UNIVERSITY OF CALIFORNIA
SANTA CRUZ

UNDERSTANDING THE PROPERTIES OF THE INTERGALACTIC MEDIUM THROUGH HIGH-PERFORMANCE COSMOLOGICAL SIMULATIONS.

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

Doctor of Philosophy

in

ASTRONOMY AND ASTROPHYSICS

by

Bruno Villasenor Alvarez

September 2022

The Dissertation of Bruno Villasenor Alvarez is approved:

Brant Robertson, Chair

Piero Madau

Evan Schneider

J. Xavier Prochaska

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Understanding the Properties of the Intergalactic Medium through High-Performance Cosmological Simulations.

by

Bruno Villasenor Alvarez

Understanding the evolution of the intergalactic medium (IGM) and how large-scale structures in the Universe develop over cosmic history represents one of the fundamental goals of modern astrophysics. The filamentary network of dark matter and gas, known as the "cosmic web" encodes information about the relative abundance of baryons and dark matter, the evolution of the radiation emitted by galaxies and quasars, the expansion history of the Universe, the nature and properties of the dark matter particle and the mass of neutrinos, among other relevant physics. The Lyman-α forest originates from the absorption signature that cosmic neutral hydrogen imprints on the spectra of distant quasars. Therefore, the forest provides a blueprint of the cosmic web, making it a primary probe of the properties of the IGM. The promise of the Lyman-α forest for constraining the nature of dark matter and dark energy has in part motivated the construction of the Dark Energy Spectroscopic Instrument (DESI), which will measure absorption line spectra from nearly a million distant quasars.

To extract the physics encoded in the observations of the Lyman-α forest, one requires sophisticated numerical simulations. For my thesis, I have extended the GPU-native hydrodynamical solver Cholla to run cosmological simulations. Using Cholla on

the largest supercomputers in the world, we have been able to run over 1500 high-resolution simulations that vary the physical models that shape the structure of the forest, allowing us to study the properties of the IGM with unprecedented detail.

This thesis presents an analysis of the impact on the IGM from different models for the photoionization and photoheating due to the metagalactic UVB radiation emitted by early galaxies and quasars. When comparing the Lyman-α forest from simulations that apply current models for the UVB to observations, we find that the models fail to reproduce the evolution of the statistical properties from the observational measurements. From a Bayesian approach where we compared the simulations to observations of the Lya forest, we were able to improve the current models and provide a better fit to the observations. Our best-fit model provides an inference of the thermal and ionization history of the IGM post-reionization consistent with other independent determinations.

Additionally, this thesis presents the results from our comparison of high-resolution observations of the Lya forest power spectrum to a massive grid of 1080 simulations that simultaneously vary the mass of a possible WDM candidate and the thermal history of the IGM. Interestingly, we find that a WDM particle mass of 4.5 keV provides a formally better fit to the small-scale power spectrum than Lambda-CDM, which motivates continued observational experiments of the Lyman-α forest.

Acknowledgments

All the work presented in this thesis was made possible only because of the joint effort with my fantastic collaborators, Brant Robertson, Piero Madau, and Evan Schneider. I am indebted for all their support, advice and encouragement.

First, I want to thank my advisor Brant Robertson. With Brant, I undertook the difficult job of extending *Cholla* to perform cosmological simulations. While this was a very challenging task, I always knew he was confident I would succeed in the end, and he was right! His consistent support as an advisor has had an enormous impact on my professional development and overall confidence. I'm also very grateful for Brant's patience, kindness, and understanding during difficult times when I doubted whether I was capable of completing my Ph.D. Your support has been key, and in a very significant part, I am where I am because of you Brant. I am also extremely grateful for everything I learned from you.

My co-advisor, Piero Madau, has also made an enormous impact not just in my work but also in my approach to solving problems. Piero re-sparked my interest and love for thinking about physics problems. I always enjoyed my regular meetings with him and Brant, during which my intuition for the physics of the IGM grew enormously.

My collaborator Evan Schneider has been a huge figure of inspiration. Her sharpness, confidence, and determination are qualities that I strive to achieve. I want to give special thanks to Evan because all my simulations are based on the amazing code that she wrote. As a developer, I appreciate the enormous effort that writing Cholla must have been, and it makes me very happy to see all the success you've earned from

it. I truly hope my contribution to Cholla will be helpful, but I know it pales compared to how much my work and career have benefited from your previous effort.

Again, I want to acknowledge how important Brant and Evan have been in impulsing my career. Thanks to them, I had the opportunity to use some of the largest supercomputers in the country, which not only was extremely fun, but I also gained invaluable experience that undoubtedly was critical for getting a great job doing what I love. Thank you Brant and Evan! I hope we continue to collaborate in developing Cholla.

I want to thank professor X Prochaska for been a valuable member of my qualifying and dissertation committee and for his science-related ideas, career advice and encouraging comments, which always made me feel more confident in my work. I want to acknowledge another important figure, Enrico Ramirez-Ruiz. When giving a workshop in Mexico with Enrico, he told me, "We have one day with the students... what can we do to change their life", this left a big impression on me. Your passion and courage are truly inspirational. Thank you Enrico for your advice and support, and thank you for our conversations from which I grew as a person.

I want to thank the Cholla-CAAR team for all the fun meetings. I want to especially acknowledge Reuben Budiardja, Trey White, and Damon McDougall, not only for their collaboration and contribution to Cholla but also for their help and guidance in writing better, faster (and stronger) code and in solving numerical problems.

I want to acknowledge my parents and my brother, who always have been an example to follow. I want to thank them for always supporting my goals. All my success

is thanks to everything I learned from my parents as a kid. Thanks to my brother for all his help when we lived in Mexico City.

Special thanks to all the great friends I made while at UCSC. Amanda, Enia, Dave, Viraj, Erica, Brittany, Zhou... Y claro también mis buenos amigos Cesar, Aldo, Nicolas y Ricardo. También muchísimas gracias a mi Chuski que quiero muchísimo. Thanks for all the laughs, the great moments, and adventures, but also thanks for the support during difficult times. You all occupy ample space in my heart.

Last but not least, I want to thank Lindsay Lauver and Libby Severson for their huge support as graduate advisors. I always felt like they have a genuine interest in the well-being and success of the students.

Published Material

The text of this dissertation includes reprints of the following previously published material, led by Bruno Villasenor Alvarez, with the permission of the listed coauthors. For all the publications, I was responsible for running the relevant simulation, analyzing the data, comparing to observations, and computing the results. All coauthors contributed to the design of the original idea, the interpretation of the physical results, and to the writing of the papers.

Chapter 2 was published in the literature as Villasenor et al. (2021a). Acknowledgment to coauthors Brant Robertson, Piero Madau, and Evan Schneider.

Chapter 3 was published in the literature as Villasenor et al. (2021b). Acknowledgment to coauthors Brant Robertson, Piero Madau, and Evan Schneider.

This work was supported in part by the UC MEXUS-CONACyT doctoral

fellowship. This work used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC05-00OR22725, using Summit allocations CSC434 and AST169. An award of computer time was provided by the INCITE program, via project AST175. I acknowledge use of the *lux* supercomputer at UC Santa Cruz, funded by NSF MRI grant AST1828315, and support from NASA TCAN grant 80NSSC21K0271.

To a life of challenge, superation, and happiness.

Chapter 1

Introduction

1.1 The Lyman-α Forest

Understanding how large-scale structures in the Universe develop over cosmic history represents one of the fundamental goals of modern astrophysics. The filamentary network of dark matter and gas known as the "cosmic web" encodes information about the relative abundance of baryons and dark matter, the evolution of the radiation emitted by galaxies and quasars, the expansion history of the Universe, the nature and properties of the dark matter particle and the mass of neutrinos, among other relevant physics (for a review, see McQuinn 2016). The Lyman-α forest provides a primary probe of the properties of the cosmic web (Gunn & Peterson 1965). The Lyman-α forest arises from a set of absorption lines observed in the spectra of distant quasars that originate when the traveling light from the quasar encounters clouds of neutral hydrogen in the intergalactic medium (IGM). Photons with the energy of the Lyman-α transition of hydrogen are scattered, producing the absorption features in the spectra

(e.g., Hernquist et al. 1996; Croft et al. 1998a; Madau et al. 1999; Meiksin 2009a; Slosar et al. 2011; McQuinn 2016; Worseck et al. 2019).

The Lyman-α forest comprises a powerful tool for cosmology as it allows us to detect the diffuse gas in the Universe, providing us with a blueprint of the cosmic web. The promise of the Lyman-α forest for constraining the nature of dark matter and dark energy has in part motivated the construction of the Dark Energy Spectroscopic Instrument (DESI), which will measure absorption line spectra from nearly a million of distant quasars enabling the measurement of the baryon acoustic oscillations in the forest with unprecedented precision (DESI Collaboration et al. 2016a,b). One of the principal statistics that can be obtained from the Lyman-α forest is the power spectrum of the fluctuations in the Lyman-α transmitted flux. On large scales, the power spectrum informs us of the ionization state of hydrogen in the IGM, while the small-scale power spectrum encodes information about the temperature of the cosmic gas, the possibility of a "warm" dark matter particle, and the mass of neutrinos from their effect suppressing small-scale structure formation (e.g., Lukić et al. 2015; Rorai et al. 2017; Sorini et al. 2018; Krolewski et al. 2018). These exciting possibilities will be measurable from the small scales of the power spectrum taken from large numbers of high-resolution spectra observed by new generation facilities like DESI.

1.2 Modeling the Properties of the IGM

To extract the physics encoded in the Lyman-α forest, one requires sophisticated cosmological hydrodynamical simulations that solve the dynamics of the dark

matter and the gas in the Universe (e.g., Cen et al. 1994; Hernquist et al. 1996). These simulations initiate with a matter distribution consistent with the statistical properties of the Cosmic Microwave Background (CMB) and evolve the gravitational collapse of the small initial density fluctuations into the large-scale matter distribution that we observe today (Planck Collaboration et al. 2020). To resolve to properties of the cosmic web in simulations, we must evolve boxes that represent large cosmological volumes ($\sim$ 100Mpc per side) with high resolution across the entire box ($\sim$ 40kpc) making this type of calculations exceptionally computationally demanding.

Comparing the statistical properties of the Lyman-α forest measured from the simulations to the observations allows us to distinguish which physical models reproduce the structure of the gas in the Universe (e.g., Oñorbe et al. 2017b; Gnedin et al. 2017; Davies et al. 2018a; Hiss et al. 2018; D'Aloisio et al. 2019; Faucher-Giguère 2020). Achieving this goal challenges our current state-of-the-art models as a plethora of yet unconstrained physical processes shape the distribution of cosmic gas, and so identifying the models that best reproduce the observations requires sampling over many simulations that span a wide variety of physical models. Running thousands of high-resolution cosmological simulations poses a colossal endeavor and can only be accomplished by making use of the largest supercomputers in the world that leverage massively parallel architectures in the form of graphical processing units (GPUs).

Using the largest supercomputers in the world requires specialized software that can run natively on GPUs. This requirement motivated the development of *Cholla* by Evan Schneider and Brant Robertson (Schneider & Robertson 2015). *Cholla* rep-

resents a state-of-the-art hydrodynamics solver designed to run entirely on GPUs. Using *Cholla* Schneider & Robertson were able to simulate a disk galaxy with unprecedented resolution and study the galactic winds around it (Schneider & Robertson 2017, 2018; Schneider et al. 2018, 2020). For my thesis, I have extended *Cholla* to run cosmological simulations. For this purpose, I developed the following physics modules into *Cholla*: (i) An FFT-based Poisson solver to compute the gravitational potential and account for self-gravity when solving the hydrodynamics. (ii) An N-Body integrator to evolve the dynamics of the dark matter particles. (iii) A comoving frame of reference to account for the expansion of the Universe when integrating the evolution of the gas and the dark matter, and (iv) a solver to evolve the ionization states of hydrogen and helium, accounting for the radiative cooling of the gas and photoheating from the radiation of the UV background. All of these components have been engineered to run entirely on GPUs, making *Cholla* about 100 times faster than regular CPU-based codes. Using Summit (Oak Ridge National Lab), which during the time that the work for this thesis took place, was the most powerful supercomputer in the United States, we have proven that *Cholla* scales well when running on more than 16,000 GPUs, and we ran over 1500 high-resolution cosmological hydrodynamical simulations, allowing us to study the properties of the intergalactic medium with unprecedented detail.

1.3 Impact of the UVB Radiation on the IGM

In the first Chapter 2 of this thesis, I present the methodology used to implement the cosmological framework into *Cholla*, and we perform a study of the impact

that different models for the evolution of the UV background have on the properties of the Lyman-α forest. The metagalactic UV background is the radiation from star-forming galaxies and active galactic nuclei (AGN) that permeates the IGM and ionizes and heats the cosmic gas during the epochs of hydrogen and helium reionization (e.g., Haardt & Madau 2012; Robertson et al. 2015; Madau & Haardt 2015; Faucher-Giguère 2020). Understanding how reionization occurred and the nature of the early sources that drove it forms one of the principal goals of modern astrophysics. Different models have been proposed for the time evolution of the UV background. Each of them results in a different reionization and thermal history of the IGM that affects the statistical properties of the Lyman-α forest. By running a couple of high-resolution simulations that adopt two models for the evolution of the UV background (Haardt & Madau (2012) and Puchwein et al. (2019)), we examined the structure of the Lyman-α forest produced by each model, and by comparing the power spectrum of the forest from the simulations to the observational determinations, we concluded that both models failed to reproduce the properties of the observed Lyman-α forest (Becker et al. 2013a; Viel et al. 2013a; Walther et al. 2018; Boera et al. 2019; Chabanier et al. 2019b). Therefore, the ionization and thermal histories of the IGM resulting from these models do not capture the real evolution of the IGM (Villasenor et al. 2021a).

Motivated by the significant results obtained from our first study, for our second project (Villasenor et al. 2021b), we aimed to modify the model for the UV background from Puchwein et al. (2019) to achieve a better match with the observed properties of the Lyman-α forest (Iršič et al. 2017b; Worseck et al. 2019; Boera et al.

2019; Chabanier et al. 2019b). To this end, we performed an extensive suite of 400 cosmological simulations that vary the photoionization and photoheating rates from the metagalactic background radiation and studied the different ionization and thermal histories of the IGM they computed. We performed a Markov Chain Monte Carlo (MCMC) analysis comparing the power spectrum of the Lyman-α forest and the effective optical depth of the He II Lyα forest from the simulation grid to several observational measurements and obtained a well-constrained model that reproduces the observations. From this analysis, we were able to infer the thermal and ionization evolution of the IGM. We obtained that hydrogen reionization completes by redshift $z \sim 6$ in agreement with recent determinations (Becker et al. 2001; Bosman et al. 2018; Becker et al. 2021; Qin et al. 2021). Additionally, our model results in gas temperatures consistent with other independent determinations from Bolton et al. (2014), Boera et al. (2019), Walther et al. (2019), Gaikwad et al. (2020), and Gaikwad et al. (2021). This analysis and its results are presented in Chapter 3 of this thesis.

1.4 Constraints on WDM from the Lyman-α Forest

It is overwhelmingly accepted that dark matter composes about 80% of all matter in the Universe, and the Λ-cold dark matter (ΛCDM) cosmological paradigm has proven immensely successful at matching a variety of observational determinations on large scales (e.g., Chabanier et al. 2019a; Planck Collaboration et al. 2020). However, on scales below a few Mpc, there exists tension between the predictions from the ΛCDM model and observations as it appears that ΛCDM may predict an excess of small- scale

systems than observed (Bullock & Boylan-Kolchin 2017). A possible solution for this discrepancy could be that the dark matter in the Universe is "warm" (Bode et al. 2001). Warm dark matter (WDM) particles with masses lighter than a few keV have significant intrinsic velocities. These velocities interfere with the gravitational collapse of systems and prevent the formation of structure on small scales, therefore suppressing small-scale fluctuations in the Lyα forest. For this reason, the forest provides excellent means to constrain models that describe the nature of dark matter (Seljak et al. 2006; Viel et al. 2013a; Baur et al. 2016; Garzilli et al. 2019, 2021). For *Villasenor et al. 2022a (in prep.)* we have run an unprecedented grid of 1,080 high-resolution cosmological simulations that vary the mass of the dark matter particle and the thermal evolution of the IGM. We compared the Lyman-α forest power spectrum from the simulations to the observed power spectrum down to the smallest scales ever probed (Boera et al. 2019). Interestingly, we find that a WDM particle mass of ~ 4.2 keV is preferred over the CDM model with a statistical significance of $2-3\,\sigma$. While our result is not significant enough to discard CDM as the prevailing model, our result motivates future experiments of the Lyman-α forest. This analysis and its results are presented in Chapter 4.

Chapter 2

Effects of Photoionization and Photoheating on Lyman-α Forest Properties from *Cholla* Cosmological Simulations

2.1 Introduction

The absorption signatures of neutral hydrogen gas provide important observational probes of cosmological structure formation (Gunn & Peterson 1965). The intergalactic medium traces the filamentary structure of the cosmic web, and the properties of H I absorption features (the "Lyman-α forest") reflect the temperature and density distribution of the medium that originate through the structure formation pro-

cess and the photoheating from ionizing sources (e.g., Madau et al. 1999). This paper presents the first results from the new *Cholla* IGM Photoheating Simulation (CHIPS) suite of cosmological simulations of the Lyman-α forest performed with the *Cholla* code (Schneider & Robertson 2015), comparing the statistics of the simulated Lyman-α forest calculated using different photoionization and photoheating histories with the available observational data at $z \sim 2 - 5$.

The Lyman-α forest originates in IGM gas that traces the matter density, and its properties inform us about the relative abundance of baryons and dark matter, the properties of dark matter including the matter power spectrum, the metagalactic radiation field, and the expansion history of the universe including the role of dark energy (for a review, see McQuinn 2016). The promise of the Lyman-α forest for constraining the nature of dark matter and dark energy has in part motivated the construction of the Dark Energy Spectroscopic Instrument, which will measure absorption line spectra backlit by quasars at $z > 2.1$ and detect baryon acoustic oscillations in the cosmic web (DESI Collaboration et al. 2016a,b).

Given its critical role as a probe of cosmic structure formation, the Lyman-α forest was an early subject of hydrodynamical cosmological simulations (e.g., Cen et al. 1994; Hernquist et al. 1996). The prospect of measuring quasar absorption spectra densely sampled on the sky over large statistical volumes has led to a resurgence of cosmic web studies in the literature (e.g., Lukić et al. 2015; Sorini et al. 2018; Krolewski et al. 2018). Owing to the power of DESI and other new spectroscopic facilities, the driving focus of theoretical efforts is to study the physics that affect the fine details of

the forest (e.g., Rorai et al. 2017). These physics include non-linear effects (Arinyo-i-Prats et al. 2015), environment (Tonnesen et al. 2017), and how the forest evolves to low redshifts (Khaire et al. 2019), but a consensus is building that the impact of IGM heating history on the temperature structure of the Lyman-α forest is the most critical to understand in detail (e.g., Hiss et al. 2018). The temperature structure affects most strongly the shape of the absorption profiles that provide information about the matter distribution, and without understanding the thermal structure of filaments the full power of Lyman-α absorption line studies cannot be realized.

The statistical properties of the Lyman-α forest are primarily measured via the "transmitted flux power spectrum", which probes fluctuations in the opacity (and therefore density, temperature and velocity field) of neutral hydrogen via transmission of flux from background quasars or galaxies. Similarly to the matter power spectrum, whose measurements extend to large (> 100 Mpc comoving) scales via galaxy spatial correlations, the baryonic acoustic oscillations are probed by the Lyman-α forest correlations at these large scales as well (e.g., Chabanier et al. 2019b). Structure in the Lyman-α forest extends down to scales of ~ 50 kpc comoving, where thermal pressure smoothing becomes important (Kulkarni et al. 2015a). Simultaneously capturing representative volumes while resolving the relevant spatial scales everywhere in the forest presents a challenging goal for cosmological simulations.

While dark matter and cosmological structure formation erect the scaffolding of the cosmic web, the temperature structure of the IGM depends on the competition between heating, radiative cooling, and adiabatic cooling via universal expansion. Heating

of the IGM predominately occurs via photoheating from excess energy deposited during the photoionization of (most importantly) the H I and He II species. Observationally, the Lyman-α forest probes redshifts $z < 6$ when H I has mostly been ionized. Between $4 \lesssim z \lesssim 6$, the IGM temperature declines from a local maximum at the end of H I reionization from which the IGM thermal structure inherits residual signatures (Oñorbe et al. 2017b; Gnedin et al. 2017; Davies et al. 2018a; D'Aloisio et al. 2019; Faucher-Giguère 2020). At redshifts $z \lesssim 4$, photoheating from the gradual ionization of He II from quasars leads to a maximum IGM temperature sometime around $z \sim 3$ (La Plante et al. 2017; Schmidt et al. 2018). The low-redshift ($2 \lesssim z \lesssim 4$) IGM is therefore heavily influenced by He II reionization (Worseck et al. 2016), and the helium Lyman-α forest (La Plante et al. 2018) provides critical information on the ionizing flux from quasars (La Plante et al. 2017; Schmidt et al. 2018).

At higher redshifts the ionizing flux from galaxies becomes increasingly important. The various transitions of the hydrogen Lyman series provide details on the ionization state of the gas, and can constrain the post-reionization ionizing background (Davies et al. 2018a). The hydrogen reionization process heats the IGM sufficiently to leave residual signatures in the structure of the filaments (D'Aloisio et al. 2019). The thermal evolution of the IGM, reflecting early $z \sim 6$ heating from galaxies during H reionization and late $z \sim 2 - 4$ heating from QSOs during helium reionization, can therefore be probed through the Lyman-α forest power spectrum (Walther et al. 2019).

By changing the thermal history of the IGM, the process of cosmic reionization at $z > 6$ couples to the observed properties of the Lyman-α forest on small scales.

Probes of reionization have become increasingly powerful, including quasar proximity zones (Eilers et al. 2017) and the IGM damping wing (Davies et al. 2018b), the high-redshift forest and post-reionization IGM (Oñorbe et al. 2017b; Gnedin et al. 2017), and Lyman-α transmission spikes (Garaldi et al. 2019; Gaikwad et al. 2020). The physics of reionization has driven a host of cosmological simulation efforts (e.g., Gnedin & Kaurov 2014; Kaurov & Gnedin 2014, 2015; Trac et al. 2015; Gnedin 2016; Oñorbe et al. 2017a; Doussot et al. 2019), but more work is required to connect these simulations to the physics of the IGM at lower redshifts. Capturing fluctuations in the metagalactic background (D'Aloisio et al. 2018) and the potential impact of rare AGN (D'Aloisio et al. 2017) require large volumes ($L \sim 100h^{-1}$Mpc), but simultaneously maintaining high spatial resolution in the IGM is computationally demanding.

Cholla models the baryionic component on a uniform Cartesian grid, and while other approaches to solve the hydrodynamics can be employed such as SPH or AMR, for which the computational power is commonly focused in solving the high density regions. These approaches present some disadvantages compared to a uniform grid when applied to the study of the IGM. For instance, the gas responsible for the Lyman-α forest is low density gas that spans over most of the volume of the box, making the use of AMR unnecessary and inefficient. Compared to an SPH approach, grid methods exhibit other advantages such as a clearly defined spatial resolution instead of a fixed mass resolution and generally a more accurate treatment of shocks and hydrodynamics. More importantly, a uniform grid achieves a better sampling of the IGM when compared to an AMR or SPH implementation that uses the same computational resources. Although,

the advantages of a uniform grid come at a high computational cost if a high resolution is maintained over large volumes.

To this end, our new CHIPS simulation suite uses the Graphics Processing Unit (GPU)-native *Cholla* code (Schneider & Robertson 2015, 2017) to perform high-resolution simulations of the cosmic web to achieve simultaneously the resolution required to model the thermal structure of the Lyman-α forest ($\Delta x \approx 35$ kpc) over large volumes ($L \approx 75$Mpc). We will study how different H I + He II photoionization and photoheating histories shape the thermal structure of the IGM.

This first paper presents Lyman-α forest results for the widely-used Haardt & Madau (2012) photoionization and photoheating model, as well as the more recent Puchwein et al. (2019) implementation that has a similar emissivity but for which H I reionization completes later ($z \sim 6$). Section 2.2 presents our numerical methodology for performing the cosmological simulations including our new extensions to the *Cholla* code that enable, self-gravity, dark matter particle integration, and coupling to the GRACKLE heating and cooling library (Smith et al. 2017). Section 2.2.9 provides a high-level summary of the algorithm used in our cosmological simulations. Section 2.3 presents several validation tests, including a new validation test for the dual-energy formalism when modeling cosmological structure formation. We present the first simulations of the CHIPS suite in §2.4, including the cosmological parameters, resolution, and box sizes. Our scientific results for the properties of the IGM are reported in §2.5. We discuss our results in §2.6, and summarize and conclude in §2.7. Finally, we demonstrate the numerical convergence of our results in Appendix 2.8 and perform a cosmological

parameter study in Appendix 2.9

2.2 Methodology

To simulate the Lyman-α -forest, we engineered substantial extensions to the *Cholla* code. These additions included implementing a cosmological framework to account for the expansion history of the universe (§2.2.1), including changes to the model of gas dynamics (§2.2.1) and the coordinate system (§2.2.1), and are discussed below. We briefly review the *Cholla* hydrodynamical integrator (§2.2.2) and the dual energy formalism (§2.2.3) that allows for accurate evolution of the gas internal energy in Eulerian cosmological simulations (e.g., Bryan et al. 1995). We present our new implementations of solvers for the gravitational force and particle motions in §2.2.4 and §2.2.5, respectively. Cooling and heating from a UV background are now treated using the GRACKLE library (Smith et al. 2017), and are detailed in §2.2.6 and §2.2.7. Adjustments to the time step calculation to account for particle motions are described in §2.2.8. We conclude the review of our methods with a summary of the overall algorithm in §2.2.9.

2.2.1 Cosmological Framework

For cosmological simulations, the gas follows the equations of hydrodynamics in a frame comoving with the expanding universe. To convert from the comoving to the physical system, the scale factor a is introduced and provides a distance transformation between the two systems, with coordinates in the proper system $\mathbf{r}$ related to comoving

coordinates $\mathbf{x}$ by $\mathbf{r} = a\mathbf{x}$. The rate of change of the scale factor corresponds to the expansion rate of the universe and follows the Friedmann equation given by

$$H = \frac{\dot{a}}{a} = H_0\sqrt{\frac{\Omega_M}{a^3} + \Omega_\Lambda + \frac{\Omega_k}{a^2}}, \tag{2.1}$$

where H is the Hubble parameter that quantifies the expansion rate of the universe $\dot{a} = da/dt$, and H_0, Ω_M, Ω_Λ and Ω_k are the cosmological parameters that correspond to the current expansion rate of the universe and its matter, dark energy, and curvature content, respectively. Given an initial value of the scale factor a, the Friedmann equation provides a relation between the scale factor and cosmic time, and therefore the scale factor can be used as a time-like variable.

Gas Dynamics

Consider the hydrodynamical quantities of comoving baryon density ρ_b, proper peculiar velocity $\mathbf{v}$, and total specific energy E in the proper frame. The relation between comoving and proper densities is $\rho = a^3\rho_{proper}$. In this system, the basic equations of hydrodynamics include the continuity equation

$$\frac{\partial\rho_b}{\partial t} = -\frac{1}{a}\nabla\cdot(\rho_b\mathbf{v}), \tag{2.2}$$

the force-momentum equation

$$\frac{\partial a\rho_b\mathbf{v}}{\partial t} = -\nabla\cdot(\rho_b\mathbf{v}\mathbf{v}) - \nabla p + \rho_b\mathbf{g}, \tag{2.3}$$

where the pressure p transforms to the proper pressure by the relation $p = a^3 p_{proper}$ and $\mathbf{g}$ is the gravitational acceleration, and the energy equation

$$\begin{aligned}\frac{\partial a^2 \rho_b E}{\partial t} &= -a\nabla \cdot (\rho_b \mathbf{v} E + p\mathbf{v}) + a\rho_b \mathbf{v} \cdot \mathbf{g} \\ &\quad + a\dot{a}[(2 - 3(\gamma - 1))\rho_b e] + a(\Gamma - \Lambda),\end{aligned} \tag{2.4}$$

where Γ and Λ correspond to the heating and cooling rates, respectively.

From the specific total energy one can obtain the specific internal energy e in the proper system by subtracting the kinetic energy per unit mass $e = E - v^2/2$. In Eulerian cosmological simulations where gas often flows supersonically, the above equations can be supplemented by a dual energy formalism (Bryan et al. 1995) in which the internal energy is additionally followed. The supplemental internal energy is then used in cells where the computation of the internal energy from the total energy is expected to be inaccurate (see §2.2.3 for details).

The evolution of the specific internal energy e is given by

$$\begin{aligned}\frac{\partial a^2 \rho_b e}{\partial t} &= -a\nabla \cdot (\rho_b \mathbf{v} e) - ap\nabla \cdot \mathbf{v} \\ &\quad + a\dot{a}[(2 - 3(\gamma - 1))\rho_b e] + a(\Gamma - \Lambda)\end{aligned} \tag{2.5}$$

The relation between the pressure and the internal energy is given by the equation of state $p = (\gamma - 1)\rho_b e$.

For simplicity, first we limit the description of the hydrodynamics solver to the adiabatic case ($\Gamma = \Lambda = 0$), and delay a description of the radiative cooling implementation to §2.2.6. In the particular case of a $\gamma = 5/3$ gas, the adiabatic energy equations

simplify to

$$\begin{aligned} \frac{\partial a^2 \rho_b E}{\partial t} &= -a\nabla \cdot (\rho_b \mathbf{v} E + p\mathbf{v}) + a\rho_b \mathbf{v} \cdot \mathbf{g}, \\ \frac{\partial a^2 \rho_b e}{\partial t} &= -a\nabla \cdot (\rho_b \mathbf{v} e) - ap\nabla \cdot \mathbf{v}. \end{aligned} \tag{2.6}$$

From Equations 2.2, 2.3, and 2.6 it follows that for a uniform expanding universe, the comoving density ρ_b will remain constant, the peculiar velocity $\mathbf{v}$ will decrease as a^{-1} and the specific energies E and e will decrease as a^{-2}.

