is to be expected, partially because these ions are not ideal for studying the WHIM, but also because this warm/warm-hot diffuse gas is optically thin and so not self-shielded. The change from assuming solar metallicity is important for only some ions as they are typically tracing less enriched gas than others (see Figure 9 for a comparison of O VI and Si II). A larger portion of O VI traces gas closer to solar metallicity than Si II does, so assuming solar metallicity has a larger effect for Si II. Generally, the assumption of solar metallicity has a smaller effect than not accounting for self-shielding, but both are on a similar scale.

In the case of the warm-hot intergalactic medium current models of optically thin gas are close to self-shielded ones, and the change introduced to self-shielding by allowing non-solar metallicity is similarly sized. However further analysis of the higher ions (e.g. O VII) used in tracing the WHIM is desirable, because as can be seen in Figure 6 the fraction of oxygen in the higher states changes more than lower ones when metallicity changes.

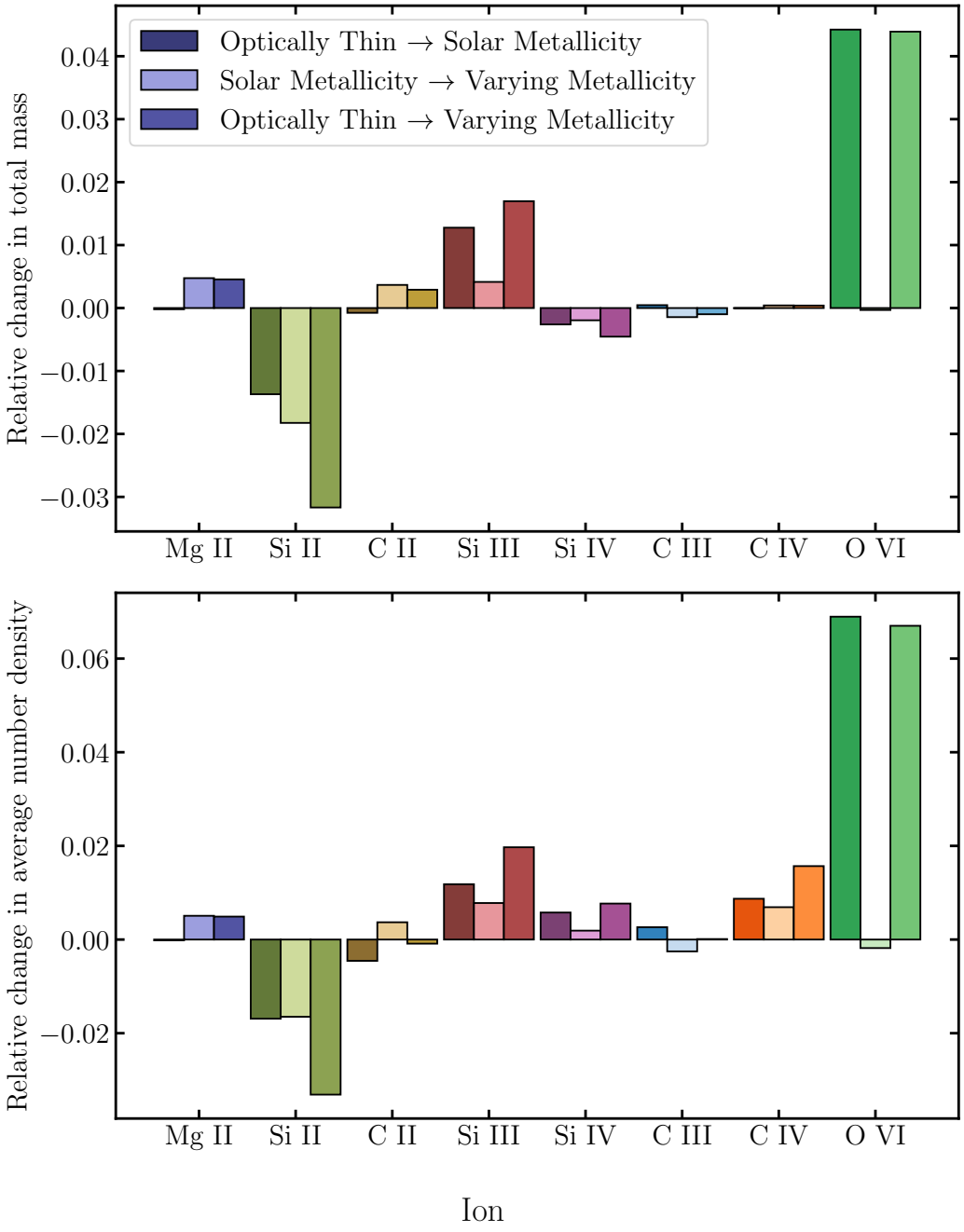

Figure 8. How total mass (top) and average number density (bottom) change as different ionization tables are used to post-process a simulation of the WHIM. The darkest and leftmost bar is the change between the optically thin case and solar-metallicity self-shielded. The middle and lightest bar is the difference between solar and non-solar self-shielded, and the rightmost bar is the difference between the optically thin and non-solar metallicity self-shielded table. Ions are organised in increasing ionization potential as a proxy for the temperature of the gas they trace. Generally, the change is small, ∼1%, and biggest for O VI, which makes sense given the ions analysed and the temperature range of the WHIM. Number density follows a similar trend to mass.