Super-comoving Coordinates

A convenient approach for the implementation of the comoving coordinate system is to define a new set of coordinates that simplify Equations 2.2, 2.3, and 2.6 such that the scale factor a does not explicitly appear. A detailed description of these "super-comoving coordinates" can be found in Martel & Shapiro (1998) and are used for cosmological simulations in the *Ramses* code by Teyssier (2002). The transformation to the new system of coordinates is given by

$$\begin{aligned} d\tilde{t} &\equiv H_0 \frac{dt}{a^2}, & \tilde{\mathbf{v}} &\equiv a\frac{\mathbf{v}}{H_0}, \\ \tilde{E} &\equiv a^2 \frac{E}{H_0^2}, & \tilde{e} &\equiv a^2 \frac{e}{H_0^2}, \\ \tilde{\rho_b} &\equiv \rho_b, & \tilde{p} &\equiv a^2 \frac{p}{H_0^2} = (\gamma - 1)\tilde{\rho_b}\tilde{e}, \\ \tilde{\phi} &\equiv a^2 \frac{\phi}{H_0^2}, & \tilde{\mathbf{g}} &= -\nabla\tilde{\phi} = a^2 \frac{\mathbf{g}}{H_0^2}. \end{aligned} \tag{2.7}$$

Throughout we will denote super-comoving variables with a tilde, e.g., $\tilde{\phi}$. After the transformation to the super-comoving system of coordinates, the equations of adiabatic

hydrodynamics for a $\gamma = 5/3$ gas can be written as

$$\frac{\partial \tilde{\rho}_b}{\partial \tilde{t}} = -\nabla \cdot (\tilde{\rho}_b \tilde{\mathbf{v}}) \tag{2.8}$$

$$\frac{\partial \tilde{\rho}_b \tilde{\mathbf{v}}}{\partial \tilde{t}} = -\nabla \cdot (\tilde{\rho}_b \tilde{\mathbf{v}} \tilde{\mathbf{v}}) - \nabla \tilde{p} + \tilde{\rho}_b \tilde{\mathbf{g}} \tag{2.9}$$

$$\frac{\partial \tilde{\rho}_b \tilde{E}}{\partial \tilde{t}} = -\nabla \cdot (\tilde{\rho}_b \tilde{\mathbf{v}} \tilde{E} + \tilde{p} \tilde{\mathbf{v}}) + \tilde{\rho}_b \tilde{\mathbf{v}} \cdot \tilde{\mathbf{g}} \tag{2.10}$$

$$\frac{\partial \tilde{\rho}_b \tilde{e}}{\partial \tilde{t}} = -\nabla \cdot (\tilde{\rho}_b \tilde{\mathbf{v}} \tilde{e}) - \tilde{p} \nabla \cdot \tilde{\mathbf{v}} \tag{2.11}$$

The set of equations resulting from the transformation are the same as the original equations of hydrodynamics in a non-expanding system. This formulation allows the extension of the hydrodynamics solver to an expanding frame system without any significant changes to the original solver.

2.2.2 Hydrodynamics Solver

Without the gravitational source terms, Equations 2.8-2.11 correspond to the conserved form of the Euler equations. A detailed description of the methodology used for solving the gravity-free fluid dynamics can be found in Schneider & Robertson (2015). The hydrodynamics solver is a Godunov-based method for which an approximation to the cell averaged values of the conserved quantities $\mathbf{U} = [\rho, \rho\mathbf{v}, \rho E, \rho e]$ are evolved using a numerical discretization of the Euler equations, given by

$$\frac{\mathbf{U}_i^{n+1} - \mathbf{U}_i^n}{\Delta t} + \frac{\mathbf{F}_{i+1/2}^{n+1/2} - \mathbf{F}_{i-1/2}^{n+1/2}}{\Delta x} = 0 \tag{2.12}$$

where $\mathbf{U}_i^n$ denotes the average value of the conserved quantities for cell i at time-step n. The change of the conserved quantities in cell i is given by the time centered fluxes across the cell interfaces $\mathbf{F}_{i\pm1/2}^{n+1/2}$. The flux components $\mathbf{F} = [\rho\mathbf{v}, \rho\mathbf{v}\mathbf{v}, (\rho E + p)\mathbf{v}, \rho e\mathbf{v}]$ are computed by solving the Riemann problem at the cell interfaces using the reconstructed values of the conserved quantities obtained via a Piecewise Parabolic Method (PPM; Colella & Woodward 1984). The PPM scheme is third-order accurate in space and second-order accurate in time.

2.2.3 Dual Energy Implementation

Owing to the supersonic flows from structure formation and adiabatic cooling of gas from universal expansion, regions where the gas kinetic energy is much larger than the internal energy are common in cosmological simulations. Under these conditions, calculation of the internal energy $E - v^2/2$ can be affected by numerical errors. These errors can be ameliorated by using a "dual energy formalism" (Bryan et al. 1995), where the internal energy is evolved separately via Equation 2.5, or the corresponding simplified Equation 2.11, and substituted for the internal energy computed from the total and kinetic energies when appropriate. The two terms on the right side of Equation 2.11 correspond to advection and compression terms, respectively. To reconcile the total internal energy E with the separately tracked internal energy e, at the end of each time step a condition is applied on a cell to cell basis to select which value of the internal energy to employ. We adopt a condition similar to that used in *Enzo* (Bryan et al. 2014), where the selection is based on the fraction of the internal energy in a given cell relative to the maximum of the total energy in a the neighborhood of the cell. Mathematically,

this condition is given by

$$e_i = \begin{cases} e_i, & \rho_i \left(E - \mathbf{v}^2/2\right)_i / \max(\rho E)_i < \eta \\ \left(E - \mathbf{v}^2/2\right)_i, & \text{otherwise} \end{cases}, \tag{2.13}$$

where $\max(\rho E)_i$ is the maximum total specific energy in the local and adjacent cells. In one dimension, $\max(\rho E)_i = \max\left[(\rho E)_{i-1}, (\rho E)_i, (\rho E)_{i+1}\right]$. At the end of every time step, after applying Equation 2.13, the total specific energy E is synchronized with the selected internal energy by setting $E = e + v^2/2$.

The value of η should be chosen carefully, as setting η too low will allow spurious heating owing to numerical errors introduced in the total energy evolution. If η is set to high then shock heating in regions where the gas flows converge could be suppressed since the advected internal energy e will be preferentially selected over the conserved internal energy $E - v^2/2$, and Equations 2.5 and 2.11 do not capture shock heating. To estimate an appropriate value for η in cosmological simulations, we developed a test to evaluate how the dual energy condition affects the average cosmic gas temperature, as described below in §2.3.4. Based on the results of this test, we set $\eta = 0.035$.

Another approach on the selection criteria for the internal energy is presented in Teyssier (2015). Here, the conserved internal energy $\rho(E - v^2/2)$ is compared to an estimate of the numerical truncation error

$$e_{\text{trunc}} \simeq \frac{1}{2}\rho(\Delta v)^2, \tag{2.14}$$

where Δv corresponds to the difference of the velocities in the neighboring cells. The selection condition for this scheme is given by

$$e_i = \begin{cases} \left(E - \mathbf{v}^2/2\right)_i & , \rho_i \left(E - \mathbf{v}^2/2\right)_i > \beta e_{\rm trunc} \\ e_i & , \text{ otherwise} \end{cases} , \tag{2.15}$$

where β is a numerical parameter with suggested value $\beta = 0.5$. We also evaluated this dual energy condition using the average cosmic temperature test described in §2.3.4. As we discuss below, we found that Equation 2.15 can be overly restrictive by predominately selecting the advected internal energy e over the conserved internal energy $E - v^2/2$, thereby suppressing shock heating inside collapsed halos and significantly lowering the average cosmic temperature.

2.2.4 Gravity

The gravitational acceleration vector $\mathbf{g}$ is computed by differentiating the gravitational potential ϕ. The potential is obtained from the solution of the Poisson equation. In the comoving coordinates, the Poisson equation is written as

$$\nabla^2 \phi = \frac{4\pi G}{a}(\rho - \bar{\rho}), \tag{2.16}$$

where G is the gravitational constant, $\rho = \rho_{\rm DM} + \rho_b$ is the total dark plus baryonic matter density, and $\bar{\rho}$ is the average value of the total density over the entire box.

Integration of Equation 2.16 can be directly performed in Fourier space. In

k-space, the Poisson equation simplifies to

$$\hat{\phi}(\mathbf{k}) = G(\mathbf{k})\hat{\rho}(\mathbf{k}), \tag{2.17}$$

where $G(\mathbf{k})$ is the Greens function, which for a second-order centered two point finite difference discretization corresponds to (Hockney & Eastwood 1988)

$$G(\mathbf{k}) = -\frac{\Delta x_h^2}{\sin^2\left(k_x \Delta x_h\right)} - \frac{\Delta y_h^2}{\sin^2\left(k_y \Delta y_h\right)} - \frac{\Delta z_h^2}{\sin^2\left(k_z \Delta z_h\right)}. \tag{2.18}$$

Here $\Delta x_h = \Delta x/2$, $\Delta y_h = \Delta y/2$, and $\Delta z_h = \Delta z/2$, where Δx, Δy, and Δz are the grid cell dimensions. To compute the three-dimensional fast Fourier transforms (FFTs) we use *PFFT* (Pippig 2013), a publicly available library for performing FFTs with a box domain decomposition.

From the potential ϕ we compute the gravitational acceleration vector $\mathbf{g} = -\nabla\phi$. The derivatives along each direction are obtained using a fourth-order centered four-point finite difference approximation. In one dimension, the derivative is given by

$$\frac{\partial \phi_i}{\partial x} = \frac{1}{12\Delta x}\left(\phi_{i-2} - 8\phi_{i-1} + 8\phi_{i+1} - \phi_{i+2}\right). \tag{2.19}$$

The terms corresponding to the gravitational sources, $\tilde{\rho}_b\tilde{\mathbf{g}}$ and $\tilde{\rho}_b\tilde{\mathbf{v}}\cdot\tilde{\mathbf{g}}$ in Equations 2.9 and 2.10, are added to the momentum and total energy after the conserved variables have been updated by the hydro solver (i.e., after Equation 2.12 has been solved). The numerical implementation for the coupling of the momentum and energy with gravity is given by

$$(\tilde{\rho}_b\tilde{\mathbf{v}})_i^{n+1} = (\tilde{\rho}_b\tilde{\mathbf{v}})_i^{n+1*} + \frac{1}{2}\Delta t^n \left(\tilde{\rho}_{b,i}^n + \tilde{\rho}_{b,i}^{n+1}\right)\tilde{\mathbf{g}}_i^{n+1/2}, \tag{2.20}$$

$$(\tilde{\rho}_b\tilde{E})_i^{n+1} = (\tilde{\rho}_b\tilde{E})_i^{n+1*} + \frac{1}{2}\Delta t^n \left[(\tilde{\rho}_b\tilde{\mathbf{v}})_i^n + (\tilde{\rho}_b\tilde{\mathbf{v}})_i^{n+1*}\right]\tilde{\mathbf{g}}_i^{n+1/2}. \tag{2.21}$$

Here the superscript $n+1*$ refers to the value of the conserved quantity after the hydrodynamics solver update and the time centered value of the gravitational field $\tilde{\mathbf{g}}_i^{n+1/2} = -\nabla\tilde{\phi}_i^{n+1/2}$. The potential $\tilde{\phi}_i^{n+1/2}$ is obtained by extrapolation from ϕ_i^n and ϕ_i^{n-1} using

$$\tilde{\phi}_i^{n+1/2} = \frac{(a^n)^2}{H_0^2}\left[\phi_i^n + \frac{\Delta t^n}{2\Delta t^{n-1}}(\phi_i^n - \phi_i^{n-1})\right]. \tag{2.22}$$

2.2.5 Dark Matter

We represent the cold dark matter as a system of discrete point-mass particles moving under the influence of gravity. Each dark matter particle is described by its mass m_i, comoving position $\mathbf{x}_i$, and peculiar velocity $\mathbf{v}_i$. The evolution of the particle trajectories in a comoving frame is described by

$$\frac{d\mathbf{x}_i}{dt} = \frac{1}{a}\mathbf{v}_i \tag{2.23}$$

$$\frac{d\left(a\mathbf{v}_i\right)}{dt} = \mathbf{g}_i, \tag{2.24}$$

where $\mathbf{g}_i$ is the acceleration vector owing to the gravitational field evaluated at the particle position $\mathbf{x}_i$.

To solve Equation 2.16, we compute the contribution of the dark matter particles to the density field by interpolating onto the same grid used to evolve the hydrodynamical quantities. The dark matter density $\rho_{\rm DM}$ is calculated via a cloud-in-cell scheme (Hockney & Eastwood 1988), for which each particle is represented as a cube having the same size as one grid cell Δx and uniform density $m_i/\Delta x^3$. The mass of the particle is distributed among the grid cells that intersect its volume such that the fraction of the particle mass δm_i deposited on a cell is equal to the fraction of its intersected volume.

In one dimension, the mass contribution of a particle to a cell at position x_c is given by

$$\delta m_{i,c} = m_i \begin{cases} 1 - |x_i - x_c|/\Delta x, & |x_i - x_c| < \Delta x \\ 0, & \text{otherwise} \end{cases} . \tag{2.25}$$

The gravitational acceleration $\mathbf{g}_i$ evaluated at a particle position $\mathbf{x}_i$ must be computed in a manner consistent with the particle density interpolation. For the cloud-in-cell scheme, to avoid self-forces on the particles each component of $\mathbf{g}$ should be interpolated with the same weights used during the density assignment calculation.

To integrate the particle trajectories we use the kick-drift-kick (KDK) method (Miniati & Colella 2007), consisting of three steps to update the particles position and velocity from time-step n to time-step $n+1$. The sequence of variable updates is

$$\mathbf{v}_i^{n+1/2} = \frac{1}{a^{n+1/2}} \left(a^n \mathbf{v}_i^n + \frac{\Delta t^n}{2} \mathbf{g}_i^n \right), \tag{2.26}$$

$$\mathbf{x}_i^{n+1} = \mathbf{x}_i + \frac{\Delta t^n}{a^{n+1/2}} \mathbf{v}_i^{n+1/2}, \tag{2.27}$$

$$\mathbf{v}_i^{n+1} = \frac{1}{a^{n+1}} \left(a^{n+1/2} \mathbf{v}_i^{n+1/2} + \frac{\Delta t^n}{2} \mathbf{g}_i^{n+1} \right). \tag{2.28}$$

The KDK scheme allows for variable timesteps, as required by cosmological simulations owing to the variation in gas and particles velocities as the simulation advances. This sympletic scheme conserves an integral of motion on average, preventing an accumulation of errors and maintaining the phase space trajectory of the particles.

2.2.6 Chemistry and Radiative Cooling

We integrated *Cholla* with the GRACKLE chemistry and cooling library (Smith et al. 2017) to solve a non-equilibrium chemical network. Currently, our method only tracks the atomic chemical species and metals, but it could be extended to include, e.g., molecular hydrogen and deuterium.

The chemical species (H I, H II, He I, He II, He III, electrons e^-, and metals Z) are advected as scalar fields alongside the gas conserved variables via Equation 2.12. For details about the implementation of GRACKLE, we refer the reader to Smith et al. (2017). During every time step, GRACKLE updates the ionization fractions and computes the net heating and cooling by sub-cycling the rate equations within one hydrodynamic step. The sub-cycling updates the chemical and thermal states of the gas on timescales smaller than the dynamical timescales. For the atomic H and He chemical network, GRACKLE directly computes the heating and cooling rates accounting for collisional excitation and ionization, recombination, free-free emission, Compton scattering from the cosmic microwave background, and photoheating from a metagalactic UV background. GRACKLE accounts for metals by using precomputed tables for the metallic cooling and heating rates.

The GRACKLE update routine is applied at the end of each time step, af-

ter the gas conserved variables have been updated by the hydro solver and additional gravitational source terms. This routine updates the ionization fraction of the chemical elements, and also adds the net cooling and heating to the internal energy by setting

$$(\rho e)_i^{n+1} \rightarrow (\rho e)_i^{n+1} + a\Delta t^n (\Gamma - \Lambda)_i^n \tag{2.29}$$

Finally, the total energy is updated to reflect the change in the internal energy due to the net cooling as

$$E_i^{n+1} = \frac{1}{2}(v_i^{n+1})^2 + e_i^{n+1}. \tag{2.30}$$

2.2.7 UVB Ionization and Heating

The non-equilibrium GRACKLE solver accounts for the ionization of the primordial chemical species owing to a uniform time-dependent UV background by loading tables of the redshift dependent photoionization and photoheating rates for H I, He I, and He II. We compute the photoionization rates from a given redshift dependent spectrum as

$$\Gamma_{\gamma i}(z) = \int_{\nu_i}^{\infty} \frac{4\pi J(\nu, z)}{h\nu} \sigma_i(\nu) d\nu, \tag{2.31}$$

where $J(\nu, z)$ is the intensity of the UV background at frequency ν (in $\mathrm{erg\,s^{-1}\,cm^{-2}\,sr^{-1}\,Hz^{-1}}$), and ν_i and $\sigma_i(\nu)$ are the threshold frequency and cross-section for photoionization of the species i, taken from Osterbrock (1989). Analogously, the photoheating rates are computed as

$$\epsilon_i(z) = \int_{\nu_i}^{\infty} \frac{4\pi J(\nu, z)}{h\nu} (h\nu - h\nu_i) \sigma_i(\nu) d\nu. \tag{2.32}$$

If metal line cooling is included, the contributions of metals to the heating and cooling rates are accounted by GRACKLE by loading precomputed density, temperature and redshift dependent lookup tables that were obtained by providing the UVB spectrum to the CLOUDY (Ferland et al. 2017) photoionization code (version 17.02). The tables for metallic heating and cooling rates were generated for solar metallicity under the assumption of ionization equilibrium and subtracting the contributions of primordial heating and cooling as described in Smith et al. (2017). The resulting tables are organized into a Hierarchical Data Format (version 5) file readable by GRACKLE.

2.2.8 Time Step Calculation

The simulation time step Δt is computed with constraints from the signal speed of the gas, the motion of the dark matter particles, and the expansion of the universe. For the gas, the time step is constrained by the gas velocities $\mathbf{v}$ and the sound speed c_s as

$$\Delta t_{\rm gas} = \alpha_{\rm gas} \min\left(\frac{a\Delta x}{|v_x| + c_s}, \frac{a\Delta y}{|v_y| + c_s}, \frac{a\Delta z}{|v_z| + c_s}\right), \tag{2.33}$$

where $\alpha_{\rm gas}$ is the CFL factor specified by the user ($\alpha_{\rm gas} = 0.3$ by default) and $|v| + c_s$ is evaluated over the entire grid for each direction to find the minimum value of $\Delta t_{\rm gas}$. For the particles, the time step is limited to avoid any displacement larger than the cell size in each direction using

$$\Delta t_{\rm DM} = \alpha_{\rm DM} \min\left(\frac{a\Delta x}{|v_x|}, \frac{a\Delta y}{|v_y|}, \frac{a\Delta z}{|v_z|}\right), \tag{2.34}$$

where $\alpha_{\rm DM}$ is analogous to the CFL factor ($\alpha_{\rm DM} = 0.3$ by default) and $|v|$ is evaluated over all the particles for each direction. The time step is also limited by the expansion of the universe by choosing $\Delta t_{\rm exp}$ such that the fractional change in the scale factor does not exceed 1% ($\Delta a_{\rm exp} = 0.01a$). The actual time step is selected by taking the smallest value

$$\Delta t = \min\left(\Delta t_{\rm gas}, \Delta t_{\rm DM}, \Delta t_{\rm exp}\right), \tag{2.35}$$

guaranteeing that all three limiting conditions described above are satisfied.

The time step Δt (Eq. 2.35) is applied globally to update all the cells and particles in the box. When running high resolution simulations, it is possible to have a situation in which for a single cell $\Delta t_{{\rm gas},i}$ is extremely small compared to the all the other cells, resulting in small values for Δt which significantly slow the entire simulation. To avoid this situation, when a cell satisfies the condition $\Delta t_{{\rm gas},i} < \Delta t_{\rm DM}/50$ then the conserved quantities of that cell (density, momentum, energy, and internal energy) are replaced with the conserved quantities averaged over the six closest neighboring cells, resulting in a larger $\Delta t_{{\rm gas},i}$ for such cell and avoiding extremely small steps. We keep track of this occurrences during the full run, and for the high resolution simulations presented in this work this situation happens less than a dozen times per simulation, ensuring that the dynamics of the gas are not significantly affected by these small interventions.

2.2.9 Algorithm Implementation

The complete method to evolve the gas and the dark matter particles from time-step n to time-step $n+1$ can be summarized by the following algorithm:

Initialization:

1. Load initial conditions for the gas conserved variables and the particle positions and velocities.

2. Obtain the dark matter density $\rho_{\rm DM}$ by interpolating the particle masses onto the grid via the Cloud-In-Cell method described in §2.2.5.

3. Compute the gravitational potential ϕ by solving the Poisson equation (Eqn. 2.16), using the dark matter density $\rho_{\rm DM}$ and the gas density ρ_b as the sources.

4. Calculate the gravitational field $\mathbf{g} = -\nabla\phi$ at the centers of the grid cells using a fourth-order finite difference scheme (Eqn. 2.19) and interpolate the acceleration vector evaluated at the particles positions $\mathbf{g_i} = \mathbf{g}(\mathbf{x}_i)$.

Time Step Update:

1. Compute the current time step Δt^n.

2. Obtain the gravitational potential at $t^{n+1/2}$ by extrapolation using ϕ^n and ϕ^{n-1}, Equation 2.22.

3. Advance the gas conserved quantities by Δt^n using the intercell fluxes $\mathbf{F}^{n+1/2}$, Equation 2.12.

4. Add the gravitational sources to the gas momentum and energy, Equation 2.20.

5. Call GRACKLE to update the ionization states of the chemical network and add the net cooling and heating to the internal energy, Equation 2.29.

6. Advance the particle velocities by $\frac{1}{2}\Delta t^n$ and use the updated velocities $\mathbf{v}_i^{n+1/2}$ to advance the particle positions by Δt^n, Equations 2.26 and 2.27.

7. Obtain the dark matter density $\rho_{\rm DM}^{n+1}$ via the CIC method.

8. Compute the gravitational potential ϕ^{n+1} by solving Equation 2.16 with $\rho_{\rm DM}^{n+1}$ and ρ_b^{n+1} as sources.

9. Obtain the gravitational field $\mathbf{g}^{n+1}$ at the cell centers and $\mathbf{g}_i^{n+1}$ evaluated at the particle positions.

10. Advance the particle velocities by $\frac{1}{2}\Delta t^n$ resulting in $\mathbf{v}_i^{n+1}$, Equation 2.28.

Currently, the extensions included into *Cholla* for cosmological simulations (the FFT based Poisson solver, the dark matter particles integrator and the chemical network solver) all are implemented to run in the host CPUs, while the hydrodynamics solver including the advection of the ionization states of H and He run in the GPUs. At the time of submission of this work, an entirely GPU based distributed FFT solver has been recently integrated into *Cholla* and development to transfer both the particle integrator and the H+He network solver to the GPUs is ongoing. Potentially, the GPU implementation of a H+He network solver analogous to GRACKLE could result in a significant performance increase since currently the GRACKLE call to update the

chemical network is the slowest step in our implementation. The current version of *Cholla* no longer uses GRACKLE. Instead, a GPU implementation H+He solver that applies the same methodology as GRACKLE has been developed into *Cholla*.

2.3 Validation

To test the extensions of the *Cholla* code for cosmological simulations, we present below a set of validation exercises including comparisons with other publicly available Eulerian codes. In §2.3.1 we present the standard Zel'Dovich (1970) test. We then compare in §2.3.2 the matter power spectra of N-body cosmological simulations performed with *Cholla* to results from the *Nyx* (Almgren et al. 2013), *Ramses* (Teyssier 2002), and *Enzo* (Bryan et al. 2014) codes using the same initial conditions, and find sub-percent-level agreement at all spatial scales when simulated with the same resolution. We extend these tests to adiabatic hydrodynamical cosmological simulations in §2.3.3, where we find agreement within a few percent. To test the dual energy formalism in cosmological simulations, we describe a new test that computes the mean gas temperature with redshift (§2.3.4), and show that our choice of dual energy paramterization and parameter values recovers model expectations. We validate our cosmological hydrodynamical simulations including cooling, chemistry, and heating against *Enzo* simulations using the same physical prescription, and find good agreement.

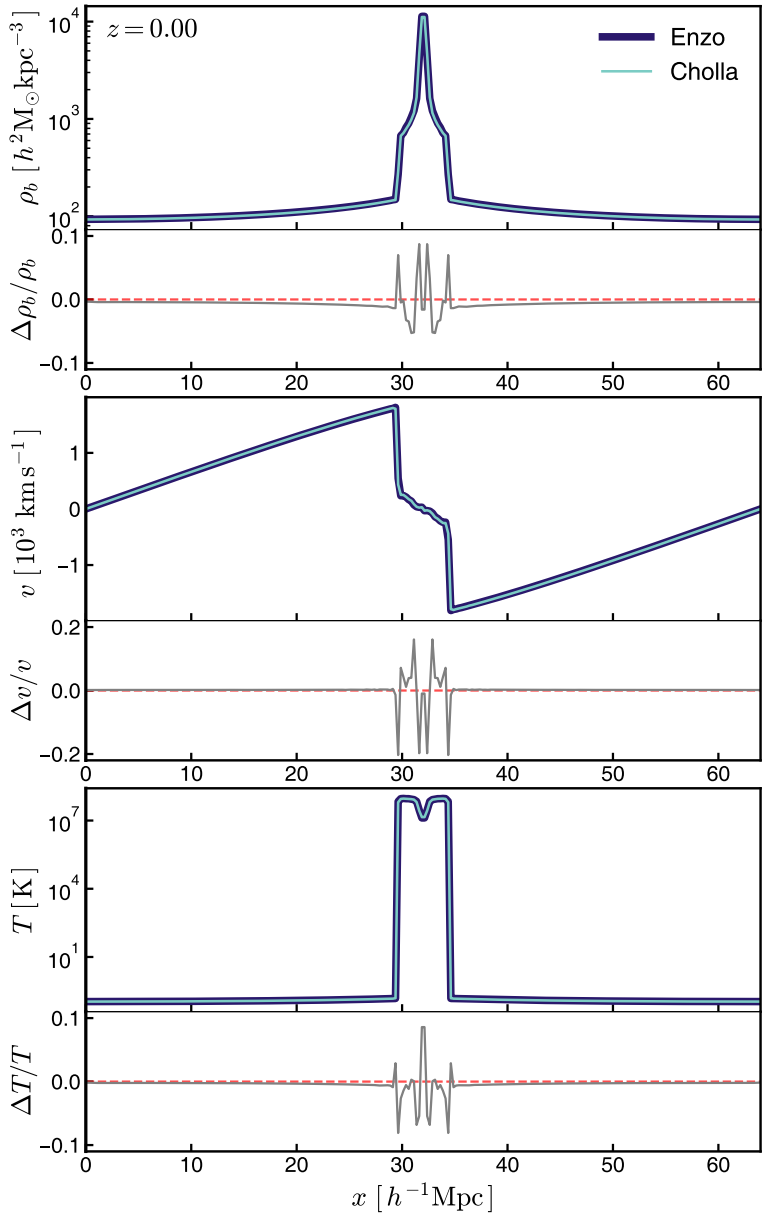

Figure 2.1: Zel'Dovich pancake test at $z = 0$ for a one dimensional 64 h^{-1}Mpc box discretized on a 256 uniform grid for a $h = 0.5$ and $\Omega_M = 1$ universe. Shown are the solutions computed using *Enzo* (purple) and *Cholla* (blue). The top, middle, and bottom main panels correspond to the density (ρ_b), velocity (v), and temperature (T), respectively, the fractional differences for each quantity are shown at the bottom of each panel (gray lines). The *Cholla* simulations resolve the central shock and overdensity, and the results are in excellent agreement with the *Enzo* simulation. Small differences $< 10\%$ for ρ_b and T are located at the sharp features of the shock, and differences $\sim 20\%$ for v at the regions where $v \sim 0$ and at the fronts of the shock.

2.3.1 Zel'Dovich Pancake

The Zel'Dovich (1970) pancake problem encompasses several of the basic components of a cosmological hydrodynamical simulation including gas dynamics, self-gravity, and an expanding frame. For this test, the evolution of a single one dimensional sinusoidal perturbation is followed to provide a useful representation of the gas evolution in a three-dimensional simulation by solving the gravitational collapse of a single mode. The initial conditions for the density, velocity and temperature on a Lagrangian frame are set as

$$\rho_b\left(x_l\right) = \rho_0\left[1 - \frac{1+z_s}{1+z}\cos\left(kx_l\right)\right]^{-1} \tag{2.36}$$

$$v\left(x_l\right) = -H_0\frac{1+z_s}{\left(1+z\right)^{1/2}}\frac{\sin\left(kx_l\right)}{k} \tag{2.37}$$

$$T\left(x_l\right) = T_0\left[\frac{\rho_b\left(x_l\right)}{\bar{\rho_b}}\right]^{2/3}, \tag{2.38}$$

where z_s is the value of the redshift at which the gravitational collapse results in the formation of a shock located at the center of the overdensity, z is the initial redshift, λ is the wavelength of the perturbation , $k = 2\pi/\lambda$ is the corresponding wavenumber, and x_l is the position of the Lagrangian mass coordinate. The conversion of the positions to the Eulerian coordinates x is given by

$$x = x_l - \frac{1+z_s}{1+z}\frac{\sin\left(kx_l\right)}{k}. \tag{2.39}$$

For this test we replicate the problem presented in Bryan et al. (2014), a one dimensional

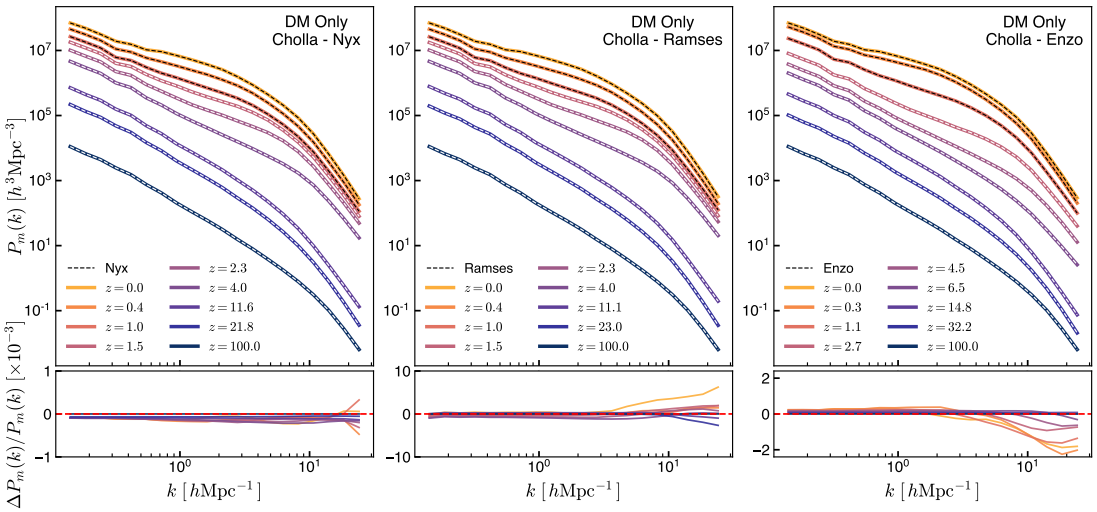

Figure 2.2: Matter power spectrum $P_m(k)$ from dark matter-only simulations. Shown are $P_m(k)$ from simulations performed with *Cholla* (colored lines) compared with analogous simulations computed with *Nyx*, *Ramses*, and *Enzo*
(left, center, and right panels; dashed lines). The bottom panels show the fractional difference between the *Cholla* simulation and each corresponding code for comparison. For each of the comparison codes, the fractional differences relative to the *Cholla* results are $< 0.1\%$ compared with *Nyx* (bottom left), $< 1\%$ compared with *Ramses* (bottom center), and $< 0.3\%$ compared with *Enzo* (bottom right).

simulation on an $L = 64h^{-1}$Mpc box with cosmological parameters $h = 0.5$ and $\Omega_M = 1$. The initial background density and temperature (ρ_0 and T_0) are set equal to the critical density $\rho_c = 3H_0^2/8\pi G$ and 100 K respectively. The redshift at which the shock develops is set as $z_s = 1$ and the wavelength of the perturbation is set as $\lambda = L = 64h^{-1}$Mpc. The simulation is initialized at $z = 20$ and runs on an $N = 256$ cell uniform grid.

Comparing the evolution of the Zel'dovich pancake problem solved using *Cholla* and *Enzo*, we find the results from the *Cholla* simulation closely resemble the *Enzo* results. Both simulations develop a central shock at $z = 1$ and the central overdensity grows at the same rate. Figure 2.3.1 compares the density ρ_b, velocity v and temperature T fields at $z = 0$ in the simulations solved with *Enzo* (purple lines) and *Cholla* (blue

lines), additionally, the fractional differences are shown at the bottom of each panel (gray lines). We measure small differences $< 10\%$ for the density and the temperature only at the sharp features of the shock. For the velocity the differences are $\sim 20\%$ at the discontinuities located at the front of the shock, and at the regions where $v \sim 0$, the other regions contained by the shock show small differences $< 10\%$. The regions outside the shock result in differences $< 1\%$ for all the fields. The small differences demonstrate an excellent agreement between the codes.

We note that we also performed the Zel'Dovich pancake test by applying Equation 2.15 for the internal energy selection in the dual energy scheme. This condition causes the code to select the advected internal energy e instead of the conserved internal energy $E - v^2/2$ during the entire simulation. Since Equation 2.5 does not captures shock heating, the central shock at $z = 1$ is suppressed and the temperature of the central region only gradually increases owing to the gas compression. This behavior results in a significantly different distribution for the density, velocity, and temperature in the central region. We discuss further ramifications of the dual energy condition in §2.3.4.

2.3.2 N-body Cosmological Simulations

To validate the results produced by *Cholla* in a realistic cosmological setting, first we compare *Cholla* dark matter-only simulations with calculations using several other well-established codes. In this test, the simulation domain consists of an $L = 50(h^{-1}\text{Mpc})^3$ box. The standard cosmological parameters are set to $\Omega_M = 0.3111$, $\Omega_\Lambda = 0.6889$, $h = 0.6766$, and $\sigma_8 = 0.8102$. Initial conditions were generated at $z = 100$

on a uniform resolution grid using the MUSIC software (Hahn & Abel 2011a). For this test, we evolve 256^3 particles on an $N = 256^3$ cell uniform grid, with the particle mass resolution equal to $m_p = 6.4345 \times 10^8 h^{-1} \mathrm{M}_\odot$.

For our validation test, we measure the matter power spectrum $P_m(k)$ evolved from identical initial conditions using *Cholla*, *Nyx* (Almgren et al. 2013), *Ramses* (Teyssier 2002), and *Enzo* (Bryan et al. 2014). To measure the matter power spectrum for each simulation, we first compute the dark matter density field by interpolating the dark matter particles onto the $N = 256^3$ uniform grid via the CIC method described in §2.2.5. The power spectrum is then computed in Fourier space by taking the FFT of the overdensity field. The density field and power spectrum are computed identically for all the comparison simulations to ensure that any power spectrum differences arise solely from differences in the evolved particle distribution.

The results of our comparison are presented in Figure 2.3.1. Each panel in the upper row shows the matter power spectrum $P_m(k)$ for several redshifts as computed by *Cholla* (colored lines), along with an overlay of the results from other codes (dashed lines). The left panel shows the comparison to the *Nyx* simulation, the center panel corresponds to the *Ramses* comparison, and the right panel shows the comparison to the *Enzo* results. The bottom row shows the fractional difference between the power spectrum of the simulation computed by *Cholla* and each comparison code. As the left lower panel shows, the power spectrum measured in the *Cholla* and *Nyx* simulations is in excellent agreement with fractional differences of $\approx 0.05\%$ at small scales, and even smaller differences of $\approx 0.02\%$ on larger scales. The comparison with *Ramses* also

shows remarkable agreement with differences of 0.1% at large scales, with the largest differences of $\sim 0.7\%$ occurring on small scales by $z = 0$. *Cholla* and *Enzo* also show excellent agreement, with differences of $< 0.1\%$ at large scales and $< 0.3\%$ on small scales.

We note that in the version of *Nyx* used in this comparison, a second order scheme was employed to compute the gravitational potential gradient,

$$\frac{\partial \phi_{i,j,k}}{\partial x} = \frac{1}{2\Delta x} \left(\phi_{i+1,j,k} - \phi_{i-1,j,k} \right), \tag{2.40}$$

instead of the fourth-order method described by Equation 2.19. To have the closest possible comparison, for the *Cholla* simulation used to compare to the *Nyx* results we used Equation 2.40 to compute the gravitational field. We note that the lower order scheme used to compute the gradient leads to significant differences on the power spectrum at small scales of about 15% relative to the same simulation employing the higher order method.

Additionally, for the comparison with *Enzo* we used the simpler kernel for the Greens function $G(\mathbf{k}) = -k^{-2}$ in our solver instead of the kernel for the discretized Poisson equation given by Equation 2.18. The choice of the kernel results in substantial differences, changing the small-scale power spectrum by as much as $\approx 28\%$ relative to *Enzo* when *Cholla* employs Equation 2.18.

2.3.3 Adiabatic Cosmological Hydrodynamical Simulation

In §2.3.1 we showed that *Cholla* accurately solved the gas dynamics in a simplified one dimensional simulation. To test the evolution of the gas in a realistic cosmological evolution, we compared the results of a *Cholla* adiabatic hydrodynamical run to a calculation with *Ramses* using identical initial conditions. The configuration for the simulation is the same as the one described in §2.3.2, but with the addition of an $\Omega_b = 0.0486$ baryonic component. For the comparison we measure the gas density fluctuation power spectrum directly from the baryon density field in both simulations. Figure 2.3.3 shows the results of the comparison, with the power spectrum measured in the *Cholla* simulation (colored lines) shown for several redshifts along with the *Ramses* simulation (dashed lines). The bottom panel shows the fractional difference between the *Cholla* and *Ramses* power spectrum measurement. On large scales the agreement is excellent ($\lesssim 1\%$), and on smaller scales there are some differences up to a maximum of $\approx 7\%$ at $z < 1$.

As described in §2.2.3, the dual energy condition used by *Ramses* (Equation 2.15) can suppress shock heating in regions where the gas is converging. This choice can have a significant effect on gas falling into dark matter potential wells, resulting in artificially low gas temperatures. A detailed study of how the dual energy condition affects cosmological gas properties is provided in §2.3.4, but for this power spectrum test we used the *Ramses* dual energy condition (Equation 2.15). We note that the lower gas temperatures computed using the *Ramses* dual energy condition result in more power on small scales relative to calculations that use Equation 2.13. The tests in §2.3.4 illustrate

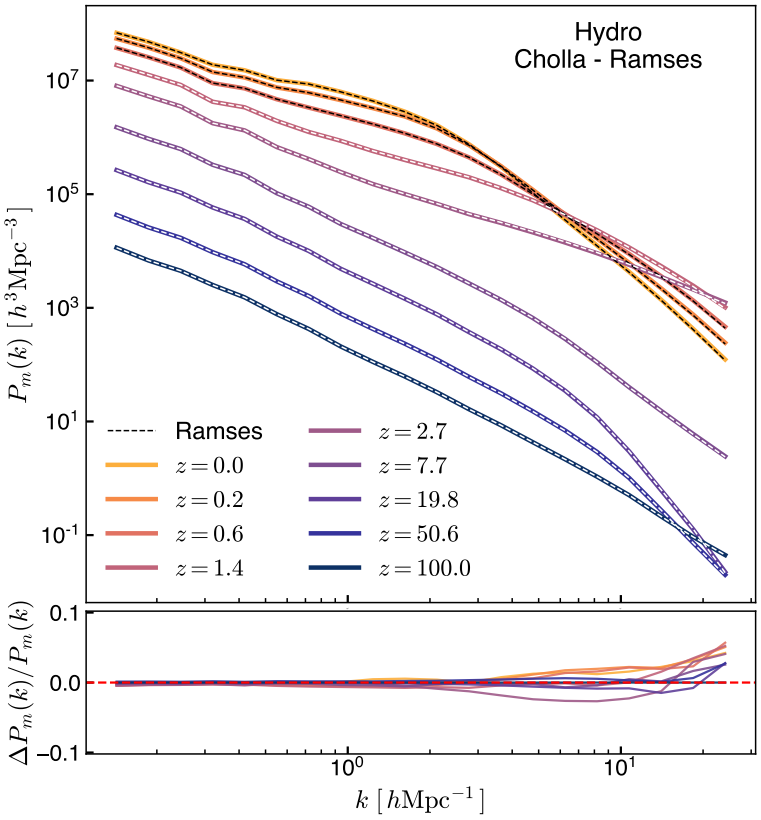

Figure 2.3: Redshift-dependent gas density fluctuation power spectra for an adiabatic cosmological simulation. Shown are the results from our *Cholla* simulation (solid lines) compared with simulation evolved using *Ramses* (dashed lines). The bottom panel shows the fractional difference between the *Cholla* and *Ramses* power spectra. The agreement is excellent on large scales, and for smaller scales the differences are $< 7\%$.

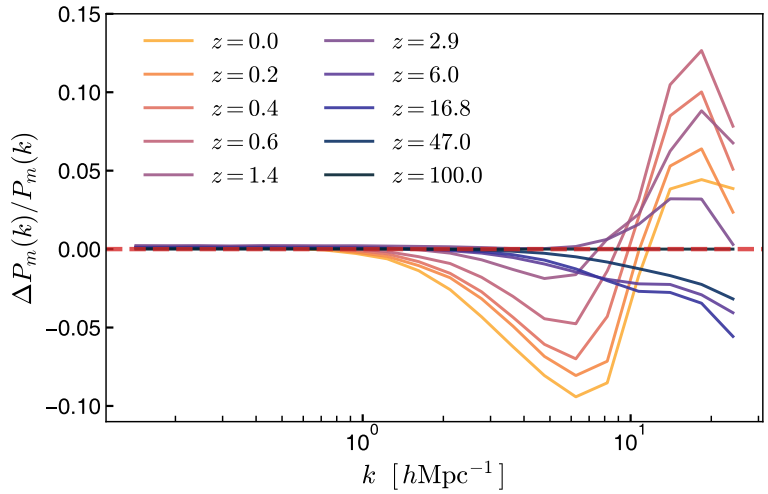

Figure 2.4: Fractional differences in the gas density fluctuation power spectra for an adiabatic cosmological *Cholla* simulation relative to an analogous *Enzo* run, the differences are $< 1\%$ for large scales and $< 15\%$ for smaller scales.

why we instead use Equation 2.13 in our CHIPS simulation suite.

Additionally we compared the power spectrum of the gas density fluctuations from a *Cholla* adiabatic cosmological simulation using Equation 2.13 for the dual energy formalism to an analogous *Enzo* run that evolved the same initial conditions. Figure 2.3.3 shows the fractional differences in $P_m(k)$ resulting from the *Cholla* simulation relative to the *Enzo* run. As shown, the agreement is excellent in the large scales with differences $< 1\%$ and the smaller scales present differences $< 15\%$. From the tests performed we note that the small scale $P_m(k)$ is highly sensible to the numerical implementation of the hydrodynamics solver and that we are not aware of a robust comparison of the gas $P_m(k)$ resulting from different codes. We argue that the small differences obtained do not represent an inaccurate evolution of the gas dynamics.

2.3.4 Average Cosmic Temperature

In §2.2.3 we discussed the dual energy formalism used when solving hydrodynamical cosmological simulations. We presented two different approaches for selecting between the advected internal energy e or the conserved internal energy $E - v^2/2$, these two methods are given by Equations 2.13 and 2.15 employed by *Enzo* and *Ramses* respectively. To test which approach best captures the shock heating of the infalling gas onto the dark matter halos when implemented in *Cholla*, we measure the mass weighted average gas temperature $\bar{T}$ in an adiabatic cosmological simulation as described in §2.3.3, and compare the simulation results to an estimate of the expected gas temperature computed from averaging the virial temperature of collapsed halos with the adiabatically-cooling IGM. We used the ROCKSTAR halo finder (Behroozi et al. 2013) to identify dark matter halos in the *Cholla* simulations, and then computed for each resolved halo a virial temperature as

$$T_{vir} = \frac{m_p}{3k_B} \frac{GM_{vir}}{R_{vir}}. \tag{2.41}$$

where m_p is the proton mass, k_B is the Boltzmann constant, and M_{vir} and R_{vir} are the virial mass and radius of the halo measured by ROCKSTAR. To compute our reference estimate of the expected mean cosmic temperature in the simulation, we take the mass in the IGM as simply total gas mass in the simulated box M_{total} minus the mass in collapsed halos $M_{halos} = \sum M_{vir}$. Then, assuming a uniform baryon fraction, the fraction of gas mass present int the IGM is simply $(M_{total} - M_{halos})/M_{total}$, and from

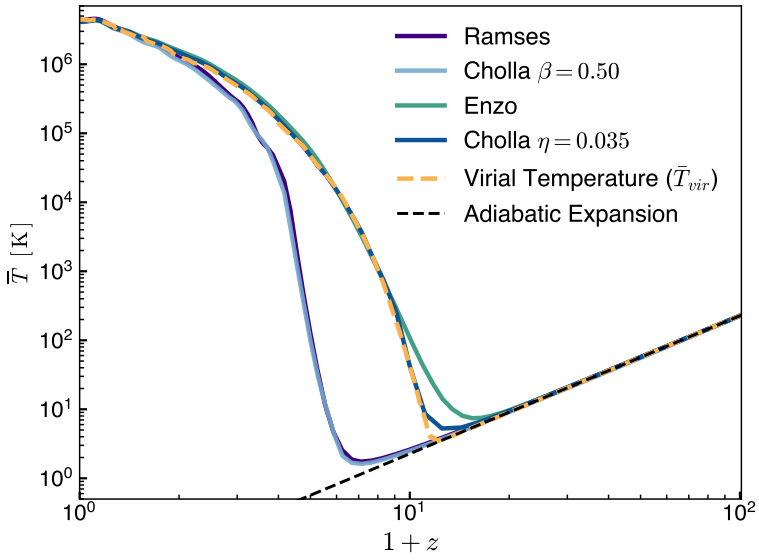

Figure 2.5: Evolution of the mass weighted average cosmic temperature $\bar{T}$ as a function of redshift for an adiabatic cosmological simulation. Shown as solid lines are the simulation results using *Ramses* (purple), *Enzo* (green), and *Cholla* simulations where Equations 2.13 (dark blue) or 2.15 (light blue) were used for the internal energy selection criteria in the dual energy formalism. The dashed lines show estimates of the temperature T expected from the viral temperature of halos (yellow) and the $T \propto a^{-2}$ dependence owing to the adiabatic expansion of the universe (black).

this we can compute the mass weighted average temperature as

$$\bar{T}_{vir} = \sum_{halos} \frac{M_{vir}}{M_{total}} T_{vir} + \left(\frac{M_{total} - M_{halos}}{M_{total}} \right) \left(\frac{a_0}{a} \right)^2 T_0, \tag{2.42}$$

where the first term corresponds to the mass weighted virial temperature of the gas present in collapsed halos and the second term corresponds to the mass weighted temperature of the gas in the IGM. The IGM temperature is taken to be the initial temperature T_0 scaled by the a^{-2} factor owing to the adiabatic expansion of the universe.

The results of this comparison are shown in Figure 2.3.4, where we plot the mass-weighted average temperature of the gas simulations evolved with *Ramses* (purple

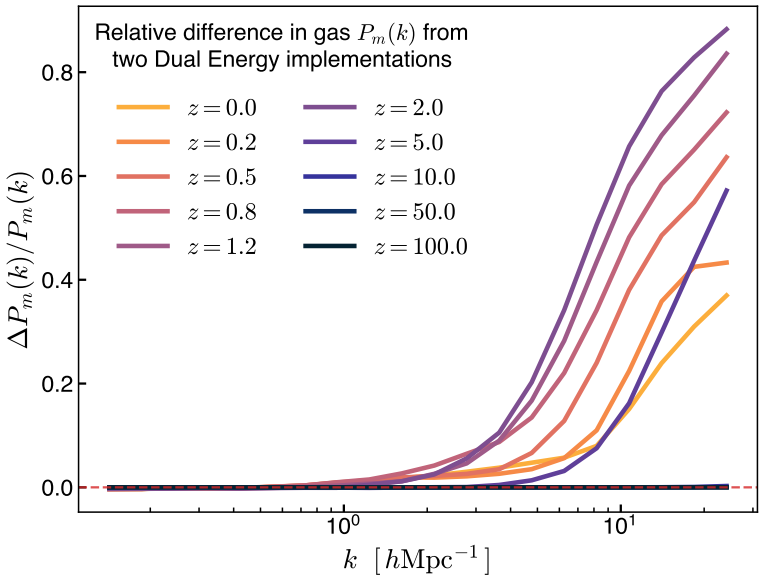

Figure 2.6: Relative fractional difference of gas density fluctuation power spectra in adiabatic cosmological simulations using Equations 2.13 or 2.15 for the dual energy implementation. The condition set by Eqn. 2.15 maintains colder gas for $1 \lesssim z \lesssim 15$, resulting in larger power at small scales. The differences range from 30% to 80% at $0 \lesssim z \lesssim 6$.

line) or *Enzo* (green line). For *Cholla* we ran two simulations with different dual energy conditions, using a criteria similar to *Ramses* (Equation 2.15, light blue line) or similar to *Enzo* (Equation 2.13, dark blue line).

All simulations start from the same initial temperature $T_0 = 230\,\mathrm{K}$ at $z = 100$. Afterward, the gas cools as $T \propto a^{-2}$ owing to universal expansion until the first halos collapse. The temperature of the infalling gas increases owing to shock heating, causing the global average temperature to increase. As Figure 2.3.4 shows, the temperature increase happens at roughly two different times for the different simulations. In the *Enzo* calculation and corresponding *Cholla* simulation that uses Equation 2.13 for the dual energy condition, shock heating in the halos becomes significant at $z \sim 15$ (green

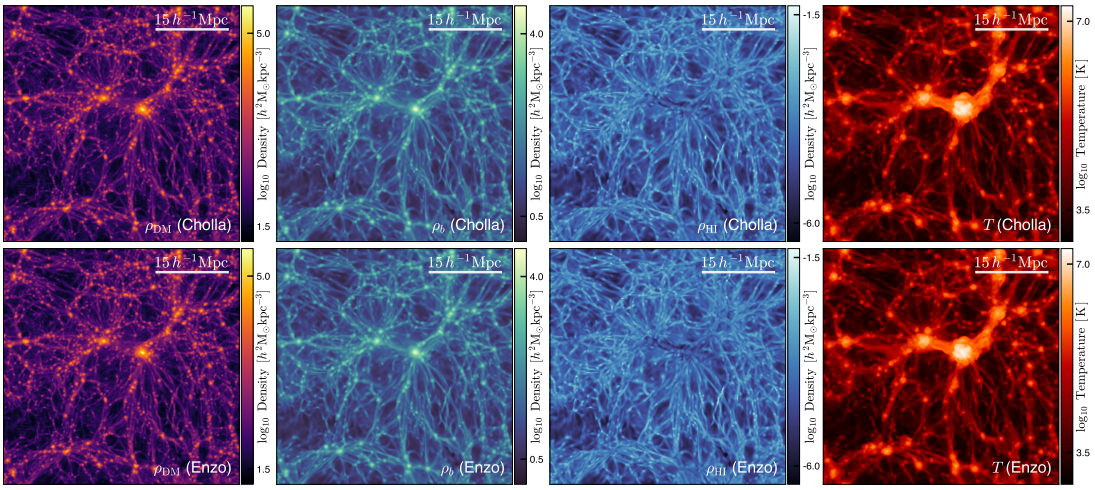

Figure 2.7: Comparison of simulations (256^3 grid and $L = 50h^{-1}$Mpc) evolved with *Cholla* (top panels) and *Enzo* (bottom panels). From left to right the columns correspond to projections of dark matter density $\rho_{\rm DM}$, gas density ρ_b, neutral hydrogen density $\rho_{\rm HI}$, and gas temperature T, all at redshift $z = 0$.

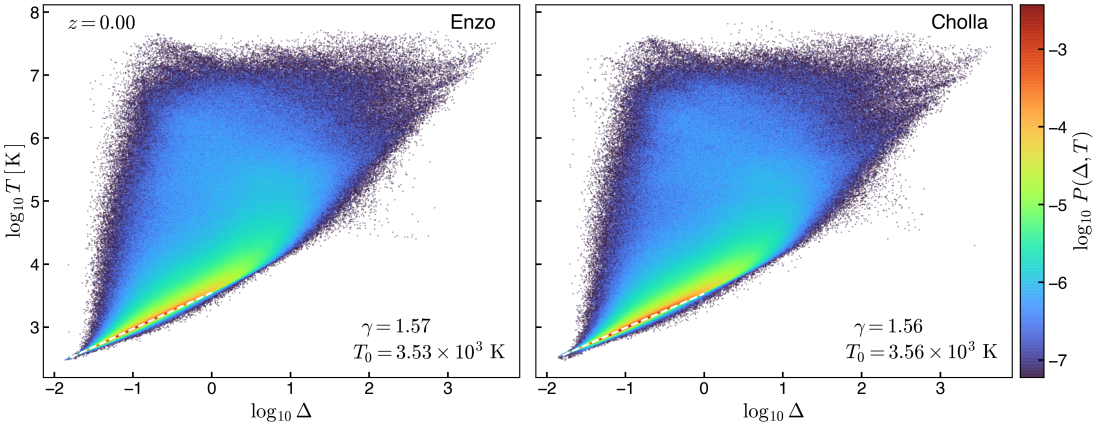

Figure 2.8: Comparison of the density-temperature ($\Delta = \rho_b/\bar{\rho}_b$) distribution at redshift $z = 0$ for analogous simulations (256^3 grid and $L = 50h^{-1}$Mpc) evolved with *Enzo* (left) and *Cholla* (right) using the HM12 photoheating and photoionzation rates. The distribution of the gas in both simulations is remarkably similar and the differences for the parameters T_0 and γ are $< 1\%$, demonstrating an excellent agreement for the gas distribution in the IGM between the codes. (The parameters T_0 and γ are defined in §2.5.)

and dark blue lines). In the *Ramses* simulation and the *Cholla* calculation using the condition given by Equation 2.15 (purple and light blue lines), the gas continues cooling owing to expansion until $z \sim 6$.

The delayed gas heating in calculations using the dual energy condition given by Eqn. 2.15 results from the advected internal energy e being dominantly selected over the conserved internal energy $E - v^2/2$ for the gas infalling into halos. In this case, the evolution of the advected internal energy e is given by Equation 2.5 that does not capture shock heating, and consequently the heating of the gas in the halos is suppressed.

In contrast, if Equation 2.13 is used for the dual energy condition in the *Cholla* and *Enzo* simulations, the resulting mean cosmic temperature closely follows the virial temperature estimate at $z < 6$. We found that for *Cholla*, adopting the parameter value $\eta = 0.035$ in Equation 2.13 results in a temperature increase that begins at $z \approx 12$, similar to our model estimate.

We note that the mean cosmic gas temperature measured in the Cholla adiabatic simulation and our $\bar{T}_{vir}$ estimate (Eq. 2.42) computed from the halo properties (Figure 5, dark blue and yellow lines respectively) display a sharp transition at $z \approx 12$, when the temperature suddenly increases. We argue that this behavior is a consequence of the limited resolution in our test simulation, as most of the low mass halos that would form at early times ($z > 12$) are not resolved and the heating of gas during their virialization is not captured. We verify this by computing an estimate of $\bar{T}_{vir}$ including early, low-mass halos predicted by an analytical mass function (Sheth & Tormen 1999). This

analytical estimate shows an earlier and more gradual increase in the mean temperature, starting at $z \sim 20$. If instead we limit the minimum halo mass used for the analytical estimate to the minimum resolved halo mass in the test simulations, the estimate follows closely the temperature evolution from the Cholla simulation including the sharp increase at $z \approx 12$, strongly suggesting that the missing unresolved halos explain the sharp feature in the temperature evolution.

Additionally, we measured the effect on the gas overdensity power spectrum of temperature differences arising from the choice of the dual energy condition. Figure 2.6 shows the fractional difference of the power spectrum measured in the simulation where Equation 2.15 was used for the dual energy selection relative to the power spectrum measured in the simulation that instead employed Equation 2.13. The comparison shows that the power spectrum in the simulation using condition set by Eqn. 2.15, where the gas remains colder for longer, is $\sim 50\%$ higher on small scales by $z \sim 5$. By $z \sim 2$ differences on small scales reach $\sim 80\%$. Afterward, the two simulations reach similar average temperatures and the differences decrease to $\sim 30\%$ by $z = 0$.

We note that the truncation error $e_{\rm trunc}$ (Eq. 2.14) is resolution dependent, and the suppressed shock heating presented in this comparison might be reduced in high resolution AMR simulations. Studying the behavior of condition 2.15 in such simulations is beyond the scope of this work.

2.3.5 Cosmological Simulation: Chemistry and UV Background

To validate our integration of the chemical network (solved by GRACKLE and advected by the hydrodynamics solver), we ran identical hydrodynamical simulations

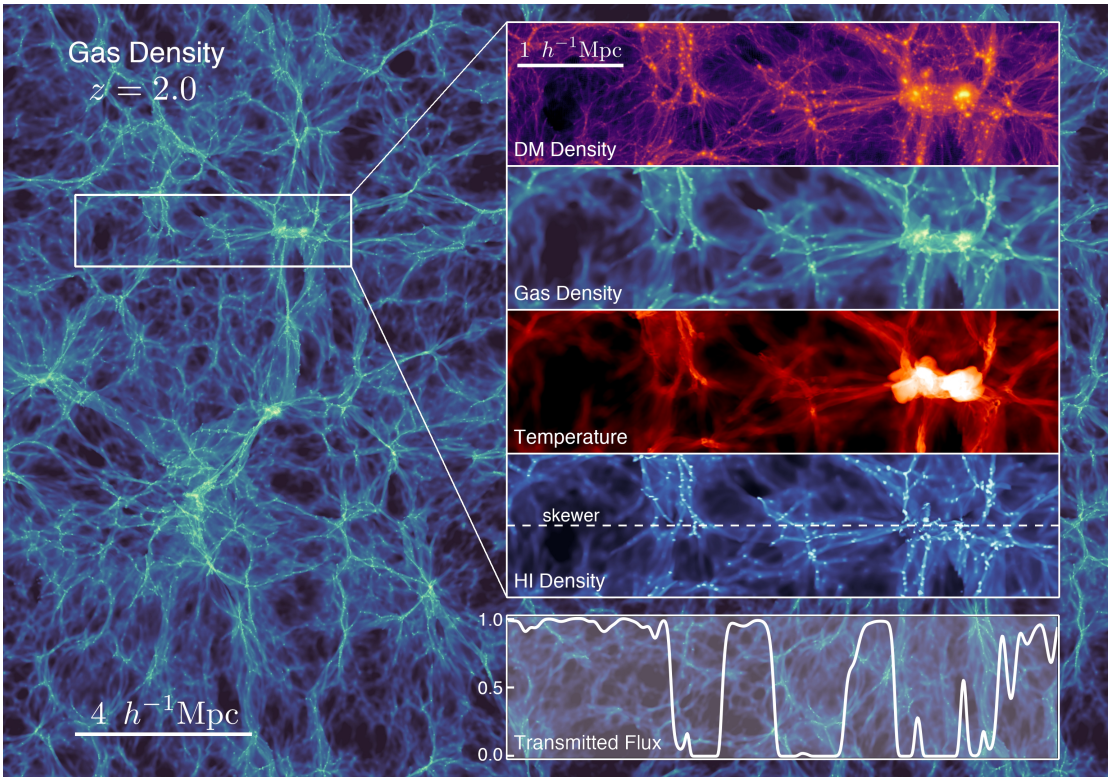

Figure 2.9: Projection of the gas density at redshift $z = 2$ from the CHIPS.P19 simulation (2048^3 grid, $L = 50\,h^{-1}$Mpc, and $1.5\,h^{-1}$Mpc projected). The zoom-in region shows the dark matter density $\rho_{\rm DM}$, gas density ρ_b, gas temperature T, and neutral hydrogen density $\rho_{\rm HI}$ from top to bottom. A skewer crossing the center of the zoom-in region is marked over the neutral hydrogen distribution, and the Lyman-α transmitted flux along the skewer is shown in the bottom panel. The scale labels refer to proper distances.

solved with *Enzo* and *Cholla* following the configuration described in §2.3.3 but including the non-equilibrium H and He network plus metals in the presence of a spatially-uniform, time-dependent UV background given by the standard HM12 (Haardt & Madau 2012) photoheating and photoionization rates. Figure 2.7 shows a comparison of projected gas quantities at $z = 0$ computed by the two codes, with *Cholla* on top and *Enzo* on the bottom. From left to right the panels show projections of dark matter density ($\rho_{\rm DM}$), gas density (ρ_b), neutral hydrogen density ($\rho_{\rm HI}$), and gas temperature (T). Figure 2.7

shows qualitatively a remarkable agreement between the results from the two codes.