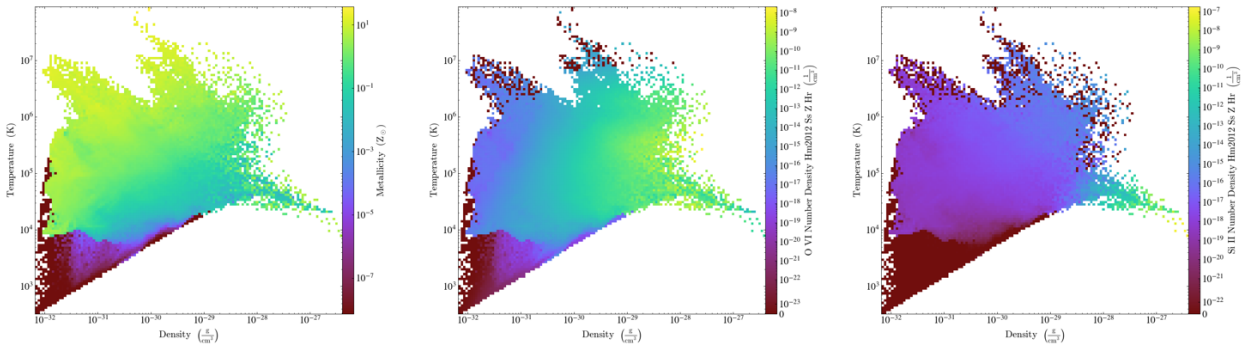

Figure 9. Phase plots of metallicity (left), O VI number density (middle), and Si II number density (right). A larger portion of O VI traces gas closer to solar metallicity than Si II, so accounting for metallicity dependence in self-shielding has a larger effect on Si II mass and number density.

4.2 Physical Properties of the CGM

To study the CGM six simulated galactic halos from the Figuring Out Gas & Galaxies in Enzo (FOGGIE) project were used. The details of the simulation are described in the first three FOGGIE papers (Peeples et al. 2019; Corlies et al. 2020; Zheng et al. 2020), and the halos are described in Simons et al. (2020), but we describe the pertinent details below. All six halos have been evolved to $z = 2$, and five to $z = 0$. These are zoom simulations run with Enzo. They differ from the simulation described in Section 4.1 in that they have updated cooling and ionization physics, using the more modern grackle code (Smith et al. 2017) and including self-shielding. The simulations also apply a novel refinement scheme, described below. A flat ΛCDM cosmology with $\Omega_\Lambda = 0.715$, $\Omega_\mathrm{m} = 0.285$, $H_0 = 69.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ was used.

The galaxies are simulated by finding an approximately Milky Way mass halo that has not undergone any significant mergers since $z = 2$ in a lower resolution cosmological simulation, then re-simulating with a two-pronged refinement approach. The first is that a $(200\,h^{-1}\,\mathrm{kpc})^3$ *forced refinement* box is drawn around the halo, forcing the maximum cell length to be 1.1 kpc (a refinement level of $n_\mathrm{ref} = 9$) in this region. Within this there is additional *cooling refinement* for especially dense and cool gas. A cell is split into eight smaller cells until it reaches either the cooling length (defined as the speed of sound × the cooling time) or the maximum level of refinement. To better resolve dense/cool gas in the forced refinement region the maximum level of refinement is increased to give a minimum cell length of 274 pc ($n_\mathrm{ref} = 11$). This increased resolution reduces artificial mixing of cool and hot gas, revealing a more complex multiphase structure of cool gas on small scales that more accurately represents the CGM (Hummels et al. 2019).

See Figures 10 and 11 for projections of the average density of the galaxies at redshifts $z = 0$ and $z = 2$, and Figure 12 for an example of the density, temperature and metallicity of a halo at both redshifts. On each plot is a white circle of radius 100 kpc, indicating the spherical region considered in the analysis.

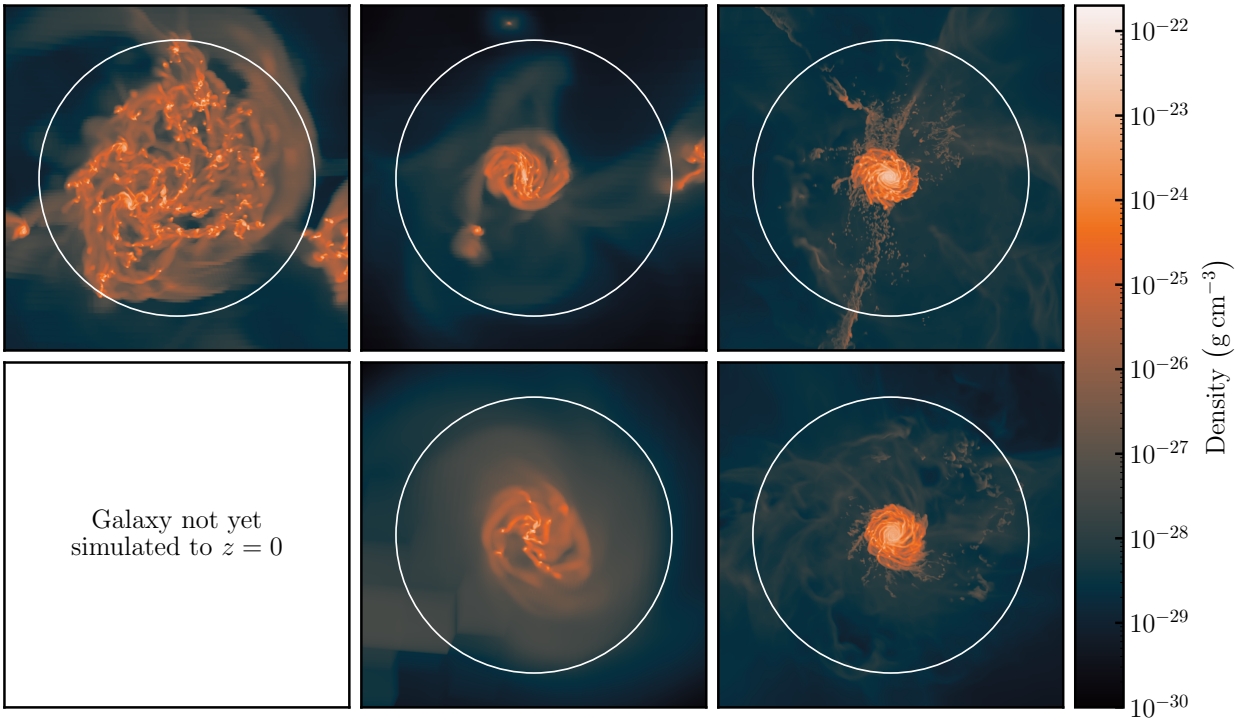

Figure 10. Average density projected along the angular momentum vector of the 5 galaxies at redshift $z = 0$. The white circles overlayed represent the 100 kpc sphere studied. Projections are 600 kpc deep and each frame is 250 kpc wide.