For a more detailed comparison, we measured the density-temperature distribution from both simulations. Figure 2.8 shows the results from the *Enzo* (left) and *Cholla* (right) runs at $z = 0$. As shown, the distribution of the gas in the simulations is remarkably similar. Additionally the parameters T_0 and γ which model the distribution of the diffuse gas in the IGM (see §2.5 for details) differ by less than 1% showing an excellent agreement between these simulations.

2.4 Simulation Suite

In this section we present the CHIPS (**CH**olla **I**GM **P**hothoheating **S**imulations) simulation suite, a set of high resolution simulations performed using the newly extended version of the *Cholla* code described above. The suite consists of a series of simulations run with a fiducial resolution of $N = 2048^3$ cells, varying the cosmic photoheating and photoionization rates from evolving UV background radiation fields and with a range of cosmological parameters. All the CHIPS simulations evolve a primordial gas composition ($X = 0.76$, $Y = 0.24$), without the inclusion of metal line cooling as star formation is not accounted for in the simulations.

Table 2.1: CHIPS Simulation Suite

Simulation	Resolution	Box Size L [h^{-1}Mpc]	Parameters [h, Ω_m, Ω_b, σ_8, n_s]	UV Background
CHIPS.HM12	$N = 2048^3$	50	$[0.6766, 0.3111, 0.0497, 0.8102, 0.9665]$	Haardt & Madau (2012)
CHIPS.P19	$N = 2048^3$	50	$[0.6766, 0.3111, 0.0497, 0.8102, 0.9665]$	Puchwein et al. (2019)
Alternative Cosmologies				
CHIPS.P19.A1	$N = 2048^3$	50	$[0.6835, 0.3010, 0.0484, 0.8098, 0.9722]$	Puchwein et al. (2019)
CHIPS.P19.A2	$N = 2048^3$	50	$[0.6917, 0.2905, 0.0477, 0.8052, 0.9783]$	Puchwein et al. (2019)
CHIPS.P19.A3	$N = 2048^3$	50	$[0.7001, 0.2808, 0.0470, 0.8020, 0.9846]$	Puchwein et al. (2019)
CHIPS.P19.A4	$N = 2048^3$	50	$[0.7069, 0.2730, 0.0465, 0.7997, 0.9896]$	Puchwein et al. (2019)
Resolution Studies				
CHIPS.P19.R1	$N = 1024^3$	50	$[0.6766, 0.3111, 0.0497, 0.8102, 0.9665]$	Puchwein et al. (2019)

Table 2.1 (cont'd): CHIPS Simulation Suite

Simulation	Resolution	Box Size L $[h^{-1}\mathrm{Mpc}]$	Parameters $[h,\ \Omega_m,\ \Omega_b,\ \sigma_8,\ n_s]$	UV Background
CHIPS.P19.R2	$N = 512^3$	50	$[0.6766, 0.3111, 0.0497, 0.8102, 0.9665]$	Puchwein et al. (2019)

Note. — Resolution refers to the number of both grid cells and dark matter particles.

Table 2.1 details the properties our initial CHIPS simulations. The primary simulations for our initial analysis of the Lyman-α forest are CHIPS.HM12 and CHIPS.P19, which use the Planck Collaboration et al. (2020) cosmological parameters and the Haardt & Madau (2012) and Puchwein et al. (2019) photoionization and photoheating rates, respectively, in an $L = 50\ h^{-1}$Mpc box. The Puchwein et al. (2019) model adopts the most recent determinations of the ionizing emissivity due to stars and AGN, as well as of the H I absorber column density distribution. Another major improvement is a new treatment of the IGM opacity for ionizing radiation that is able to consistently capture the transition from a neutral to ionized IGM. For these fiducial runs, we output 150 snapshots over the redshift range $z = [16, 2]$, spacing the time between snapshots at $\Delta a = 1.83 \times 10^{-3}$ intervals. In each snapshot, we record the conserved fluid quantities (ρ, ρv_x, ρv_y, ρv_z, E), the gas internal energy u, the neutral hydrogen H I, neutral helium He I, singly-ionized helium He II, and electron n_e densities, and the gravitational potential ϕ on the simulation grid. We also record all the dark matter particle positions and velocities. The detailed analyses performed on the simulation outputs are described in §2.5.

We complement the fiducial models with four additional simulations (CHIPS.P19.[A1-A4]) that use the Puchwein et al. (2019) photoionization and photoheating rates but vary the cosmological parameters h, Ω_m, Ω_b, σ_8, and n_s within the uncertainties reported by Planck Collaboration et al. (2020). For each simulation, a flat cosmology is assumed and we set $\Omega_\Lambda = 1 - \Omega_m$.

Table 2.1 also lists properties of the additional $N = 1024^3$ (CHIPS.P19.R1)

and $N = 512^3$ (CHIPS.P19.R2) simulations used in our resolution study to demonstrate the numerical convergence of our results (see Appendix 2.8).

The CHIPS simulation suite was run on the Summit system (Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory), each of the 2048^3 simulations ran in 512 GPUs for $\sim$11 hours costing $\sim$1000 node-hours. As described in §2.2.9, the slowest component of the simulation is the GRACKLE update step, consuming about half of the computational time for these runs. This motivates the ongoing development of a H+He network solver implemented to run in the GPUs which will potentially reduce this bottleneck. The subsequent analysis of the simulation output data was performed using the *lux* supercomputer at UC Santa Cruz.

2.5 Evolution of the IGM for Two Photoheating Histories

Redshift-dependent photoionization and photoheating rates of intergalactic gas substantially affect IGM properties. By comparing the CHIPS.HM12 with the CHIPS.P19 simulation, we can learn about how detailed differences in photoheating history lead to observable differences in the Lyman-α forest and potentially discriminate between them by further comparisons with data. We first compare the thermal history of the diffuse IGM between the simulations (§2.5.1). The redshift-dependent thermal properties of the IGM in the models provide a context for interpreting measurements of the simulated forest. We discuss our methods for generating mock Lyman-α absorption spectra in §2.5.2. These simulated spectra then provide estimates of the Lyman-α forest optical depth (§2.5.3) and transmitted flux power spectra (§2.5.4).

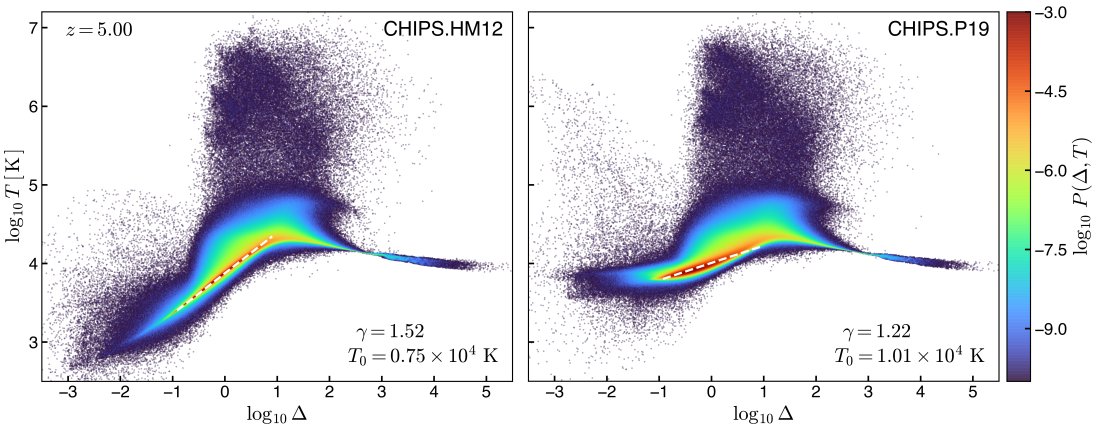

Figure 2.10: Volume-weighted density-temperature distribution of gas at redshift $z = 5$ in our two fiducial simulations (CHIPS.HM12, left; CHIPS.P19, right). The low density gas ($\Delta < 10$) is colder for the Haardt & Madau (2012) model owing to H reionization ending earlier. Dashed lines show the best fit power-law Δ-T relation for the parameters T_0 and γ.

2.5.1 Thermal History of the Diffuse IGM

The gas in the diffuse IGM comprises most of the baryons in the universe and follows a well defined density-temperature power-law relation (Hui & Gnedin 1997; McQuinn 2016; Puchwein et al. 2015) given by

$$T(\Delta) = T_0 \Delta^{\gamma - 1}, \tag{2.43}$$

where $\Delta = \rho_b / \bar{\rho}_b$ is the gas overdensity, T_0 is the temperature at the mean cosmic density $\bar{\rho}_b$, and $\gamma - 1$ corresponds to the power-law index of the relation. The time evolution of the parameters T_0 and γ is determined by the photoheating from to hydrogen and helium ionization, cooling owing to the expansion of the universe, and inverse Compton cooling, recombination, and collisional processes.

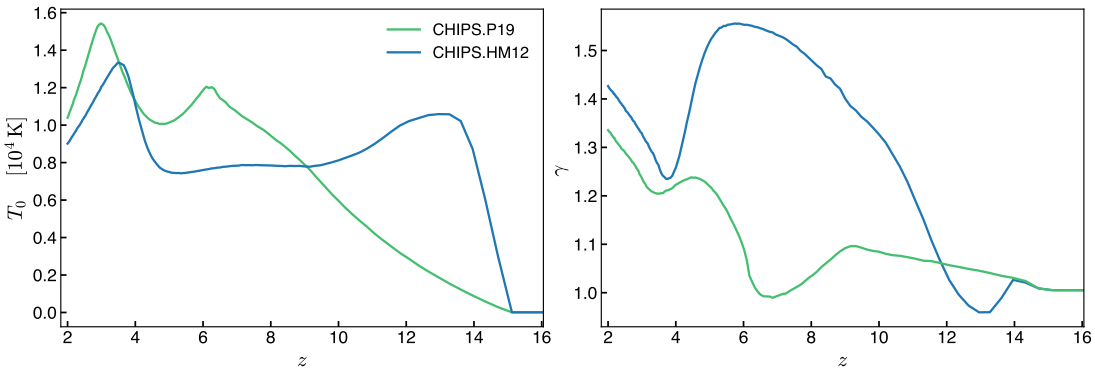

Figure 2.11: Redshift evolution of the gas temperature T_0 at mean density and the index γ of the density-temperature relation. Shown are results from our reference simulations CHIPS.HM12 (blue lines) and CHIPS.P19 (green lines). H I reionization ends earlier ($z \sim 13$) for the HM12 model, compared with $z \sim 6$ for the P19 model, allowing more time for the low density gas to cool. He II reionization begins at $z \sim 4.5$ in both models, but the lower He II photoionization rates for the P19 model result in He II being fully ionized at a later time ($z \sim 3$) compared with $z \sim 3.8$ for the HM12 run.

Figure 2.10 shows the density-temperature distribution of the gas in our simulations at redshift $z = 5$, with CHIPS.HM12 shown on the left and CHIPS.P19 shown on the right. The distributions resulting from the two UVB models are similar for gas collapsed into resolved structures ($\Delta > 10$), but for low density gas ($\Delta < 10$) the temperatures in the HM12 model are significantly lower owing to the earlier completion of hydrogen reionization ($z \sim 13$). The gas temperature in this model has had subsequently more time to decrease owing to cooling processes and adiabatic expansion. For the P19 model, where reionization ends at $z \sim 6$, there has been less time to cool by $z = 5$, resulting in a higher T_0 and lower γ at this epoch.

For each snapshot of the simulation, we determined the parameters T_0 and γ by fitting Equation 3.7 to the low density ($-1 < \log_{10} \Delta < 1$) region of the density-temperature distribution. We divided the selected interval into fifty equal bins in

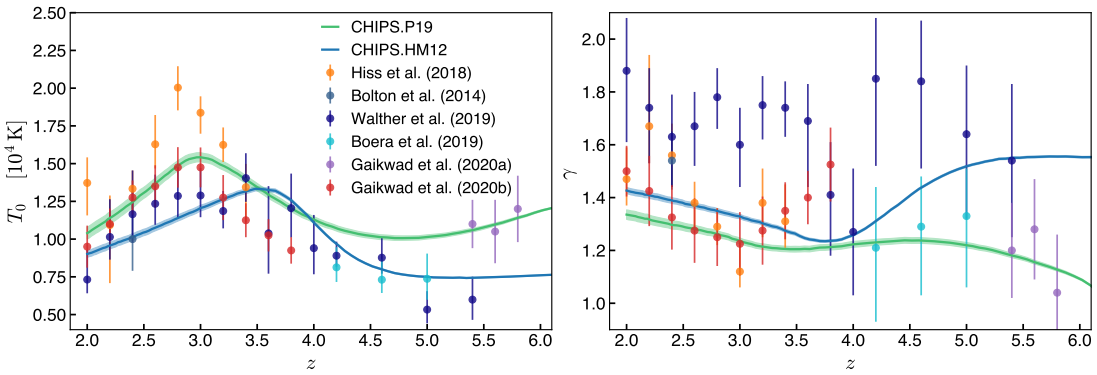

Figure 2.12: Redshift evolution of the density-temperature parameters T_0 and γ from the CHIPS.HM12 (blue) and CHIPS.P19 (green) simulations. The shaded regions show the uncertainty resulting from the power-law fitting procedure. Points show the observational results from Hiss et al. (2018), Bolton et al. (2014), Walther et al. (2019), Boera et al. (2019), Gaikwad et al. (2020), and Gaikwad et al. (2021).

$\log_{10}(\Delta)$, and for each bin i the maximum of the marginal temperature distribution $P(T|\Delta_i)$ and the temperature range containing the 68% highest probability density was used to define the bin temperature T_i and its corresponding uncertainty δT_i. The coordinates (Δ_i, T_i) and uncertainty values δT_i were used as input to a Monte Carlo Markov Chain that sampled the parameters T_0 and γ, initialized from uniform prior distributions, and returned posterior distributions for the thermal parameters that best match the density-temperature relation measured from the simulations. From the posterior distributions we extracted best-fit values for T_0 and γ and corresponding parameter uncertainties δT_0 and $\delta\gamma$.

The redshift evolution of the parameters T_0 and γ for the two UVB models is shown in Figure 2.11, where the effects of the two ionization histories on the thermal structure of the diffuse IGM are illustrated. The evolution of the IGM follows several phases. First, before H I reionization is complete, the photoheating owing to hydrogen

ionization increases the temperature to an early local maximum. After all the H I is ionized, the diffuse IGM cools by adiabatic expansion and inverse Compton cooling until the onset of He II reionization reheats the IGM to a global maximum. Once He II is fully ionized, the IGM again cools adiabatically to the present day.

The HM12 UVB (blue line) causes a quick H I reionization around redshift $z \sim 14$ and cools afterward. In the interval between $z \sim 6 - 10$, the temperature at mean density T_0 plateaus at $T \sim 8 \times 10^3\,\mathrm{K}$ while underdensities ($\Delta < 1$) keep cooling, mostly due to adiabatic expansion, this results in an increasing γ until $z \sim 5$.

For $z \leq 15$ the ionization rates for the late-reionization P19 model are significantly lower until $z \sim 6.5$, resulting in a gradual heating of the IGM. The IGM remains close to isothermal ($\gamma \sim 1$) until H I reionization completes at $z \sim 6.2$. Intergalactic gas then cools just for short period before He II reheating, resulting in a higher T_0 than in the HM12 model.

While in both runs helium reheating starts around $z \sim 4.5$, the reionization of He II is completed earlier in CHIPS.HM12. At $z < 3$, the residual heating from photoionization of recombining atoms is inefficient, and the IGM continues to cool all the way down to $z = 0$, decreasing T_0 and increasing γ.

Figure 2.12 shows a comparison between the thermal parameters $T_0(z)$ and $\gamma(z)$ from our simulations and previous observational inferences (Bolton et al. 2014; Hiss et al. 2018; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2020, 2021). The shaded regions correspond to the uncertainty in T_0 and γ resulting from our power-law fitting procedure. For the observations, the values of T_0 and γ are determined in

different ways.

Walther et al. (2019) and Boera et al. (2019) both follow a similar approach by generating Lyman-α flux power spectra from simulations evolved with different thermal histories, resulting in multiple trajectories of T_0 and γ. For each simulation snapshot, they determine the best fit T_0, γ, and mean transmitted flux $\bar{F}$ by performing Bayesian inference comparing the generated flux power spectra from the different simulations to observations.

Hiss et al. (2018) and Bolton et al. (2014) measure the b-$N_{\rm HI}$ distribution obtained from decomposing the Lyman-α forest spectra into a collection of Voigt profiles, and then infer thermal parameters by matching the b-$N_{\rm HI}$ distribution from their simulations to the observed distribution. Gaikwad et al. (2021) follows an analogous approach by comparing simulation results to Voigt profiles fitted to transmission spikes in the inverse transmitted flux $1 - F$ in $z > 5$ spectra. In a recent analysis, Gaikwad et al. (2020) report more precise results by inferring T_0 and γ from combined constraints obtained through a comparison between simulated and observed Lyman-α forest flux power spectra, b-$N_{\rm HI}$ distributions, wavelet statistics, and curvature statistics.

During the epoch of He II reionization and afterward, Hiss et al. (2018) infer a peak in T_0 ($z \sim 2.8$) that is significantly higher than the results from all the other analysis, while the measurements from Gaikwad et al. (2020) are mostly higher than those obtained by Walther et al. (2019) and Bolton et al. (2014) their results still are consistent within 1σ of each other. Compared to the simulations, both the P19 and HM12 models result in a increase of T_0 due to He II photoheating that begins too

early ($z \sim 4.2$) to be consistent with the measurements from Gaikwad et al. (2020) and Walther et al. (2019) simultaneously. The heating from He II reionization could be delayed in the P19 model by decreasing the He II photoheating and photoionization rates, effectively also slightly decreasing the peak of T_0 at $z \sim 2.8$, this would produce a trajectory for T_0 at $z \lesssim 4$ that better matches the results from Gaikwad et al. (2020) and Walther et al. (2019).

For $z > 4$, the results from Walther et al. (2019) and Boera et al. (2019) measure temperatures lower than those produced by the Puchwein et al. (2019) model. In particular, at $z \geq 5$ Walther et al. (2019) find remarkably low temperatures, which may be related to the strong correlation between T_0 and $\bar{F}$ in their analysis.

The γ values inferred by Walther et al. (2019) are overall higher than all the other data sets. In the redshift range $2.6 \lesssim z \leq 5$, both our CHIPS.HM12 and CHIPS.P19 simulations produce $\gamma(z)$ consistent with the measurements from Hiss et al. (2018) and Boera et al. (2019), within their respective uncertainties. Additionally both models result in $\gamma(z)$ consistent with the results from Gaikwad et al. (2020) only for $z \lesssim 3.3$ as their observational inference results in higher values ($\gamma \sim 1.5$) at $z \sim 3.8$ compared to $\gamma \sim 1.3$ produced in both simulations at this redshift. Delaying the heating from He II reionization would allow more time for the diffuse gas to cool after H I reionization, effectively increasing γ produced by the models at $z \sim 4$, to be in better agreement with the results from Gaikwad et al. (2020).

The CHIPS.P19 results are consistent with constraints on T_0 and γ for $z > 5.3$ from Gaikwad et al. (2021), likely because H I reionization ends at $z \sim 6$ in the Puchwein

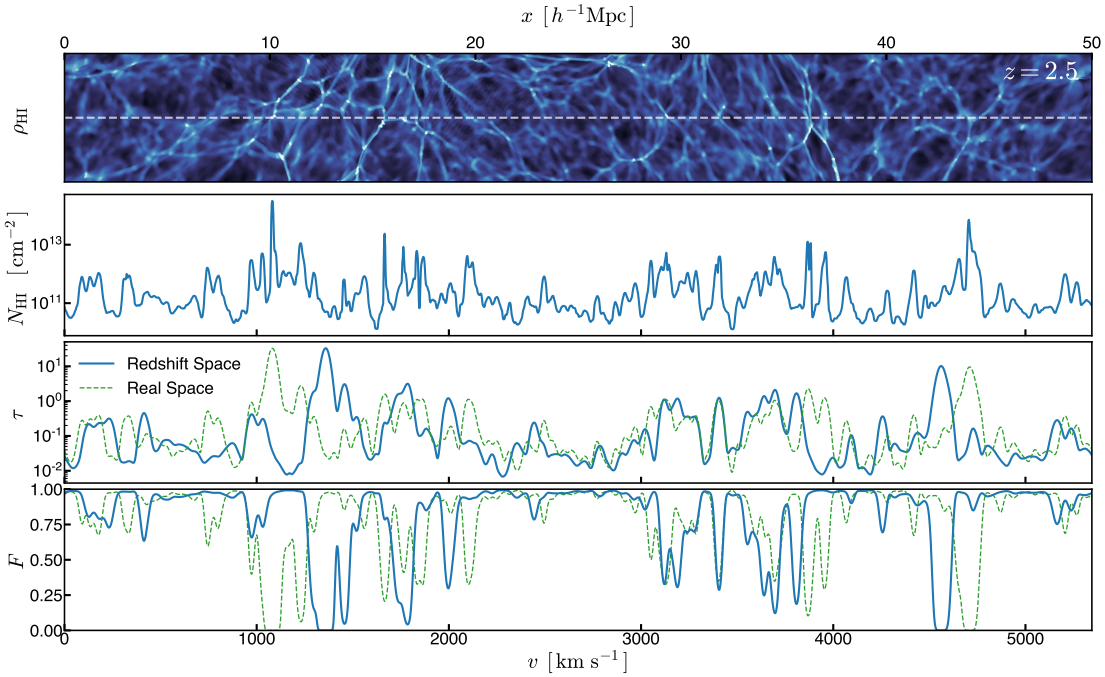

Figure 2.13: Elements for the calculation of the transmitted flux along a single skewer crossing the CHIPS.P19 simulation box at redshift $z = 2.5$. Shown in the panels from top to bottom are the neutral hydrogen density $\rho_{\rm HI}$ in a region surrounding the skewer, neutral hydrogen column density $N_{\rm HI}$ along the skewer, optical depth τ computed via Equation 2.48, and transmitted flux $F = \exp(-\tau)$ along the line of sight.

et al. (2019) photoionization and photoheating model. Although, reconciling the high $\gamma \sim 1.5$ at $z \sim 3.8$ from Gaikwad et al. (2020) with the low $\gamma \sim 1.2$ at $z \sim 5.5$ inferred by Gaikwad et al. (2021) could require the low density gas to cool faster than physically possible in a spatially-uniform UV background model. However, near H I reionization a non-uniform UV background may be required for the simulations to model accurately the effects of a "patchy" reionization (Keating et al. 2020).

2.5.2 Synthetic Lyman-α Forest Spectra

The Lyman-α forest is a sensitive probe of the diffuse baryons in the IGM, as the amplitude and width of the absorption lines in the forest trace the neutral hydrogen density and the gas temperature. The observed Lyman-α forest global statistics, such as the effective optical depth and the transmitted flux power spectra, constrain the thermal state of the IGM. To compare directly with the observed effective optical depth and the transmitted flux power spectra, we compute synthetic Lyman-α forest spectra from our high resolution simulations. In total we drew 60000 skewers through the simulation volume, located in random positions and aligned parallel with the three box axes (20000 skewers for each axis). Along each skewer, the neutral hydrogen density, gas temperature, and the component of the velocity parallel to the line of sight are sampled at the native resolution of the simulation, rendering 2048 uniformly distributed pixels for each skewer. The optical depth as a function of frequency τ_ν along the skewer is computed by integrating the Lyman-α interaction cross section σ_ν and the neutral hydrogen number density $n_{\rm HI}$ along the line of sight, following

$$\tau_\nu = \int n_{\rm HI}\sigma_\nu dr, \tag{2.44}$$

where dr is the physical length of the path element. Assuming a Doppler profile for the absorption line, the optical depth at frequency ν_0 is given by

$$\tau_{\nu_0} = \frac{\pi e^2}{m_e c} f_{12} \int \frac{n_{\rm HI}}{\sqrt{\pi}\Delta\nu_{\rm D}} \exp\left[-\left(\frac{\nu-\nu_0}{\Delta\nu_{\rm D}}\right)^2\right] dr, \tag{2.45}$$

where f_{12} is the Lyman-α transition upward oscillator strength, $\Delta\nu_{\rm D} = (b/c)\nu_0$ is the absorption width owing to Doppler shifts and $b = \sqrt{2k_{\rm B}T/m_{\rm H}}$ corresponds to the thermal velocity of the gas. The shift in the frequency of absorption along the skewer is given by the Doppler shift from the change in the gas velocity along the line of sight,

$$\nu = \nu_0 \left(1 - \frac{u - u_0}{c}\right). \tag{2.46}$$

Applying a variable transformation from frequency to velocity space and the expansion relation $du = Hdr$, the optical depth as a function of velocity is expressed as

$$\tau_{u_0} = \frac{\pi e^2 \lambda_0}{m_e c H} f_{12} \int \frac{n_{\rm HI}}{\sqrt{\pi} b} \exp\left[-\left(\frac{u - u_0}{b}\right)^2\right] du. \tag{2.47}$$

Following the method described by Lukić et al. (2015), we solved the Gaussian integral analytically and computed the optical depth along the discretized line of sight using

$$\tau_j = \frac{\pi e^2 \lambda_0 f_{12}}{m_e c H} \sum_i \frac{1}{2} n_{{\rm HI},i} \left[\mathrm{erf}\left(y_{j+1/2,i}\right) - \mathrm{erf}\left(y_{j-1/2,i}\right)\right], \tag{2.48}$$

where the argument to the error function is

$$y_{j\pm1/2,i} = \left(v_{{\rm H},j\pm1/2} - v_{{\rm H},i} - v_{{\rm LOS},i}\right)/b_i. \tag{2.49}$$

The term $v_{{\rm H},j\pm1/2}$ corresponds to the Hubble flow velocity at the interfaces of cell j and the terms $v_{{\rm H},i}$ and $v_{{\rm LOS},i}$ represent the centered values of Hubble velocity and the line of sight component of the peculiar velocity of the gas at cell i. Note that the

factor of 1/2 in Equation 2.48 comes from the definition used for the error function, $\mathrm{erf}(x) = 1/\sqrt{\pi}\int_{-x}^{x}\exp(-t^2)dt$.

The calculation of the optical depth τ and the transmitted flux for a single skewer spanning across the length of the CHIPS.P19 box at redshift $z = 2.5$ is illustrated in Figure 2.13. The figure panels from top to bottom show the distribution of the neutral hydrogen density in the neighborhood of the skewer (ρ_{HI}), the 1D neutral hydrogen column density integrated over the length of the cell across the skewer (N_{HI}), the optical depth τ in redshift space computed via Equation 2.48 (blue line) and ignoring the shift of of the absorption lines due to peculiar real space velocities (green), and the transmitted flux $F = \exp(-\tau)$ along the skewer in both redshift (blue) and real (green) space.

2.5.3 Evolution of the Lyman-α Effective Optical Depth

The Lyman-α effective optical depth τ_{eff} is a measure of the overall H I content of the gas in the IGM. Hence, τ_{eff} tracks the ionization state of hydrogen and the intensity of the ionizing UV background. To compare with observational measurements of τ_{eff}, we computed synthetic Lyman-α absorption spectra from all the outputs of our two simulations using the method described in §2.5.2. From the large sample of skewers, the effective optical depth is computed as $\tau_{eff} = -\log(\bar{F})$, where $\bar{F}$ is the transmitted flux averaged over all skewers.

The redshift evolution of the effective optical depth τ_{eff} for our two simulations is shown as colored lines in Figure 2.14 (blue and green for the HM12 and P19 models, respectively) and the shaded region shows the variability of τ_{eff} measured over the different skewers. For each redshift the shaded interval corresponds to the optical depth

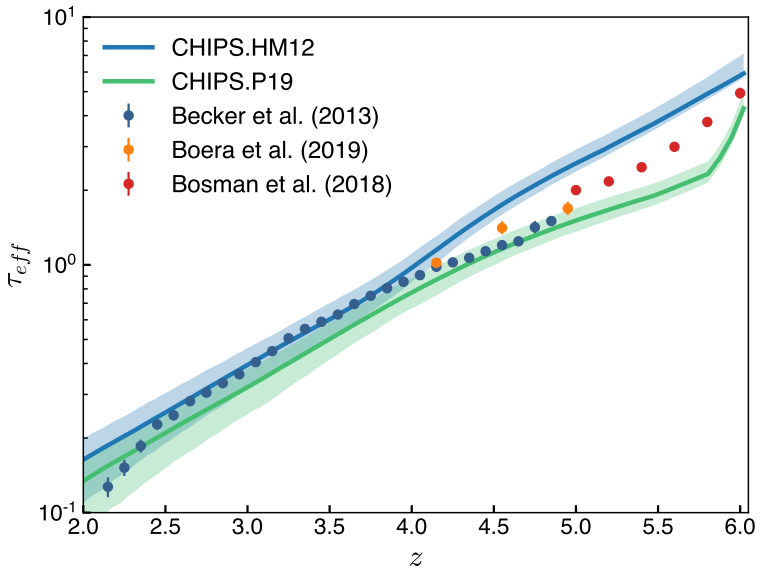

Figure 2.14: Redshift evolution of the effective optical depth τ_{eff}. Shown are our simulated measurements from CHIPS.HM12 (blue) and CHIPS.P19 (green), compared to data from Becker et al. (2013a), Boera et al. (2019) and Bosman et al. (2018) (data points). Measurements of τ_{eff} from the HM12 model match the observations at $2.5 \lesssim z \lesssim 4.2$, but display higher amounts of H I for $z \gtrsim 4.2$ owing to the low temperatures at those redshifts. The P19 model produces values of τ_{eff} slightly lower than the observations for $2.5 \lesssim z \lesssim 4$, suggesting that He II reionization is overheating the IGM at these epochs. For $5 < z < 5.8$ the P19 model results in measurements of τ_{eff} significantly lower than the observations. This discrepancy could result from the relatively hot IGM produced by the P19 model at this epoch and may be addressed by introducing a non-uniform UVB in the simulations.

computed from the highest probability interval that encloses 68% of the distribution of the transmitted flux averaged over individual skewers.

The effective optical depth resulting from the HM12 UVB model (blue) shows good agreement with the observed data (Becker et al. 2013a) for $z < 4$, but underestimates the ionization fraction at $z > 4$. At the higher redshifts, the model produces an excess of neutral hydrogen, likely because the early H I reionization renders the gas too cold, this results in an effective optical depth higher than estimated by Boera et al. (2019) and measured by Bosman et al. (2018).

The P19 model (green) has too high an ionization fraction, resulting in an optical depth that is slightly lower than the observations. In the redshift range $2.5 \lesssim z \lesssim 4$, He II reionization in the P19 model is overheating the IGM. For $5 < z < 5.8$, the P19 model produces τ_{eff} that are 10 to 25% lower than the observations, suggesting that the temperatures of the IGM in this model at $z \sim 5.2$ are higher than those in reality. At these high redshifts ($z \gtrsim 5.4$), a non-uniform UV background may be required to accurately represent the effects of a "patchy" reionization in τ_{eff} (Keating et al. 2020), and the inclusion of a non-uniform UVB could reduce the discrepancies between the data and the P19 model at these times.

2.5.4 Lyman-α Transmitted Flux Power Spectrum

On scales of a few Mpc, the Lyman-α flux power spectrum (FPS) is an excellent probe of the thermal properties of the photoionized IGM and can constrain the IGM temperature at various epochs. On scales below $\sim$ 100kpc, the FPS exhibits a cutoff beyond which the Lyman-α forest has suppressed structure owing to both the pressure

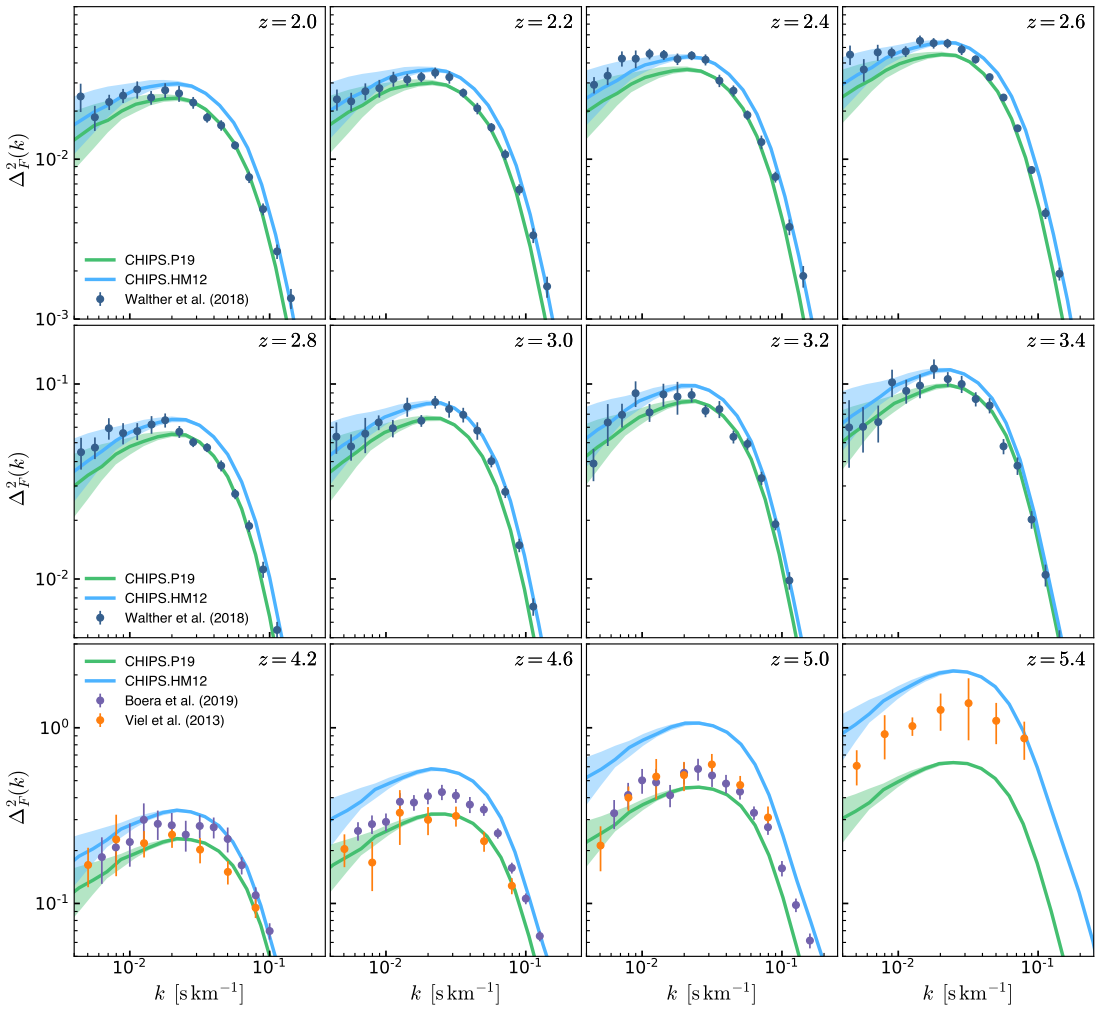

Figure 2.15: One dimensional power spectra of the Lyman-α transmitted flux fluctuations (FPS) from our two simulations CHIPS.HM12 (blue) and CHIPS.P19 (green), compared with the observational measurements from Walther et al. (2018), Boera et al. (2019) and Viel et al. (2013a) (data points). The colored lines show the FPS averaged over all the skewers, and the shaded regions show the $\sigma(k)/\sqrt{N_{\rm ind}}$ region where $\sigma(k)$ is the standard deviation of the distribution $P(\Delta_F^2)$ obtained from the FPS of the all the individual skewers and $N_{\rm ind}(k)$ is the number of independent skewers that can be drawn from the simulation grid for each axis. For $k \gtrsim 0.01$ s km^{-1} the agreement between CHIPS.P19 and the observational measurement of $P(k)$ for $2 \lesssim z \lesssim 4.5$ is relatively good (time-averaged $\langle\chi^2_\nu\rangle \sim 2$), as compared with CHIPS.HM12 ($\langle\chi^2_\nu\rangle \sim 8$). For $z \gtrsim 5$ the amplitude of $P(k)$ for the P19 UVB model is lower than the data, reflecting the lower estimate of τ_{eff} from the P19 model relative to the observations. For $k \lesssim 0.01$ s km^{-1} the resulting FPS from the HM12 model is in better agreement with the data from Walther et al. (2018) at $z \lesssim 3$ ($\langle\chi^2_\nu\rangle \sim 1$) in contrast to the P19 model ($\langle\chi^2_\nu\rangle \sim 3$).

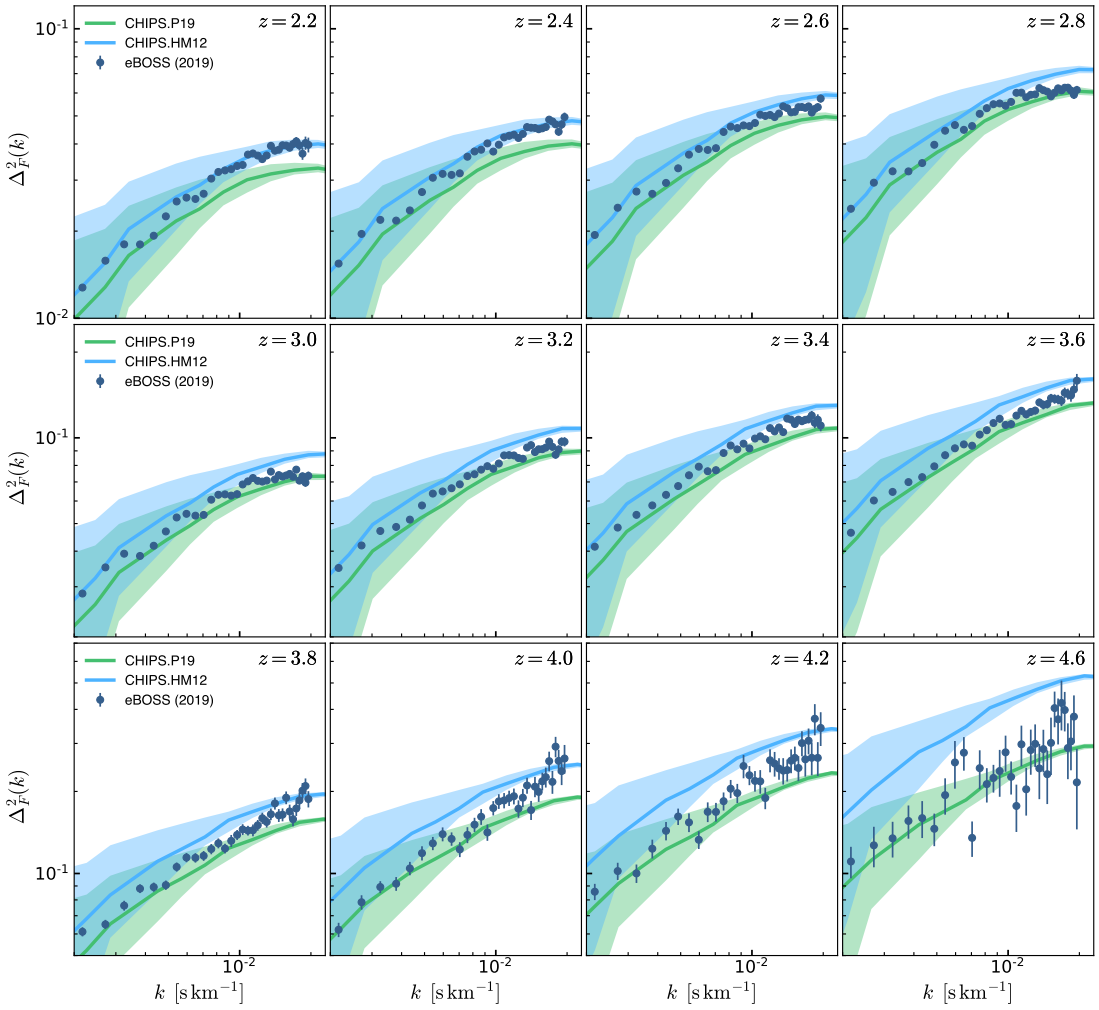

Figure 2.16: One dimensional power spectra of the Lyman-α transmitted flux fluctuations (FPS) from our two simulations CHIPS.HM12 (blue) and CHIPS.P19 (green), compared with the large scale power spectra data from the eBOSS experiment (Chabanier et al. 2019b). The colored lines show the FPS averaged over all the skewers, and the shaded regions show the $\sigma(k)/\sqrt{N_{\rm ind}}$ region where $\sigma(k)$ is the standard deviation of the distribution $P(\Delta_F^2)$ obtained from the FPS of the all the individual skewers and $N_{\rm ind}(k)$ is the number of independent skewers that can be drawn from the simulation grid for each axis. From $z = 2.2$ to $z = 4.6$ the amplitude of $P(k)$ for the models increases faster than the data, and for the range $2.6 \lesssim z \lesssim 4.2$ the data lie in between the models. The P19 and HM12 models result in $P(k)$ that match the data at $z = 4.6$ and $z = 2.2$, respectively.

smoothing of the gas distribution (which sets the pressure smoothing scale λ_P) as well as the thermal Doppler broadening of the absorption lines. To compare the FPS produced by the UVB models in the simulations with observation, we measured the Lyman-α transmitted flux power spectra from the large sample of skewers computed as described in §2.5.2. The FPS as a function of velocity u is obtained from the flux fluctuations as

$$\delta_F(u) \equiv \frac{F(u) - \bar{F}}{\bar{F}}, \tag{2.50}$$

where $F(u)$ if the transmitted flux along the skewer over the velocity interval [0, u_{max}] and $\bar{F} = \exp(-\tau_{eff})$ is the transmitted flux averaged over all the skewers. The FPS is commonly expressed in terms of the dimensionless quantity Δ^2_F, defined as

$$\Delta^2_F(k) = \frac{1}{\pi} k P(k). \tag{2.51}$$

The transmitted flux power spectrum $P(k)$ is computed as

$$P(k) = u_{\rm max} \left\langle \left| \tilde{\delta}_F(k) \right|^2 \right\rangle \tag{2.52}$$

$$\tilde{\delta}_F(k) = \frac{1}{u_{\rm max}} \int_0^{u_{max}} e^{-iku} \delta_F(u) du, \tag{2.53}$$

where $k = 2\pi/u$ corresponds to the wavenumber associated to the velocity u and has units of $\mathrm{s\,km^{-1}}$ and $u_{\rm max}$ is the Hubble flow velocity for the box length $u_{\rm max} = HL/(1+z)$. From our simulations, we measured the FPS resulting from both UVB models and compared with the analogous observational measurements. Our results are

shown in Figures 2.15 and 2.16, where the lines show the FPS averaged over all the skewers and the shaded bars show the $\sigma(k)/\sqrt{N_{\rm ind}(k)}$ region. Here $\sigma(k)$ corresponds to the standard deviation from the distribution of $P(k)$ measured over all individual skewers and $N_{\rm ind}(k) = (2\pi)^{-2}(ku_{max})^2$ is the number of independent skewers that can be drawn from the simulation grid for each axis. Applying the Central Limit Theorem, $\sigma(k)/\sqrt{N_{\rm ind}(k)}$ is analogous to the standard deviation $\sigma_s(k)$ of the distribution that results from sampling the mean FPS ($\overline{P(k)}$) over all the possible groups of independent skewers, therefore $\sigma_s(k) \simeq \sigma(k)/\sqrt{N_{\rm ind}(k)}$ provides the uncertainty in the mean FPS measured from the simulation grid.

The flux power spectrum on scales of $0.004 < k < 0.2\,{\rm s\,km^{-1}}$ is presented in Figure 2.15, as shown, the uncertainty bars in the in the power spectrum from the simulations are larger for smaller k values as the number of independent skewers that can be drawn from the grid decreases as the size of the probed fluctuation increases. We compare the simulation results against observational measurements from Walther et al. (2018) at $2 < z < 3.4$, and higher redshift measurements from Boera et al. (2019) and Hiss et al. (2018) for $4 \lesssim z \leq 5.4$. To assess the performance of both photoionization and photoheating models to reproduce the observed FPS, we quantify the differences in $P(k)$ with the statistic $\chi^2_\nu = \chi^2/N$, where

$$\chi^2 = \sum_i^N \left[\frac{P(k_i)^{\rm obs} - P(k_i)^{\rm model}}{\sigma_i^{\rm obs}}\right]^2, \tag{2.54}$$

and N is the number of observed data points measured at the wavenumbers k_i and having uncertainties σ_i. We will use χ^2 to denote the statistic computed by Equation

2.54 for multiple scales at the same redshift, and $\langle\chi^2\rangle$ to denote a "time-averaged" statistic when using multiple scales over multiple redshifts.

For $k \gtrsim 0.01$ s km^{-1}, the agreement between the CHIPS.P19 simulation (green) and the observational measurement of $P(k)$ for $2 \lesssim z \lesssim 4.5$ is relatively good as the time-averaged differences are $\langle\chi^2_\nu\rangle \sim 2$, compared to $\langle\chi^2_\nu\rangle \sim 8$ for CHIPS.HM12 (blue). At high redshifts ($z \gtrsim 5$), the observational data lie between the predictions from the two models. This result is consistent with the behavior of τ_{eff} in Figure 2.14, since the normalization of the transmitted flux fluctuations δF is determined by $\bar{F}$ and an underestimate of τ_{eff} will result in a lower normalization in $P(k)$. For $k \lesssim 0.01$ s km^{-1}, the HM12 UVB model produces FPS that agree better with the data from Walther et al. (2018) at $z \lesssim 3$ ($\langle\chi^2_\nu\rangle \sim 1$) than the P19 model ($\langle\chi^2_\nu\rangle \sim 3$). At high redshift ($z \gtrsim 4.6$), the HM12 model results in a FPS higher than the observations, and this discrepancy is again consistent with the higher values of τ_{eff} produced by the model compared with the observations, as shown in Figure 2.14.

For larger scales $0.002 < k < 0.02$ s km^{-1}, the flux power spectrum shown in Figure 2.16 is compared with the observational measurements from the eBOSS experiment presented in Chabanier et al. (2019b). The evolution of $P(k)$ in CHIPS.HM12 (blue) and CHIPS.P19 (green) differs from the observed data, with the amplitude of $P(k)$ in the models increasing faster than the data at $z = 2.2 - 4.6$. Figure 2.16 shows that at higher redshift ($z \sim 4.6$) the P19 model $P(k)$ matches the observations ($\chi^2_\nu = 1.4$), while for the redshift range $2.6 \lesssim z \lesssim 4.2$ the data lies in between the models. For $z \lesssim 2.4$, the HM12 model agrees better with the data measured by the eBOSS

experiment but the small uncertainties in the observational measurements result in large values of χ^2_ν regardless. Since the temperature of the IGM at $z \lesssim 4$ is primarily set by He II reionization, we argue that the discrepancies between the P19 results and the observed data could be alleviated by changing the He II photoheating rate associated with active galactic nuclei to reduce the IGM temperature at $z \sim 3$.

Comparing Lyman-α flux power spectrum between simulations and observation offers a direct way to assess the performance of the chosen photoionization and photoheating rates in reconstructing the thermal history of the IGM. Figures 2.15 and 2.16 show that both photoionization and photoheating models used for this work, Haardt & Madau (2012) and Puchwein et al. (2019), fail to recover the observed $P(k)$ on scales of $0.002 < k < 0.2\,\mathrm{s\,km^{-1}}$ over the redshift range $2 \lesssim z \lesssim 5$. The observed Lyman-α forest statistics τ_{eff} and $P(k)$ lie in between the results produced by the two models. This tension motivates further studies using cosmological simulations with modified photoionization and photoheating rates that result in lower IGM temperatures at $2 < z \lesssim 3$ and increase the amplitude of the FPS on large scales (e.g., $0.002 \lesssim k \lesssim 0.02\,\mathrm{s\,km^{-1}}$) to match better the observations.

2.6 Discussion

By comparing the results of our simulations with observations, we have demonstrated the broad properties of the forest are reproduced by models using either the Haardt & Madau (2012) or Puchwein et al. (2019) photoheating and photoionization rates. The Puchwein et al. (2019) rates in particular lead to realistic small-scale struc-

ture in the forest over a range in redshift. However, both models fail to recover the detailed shape of the transmitted flux $P(k)$ or the magnitude of the optical depth at all redshifts or spatial scales. The physical reasons for these inadequacies likely also depend on redshift and spatial scale, and are discussed below.

2.6.1 What is the IGM Photoheating History?

As Figures 2.11 and 2.12 illustrate, the observational inferences on the evolution of the IGM mean temperature and density-temperature relation are currently widely-varying (Hiss et al. 2018; Bolton et al. 2014; Walther et al. 2018; Boera et al. 2019; Gaikwad et al. 2020, 2021). Deriving these observed properties requires assistance from simulations, for instance by generating model spectra or aiding the interpretation of transmission spikes. As a result, discriminating between early-reionization (Haardt & Madau 2012) and late-reionization (Puchwein et al. 2019) based on $T_0(z)$ and $\gamma(z)$ remains hazardous. The Lyman-α optical depth and $P(k)$ evolution show that the observations mostly reside in between the model predictions. At redshifts $z \sim 2-3$ the Puchwein et al. (2019) rates produce structure in the forest that agrees better with the data on small scales, but on larger scales Haardt & Madau (2012) performs better. The IGM structure at these redshifts is heavily influenced by He II photoheating powered by active galactic nuclei, and both the relative spectral slopes of the AGN spectral energy distribution and the difference in emissivity with redshift in these models could affect their scale-dependent relative agreement. Finding the He II photoheating and photoionization rates that result in agreement across all scales at these redshifts will require future work and more simulations.

Close to the hydrogen reionization era, in simulations using the Haardt & Madau (2012) rates the Lyman-α $P(k)$ amplitude increases more rapidly with redshift than in the Puchwein et al. (2019) model. Reionization occurs very early in the Haardt & Madau (2012) UVB, and the $P(k)$ amplitude evolution reflects the progressively colder IGM in this scenario. The hydrogen ionization and temperature states of the IGM in the Puchwein et al. (2019) model are apparently too high, and produce a lower $P(k)$ than seen in the observation. Balancing the high temperature resulting from the late reionization in this model and the larger $P(k)$ may require a patchy reionization process, as noted previously (Keating et al. 2020).

2.6.2 IGM Thermal History vs. Instantaneous Properties

The importance of developing self-consistent histories for IGM properties, including the phase structure, flux power spectra, and Lyman-α optical depth, cannot be overstated for interpreting observations. The power of the Lyman-α forest for constraining the small-scale physics of structure formation, including the possible presence of warm dark matter (e.g., Viel et al. 2013a; Iršič et al. 2017a) and the importance of neutrinos (e.g., Chabanier et al. 2019b), is limited by the imprecisely known thermal properties that can impact such scales. These uncertainties on the thermal properties, often characterized by the temperature T_0 at the mean density and the slope γ of the IGM temperature-density relation, have frequently been treated as nuisance parameters when developing cosmological parameter constraints from the Lyman-α forest. Analyses typically marginalize over the uncertainties in the thermal structure of the forest to arrive at, e.g., the possible contribution of warm dark matter to the suppression of

small-scale power.

One complication with these analyses that our simulations highlight is how the history and evolution of the thermal properties influence the structure of the forest. The characteristics of the forest at one redshift cannot be disentangled entirely from its properties at similar redshifts. The response of the forest, in terms of both its thermal and ionization structure, depends on the evolving photoheating and photoionization rates, and the values of both T_0 and γ change along tracks with redshift. These properties are not independent, and have redshift correlations that cannot be ignored by separately marginalizing over their properties independently at an array of redshifts. Instead, when marginalizing over the thermal structure of the forest to infer constraints from $P(k)$, full simulated histories of the forest properties are required with marginalization occurring simultaneously over the forest structure at all redshifts where observations are available. For instance, synthesizing the results shown in Figures 2.11, 2.12, 2.15, we find that the consistently lower $P(k)$ amplitude of the forest at $z = [4.2, 4.6, 5.0, 5.4]$ when using the Puchwein et al. (2019) rates results from the hotter IGM induced by the late global reionization in this model. The temperature at these times is not independent across redshift, and varying models over a range of IGM temperature and Lyman-α $P(k)$ at a given redshift amounts to direct assumptions on those properties at adjacent redshifts where the dominant photoheating mechanisms are the same.

Addressing this issue requires a potentially large number of hydrodynamical simulations of cosmological structure formation that capture various photoheating histories and manage to resolve the Lyman-α forest structure robustly on small scales.

Our CHIPS simulations represent a first step in this direction, and the computational efficiency of the *Cholla* code will enable us to realize the required number of simulations with moderate additional effort.

2.7 Summary

Motivated by new observational efforts that will provide unprecedented detail on the properties of the Lyman-α forest (e.g., DESI Collaboration et al. 2016a), we have initiated the CHIPS series of hydrodynamical simulations of cosmological structure formation. Our simulations use the GPU-native *Cholla* code to maintain exquisite spatial resolution in the low-density intergalactic medium throughout cosmological volumes. In this first paper, we conduct $N = 2048^3$ resolution simulations to compare the thermal history and physical properties of the Lyman-α forest using two models for the photoheating and photoionization rates induced by an evolving ultraviolet background (Haardt & Madau 2012; Puchwein et al. 2019). A summary of our efforts and conclusions follows.

- We extended the *Cholla* code to perform cosmological simulations by engineering gas self-gravity, a Fourier-space Poisson solver, a particle integrator, a comoving coordinate scheme, and a coupling to the GRACKLE heating and cooling library (Smith et al. 2017).

- We provided extensive tests of our cosmological simulations, including the Zel'Dovich (1970) pancake test and comparisons with the results of other cosmological simulation codes. *Cholla* agrees with *Nyx* (Almgren et al. 2013), *Ram-*

ses (Teyssier 2002), and *Enzo* (Bryan et al. 2014) to sub-percent accuracy on all spatial scales for N-body cosmological simulations, and matches *Ramses* and *Enzo* results to within a few percent for adiabatic hydrodynamical simulations.

- We provide a new method for testing the dual energy formalism (Bryan et al. 1995) of Eulerian hydrodynamical codes for cosmological simulations by matching the mean cosmic gas temperature in adiabatic simulations, and show that *Cholla* recovers the expected results.

- In accordance with prior results (e.g., Puchwein et al. 2019), we find that after hydrogen reionization, simulations using the Haardt & Madau (2012) photoheating rates predict a cooler IGM temperature than the Puchwein et al. (2019) UVB model. At redshfits $z \sim 4-6$, the Puchwein et al. (2019) model is hotter owing to hydrogen reionization completing later than in the Haardt & Madau (2012) scenario. At redshifts $z \lesssim 4$, the IGM is hotter in the Puchwein et al. (2019) model owing to the He II photoheating rates powered by active galactic nuclei.

- We compare the Lyman-α transmitted flux power spectra $P(k)$ computed for these simulations with observations. We find that at redshifts $2 \lesssim z \lesssim 5.5$ the performance of the models varies with scale. Using the Puchwein et al. (2019) photoheating rates results in good agreement with the observed $P(k)$ on $k \gtrsim 0.01\ \mathrm{s\,km^{-1}}$ at $2 \lesssim z \lesssim 4.5$.

- On larger scales, the amplitude of the observed $P(k)$ increases faster from $z \sim 2.2$ to $z \sim 4.6$ than the structure in the simulated forest. The observations appear

intermediate between the simulation results using the Haardt & Madau (2012) and Puchwein et al. (2019) photoheating rates at $2.2 \lesssim z \lesssim 4.6$ for $k \approx 0.002 - 0.01$ s km^{-1}.

- At higher redshifts $z \gtrsim 4.5$, as the epoch of hydrogen reionization is approached, the $P(k)$ amplitude in the simulations increase at rates that bracket the observed flux $P(k)$. The observed Lyman-α optical depth also lies in between the model predictions at these redshifts.

- We show that our results are insensitive to small changes in the cosmological parameters comparable to the Planck Collaboration et al. (2020) uncertainties, and demonstrate our results for the flux power spectra have converged with resolution studies.

These initial CHIPS simulations demonstrate that commonly used models for the photoheating and photionization rates (Haardt & Madau 2012; Puchwein et al. 2019) broadly reproduce the observed thermal history and transmitted flux power spectra of the Lyman-α forest. However, in detail the agreement with the observations can be improved, including better recovering the redshift- and scale-dependence of the the flux $P(k)$ and the evolution in the Lyman-α optical depth. Matching these observations more completely will require changing the photoionization and photoheating rates for both hydrogen and helium. We will explore these improvements using additional large-scale cosmological simulations in future work.

2.8 Resolution Convergence Analysis

To assess the numerical convergence of our results, we performed runs with different resolutions. Each simulation was run with the same box size ($L = 50h^{-1}$Mpc) and identical cosmological parameters (Planck Collaboration et al. 2020), but with differing resolutions of $N = 512^3$, $N = 1024^3$, and $N = 2048^3$ cells and dark matter particles. These simulations have comoving spatial resolutions of $\Delta x = 98$, 49, and 24 h^{-1} kpc, respectively. The initial conditions were generated to preserve the large-scale modes in common to each simulation, such that the properties of the simulations could be compared directly on shared spatial scales. We measured the Lyman-α effective optical depth and transmitted flux power spectrum at each resolution.

The left panel of Figure 3.17 shows the convergence of the Lyman-α transmitted flux power spectra with resolution. As the resolution increases, the large-scale $P(k)$ decreases while the small-scale power increases. This progression reflects the structure of the forest becoming better resolved as the number of cells increases. The $N = 1024^3$ and $N = 2048^3$ simulations agree well over most spatial scales, with little evidence that the $N = 2048^3$ simulation requires further refinement on scales $k \gtrsim 0.007$ s km^{-1}.