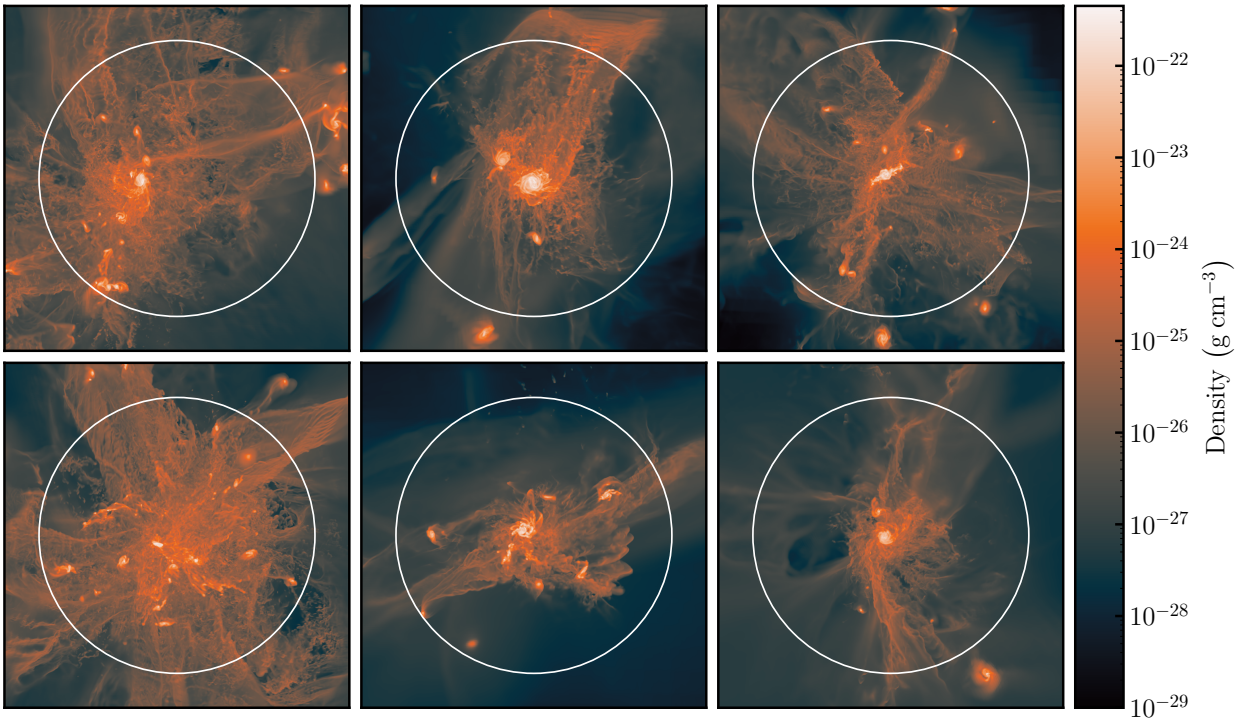

Figure 11. Average density projected along the angular momentum vector of the galaxies studied, at redshift $z = 2$. Projections are 600 kpc deep and each frame is 250 kpc wide.

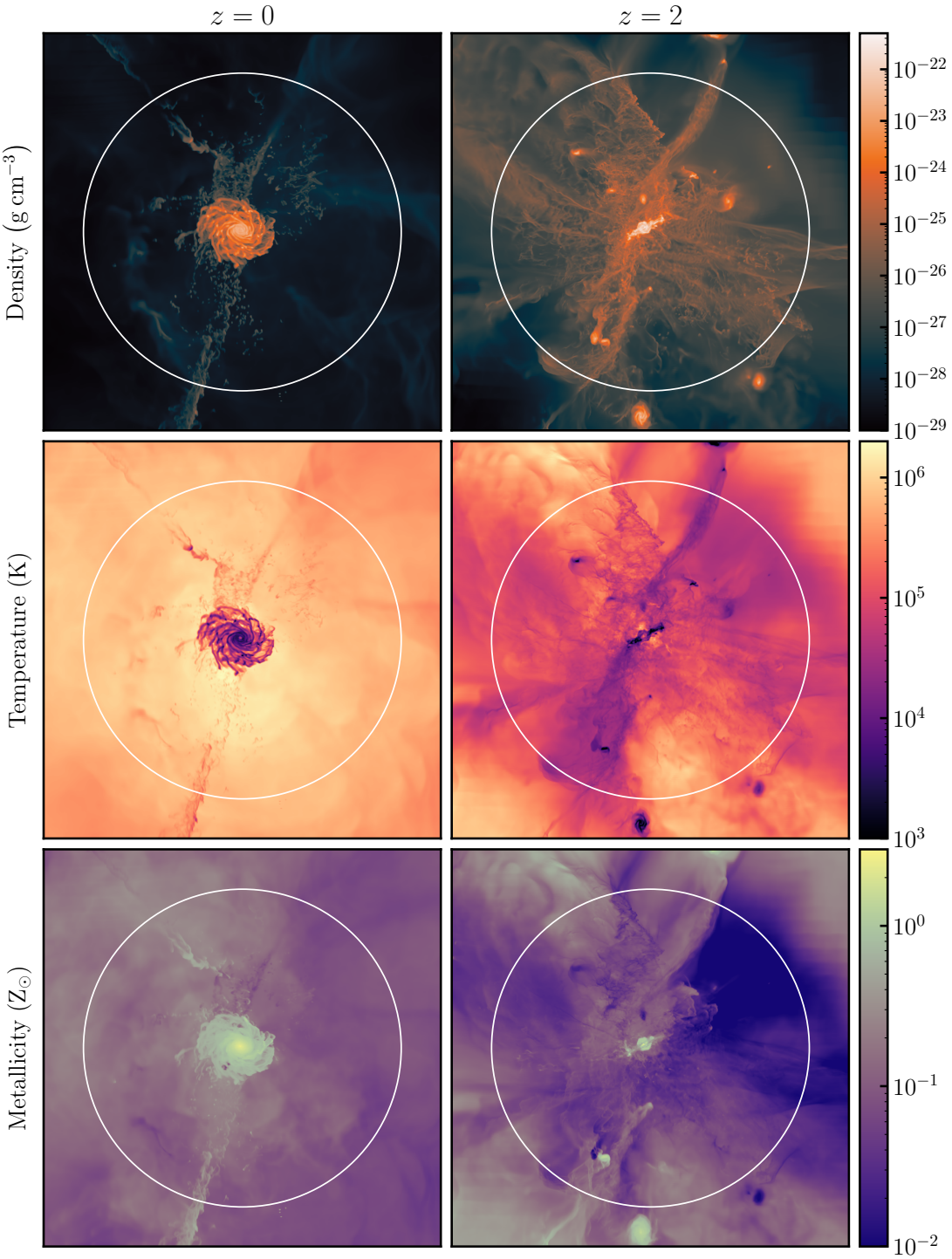

Figure 12. An example of average density, temperature, and metallicity of a halo at redshifts $z = 0$ and $z = 2$. This galaxy is top-right in Figures 10 and 11, and the length scales are same.

The same analysis as before—post-processing with different ionization tables and comparing how mass and number density change between the three table pairings—was carried out. Figures 13 and 16 are both similar to Figure 8, plotting the relative change of total mass and average number density of ions for redshifts $z = 0$ and $z = 2$, but instead of a bar chart a point is plotted for each halo to communicate the average and spread.

4.2.1 Redshift $z = 0$

At redshift $z = 0$ (Figure 13) it is clear there is a large ($\sim 20\%$) change in the total mass of certain ions (Mg II, C II, O VI) while others are not affected as much, only changing by a few percent. In all cases the effect of introducing metallicity dependence is much smaller than the effect of introducing self-shielding, implying that the assumption of solar metallicity is accurate on the scale of the total mass of a galaxy. The number density follows a similar trend, but with considerably more spread between galaxies.

Generally, the effect of considering self-shielding dominates over the effect of varying metallicity. However, the metallicity is more important when considering the average number density of the higher ions C IV and O VI. These ions tend to trace the hotter and lower metallicity gas that exists further away from the centre of the galaxy, while the lower ions trace cooler gas at smaller radii. Using Mg II and O VI as examples of low and high ions (low and high referring to the amount of ionization energy needed to reach that state) respectively, the spatial correlation can be seen by comparing the metallicity in Figure 12 with the column densities plotted in Figure 14. Figure 15 demonstrates the correlation in phase-space, plotting the average metallicity and Mg II and O VI number densities as functions of density and temperature.

Comparing the $z = 0$ CGM and WHIM it can be seen that the scale of changes is typically much larger in the CGM. This is especially true in number density, reaching a maximum difference of 150% compared to 6%. This is because the CGM consists of cooler and denser clouds than the WHIM, where self-shielding is much more important to consider. The relative importance of metallicity and self-shielding was higher in the WHIM because of the lower metallicities.

4.2.2 Redshift $z = 2$

At redshift $z = 2$ (Figure 16) there is a similar distribution of ions that have a small change and ions that have a larger change. Mg II, C II, and O VI have changes typically $\gtrsim 5\%$, but smaller than their changes at $z = 0$. Si II, Si III, Si IV, C III, and C IV all have similarly sized changes of $\lesssim 5\%$ at both redshifts, reflecting that self-shielding is not important for these ions. If it were then the change in redshift, and hence UVB, would change the mass and number density.

The effect of metallicity is again most noticeable for the high C IV and O VI ions, and the metallicity dependence for the highest O VI ion is much stronger than at $z = 0$. This could