The right panel of Figure 3.17 details the redshift evolution of the Lyman-α optical depth measured in the convergence study. As the structure of the forest becomes better resolved, the optical depth lowers. This decline reflects the decrease in large-scale power as the resolution improves, as illustrated in the left panel of Figure 2.19. As with the $P(k)$ analysis, the $N = 1024^3$ and $N = 2048^3$ simulations agree well in their Lyman-α optical depth evolution. The large-scale structure of the forest also

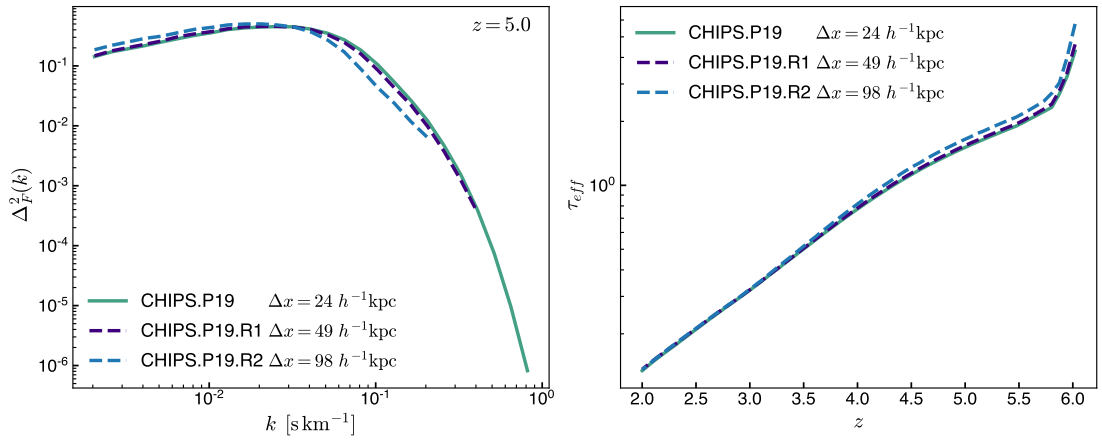

Figure 2.17: Lyman-α forest statistics measured for similar simulations with different spatial resolutions $\Delta x = 98$, 49 and 24 h^{-1}kpc (comoving). All simulations have $L = 50h^{-1}$Mpc and use the Puchwein et al. (2019) UVB model and Planck Collaboration et al. (2020) cosmological parameters. The transmitted flux power spectrum Δ_F^2 at $z = 5$ is shown in the left panel and the redshift evolution of the effective optical depth τ_{eff} is shown on the right. Both measurements demonstrate that the relevant statistics of the Lyman-α forest have converged for the high resolution simulations used in this work.

agrees well between these simulations, demonstrating that the $N = 2048^3$ simulations have converged on scales that most contribute to the optical depth.

2.9 Cosmological Parameter Study

The relative difference in the Lyman-α forest properties between simulations using the Haardt & Madau (2012) or Puchwein et al. (2019) rates is substantial, but understanding whether these differences are large compared differences in the forest resulting from cosmological parameter variations requires further simulation. To answer this question we repeated our highest resolution simulation using the Puchwein et al. (2019) rates with a range of cosmological parameters, varying over the reported un-

Table 2.2: Mean Gas Density Comparison for Alternative Cosmologies

Simulation	Mean baryon density $\bar{\rho}_b$ [$M_\odot kpc^{-3}$]	$\Delta\bar{\rho}_b/\bar{\rho}_b$ [$\times 10^{-2}$]
CHIPS.P19	6.315	
CHIPS.P19.A1	6.275	-0.62
CHIPS.P19.A2	6.334	0.31
CHIPS.P19.A3	6.394	1.25
CHIPS.P19.A4	6.449	2.12

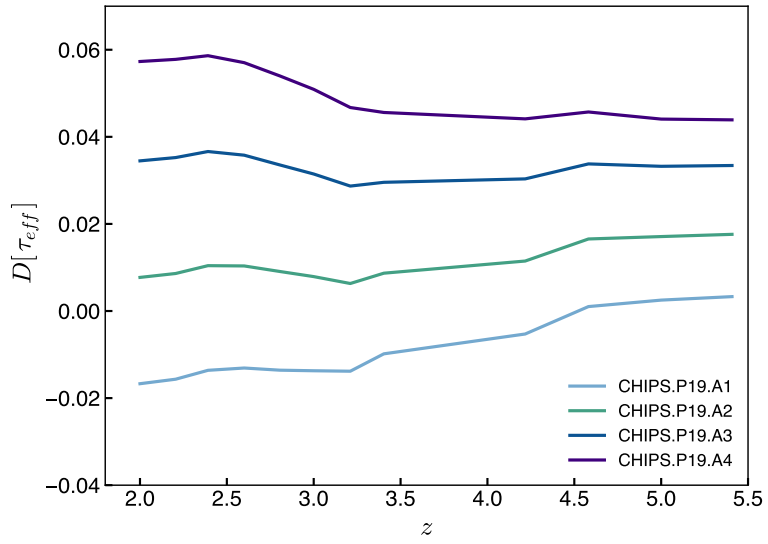

Figure 2.18: Redshift-dependent fractional difference of the effective optical depth τ_{eff} in the alternative cosmology simulations, measured with respect to the fiducial CHIPS.P19 simulation that evolves a Planck Collaboration et al. (2020) cosmology. The variations in τ_{eff} reflect differences in the mean baryonic density $\bar{\rho}_b$ for the alternative cosmologies shown in Table 2.2.

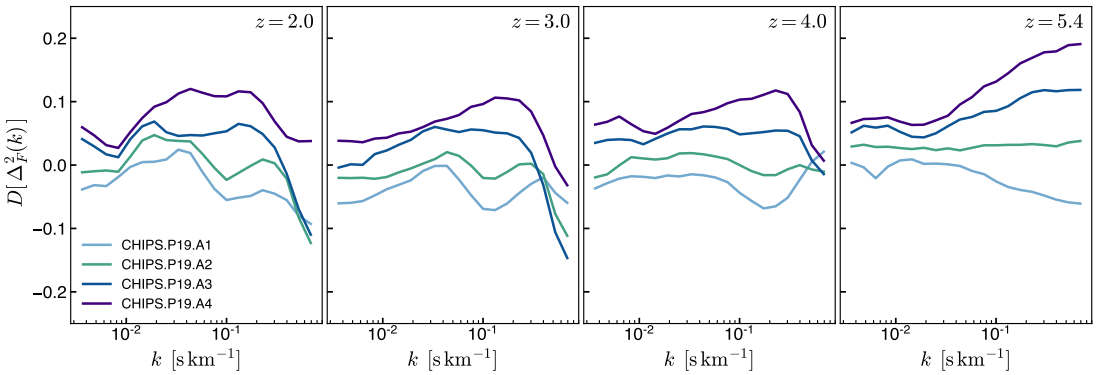

Figure 2.19: Fractional differences in the flux power spectra Δ_F^2 of the alternative cosmology simulations. The differences are measured with respect to the fiducial CHIPS.P19 simulation that evolves a Planck Collaboration et al. (2020) cosmology (shown for redshifts $z = [2, 3, 4, 5.4]$). The differences are scale dependent, and the overall variations in normalization reflect the differences in τ_{eff} shown in Figure 2.18.

certainty in the Planck Collaboration et al. (2020) analysis. We ran four simulations where we varied the Hubble parameter H_0, matter density Ω_m, baryon density Ω_b, RMS fluctuations on 8 h^{-1}Mpc scales σ_8, and the spectral slope of initial perturbations n_s. The chosen numerical values of the cosmological parameters are reported in Table 2.1.

Table 2.2 shows the mean comoving baryonic density $\bar{\rho}_b$ for the reference simulation CHIPS.P19 and the simulations with alternative cosmologies. Additionally, for the alternative cosmologies the fractional difference of the average density $\Delta\bar{\rho}_b/\bar{\rho}_b$ relative to the reference CHIPS.P19 is shown in the second column. The differences in the mean gas density are approximately a few percent and reflect differences in the mean optical depth measured in the alternative cosmology simulations with respect to the CHIPS.P19 simulation. The differences in the effective optical depth $D[\tau_{eff}]$ relative to CHIPS.P19 as a function of redshift are shown in Figure 2.18. The differences in τ_{eff} range from -2% to 6%, exhibit little evolution with redshift, and cannot, e.g., ac-

count for the deviations of τ_{eff} resulting from the Puchwein et al. (2019) model and the observational measurements shown in Figure 2.14.

Figure 2.19 shows the resulting variation in the transmitted flux $P(k)$ induced by the small variations introduced in the cosmological parameters. The differences in the power spectrum $D[\Delta_F^2]$ relative to the reference CHIPS.P19 are shown for four snapshots at redshifts between $z = 2.0$ and $z = 5.4$. The differences in $P(k)$ are scale dependent, but are consistent with the differences shown in Figure 2.18 for τ_{eff} as the mean transmitted flux $\bar{F}$ sets the normalization of $P(k)$. To disentangle the effect that varying each cosmological parameter has on the Lyman-α statistics a more extensive study would have to be performed, but we conclude the overall effect of the cosmological variations is small and as Figure 2.19 demonstrates, the difference in physical structure between simulations conducted with the Haardt & Madau (2012) or Puchwein et al. (2019) rates cannot be easily mimicked with cosmological parameter variations allowed by experimental constraints. This result also emphasizes the need to account for the IGM thermal history when inferring cosmological properties from the forest.

We also compared the evolution of the thermal parameters T_0 and γ from the alternative cosmology simulations to the reference CHIPS.P19 simulation, and found that the differences in T_0 and γ relative to the Planck Collaboration et al. (2020) cosmology were $\lesssim 3\%$. Since these differences are of the order of the uncertainties resulting from modeling the density-temperature distribution as a power-law relation (Equation 3.7), we do not report any significant variation in the thermal history of the IGM owing to small variations of the cosmological parameters.

Chapter 3

Inferring the Thermal History of the Intergalactic Medium from the Properties of the Hydrogen and Helium Lyman-α Forest

3.1 Introduction

The neutral hydrogen and singly-ionized helium components of gas near the cosmic mean density trace the distribution of matter in between galaxies and produce a "forest" of detectable Lyman-α absorption features in the spectra of distant quasars (e.g., Hernquist et al. 1996; Croft et al. 1998a; Meiksin 2009a; Slosar et al. 2011; McQuinn 2016; Worseck et al. 2019). The depth, shape, and location of absorption lines

in the Lyman-α forest depend on the ionization degree and thermal state of this intergalactic medium (IGM), which are controlled by the uncertain UV radiation background produced by star-forming galaxies and active galactic nuclei (AGNs) (e.g., Haardt & Madau 2012; Robertson et al. 2015; Madau & Haardt 2015; Faucher-Giguère 2020), and on its density and peculiar velocity fields shaped by gravity (Cen et al. 1994). Dark matter provides the backbone of large-scale structure in the Universe, a web-like pattern present in embryonic form in the overdensity motif of the initial fluctuation field and sharpened by non-linear gravitational dynamics (Bond et al. 1996). The Lyman-α forest traces this underlying "cosmic web" on scales and at redshifts that cannot be probed by any other observable. Because of its long cooling time, low-density gas at $z \sim 2$–5 that traces the underlying matter distribution retains some memory of when and how it was reheated and reionized at $z \gtrsim 6$ (Miralda-Escudé & Rees 1994). The physics that governs the properties of the IGM throughout these epochs remain similar, as the evolving cosmic UV emissivity and the transfer of that radiation through a medium made clumpy by gravity determine both the details of the reionization process and the thermodynamics of the forest.

Understanding how reionization occurred, the nature of the early sources that drove it, the thermal history and fine-grained properties of hydrogen gas in the cosmic web, and how to extract crucial information on the cosmological model from observations of Lyman-α absorption are among the most important open questions in cosmology and key science drivers for numerous major new instruments and facilities. The promise of the Lyman-α forest for constraining cosmological physics including the nature of dark

matter and dark energy has motivated in part the construction of the Dark Energy Spectroscopic Instrument (DESI Collaboration et al. 2016a), which measures absorption line spectra backlit by nearly a million quasars at $z > 2$, and the WEAVE survey (Pieri et al. 2016) which will observe more than 400,000 high-redshift quasars at $z > 2$. Interpreting such observations requires detailed cosmological hydrodynamical simulations that cover an extensive range of uncertain photoionization and photoheating histories and consistently maintain high resolution throughout a statistically representative sub-volume of the Universe.

This paper extends research efforts directly focused on advancing the state-of-the-art in modeling the IGM physical structure in cosmological simulations while still achieving high computational efficiency, thereby providing higher fidelity physical models for interpreting Lyman-α forest data. In Villasenor et al. (2021a) we introduced the Cholla IGM Photoheating Simulations (CHIPS) to investigate how different photoheating histories and cosmological parameters impact the structure of the forest. Here, we use a massive suite of more than 400 CHIPS simulations to study the IGM at a resolution of $49h^{-1}$ ckpc maintained over $(50h^{-1}$ cMpc$)^3$ volumes. Performed with the GPU-native MPI-parallelized code *Cholla* (Schneider & Robertson 2015), these simulations span different amplitudes and peak redshifts of the H I and He II photoionization and photoheating rates.

To anticipate the results of our likelihood analysis constrained by the 1D flux power spectra $P(k)$ measured in eBOSS, Keck, and VLT data and the observed He II Lyman-α forest, we find that scenarios where hydrogen in the cosmic web was fully

reionized by star-forming galaxies by redshift $z_{\rm R} \approx 6.0$ and the double reionization of helium was completed by quasar sources about 1.2 billion years later are strongly favored by the data. Models that reionize hydrogen or helium at earlier or later cosmic times produce too much or too little cold gas, and appear to be inconsistent with the observed $P(k)$ and He II Lyman-α opacity. Our approach differs from previous work in this field in the following aspects:

1. The simulation grid captures a wide range of possible thermal histories via a four-parameter scaling of the amplitude and timing of the (spatially uniform) meta-galactic UVB responsible for determining the ionization states and temperatures of the IGM (cf. Nasir et al. 2016b; Oñorbe et al. 2017a). We use the physically-motivated model of Puchwein et al. (2019) as a template, and vary the strength and redshift-timing of their ionization and heating rates.

2. We do not modify, in post-processing, the mean transmitted flux $\langle F \rangle$ in the forest by recalibrating the Lyman-α optical depth, nor do we assume or rescale an instantaneous gas temperature-density relation (cf. Viel et al. 2013a; Iršič et al. 2017a; Boera et al. 2019; Walther et al. 2021). Indeed, we find from our simulations that the often assumed perfect power-law relationship between the temperature and density of the IGM does not provide a good approximation over the relevant density and redshift intervals.

3. Our likelihood analysis evaluates the performance of a given model in matching the observations over the complete self-consistently evolved reionization and thermal history of the IGM, i.e. over the full redshift range $2.2 \leq z \leq 5.0$ for the observed

1D flux power spectrum of hydrogen and over the redshift range $2.4 < z < 2.9$ for the Lyman-α opacity of He II. Since the properties of the gas at one redshift cannot be disentangled from its properties at previous epochs and the thermal and ionization structure of the IGM evolve with cosmic time along continuous trajectories, the marginalization over the parameter posterior distributions should not be performed independently at each redshift (cf. Bolton et al. 2014; Nasir et al. 2016b; Hiss et al. 2018; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021).

This paper aims to find the optimal photoionization and photoheating rates that reproduce the observed properties of the hydrogen and helium Lyman-α forest. In Section 3.2 we describe the simulations used for this work, how we apply transformations to the UV background (UVB) model from Puchwein et al. 2019 to generate our range of photoionization and photoheating rates, and the impact of the different UVB models on the statistics of the forest and the properties of the IGM. We follow by presenting the observational data and the methodology for the Bayesian Markov Chain Monte Carlo (MCMC) inference used to constrain the model. Section 3.3 presents our result for the best-fit model and the comparison of the resulting properties of the forest and the thermal evolution of the IGM to the observational determinations and previous inferences. We summarize our results and conclusions in §3.4. In Appendix 3.5 we discuss resolution effects on the Lyman-α power spectrum $P(k)$ from our simulations. A quantitative study of the impact on $P(k)$ from rescaling the effective optical of the skewer sample is presented in Appendix 3.6. In Appendix 3.7 we show the variation in the covariance matrix of the Lyman-α power spectrum from our simulations. We

discuss in Appendix 3.8 how possible alterations to our model can modify the predicted temperature history of the IGM. Finally, Appendix 3.9 analyzes the accuracy of assuming a power-law relation for the density-temperature distribution of the gas in our simulations.

3.2 Methodology

For the study presented here, we compare the observed statistics of the Lyman-α forest to simulations that apply different models for the metagalactic UVB. In this section we briefly describe our simulation code and the method to extract Lyman-α spectra from the simulations. We then describe our simulation grid and the effects that the different UVB models have on the properties of the IGM. Finally we present the observational measurements and the inference method used to constrain our model for the UVB photoionization and photoheating rates.

3.2.1 Simulations

The simulations used for this work were run with the cosmological hydrodynamics code *Cholla* (Schneider & Robertson 2015; Villasenor et al. 2021a). *Cholla* evolves the equations of hydrodynamics on a uniform Cartesian grid using a finite volume approach with a second-order Godunov scheme (Colella & Woodward 1984). The simulations track the ionization states of hydrogen and helium given by the photoionization from the UVB, recombination with free electrons and collisional ionization. The non-equilibrium H+He chemical network is evolved simultaneously with the hydrodynamics

using the GRACKLE library (Smith et al. 2017). We assume a spatially uniform, time-dependent UVB in the form of redshift-dependent photoionization rates per ion Γ and photoheating rates per ion $\mathcal{H}$ for neutral hydrogen H I, neutral helium He I, and singly ionized helium He II. For a detailed description of the simulation code we refer the reader to the methodology section presented in Villasenor et al. (2021a).

The initial conditions for our simulations were generated using the MUSIC code (Hahn & Abel 2011b) for a flat ΛCDM cosmology with parameters $H_0 = 67.66$ km s^{-1}, $\Omega_m = 0.3111$, $\Omega_\Lambda = 0.6889$, $\Omega_b = 0.0497$, $\sigma_8 = 0.8102$, and $n_s = 0.9665$, consistent with the constraints from Planck Collaboration et al. (2020). In future work, we plan to extend our analysis and include variation of the cosmological parameters (Bird et al. 2019; Ho et al. 2022). Unless otherwise stated the volume and numerical size of our simulations correspond to $L = 50\ h^{-1}$Mpc and $N = 2{\times}1024^3$ cells and particles. The initial conditions for all runs were generated from identical random number seeds to preserve the same amplitude and phase for all initial Fourier modes across the simulation suite.

3.2.2 Synthetic Lyman-α Spectra

The Lyman-α forest sensitively probes the state of the baryons in the IGM, and absorption lines from the forest reflect the H I content and the temperature of the gas in the medium. To compare the properties of the IGM in our simulations directly to observations, we extract synthetic hydrogen Lyman-α forest spectra from the simulated boxes by measuring the H I density, temperature, and peculiar velocity of the gas along 12,228 skewers through the simulation volume, using 4096 skewers along each axis of the

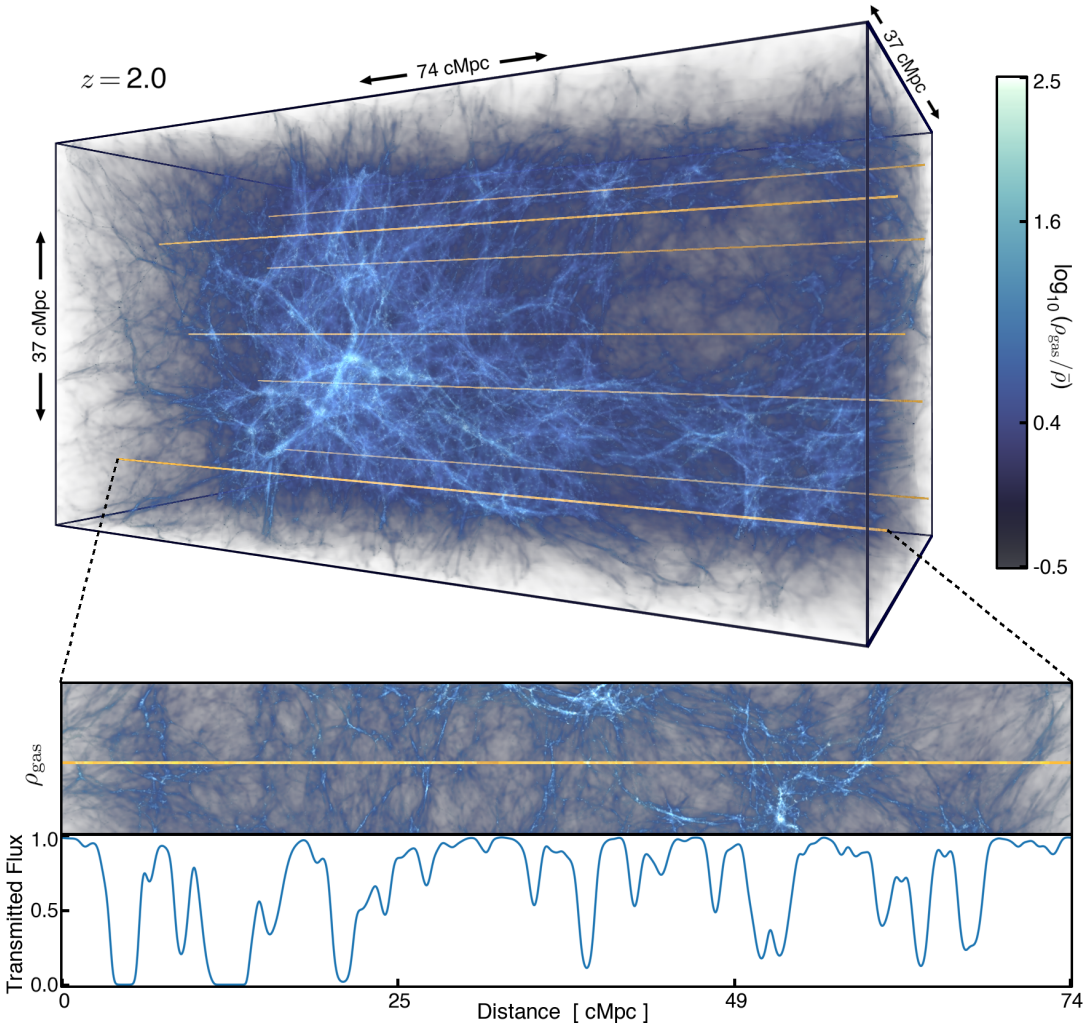

Figure 3.1: Large-scale distribution of gas density (top) from one of our highest-resolution cosmological simulations (L=50 h^{-1}Mpc, N=2×2048^3 cells and particles) at redshift z=2 and a set of skewers crossing the simulated box (yellow lines). The bottom panels show the density of the gas surrounding a selected line of sight and the Lyman-α transmitted flux along the skewer. Absorption lines in the forest probe the H I column density, the peculiar velocity, and the temperature of the gas along the line of sight.

box. The optical depth τ as a function of velocity u along each skewer is computed by integrating the product of the Lyman-α scattering cross section and the number density of neutral hydrogen along the line of sight as described in Villasenor et al. (2021a).

The transmitted flux F is computed from the optical depth τ along the skewers according to $F = \exp(-\tau)$. The power spectrum of the transmitted flux $P(k)$ is calculated as the average amplitude of the one-dimensional Fourier transform of the flux fluctuations $\delta_F(u)$,

$$\delta_F(u) \equiv \frac{F(u) - \langle F \rangle}{\langle F \rangle}, \tag{3.1}$$

where $\langle F \rangle$is the average transmitted flux over the skewer sample at a given redshift (see §5.4 from Villasenor et al. 2021a for a detailed description). Similarly, we extract the flux $F_{\rm HeII}$ transmitted through the He II Lyman-α forest from the simulations, and compute the He II effective optical depth as $\tau_{\rm eff,HeII} = -\ln \langle F_{\rm HeII} \rangle$.

Figure 3.1 (top) shows the gas density distribution at redshift $z = 2$ from a section taken from one of our highest-resolution (L=50 h^{-1}Mpc, $N = 2{\times}2048^3$ cells and particles) simulations, where several skewers crossing the simulated box are shown as yellow lines. The bottom panels show the gas density surrounding a selected line of sight and the transmitted hydrogen Lyman-α flux along the skewer. The absorption lines in the forest probe the H I column density, the peculiar velocity, and the temperature of the gas along the line of sight.

3.2.3 Photoionization and Photoheating Rates

The ionization and thermal evolution of the IGM is primarily determined by the radiation emitted by star-forming galaxies and AGNs over cosmic history (McQuinn 2016; Upton Sanderbeck et al. 2016; Oñorbe et al. 2017a). The photoionization and photoheating rates adopted in our simulations are computed from the intensity of the background radiation field, which is in turn determined by the emissivity of the radiating sources and the opacity of the IGM to ionizing photons. Recent models of the UVB (Puchwein et al. 2019; Khaire et al. 2019; Faucher-Giguère 2020), when applied to cosmological simulations, result in a hydrogen reionization era that completes by $z \sim 6$–8 in agreement with observational constraints (Davies et al. 2018b; Planck Collaboration et al. 2020).

The updated model for the photoheating and photoionizing background presented in Puchwein et al. (2019, hereafter P19) adopts an improved treatment of the IGM opacity to ionizing radiation that consistently captures the transition from a neutral to an ionized IGM. To compute the intensity of the background radiation, the P19 model employs recent determinations of the ionizing emissivity due to stars and AGNs and of the H I absorber column density distribution, and assumes an evolving escape fraction of ionizing radiation from galaxies into the IGM that reaches 18%. When the P19 model is applied in cosmological simulations, hydrogen reionization completes at $z \sim 6$ consistently with recent measurements (Becker et al. 2001; Bosman et al. 2018; Becker et al. 2021; Qin et al. 2021). However, the subsequent evolution of the Lyman-α forest spectra measured in simulations that use the P19 model fail to reproduce the

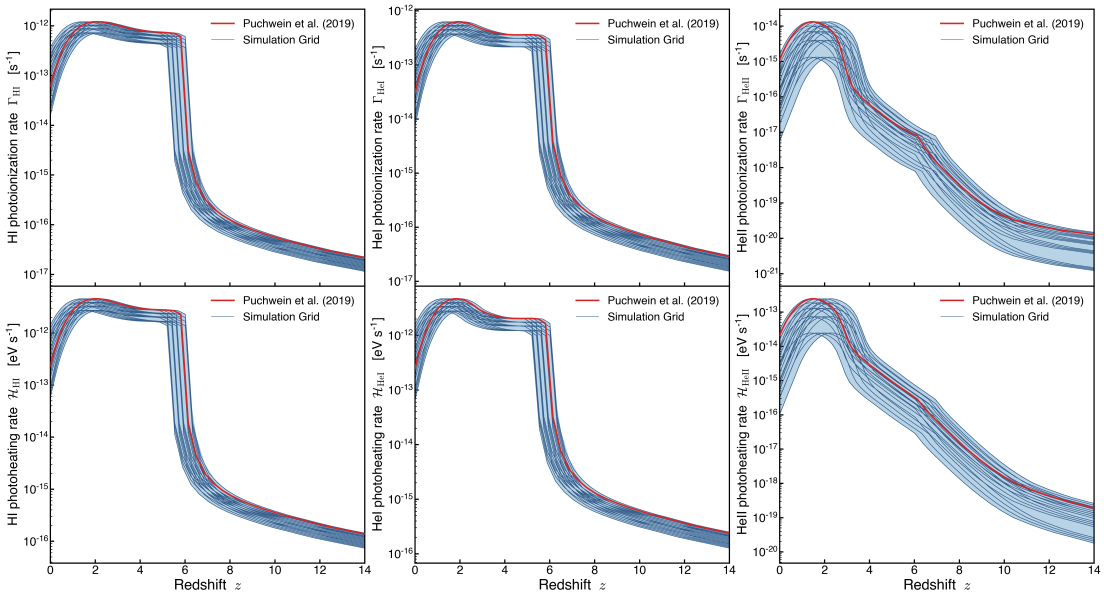

Figure 3.2: Photoionization (Γ, top) and photoheating ($\mathcal{H}$, bottom) rates for neutral hydrogen (H I, left), neutral helium (He I, center) and singly ionized helium (He II, right) from the reference P19 model (Puchwein et al. 2019) (red line) along with the photoionization and photoheating rates (blue lines) used for the 400 simulations of the CHIPS grid. The modified rates are generated by rescaling and shifting the reference P19 model as described in §3.2.4.

observed properties of the forest (Villasenor et al. 2021a) and, in particular, do not agree with the observed power spectrum of the Lyman-α transmitted flux over the redshift range $2.2 \leq z \leq 5.0$. This work aims to present a new model photoionization and photoheating rates that result in an evolution of the IGM consistent with the observational measurements of the Lyman-α flux power spectrum and the He II effective optical depth.

3.2.4 Simulation Grid

To determine ionization and heating histories that result in properties of the IGM consistent with the observed Lyman-α flux power spectrum and He II effective

opacity, we perform an unprecedented grid consisting of 400 cosmological simulations as a direct extension of the Cholla IGM Photoheating Simulations (CHIPS) suite originally presented in Villasenor et al. (2021a). Each simulation in the CHIPS grid applies different photoionization and photoheating rates to model a variety of reionization and thermal histories, and thereby produce different statistical properties for the Lyman-α forest. To generate different representations of the UVB, we modify the reference model from Puchwein et al. (2019) by rescaling the photoionization and photoheating rates (Γ and $\mathcal{H}$ respectively) by a constant factor β and shifting the redshift dependence of the rates by an offset Δz. The two transformations are expressed as

$$\begin{aligned} \Gamma(z) &\rightarrow \beta\, \Gamma^{\mathrm{P19}}\,(z - \Delta z), \\ \mathcal{H}(z) &\rightarrow \beta\, \mathcal{H}^{\mathrm{P19}}(z - \Delta z). \end{aligned} \tag{3.2}$$

Since the photoionization and photoheating rates for both H I and He I are dominated by the same sources, namely star-forming galaxies at $z \gtrsim 5$ and AGNs at lower redshifts, and the radiation that ionizes both species is absorbed by intergalactic hydrogen, we modify the H I and He I photoionization and photoheating rates jointly by applying the transformations described by Eqs. (4.14), scaling and shifting by the parameters β_{H} and Δz_{H} respectively. He II is reionized later in cosmic history primarily by the extreme UV radiation emitted by AGNs, and we rescale and redshift-offset the photoionization and photoheating rates associated with He II by a second set of parameters β_{He} and Δz_{He}. Hence, each modified UVB model is characterized by the parameter vector $\theta = \{\beta_{\mathrm{H}},\, \Delta z_{\mathrm{H}},\, \beta_{\mathrm{He}},\, \Delta z_{\mathrm{He}}\}$. The different photoionization and photoheating histories

span all the combinations of the parameter values presented in Table 3.1.

The rescaling parameters $\beta_{\rm H}$ and $\beta_{\rm He}$ control the intensity of the background radiation, determine the efficiency with which H I and He II become ionized, and govern energy input into the IGM in the form of photoheating during the epochs of non-equilibrium reionization for hydrogen and helium. After reionization completes and the gas reaches photoionization equilibrium, the balance between ionizations from the background radiation and recombinations with free electrons determines the ionization state of H I and He II. At equilibrium, the ionized fraction of H I and He II is proportional to the photoionization rates $\Gamma_{\rm HI}$ and $\Gamma_{\rm HeII}$ respectively, and inversely proportional to the temperature-dependent radiative recombination rates $\alpha_{\rm HII}(T)$ and $\alpha_{\rm HeIII}(T)$. Therefore, by rescaling the photoionization and photoheating rates, we modify the evolution of the temperature and the ionization state of the gas in the IGM during and after H I and He II reionization.