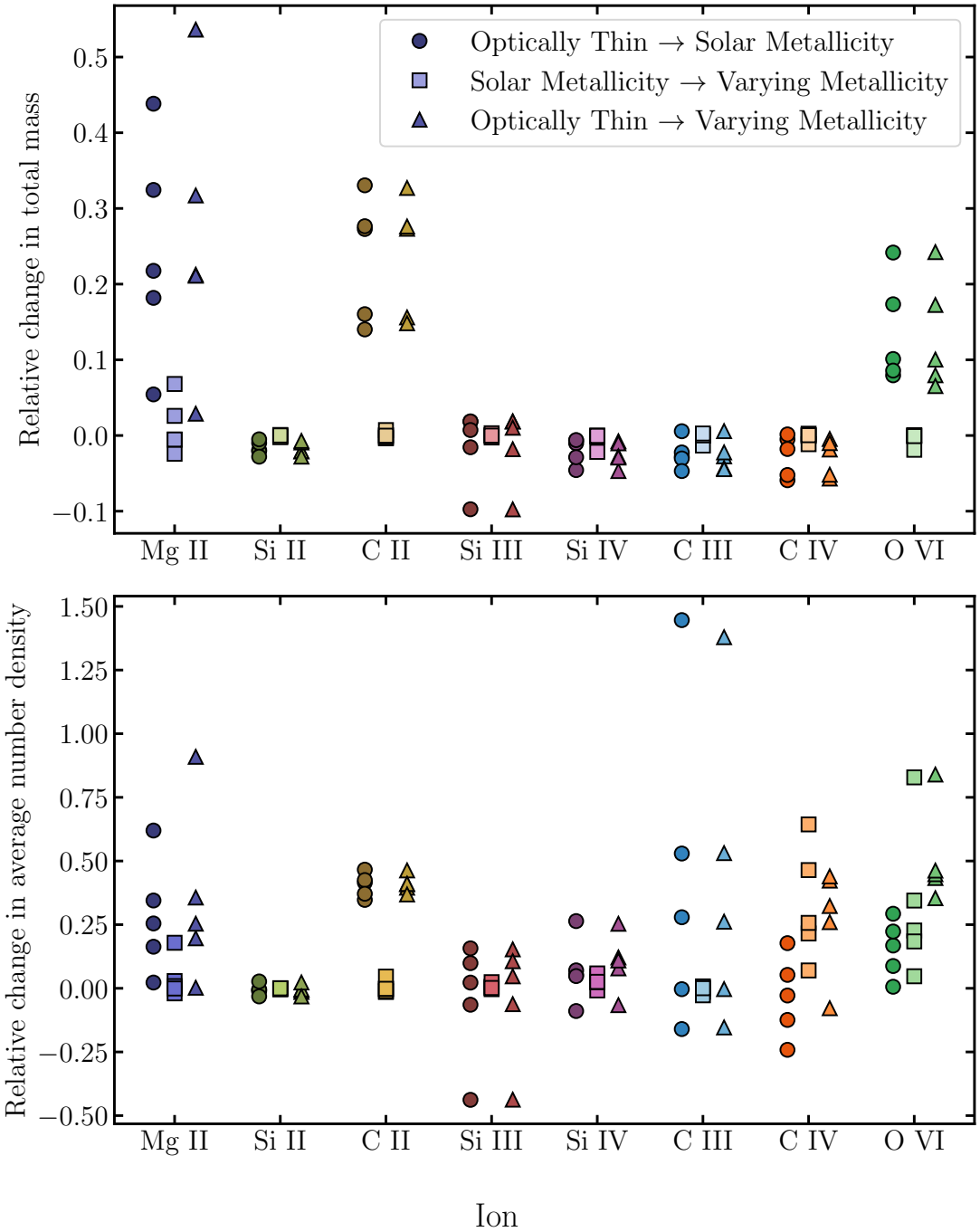

Figure 13. Plot of the relative change in the total mass (top) and average number density (bottom) of an ion at redshift $z = 0$ when post-processing using different ionization tables. Colours and positions have the same meaning as Figure 8.

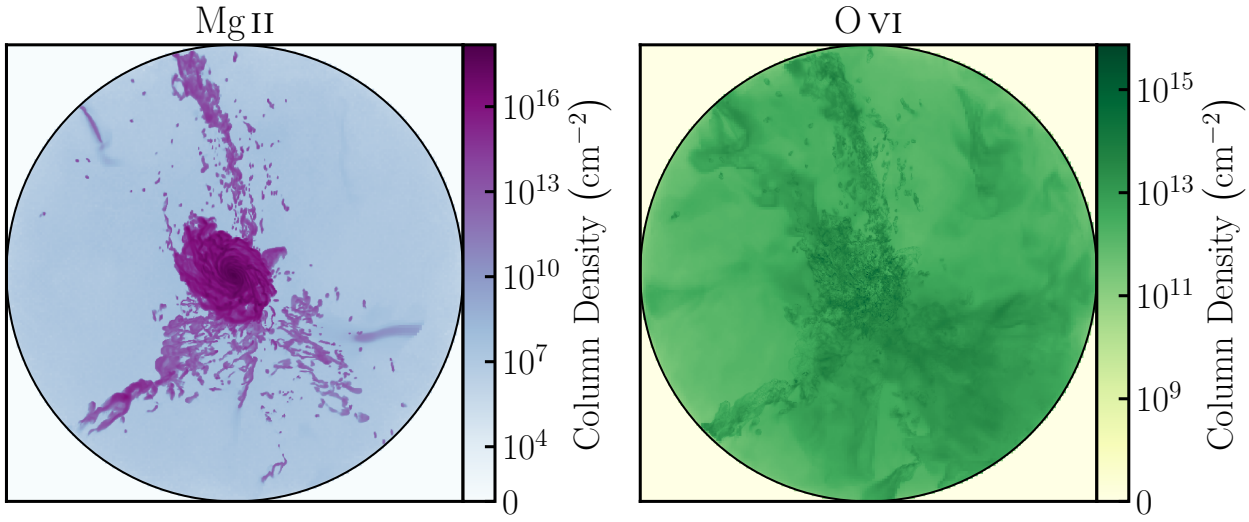

Figure 14. Column densities at $z = 0$ of Mg II (left) and O VI (right). Comparing the spatial extents they cover to Figure 12 explains why metallicity has a larger effect in the high C IV and O VI ions than the lower ions. The black circle is the same 100 kpc sphere as in previous figures, but the depth is only 200 kpc.

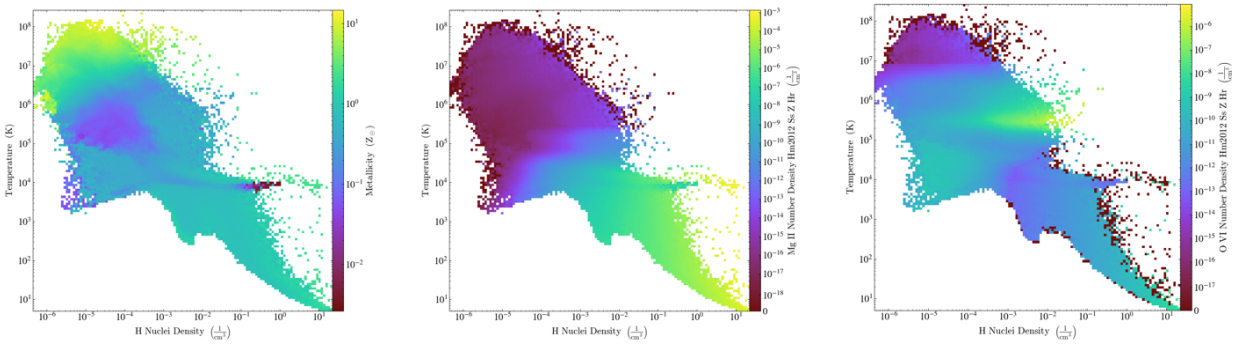

Figure 15. Phase plots at $z = 0$ of metallicity (left), Mg II number density (middle), and O VI (right) number density. This shows O VI traces non-solar metallicity gas, explaining why metallicity has a larger effect in O VI than other ions.

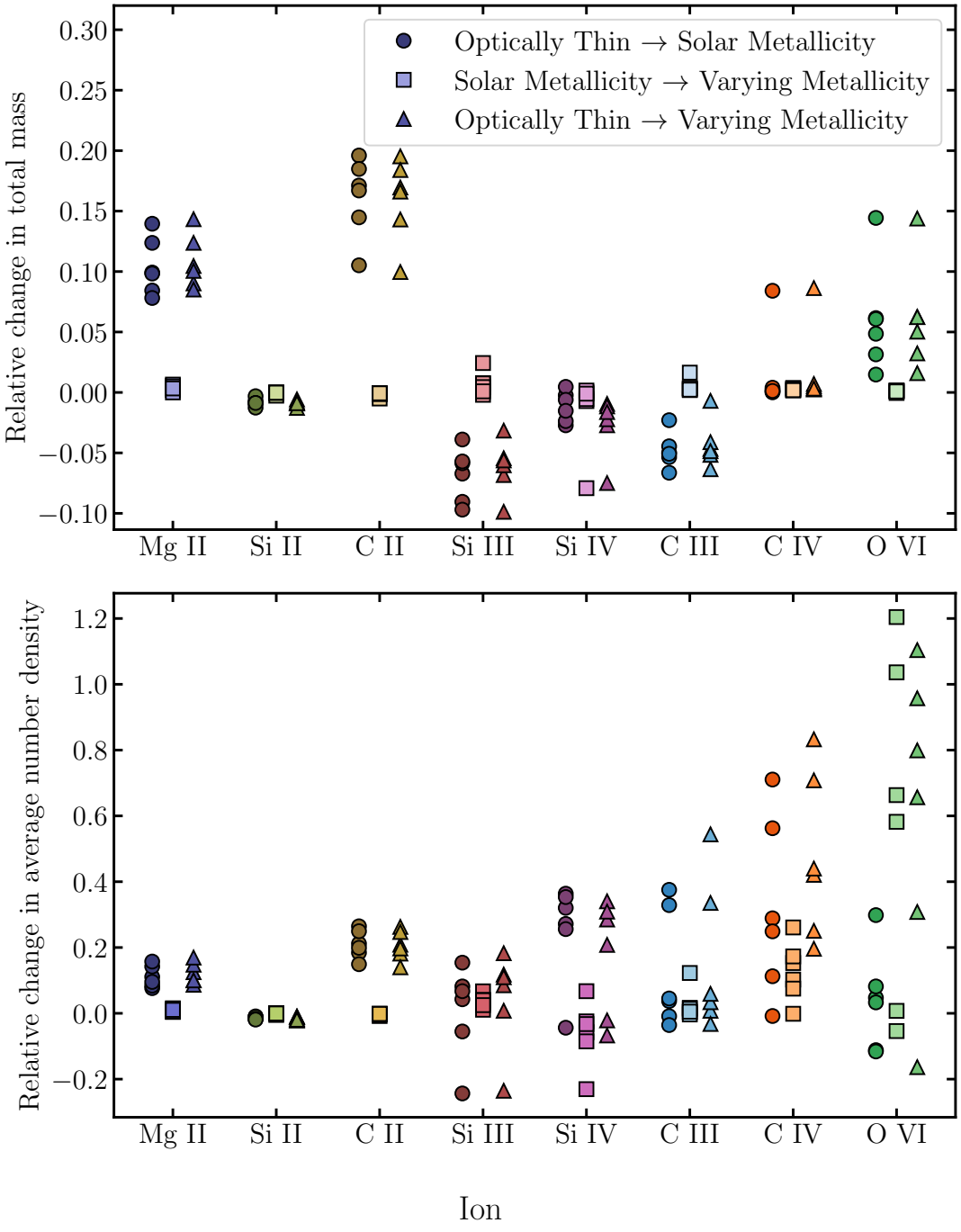

Figure 16. Plot of the relative change in the total mass (top) and average number density (bottom) of an ion at redshift $z = 2$ when post-processing using different ionization tables. Colours and positions have the same meaning as Figures 8 and 13.

be either because of the lower metallicity, as the galaxy has not had as long to enrich the CGM, or the higher UVB meaning photoionization is more important. The relationship is complicated by the fact that C IV shows the opposite effect, where metallicity is less important at higher redshift, but this difference could be explained by the lower ionization potential.

The change in the simulated CGM is sensitive to which ions in particular one is studying, but the effects of self-shielding are non-negligible so should be accounted for. Allowing metallicity to vary from the solar value is also important for specific ions that trace hotter gas in the CGM, so studies of the hot phases of gas should include this.

4.3 Observational Properties of the CGM

The column density, N, is the integral of number density along a line of sight, providing the number of particles per unit area. It is derived from observations by fitting line profiles to spectra, and typically for simulated data it is derived by generating synthetic spectra (e.g. with TRIDENT) and fitting them using the same method as for observed spectra, to compare simulated and observed column densities. However, it can also be calculated from simulation data directly by integrating through the simulation domain. Finding the column density of a galaxy along the three simulation axes x, y, and z with a 1600×1600 pixel resolution image creates on the order of 1 million column density values, sampling a galaxy at different points with different orientations. Calculating column densities for all five $z = 0$ halos and all six $z = 2$ halos creates a population of column densities along random sight lines with respect to different galaxies that is a rough approximation of CGM surveys (e.g., Tumlinson et al. 2013).

We analysed the difference between the column densities calculated with the three ionization tables with the aim of determining if the three different models could be distinguished observationally. The method below is described for a single ion, but was carried out on each of the four ions Si II, Si IV, C IV, and O VI. Two column densities calculated with different ionization tables, denoted as N_1, N_2, were said to be observably different if (i) the logarithm of the column densities $\log N_1$ and $\log N_2$ were both greater than the minimum $\log N$ observed in Peeples et al. (2019); and (ii) the difference $\log N_2 - \log N_1$ was greater than the median error on the fitted $\log N$ in the aforementioned paper. Counting the number of observably different column densities and dividing by the total amount provides a fraction of sight lines that have been affected by using a different model; alternatively it can be thought of as the probability that a random spectrum from a galaxy could distinguish between the two models.