The parameters $\Delta z_{\rm H}$ and $\Delta z_{\rm He}$ shift the redshift dependence of the photoionization and photoheating rates by a constant offset, affecting the timing of H I and He II reionization. In general, an offset of $\Delta z_{\rm H} > 0$ or $\Delta z_{\rm He} > 0$ moves H I or He II reionization to higher redshift and earlier cosmic time relative to the reference P19 model. Negative values of $\Delta z_{\rm H}$ or $\Delta z_{\rm He}$ shift reionization to lower redshift and later cosmic times. The offset in redshift of the models also affect the properties of the IGM after H I and He II reionization complete, as the photoheating and photoionization rates at a given redshift are generally modified when $\Delta z_{\rm H} \neq 0$ or $\Delta z_{\rm He} \neq 0$.

Figure 3.2 shows the photoionization and photoheating rates from the reference

Table 3.1: CHIPS Simulation Grid

Parameter	Parameter Values
$\beta_{\rm H}$	0.60, 0.73, 0.86, 1.00
$\Delta z_{\rm H}$	-0.6, -0.4, -0.2, 0.0, 0.2
$\beta_{\rm He}$	0.10, 0.30, 0.53, 0.76, 1.00
$\Delta z_{\rm He}$	-0.1, 0.2, 0.5, 0.8

Note. — The parameters $\beta_{\rm H}$ and $\Delta z_{\rm H}$ determine the amplitude and redshift offset of the H I and He I photoionization and photoheating rates, while $\beta_{\rm He}$ and $\Delta z_{\rm He}$ rescale and offset the He II rates.

model by Puchwein et al. (2019) together with the modified rates adopted in the 400 simulations of the CHIPS grid. In Villasenor et al. (2021a), we presented a comparison of the statistical properties of the Lyman-α forest and the thermal history of the IGM that result from a high-resolution simulation using the UVB model from Puchwein et al. (2019). We concluded that, in general, the gas in the simulation was too highly ionized after hydrogen reionization and possibly too hot during the epoch of helium reionization to be compatible with the observed statistics of the forest and other inferences of the thermal state of the IGM. We therefore do not include values of $\beta_{\rm H} > 1$ or $\beta_{\rm He} > 1$ in our grid, as such models would result in overall higher ionization fractions and temperatures of the IGM compared with the P19 case.

The simulations were run on the Summit system (Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory). Each simulation was performed on 128 GPUs and completed in less than two wall clock hours. The cost of the entire grid of computations was only ∼16,000 node hours. This work demonstrates that by

taking advantage of an efficient code like *Cholla* and a capable system like Summit, future studies of the IGM using thousands of cosmological simulations are now possible.

3.2.5 Effects of UVB Models on the IGM Properties

The different photoionization and photoheating histories adopted in our simulations affect the ionization state of hydrogen and helium and the temperature of the IGM. Figure 3.3 shows the redshift evolution of the global properties of the IGM for each of the simulated histories. The top panels show the temperature of gas at mean density T_0 (left) and the index γ (right) of the power-law density-temperature relation $T(\Delta) = T_0 \Delta^{\gamma-1}$, where $\Delta = \rho_{\rm gas}/\bar{\rho}$. The bottom panels show the volume weighted average fraction of neutral hydrogen $x_{\rm HI}$ (left) and singly ionized helium $x_{\rm HeII}$ (right).

As hydrogen becomes ionized at $z \gtrsim 5.5$ the gas in the IGM experiences a monotonic increase of T_0 while showing a close to isothermal distribution $\gamma \sim 1$. After hydrogen reionization ends at $z \sim 5.5$–6.5, the gas cools primarily through the adiabatic expansion of the Universe. During this period the low-density gas cools faster and γ increases. This first epoch of cooling ends with the onset of helium (He II) reionization from the extreme UV radiation emitted by AGNs at $z \lesssim 4-5$ which reheats the IGM, increasing T_0 and decreasing γ. After the double reionization of helium completes ($z \sim 2.5$–3.5) the IGM cools monotonically by adiabatic expansion increasing γ for a second time. Because of these two distinct photoheating epochs, the thermal state of the IGM in our simulations is more sensitive to variations in the hydrogen photoheating/photoionization parameters $\beta_{\rm H}$ and $\Delta z_{\rm H}$ at $z \gtrsim 5$, and more sensitive to the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ at $z \lesssim 5$ during the epoch of helium reionization.

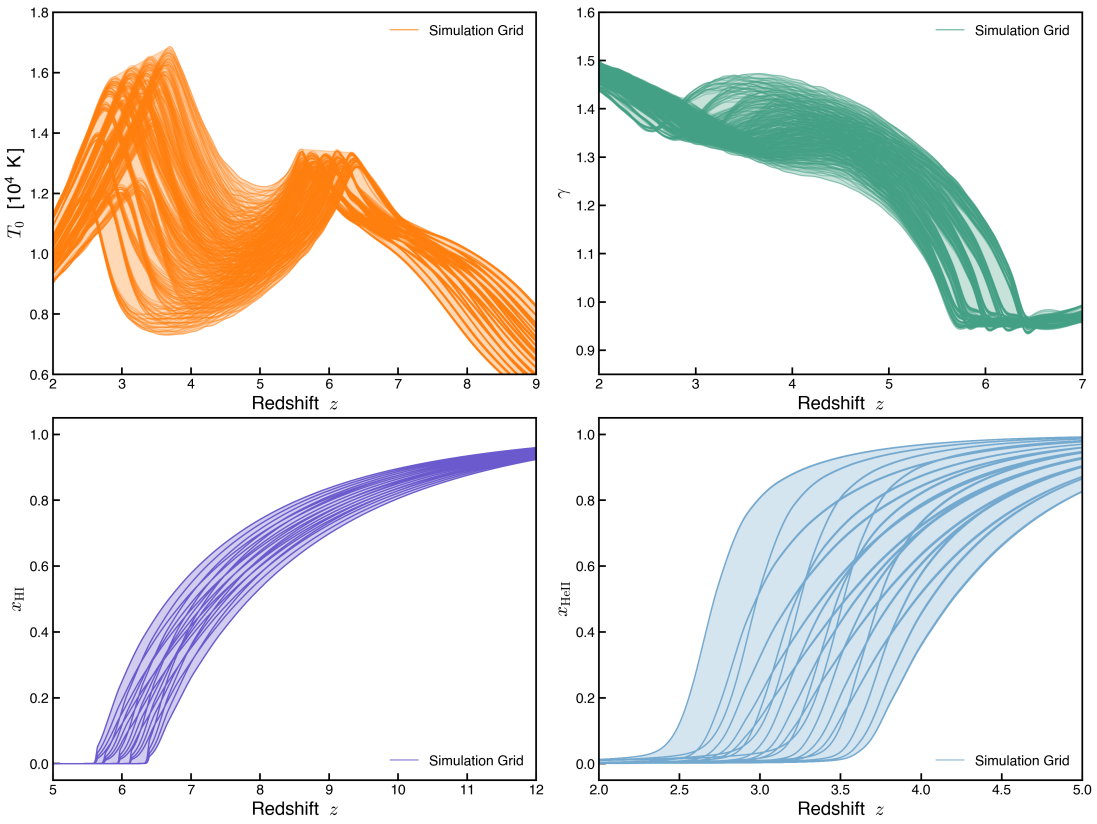

Figure 3.3: Evolution of global properties of the IGM computed form the 400 CHIPS simulations. The simulations evolve under different photoionization and photoheating rates resulting in a large variety of ionization and thermal histories of the IGM. The top panels show the the temperature, T_0, of intergalactic gas at the mean density (left) and the index γ from the power-law density-temperature relation $T(\Delta) = T_0\Delta^{\gamma-1}$ (right). The bottom panels show the volume-weighted average of the neutral hydrogen fraction $x_{\rm HI}$ (left) and the singly ionized helium fraction $x_{\rm HeII}$ (right). The amplitude and timing of the rates impact the thermal state of the IGM during H I and He II reionization. Simulations with higher values of $\beta_{\rm He}$ result in a higher temperature peak during He II reionization ($2.5 \lesssim z \lesssim 3.8$) and for simulations with $\Delta z_{\rm He} > 0$ the epoch of He II reionization is shifted to earlier epochs. Analogously, negative values of $\Delta z_{\rm H}$ move the timing of H I reionization to later epochs and simulations with different $\beta_{\rm H}$ show a different temperature peak during H I reionization at $z \sim 5.6 - 6.3$.

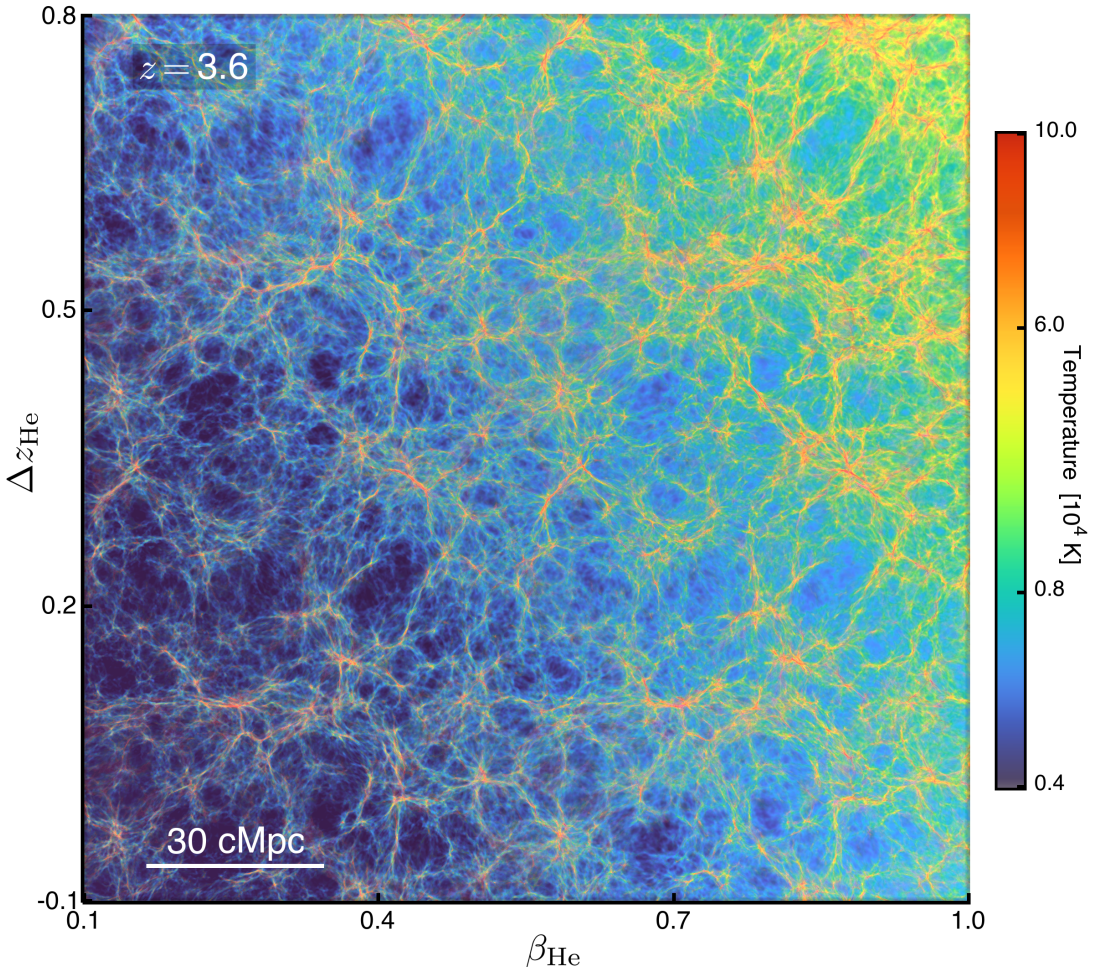

Figure 3.4: Gas temperature from a slice through the IGM at $z = 3.6$ in a subset of 20 simulations with different He II reionization scenarios. An increase in the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ corresponds to higher He II photoheating and a shift of the He II reionization epoch to earlier cosmic times (closer to $z \sim 3.6$) respectively. Either effect increases the temperature of the IGM at $z \sim 3.6$.

For simulations with $\Delta z_{\rm H} < 0$ the temperature peak from hydrogen reionization is shifted to later times (lower redshift) and the amplitude of the temperature peak depends on the value of $\beta_{\rm H}$. Analogously, the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ determine the amplitude and timing of the second peak in T_0 caused by helium reionization. Positive values of $\Delta z_{\rm He}$ move helium reionization to higher redshifts compared with the reference P19 model, and higher values of $\beta_{\rm He}$ produce a higher peak in T_0 during the epoch $2.5 \lesssim z \lesssim 3.8$.

Variation in the timing of H and He reionization changes the cooling periods during which the power-law index γ increases. The different tracks of γ in our simulation grid then arise primarily from the different values of $\Delta z_{\rm H}$ and $\Delta z_{\rm He}$ adopted. In future work we plan to expand the flexibility of our simulations to sample the thermal state of the IGM by introducing density dependent photoheating rates which will allow for larger variation in the evolution of γ.

The effects on the temperature of the IGM from the different helium reionization scenarios in our simulations are illustrated in Figure 3.4. The image displays the gas temperature of a slice through the IGM at $z = 3.6$ generated from a subset of 20 simulations that vary the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ controlling the He II photoionization and photoheating rates. Increases in $\beta_{\rm He}$ and $\Delta z_{\rm He}$ correspond to a larger extreme UV background from AGNs and to a shift of the epoch of helium reionization to earlier cosmic times, respectively. Either effect causes the temperature of the IGM to increase at $z \sim 3.6$. Decreasing the He II photoheating rates or shifting helium reionization to later cosmic times (toward $z \sim 2.8$) decreases the temperature of IGM gas at $z \sim 3.6$.

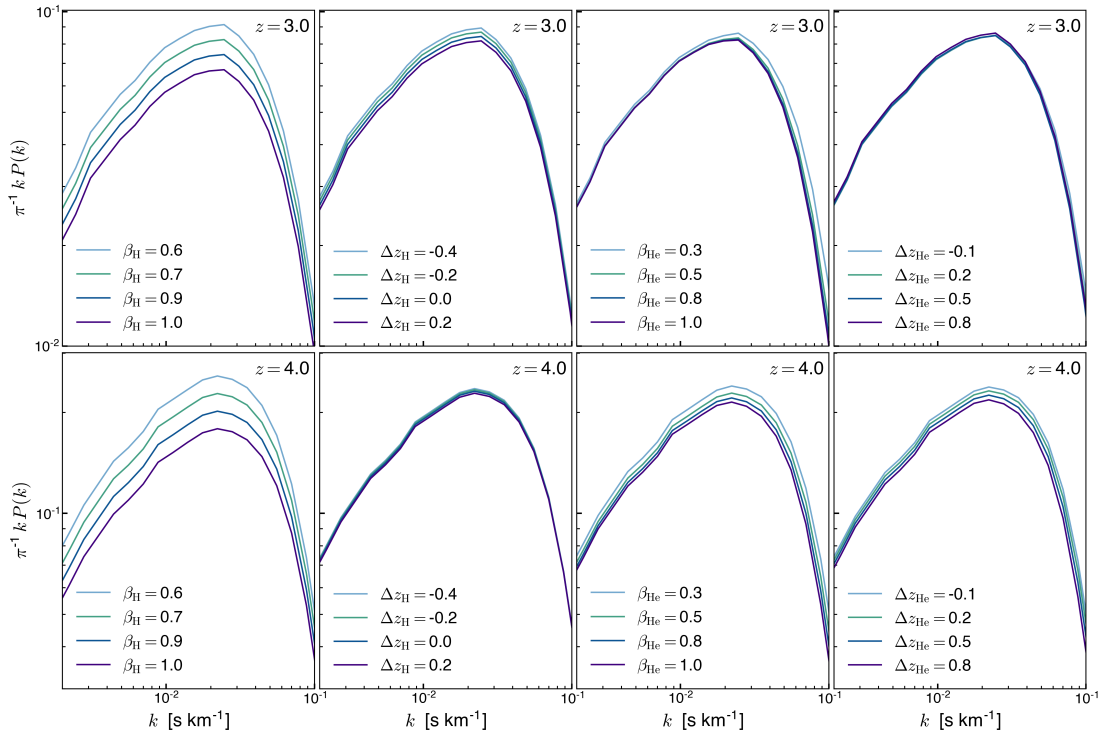

Figure 3.5: Sensitivity of the Lyman-α flux power spectrum $P(k)$ to independent variations of the parameters $\theta = \{\beta_{\rm H},\, \Delta z_{\rm H},\, \beta_{\rm He},\, \Delta z_{\rm He}\}$ for redshifts $z = 3$ (top) and $z = 4$ (bottom). Independent changes of each parameter have different effects on the redshift-dependent $P(k)$. After hydrogen reionization completes, differences in the power spectrum at $z \lesssim 5.5$ arise from changes in the ionization state and temperature of the IGM. Variation of the parameters $\beta_{\rm H}$ and $\Delta z_{\rm H}$ mostly affect the ionization state of hydrogen and therefore the overall normalization of $P(k)$. Changes in the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ impact $P(k)$ through their effect on the temperature of the gas during and after helium reionization, as variations in the thermal state of the IGM control the ionization fraction of hydrogen by its effect on the recombination rate $\alpha_{\rm HII}(T)$, and lead to the Doppler broadening of absorption lines and the smoothing of density fluctuations that suppress small-scale power ($k \gtrsim 0.02$ s km^{-1}).

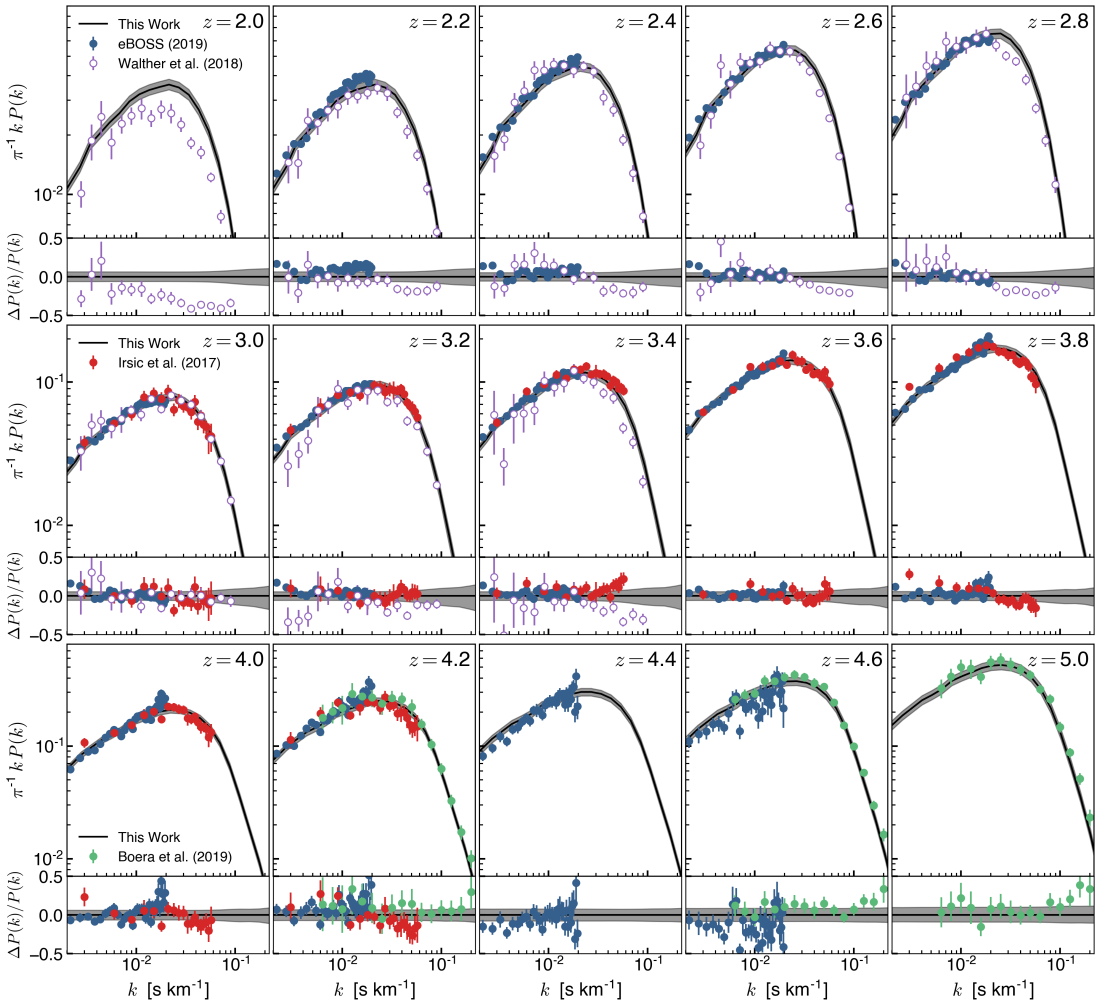

Figure 3.6: The transmitted flux power spectrum $P(k)$ from observations by eBOSS (Chabanier et al. 2019c), Keck Observatory and the Very Large Telescope (Iršič et al. 2017b; Boera et al. 2019) used to constrain models of the cosmic photoionization and photoheating history. The best-fit evolution of $P(k)$ marginalized over the posterior distribution of the parameters $\theta = \{\beta_{\rm H}, \Delta z_{\rm H}, \beta_{\rm He}, \Delta z_{\rm He}\}$ is shown with black curves, along with 95% confidence intervals (shaded bands). The fractional differences from the observations and the best-fit model are shown in the bottom part of each panel. Overall, the best-fit $P(k)$ is in good agreement with the large-scale power spectrum from eBOSS for $2.4 \leq z \leq 4.2$, and with the intermediate scales data from Iršič et al. (2017b) at $3.0 \leq z \leq 4.2$. Our best-fit results also agree with the measurements from Boera et al. (2019) at $4.2 \leq z \leq 5.0$, showing 10–30% differences mostly on the smallest scales ($0.1 - 0.2$ s km^{-1}) and suggesting that the temperature of the IGM at this epoch could be slightly overestimated by the model. We also show the $P(k)$ determinations by Walther et al. (2018) for comparison. Owing to discrepancies with the eBOSS results on large scales, we have not included the Walther et al. (2018) data points in our MCMC analysis.

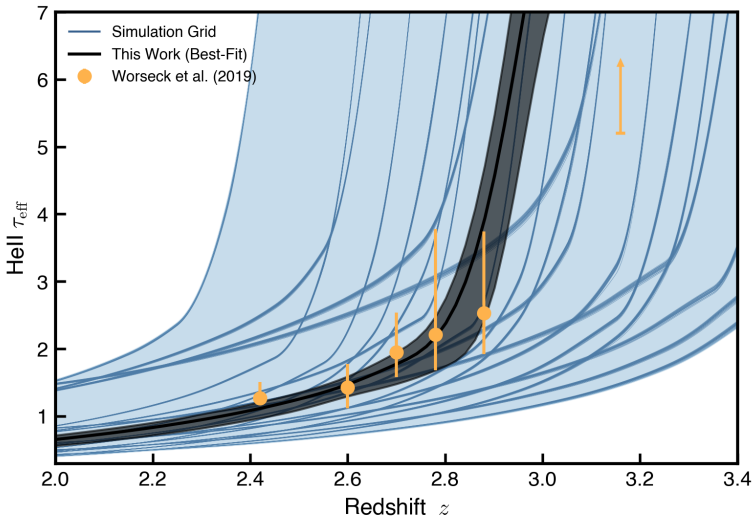

Figure 3.7: Evolution of the singly ionized helium (He II) effective optical depth $\tau_{\rm eff,HeII}$ from our simulation grid (blue lines), along with the best-fit model (black line) and the 95% confidence interval (gray area) obtained from our MCMC marginalization. The orange points show the observational measurements of $\tau_{\rm eff,HeII}$ (Worseck et al. 2019). While only data in the redshift range $2.4 \lesssim z \lesssim 2.9$ were used as constraints for our statistical analysis, the observed lower limits at $z > 3$ are consistent with the model results.

3.2.6 Effects of UVB Models on the Lyman-α Forest Power Spectrum

The statistical properties of the Lyman-α forest provide insight into the state of the baryons in the IGM. The effective optical of the forest $\tau_{\rm eff,H} = -\ln\langle F\rangle$ provides a global measurement of the overall H I content of the gas in the IGM, probes the hydrogen ionization fraction, and allows for estimates of the intensity of the ionizing background radiation. The power spectrum $P(k)$ of the flux transmitted through the forest contains more information encoded across different spatial scales. On scales larger than a few Mpc the $P(k)$ is sensitive to the ionization fraction of hydrogen in the IGM and provides information similar to $\tau_{\rm eff,H}$. This connection makes $P(k)$ and $\tau_{\rm eff,H}$ a dependent pair of measurements, and §3.6 presents a detailed analysis about the effects that variations in $\tau_{\rm eff,H}$ induce in $P(k)$. On scales smaller than a few comoving Mpc, structure in the forest is suppressed by pressure smoothing of the gas density fluctuations as well as Doppler broadening of the absorption lines. These effects cause a cutoff in the dimensionless power spectrum $\Delta^2(k) = \pi^{-1}kP(k)$ for $k \gtrsim 0.02\,{\rm s\,km^{-1}}$, making the flux power spectrum at intermediate and small scales a sensitive probe of the thermal state of IGM gas.

The different ionization and thermal histories produced by the range of photoionization and photoheating rates adopted in our simulations manifest as variations in the flux power spectrum of the Lyman-α forest. The effects on $P(k)$ from changing each of the four parameters $\beta_{\rm H}$, $\Delta z_{\rm H}$, $\beta_{\rm He}$, or $\Delta z_{\rm He}$ independently is shown in Figure 3.5 for redshifts $z = 3$ (top) and $z = 4$ (bottom). The variation in $P(k)$ measured from our simulation grid over the redshift range $2 \lesssim z \lesssim 5$ can be attributed mainly to

three physical effects. First, since hydrogen is in photoionization equilibrium after H I reionization, changes to the photoionization rate $\Gamma_{\rm HI}$ from rescaling by $\beta_{\rm H}$ or applying a shift $\Delta z_{\rm H}$ alter the ionization fraction of hydrogen. This alteration globally affects the hydrogen effective optical depth $\tau_{\rm eff,H}$ and, as a result, the overall normalization of $P(k)$ changes. Second, changes in the temperature of the IGM from the different hydrogen and helium reionization scenarios alter the recombination coefficient $\alpha_{\rm HII}(T)$ in the IGM. In turn, changes to the recombination rate adjust the ionization fraction of hydrogen in the IGM and thereby the normalization of $P(k)$. Third, the different thermal histories of the IGM affect $P(k)$ on small scales through Doppler broadening of the absorption lines and the pressure smoothing of the density fluctuations. As shown in Figure 3.5, the parameters $\beta_{\rm H}$ and $\Delta z_{\rm H}$ mainly influence the normalization of $P(k)$ by changing the overall ionization fraction in the IGM, while the parameters $\beta_{\rm He}$ and $\Delta z_{\rm He}$ change the temperature of the IGM and thereby affect both the normalization and small-scale shape of $P(k)$.

3.2.7 Observational Data

For comparison with our simulations, we use the observational determinations of the flux power spectrum measured by the extended Baryon Oscillation Spectroscopy Survey (eBOSS; Chabanier et al. 2019c) and separate measurements with the Keck Observatory and the Very Large Telescope (Iršič et al. 2017b; Boera et al. 2019). The power spectrum estimates from Chabanier et al. (2019c) probe mostly large scales ($0.001 \lesssim k \lesssim 0.02$ s km^{-1}) in the redshift range $2.2 < z < 4.6$. The determinations from Iršič et al. (2017b) overlap with the eBOSS measurements on the large scales, albeit

with lower precision, and extend to intermediate scales ($0.003 \lesssim k \lesssim 0.06\ \mathrm{s\,km^{-1}}$) for redshifts $3.0 < z < 4.2$. The data from Boera et al. (2019) cover intermediate to small scales ($0.006 \lesssim k \lesssim 0.2\ \mathrm{s\,km^{-1}}$) over the redshift range $4.2 < z < 5.0$. The combined data set spans a large redshift range from $z = 2.2$ to $z = 5.0$ and a wide range of scales, and is shown along with our best-fit model $P(k)$ in Figure 3.6.

Figure 3.6 also shows the observational measurements of $P(k)$ presented by Walther et al. (2018, purple empty points) for the redshift range $3.0 \leq z \leq 3.4$. We find that, in the overlapping range of scales ($0.003 \lesssim k \lesssim 0.02\ \mathrm{s\,km^{-1}}$) and redshift ($2.2 \lesssim z \lesssim 3.4$), the estimates from Walther et al. (2018) show significant differences with those from eBOSS (Chabanier et al. 2019c). The normalization and, in some cases, the shape of the large-scale $P(k)$ appear to be inconsistent between the two datasets. For several redshift bins (e.g. $z = 2.4$ and $z = 3.2$), a simple renormalization applied to the Walther et al. (2018) power spectrum would not be sufficient to match the large-scale measurements from eBOSS. Because of this discrepancy, we have not included the Walther et al. (2018) $P(k)$ determinations in our MCMC analysis, and we show them in Figure 3.6 only for comparison with our modeling and other data sets.

To obtain a better determination of the He II photoionization and photoheating rates, we complement the power spectrum comparison with observational measurements of the helium effective optical depth $\tau_{\mathrm{eff,HeII}}$ (Worseck et al. 2019) over the redshift range $2.4 \lesssim z \lesssim 2.9$ as additional constraints on our model. The data are shown in Figure 3.7 along with the corresponding evolution of $\tau_{\mathrm{eff,HeII}}$ from our simulation grid and the best-fit model from our analysis. We do not include the observational lower limits at

$z > 3$ as constraints in our MCMC analysis, but our best-fit model is consistent with those limits.