These fractions/probabilities are presented as percentages in Table 2. It convincingly shows that the column densities of C IV and Si IV are not affected by assumptions about the gas being optically thin or at solar metallicity at both redshift $z = 0$ and $z = 2$. It also shows that the assumption of solar metallicity does not affect spectra at either redshift, with a maximum 2.6% of column densities changing, but a more typical value being $< 0.02\%$. O VI has the

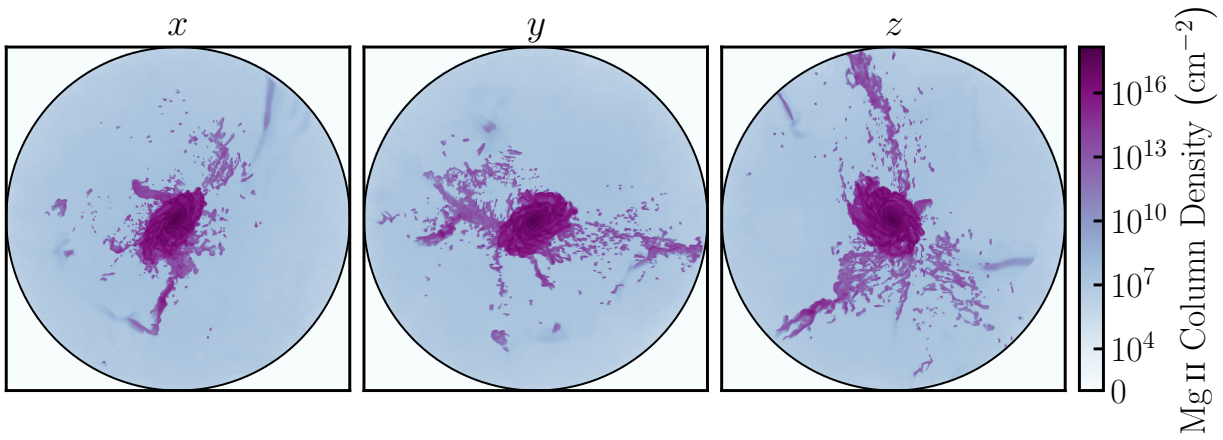

Figure 17. Column densities of Mg II at $z = 0$, projected along the three simulation axes x, y, z. Each pixel represents the column density along a line of sight, repeating this for each galaxy generates a large number of random sight lines, mimicking an observational survey. The black circle is the same 100 kpc sphere as in previous figures, and the depth is 200 kpc.

largest change overall, which is unsurprising as it has the largest change in number density of the four ions considered here. The indistinguishability of the three different ionization tables is unproblematic as these ions, excluding O VI, also do not have a large change in the total mass, so would not be able to account for any missing baryons.

There are two obvious areas for improvement: considering more ions, and generating synthetic spectra. Considering more ions, especially those with large changes in mass and number density like Mg II and C II, will lead to a better understanding of the physical processes causing the difference that self-shielding and varying metallicity have. Generating synthetic spectra would be a more accurate way to test if it is possible to observe a difference between these models, as it could test if the profile fits of spectra themselves could be distinguished at a statistically significant level.

5 Conclusions and Future Work

In this work we introduce a new 4D ionization table for use in post-processing simulations while taking into account the effects of self-shielding in non-solar metallicity gas. This is an important consideration for studies of the small dense clouds that populate the circumgalactic medium. We use this table alongside two existing tables—one that assumes gas is optically thin to ionizing radiation; and the other assumes self-shielding at solar metallicity—to study the effect of varying metallicity on mass and number density of several low ions by post-processing simulations of the WHIM and CGM. We study the WHIM in a large-scale simulation, and the CGM in a series of cutting-edge zoom simulations of Milky Way sized halos. Our results are as follows:

1. The effect of self-shielding in the WHIM is small because it is diffuse and hot enough

to be optically thin to ionizing radiation. The low metallicity of the gas means that allowing metallicity to vary is as important as considering self-shielding.

2. Self-shielding has a large effect on the mass and number density of some ions (Mg II, C II, O VI), but not others (C III, C IV, Si II, Si III, Si IV). When considering mass, changing metallicity has a small effect across all ions, but is important to consider for higher ions that trace hotter gas. Increasing redshift does not have a clear trend; it increases the metallicity dependence for O VI, but decreases it for C IV.

3. Assumptions about self-shielding and metallicity have very little effect on the column densities at redshifts $z = 0$ and $z = 2$, implying constraining different models observationally would prove difficult.

There are many interesting avenues for future work. Expanding the analysis to a wider range of ions and redshifts would enable us to better understand how self-shielding at different metallicities affects different phases of gas. Separating the effects of the chemical evolution of galaxies and the evolution of the UVB would shed light on how estimates of ion masses and densities vary with cosmic time. Considering more ions and generating synthetic spectra to fit line profiles to would allow us to test how these models can be constrained, and if it would affect current estimates of the mass of baryons in galaxies.

References

Anderson, M. E. and J. N. Bregman, 'Do Hot Halos Around Galaxies Contain the Missing Baryons?', ApJ **714**, 320 (2010).

Appleby, S. et al., 'The physical nature of circumgalactic medium absorbers in SIMBA', MNRAS **519**, 5514 (2023).

Bordoloi, R. et al., 'The COS-Dwarfs Survey: The Carbon Reservoir around Sub-L* Galaxies', ApJ **796**, 136 (2014).

Brummel-Smith, C. et al., 'ENZO: An Adaptive Mesh Refinement Code for Astrophysics (Version 2.6)', The Journal of Open Source Software **4**, 1636 (2019).

Bryan, G. L. et al., 'ENZO: An Adaptive Mesh Refinement Code for Astrophysics', ApJS **211**, 19 (2014).

Bahcall, J. N. and J. Spitzer Lyman, 'Absorption Lines Produced by Galactic Halos', ApJ **156**, L63 (1969).

Cen, R. and T. Fang, 'Where Are the Baryons? III. Nonequilibrium Effects and Observables', ApJ **650**, 573 (2006).

Corlies, L. et al., 'Figuring Out Gas & Galaxies in Enzo (FOGGIE). II. Emission from the z = 3 Circumgalactic Medium', ApJ **896**, 125 (2020).

Davé, R. et al., 'Baryons in the Warm-Hot Intergalactic Medium', ApJ **552**, 473 (2001).

Davé, R. et al., 'SIMBA: Cosmological simulations with black hole growth and feedback', MNRAS **486**, 2827 (2019).

Danforth, C. W. and J. M. Shull, 'The Low-z Intergalactic Medium. III. H I and Metal Absorbers at z < 0.4', ApJ **679**, 194 (2008).

Emerick, A. et al., 'Simulating an isolated dwarf galaxy with multichannel feedback and chemical yields from individual stars', MNRAS **482**, 1304 (2019).
Faucher-Giguère, C.-A. et al., 'A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization', ApJ **703**, 1416 (2009).
Ferland, G. J. et al., 'The 2017 Release Cloudy', Rev. Mexicana Astron. Astrofis. **53**, 385 (2017).
Ferland, G. J. et al., 'CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra', PASP **110**, 761 (1998).
Fukugita, M. et al., 'The Cosmic Baryon Budget', ApJ **503**, 518 (1998).
Fukugita, M. and P. J. E. Peebles, 'The Cosmic Energy Inventory', ApJ **616**, 643 (2004).
Hahn, O. and T. Abel, 'Multi-scale initial conditions for cosmological simulations', MNRAS **415**, 2101 (2011).
Haardt, F. and P. Madau, 'Radiative Transfer in a Clumpy Universe. IV. New Synthesis Models of the Cosmic UV/X-Ray Background', ApJ **746**, 125 (2012).
Hummels, C. B. et al., 'Trident: A Universal Tool for Generating Synthetic Absorption Spectra from Astrophysical Simulations', ApJ **847**, 59 (2017).
Hummels, C. B. et al., 'The Impact of Enhanced Halo Resolution on the Simulated Circumgalactic Medium', ApJ **882**, 156 (2019).
Johnson, S. D. et al., 'The Physical Origins of the Identified and Still Missing Components of the Warm-Hot Intergalactic Medium: Insights from Deep Surveys in the Field of Blazar 1ES1553+113', ApJ **884**, L31 (2019).
Lanzetta, K. M. et al., 'The Gaseous Extent of Galaxies and the Origin of Lyman-Alpha Absorption Systems: A Survey of Galaxies in the Fields of Hubble Space Telescope Spectroscopic Target QSOs', ApJ **442**, 538 (1995).
Macquart, J. .-. et al., 'A census of baryons in the Universe from localized fast radio bursts', Nature **581**, 391 (2020).
McGaugh, S. S. et al., 'The Baryon Content of Cosmic Structures', ApJ **708**, L14 (2010).
Madau, P. and M. Dickinson, 'Cosmic Star-Formation History', ARA&A **52**, 415 (2014).
Nicastro, F. et al., 'Observations of the missing baryons in the warm-hot intergalactic medium', Nature **558**, 406 (2018).
Peeples, M. S. et al., 'Figuring Out Gas & Galaxies in Enzo (FOGGIE). I. Resolving Simulated Circumgalactic Absorption at $2 \leq z \leq 2.5$', ApJ **873**, 129 (2019).
Pitrou, C. et al., 'Precision big bang nucleosynthesis with improved Helium-4 predictions', Phys. Rep. **754**, 1 (2018).
Planck Collaboration et al., 'Planck 2018 results. VI. Cosmological parameters', A&A **641**, A6 (2020).
Prochaska, J. X. et al., 'The COS-Halos Survey: Metallicities in the Low-redshift Circumgalactic Medium', ApJ **837**, 169 (2017).
Persic, M. and P. Salucci, 'The baryon content of the universe', MNRAS **258**, 14P (1992).
Rahmati, A. et al., 'On the evolution of the H I column density distribution in cosmological simulations', MNRAS **430**, 2427 (2013).
Shull, J. M., 'Where do Galaxies End?', ApJ **784**, 142 (2014).
Simons, R. C. et al., 'Figuring Out Gas & Galaxies in Enzo (FOGGIE). IV. The Stochasticity of Ram Pressure Stripping in Galactic Halos', ApJ **905**, 167 (2020).

Smith, B. D. et al., 'The Nature of the Warm/Hot Intergalactic Medium. I. Numerical Methods, Convergence, and O VI Absorption', ApJ **731**, 6 (2011).

Smith, B. D. et al., 'GRACKLE: a chemistry and cooling library for astrophysics', MNRAS **466**, 2217 (2017).

Sołtan, A. M., 'Warm-hot intergalactic medium contribution to baryonic matter', A&A **460**, 59 (2006).

Spitzer Lyman, J., 'On a Possible Interstellar Galactic Corona.', ApJ **124**, 20 (1956).

Smith, B. et al., 'Metal cooling in simulations of cosmic structure formation', MNRAS **385**, 1443 (2008).

Shull, J. M. et al., 'The Baryon Census in a Multiphase Intergalactic Medium: 30% of the Baryons May Still be Missing', ApJ **759**, 23 (2012).

Tumlinson, J. et al., 'The Large, Oxygen-Rich Halos of Star-Forming Galaxies Are a Major Reservoir of Galactic Metals', Science **334**, 948 (2011).

Tumlinson, J. et al., 'The COS-Halos Survey: Rationale, Design, and a Census of Circumgalactic Neutral Hydrogen', ApJ **777**, 59 (2013).

Turk, M. J. et al., 'yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data', ApJS **192**, 9 (2011).

Werk, J. K. et al., 'The COS-Halos Survey: An Empirical Description of Metal-line Absorption in the Low-redshift Circumgalactic Medium', ApJS **204**, 17 (2013).

Werk, J. K. et al., 'The COS-Halos Survey: Physical Conditions and Baryonic Mass in the Low-redshift Circumgalactic Medium', ApJ **792**, 8 (2014).

Zheng, Y. et al., 'Figuring Out Gas & Galaxies in Enzo (FOGGIE). III. The Mocky Way: Investigating Biases in Observing the Milky Way's Circumgalactic Medium', ApJ **896**, 143 (2020).

	Ion	Optically Thin → Solar Metallicity	Solar Metallicity → Varying Metallicity	Optically Thin → Varying Metallicity
Redshift $z = 0$	C IV	0.007	0.020	0.013
	O VI	17.507	0.022	17.509
	Si II	0.552	0.000	0.574
	Si IV	0.014	0.002	0.015
Redshift $z = 2$	C IV	0.012	0.003	0.013
	O VI	24.395	2.331	24.622
	Si II	9.703	2.604	10.398
	Si IV	0.012	0.002	0.011

Table 2. The fraction of column densities that differ by more than typical observational error as different ionization tables are considered. As an example, 24.4% of O VI column densities at $z = 2$ are detectably different when an optically thin ionization table is replaced by a table that considered self-shielding at solar metallicity.