3.2.8 Systematic Uncertainties

When comparing models to observations, we include systematic uncertainties owing to cosmological parameter variations and possible resolution limitations of the simulations. In Villasenor et al. (2021a), we performed a study of the changes in the Lyman-α flux power spectrum $P(k)$ induced by small variations of the cosmological parameters within the constraints from Planck Collaboration et al. (2020). Our results suggested that uncertainties in the cosmological parameters could cause a fractional change of $\lesssim 5\%$ on the hydrogen effective optical depth in the redshift range $2 \lesssim z \lesssim 5$ and a similar $\lesssim 5\%$ effect in $P(k)$ for scales $0.002 \lesssim k \lesssim 0.2\,\mathrm{s\,km^{-1}}$ and redshifts $2 \lesssim z \lesssim 5$. For this reason, we include here an additional systematic uncertainty σ_{cosmo} of 5% to the observational determinations of the Lyman-α power spectrum. For the He II effective optical depth, we estimate similar variations of $\sim 5\%$ at $2 \lesssim z \lesssim 3$ from differences in the mean baryonic density of different cosmologies. We therefore include a $\sigma_{\mathrm{cosmo}} = 5\%$ to the measurements of $\tau_{\mathrm{eff,HeII}}$ as well.

In Appendix 3.5 we present a resolution convergence study where we compare the forest flux power spectrum from three simulations performed with the same cosmological parameters and photoionization and photoheating histories. The initial conditions used for the runs were generated to preserve the large-scale modes in common to each simulation, such that the properties of the simulations could be compared directly on shared spatial scales. The three simulations model a box of size $L = 50h^{-1}$Mpc and

$N = 512^3$, $N = 1024^3$, or $N = 2048^3$ cells and particles. The corresponding spatial resolutions are $\Delta x \simeq 98\,h^{-1}$Mpc, $\Delta x \simeq 49\,h^{-1}$Mpc, and $\Delta x \simeq 24\,h^{-1}$Mpc, respectively. In comparing the moderate resolution ($\Delta x \simeq 49\,h^{-1}$Mpc) and high resolution ($\Delta x \simeq 24\,h^{-1}$Mpc) simulations, we measure small fractional differences $\Delta P(z,k)/P(z,k)$ of $\lesssim 5\%$ for the large scales $k \lesssim 0.02$ s km^{-1}. On small scales, $0.02 \lesssim k \lesssim 0.2$ s km^{-1}, the fractional differences are slightly larger ($\lesssim 12\%$).

To approximate resolution effects from the grid of simulations used for our analysis ($N = 1024^3$, $\Delta x \simeq 49\,h^{-1}$Mpc), we add an additional systematic uncertainty $\sigma_{\rm res}$ to the observational determinations of the flux power spectrum and the He II effective optical depth. For $P(k)$, the additional uncertainty $\sigma_{\rm res}(z,k) = \Delta P(z,k)$ is set equal to the difference between $P(k)$ from the $N = 1024^3$ and $N = 2048^3$ reference simulations used for our resolution study. For the He II effective optical depth the impact of resolution is a small increase of $\lesssim 3\%$ from the $N = 1024^3$ box to the $N = 2048^3$ run at $z < 3$; we then add an uncertainty of $\sigma_{\rm res}(z) = 3\%$ to the estimate of $\tau_{\rm eff,HeII}$. We note that the systematic errors added to $\tau_{\rm eff,HeII}$ are significantly smaller than the observational uncertainties $\sigma_{\rm obs} \sim 12 - 45\%$ of Worseck et al. (2019).

The total uncertainty applied to the observational determinations of $P(k)$ and $\tau_{\rm eff,HeII}$ is finally given by the quadrature sum of the errors as

$$\sigma_{\rm total} = \sqrt{\sigma_{\rm obs}^2 + \sigma_{\rm cosmo}^2 + \sigma_{\rm res}^2} \tag{3.3}$$

where $\sigma_{\rm obs}$ is the reported observational uncertainty in the flux power spectrum or helium opacity respectively.

In their study, Wolfson et al. (2021) showed the importance of using the covariance matrix when inferring the temperature of the IGM from measurements of the Lyman-α power spectrum and wavelet statistics. For our MCMC analysis we use the covariance matrices of $P(k)$ in the likelihood calculation (see §3.2.9). To reflect the increased uncertainty from Eq. 3.3, we rescale the elements of the covariance matrices according to

$$\mathbf{C}[i,j] \;=\; \mathbf{C}_{\rm obs}[i,j]\frac{\sigma_{{\rm total},i}\;\sigma_{{\rm total},j}}{\sigma_{{\rm obs},i}\;\sigma_{{\rm obs},j}}, \tag{3.4}$$

where $\mathbf{C}_{\rm obs}$ is the reported covariance matrix of $P(k)$ taken from the published observational datasets used for our analysis.

3.2.9 Inference of the UVB Model

To find the photoionization and photoheating rates that best reproduce the properties of the IGM encoded in the flux power spectrum of the Lyman-α forest $P(k)$ and the helium effective optical depth $\tau_{\rm eff,HeII}$, we apply an MCMC sampler to compare the simulated $P(k)$ and $\tau_{\rm eff,HeII}$ to the observational measurements over the redshift and frequency range where data are available. The likelihood function for the model given by the parameters $\theta = \{\beta_{\rm H},\, \Delta z_{\rm H},\, \beta_{\rm He},\, \Delta z_{\rm He}\}$ is evaluated as:

$$\begin{aligned}\ln \mathcal{L}(\theta) = &-\frac{1}{2}\sum_{z}\left[\left(\frac{\tau_{\rm obs}(z)-\tau(z|\theta)}{\sigma_{\tau}(z)}\right)^{2} + \ln\left(2\pi\sigma_{\tau}(z)^{2}\right)\right] \\ &-\frac{1}{2}\sum_{\rm datasets}\sum_{z}\left[\mathbf{\Delta}^{T}\mathbf{C}^{-1}\mathbf{\Delta} + \ln\det(\mathbf{C}) + N\ln 2\pi\right],\end{aligned} \tag{3.5}$$

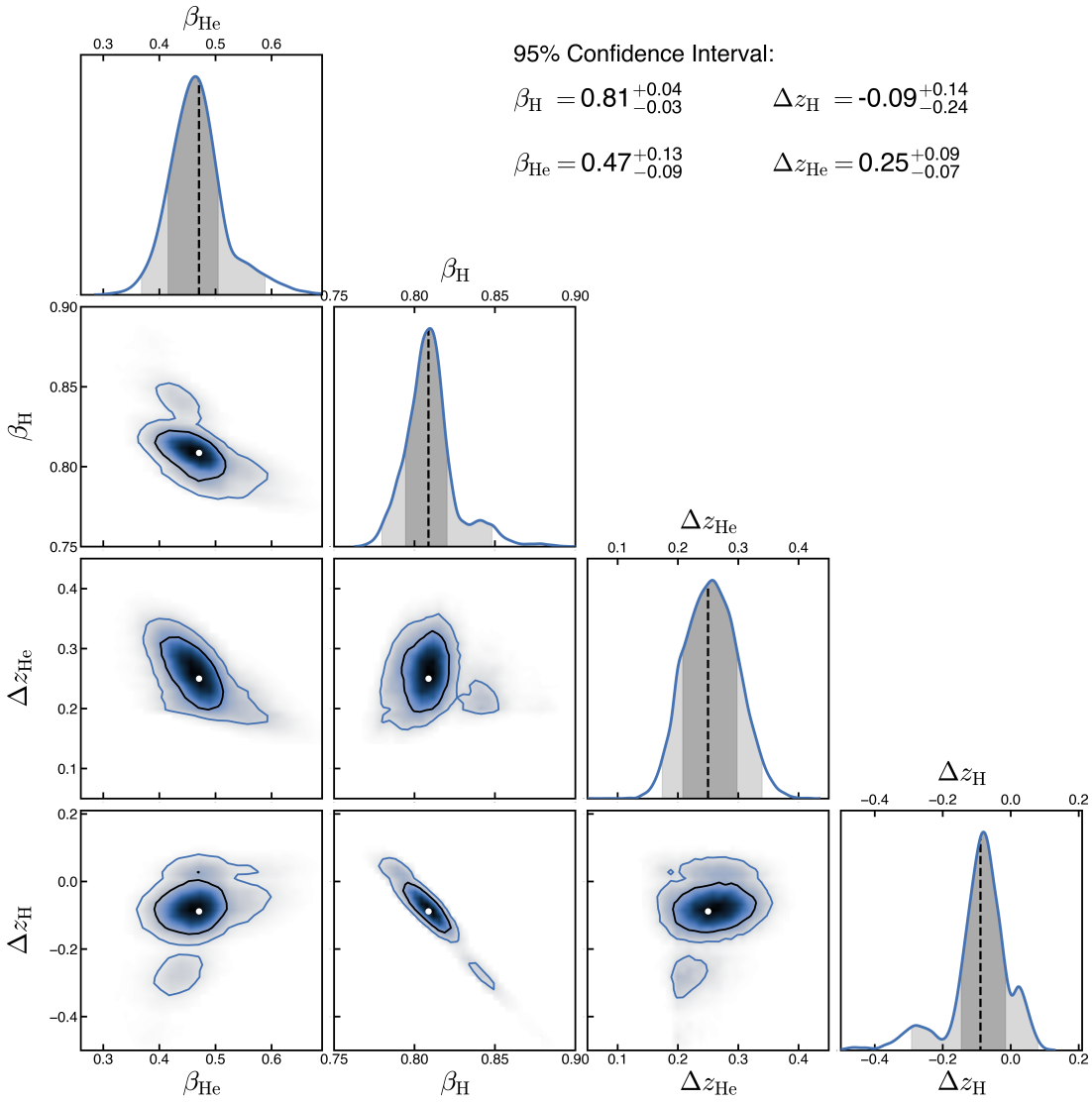

Figure 3.8: Results from the Bayesian inference procedure, showing one- and two-dimensional projections of the posterior distributions for the parameters $\theta = \{\beta_{\rm H}, \Delta z_{\rm H}, \beta_{\rm He}, \Delta z_{\rm He}\}$. The parameter constraints were obtained by fitting the observed flux power spectrum of the Lyman-α forest and the He II effective optical depth (Chabanier et al. 2019c; Boera et al. 2019; Worseck et al. 2019; Iršič et al. 2017b) with a grid of CHIPS simulations. The posterior distribution shows a clear global maximum, and while other local maxima are present their peak likelihoods are significantly lower than the global maximum. The resulting best-fit parameters and their 95% confidence intervals are shown in the top right corner.

where the first term compares the He II effective optical depth measured from our simulations $\tau(z|\theta)$ for a given photoionization and photoheating model represented by the vector θ to the observational measurement $\tau_{\rm obs}(z)$ from Worseck et al. (2019) with total (observational + systematic) uncertainty $\sigma_\tau(z)$. The second term compares the Lyman-α power spectrum, with $\mathbf{\Delta}$ denoting the difference vector between the observations and the model $\mathbf{\Delta} = P_{\rm obs}(z,k) - P(z,k|\theta)$. Here, $\mathbf{C}$ corresponds to the covariance matrix of size $N \times N$ associated with the observational determination, where N is the number of points of each measurement. To compute $P(z,k|\theta)$ and $\tau(z|\theta)$ for arbitrary values of the parameters θ not directly simulated by our grid, we perform a four-dimensional linear interpolation of the sixteen neighboring simulations in parameter space.

As described in §3.2.7, we employ the datasets from Chabanier et al. (2019c) ($2.2 \leq z \leq 4.6$), Iršič et al. (2017b) ($3.0 \leq z \leq 4.2$), and Boera et al. (2019) ($4.2 \leq z \leq 5.0$) for the observational measurements of the power spectrum used in our analysis. While there is some overlap in the measurements from the datasets, in general their determinations are consistent with each other. For this reason we include all the data points from each dataset for the likelihood calculation. The only significant discrepancy is at $z = 4.6$ where $P(k)$ from Chabanier et al. (2019c) is lower to the determination from Boera et al. (2019). We repeated our the analysis excluding the $z = 4.6$ measurement from Chabanier et al. (2019c) and obtained similar posterior distributions. We conclude that this difference does not impact our result.

The contribution from each redshift bin to the total log likelihood $\ln \mathcal{L}$ (Eq.

Table 3.2: Redshift Bin Contribution to the Likelihood

Type	z	$-\Delta \ln \mathcal{L}$	Type	z	$-\Delta \ln \mathcal{L}$
$P(k)$	2.2	330.6	$P(k)$	4.2	489.2
$P(k)$	2.4	363.3	$P(k)$	4.4	135.7
$P(k)$	2.6	229.2	$P(k)$	4.6	190.2
$P(k)$	2.8	297.0	$P(k)$	5.0	40.0
$P(k)$	3.0	215.1	$\tau_{\rm eff,HeII}$	2.30	0.5
$P(k)$	3.2	134.4	$\tau_{\rm eff,HeII}$	2.54	0.2
$P(k)$	3.4	113.8	$\tau_{\rm eff,HeII}$	2.66	0.3
$P(k)$	3.6	84.1	$\tau_{\rm eff,HeII}$	2.74	1.0
$P(k)$	3.8	137.1	$\tau_{\rm eff,HeII}$	2.82	2.3
$P(k)$	4.0	180.3			

4.15) from $P(k)$ and $\tau_{\rm eff,HeII}$ in our analysis is presented in Table 3.2.9. The quantity $\Delta \ln \mathcal{L}$ is evaluated as the first and second terms of Eq. 4.15 for $\tau_{\rm eff,HeII}$ and $P(k)$, respectively, for each redshift bin. The power spectrum mostly strongly influences the log likelihood, with data from redshifts $z = 2.4$ and $z = 4.2$ inducing the largest fractional changes in the likelihood.

The covariance matrices of $P(k)$ are taken from the published observations. We note that Iršič et al. (2017b) provides the complete covariance of $P(k)$ across the seven redshift bins of their measurement. For this dataset we employ the reported full covariance and the residual vector $\mathbf{\Delta}$ consists of the $P(k)$ difference from the model and observation concatenated over the seven redshift bins.

While our likelihood analysis uses the reported covariance matrices from the observations, in Appendix 3.7 we present the covariance of $P(k)$ measured from a subset of our simulations to quantify the differences induced by variation of our four model parameters. We show that the structure of the covariance is mantained across our simulations and we measure relatively small variations between the different models.

We emphasize that our approach differs from previous studies of the thermal history of the IGM (e.g., Bolton et al. 2014; Nasir et al. 2016b; Hiss et al. 2018; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021) in an important aspect. Typically, the method adopted to infer the thermal state of the IGM from observations of the Lyman-α forest involves marginalizing over the thermal parameters T_0 and γ in the approximate power-law density-temperature relation (Hui & Gnedin 1997) $T(\Delta) = T_0 \Delta^{\gamma-1}$, where $\Delta = \rho_{\rm gas}/\bar{\rho}$ is the gas overdensity. This marginalization is often performed independently for each redshift. Instead, our approach to find the optimal photoionization and photoheating rates that best reproduce the observational measurements is to compare the simulated $P(k)$ and $\tau_{\rm eff,HeII}$ to the observations over the full redshift range where data is available, namely $2.2 \leq z \leq 5.0$ for $P(k)$ and $2.2 < z < 3.0$ for $\tau_{\rm eff,HeII}$.

In our approach, the performance for a given UVB model to match the observations is evaluated over the complete self-consistently evolved reionization and thermal history of the IGM that results from that model. Since the properties of the gas at one redshift cannot be disentangled from its properties at previous epochs, the thermal and ionization structure of the forest depends on the time-dependent photoheating and photoionization rate. Both T_0 and γ evolve along continuous trajectories with redshift, and we therefore marginalize over the full simulated histories of IGM properties.

Our simulations span a wide range of reionization histories for hydrogen in the IGM. Instead of following the common practice of rescaling the optical depth of the simulated skewers in post-processing to match the observed mean transmission of

the forest, our method self-consistently follows the ionization evolution of hydrogen and the effective optical depth $\tau_{\rm eff,H}$ encoded in the redshift-dependent power spectrum of the transmitted flux. Furthermore, during our inference procedure, we do not assume a power-law approximation for the density-temperature distribution of IGM gas or apply a post-processing procedure that artificially modifies the temperature of the gas in the simulations. Instead, our synthetic Lyman-α spectra reflect the real $\rho_{\rm gas}-T$ distribution from the simulations. This improvement proves relevant, as we find that a single power law is not a good fit over the full range of gas densities responsible for the bulk of the Lyman-α absorption signal (see Appendix 3.9).

The posterior distribution for our parameters $\theta = \{\beta_{\rm H},\, \Delta z_{\rm H},\, \beta_{\rm He},\, \Delta z_{\rm He}\}$ resulting from the Bayesian inference procedure is shown in Figure 3.8. A clear global maximum of the posterior distribution is observed, and while the posterior shows other local maxima their likelihoods are significantly lower than the global peak. The four model parameters are well constrained and show only small correlations that arise from the weak degeneracies in the resulting ionization and thermal histories produced by the different photoionization and photoheating rates. Our best-fit parameters and their 95% confidence limits are

$$\begin{aligned} \beta_{\rm H} &= 0.81^{+0.04}_{-0.03} \qquad \Delta z_{\rm H} &&= -0.09^{+0.14}_{-0.24} \\ \beta_{\rm He} &= 0.47^{+0.13}_{-0.09} \qquad \Delta z_{\rm He} &&= 0.25^{+0.09}_{-0.07}. \end{aligned} \tag{3.6}$$

To measure the properties of the IGM that result from our best-fit distribution, we sample $P(k)$, $\tau_{\rm eff,H}$, and $\tau_{\rm eff,HeII}$, together with the thermal parameters T_0 and γ, over

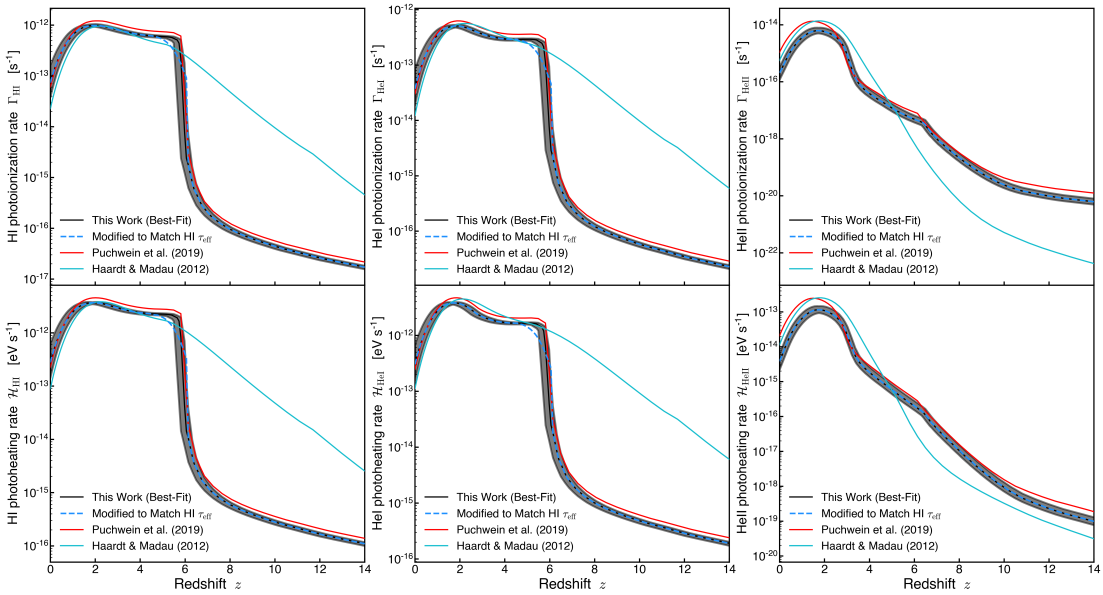

Figure 3.9: Best-Fit (black lines) and 95% confidence intervals (grey bands) for the photoionization (Γ, top) and photoheating ($\mathcal{H}$, bottom) rates for neutral hydrogen (H I, left), neutral helium (He I, center), and singly ionized helium (He II, right) obtained from our MCMC analysis. The *modified* H I and He I photoionization and photoheating rates (dashed blue lines) are identical to the reference best-fit model except for the redshift range $4.8 \leq z \leq 6.1$ where they have been modified to produce an evolution of the hydrogen effective optical depth consistent with the observational determinations of Bosman et al. (2018) for $z > 5$ (see §3.3.4 and §3.3.5 for details). For reference, we also show the models from Puchwein et al. (2019) (red) and Haardt & Madau (2012) (cyan).

the posterior distribution of the parameter vector θ, resulting in determinations of the highest-likelihood and 95% confidence interval for the forest statistics and thermal history. When necessary, we interpolate results for values of θ not directly simulated by our grid.

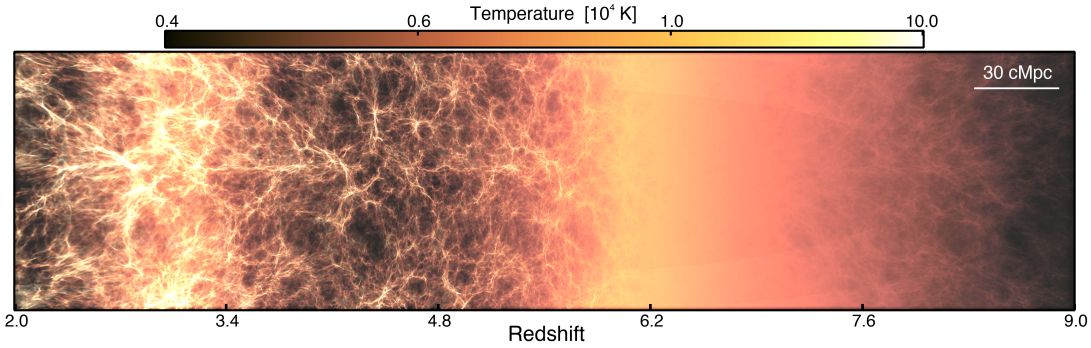

Figure 3.10: Redshift evolution of the gas temperature from a high-resolution simulation ($L = 50\,h^{-1}$Mpc, $N = 2048^3$ cells and particles) that employed our best-fit model for the photoheating and photoionization rates. The image displays the monotonic increase in the temperature of the IGM due to hydrogen reionization for $z \gtrsim 6.0$ followed by an epoch of cooling of the IGM due to cosmic expansion. The onset of helium reionization ($z \sim 4.5$) initiates a second epoch of heating of the IGM that ends at $z \sim 3$ when He II reionization completes. A second epoch of cooling due to cosmic expansion then follows. The temperature increase of gas collapsing into the filamentary cosmic web as large-scale structure develops is also visible in the image.

3.3 Results and Discussion

By comparing the flux power spectrum and the He II effective opacity in our CHIPS simulation grid to observational determinations, we can infer a set of photoionization and photoheating histories that, when input in cosmological hydrodynamical simulations, result in statistical properties of the Lyman-α forest that are consistent with observations. In this section, we present the best-fit rates obtained from our inference procedure, as well as the Lyman-α forest statistics and thermal evolution of the IGM produced by our best-fit UVB model. We compare our results to previous work and finalize our discussion by describing the limitations of our method.

3.3.1 Best-Fit Photoionization and Photoheating Rates

Figure 3.9 shows our best-fit model for the photoionization and photoheating rates along with the corresponding 95% confidence interval that results from our MCMC marginalization of the UVB rates over the posterior distribution of the model parameters obtained from our MCMC analysis. We note that the transformations applied in this work to generate new photoionization and photoheating rates from the reference model (Puchwein et al. 2019) are relatively simple and preserve the functional form of the P19 model. While we allow for orders of magnitude variations in the rates, the flexibility of the ionization and thermal histories sampled here is limited by the fixed shape of the UVB model employed in our simulation grid. A study that allows for more flexibility in the photoionization and photoheating rates of hydrogen and helium will be the scope of future work.

3.3.2 $P(k)$ Model Comparison with the Data

Figure 3.6 shows the evolution of the best-fit flux power spectrum and 95% confidence intervals over the redshift range $2.2 \leq z \leq 5.0$ that result from marginalizing $P(k)$ over the posterior distribution of model parameters $\theta = \{\beta_{\rm H}, \Delta z_{\rm H}, \beta_{\rm He}, \Delta z_{\rm He}\}$. Our best-fit synthetic power spectrum shows good agreement with the large scale $P(k)$ measured by the eBOSS experiment (Chabanier et al. 2019c) in the range $2.4 \lesssim z \lesssim 4.2$, suggesting that the mean transmission $\langle F \rangle$of the forest inferred by our analysis is consistent with the measurements by Chabanier et al. (2019c). Only for $z = 2.2$ and $z = 4.4 - 4.6$ do our results show significant differences with the eBOSS data set.

At $z = 2.2$, the $P(k)$ from eBOSS is higher than our results by $\sim 8 - 20\%$ on scales $0.008 \lesssim k \lesssim 0.02\ \mathrm{s\,km^{-1}}$. This modest tension may suggest that the hydrogen opacity $\tau_{\mathrm{eff,H}}$ is underestimated by $\sim 10\%$ in our modeling relative to eBOSS. At $z = 4.4$ and $z = 4.6$ the opposite is true, and our best-fit $P(k)$ on large scales is $\sim 15\%$ and $\sim 25\%$ higher than the eBOSS measurements, respectively. These small discrepancies could be alleviated, e.g., by a small 15% decrease of the H I photoionization rate at $z = 2.2$ and by a comparable small increase in the same quantity at $z = 4.4 - 4.6$ by $\sim 10 - 20\%$.

Our results also agree on large and intermediate scales ($0.003 \lesssim k \lesssim 0.06\ \mathrm{s\,km^{-1}}$) with the estimates of Iršič et al. (2017b). The best-fit model reproduces the turnover in the observed dimensionless power spectrum $\Delta^2(k) = \pi^{-1} k P(k)$ at $k \sim 0.02 - 0.03\ \mathrm{s\,km^{-1}}$, and generally lies within the observational uncertainties at intermediate scales $0.01 \lesssim k \lesssim 0.06\ \mathrm{s\,km^{-1}}$. Only at redshift $z = 3.4$ and $z = 3.8$ the $P(k)$ measurements show some differences relative to the model. At $z = 3.4$ the data are higher than the model by $\sim 5 - 20\%$. A similar discrepancy is observed when comparing Iršič et al. (2017b) with the determinations by eBOSS at the same redshift, suggestive of a slightly higher H I opacity $\tau_{\mathrm{eff,H}}$ in the former sample. Differences with the model are more significant at $z = 3.8$, where on intermediate scales ($k \gtrsim 0.2\ \mathrm{s\,km^{-1}}$) the measurements of Iršič et al. (2017b) are lower than the model by $\sim 10 - 20\%$, while on large scales ($k \lesssim 0.2\ \mathrm{s\,km^{-1}}$) their estimates are higher than both the model and the determinations by eBOSS by $\sim 5 - 30\%$.

Our model is in good agreement with the high-redshift measurements of $P(k)$ by Boera et al. (2019), with minor differences that could be addressed by small modifi-

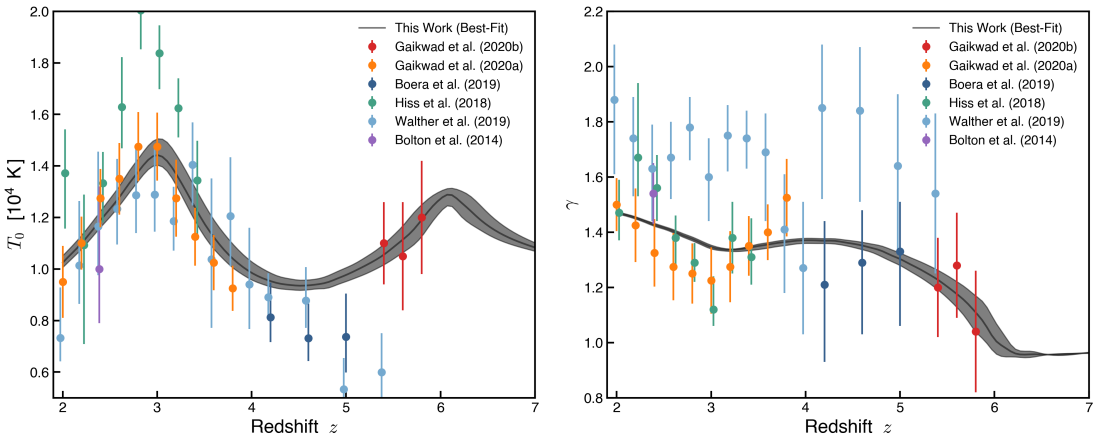

Figure 3.11: Redshift evolution of the parameters T_0 and γ [Eq. (3.7)] from the best-fit model (black lines) and 95% confidence interval (gray band) obtained from our MCMC analysis. The data points show the values of T_0 and γ inferred from observations of the Lyman-α forest by Bolton et al. (2014); Hiss et al. (2018); Boera et al. (2019); Walther et al. (2019); Gaikwad et al. (2020, 2021). Our results reveal two peaks in the evolution of T_0 due to hydrogen reionization at $z \sim 6$ and helium reionization at $z \sim 3$, and are consistent with previous measurements from Gaikwad et al. (2020, 2021).

cations to the early photoheating history. At $z = 4.2$, $z = 4.6$, and $z = 5.0$, our best-fit $P(k)$ is consistent with their data points on large scales $k \lesssim 0.02\ \mathrm{s\,km^{-1}}$, suggesting that our inferred IGM H I opacity matches that measured by Boera et al. (2019). The model also reproduces the cutoff in $\Delta^2(k)$ at $k \sim 0.02 - 0.03\ \mathrm{s\,km^{-1}}$ and the consistency with the observations extends to small scales $k \lesssim 0.1\ \mathrm{s\,km^{-1}}$. Discrepancies appear only on the smallest scales $0.1 \lesssim k \lesssim 0.2\ \mathrm{s\,km^{-1}}$ where the model has less power ($\sim 10 - 30\%$) than Boera et al. (2019). This may suggest that the temperature of the IGM has been overestimated by the model in the redshift range $4 \lesssim z \lesssim 5$ (see §3.3.3 for a discussion of this issue).

3.3.3 Evolution of the IGM Temperature

The flux power spectrum and helium opacity tightly constrain the time-dependent photoionization and photoheating rates, which in turn determine the IGM ionization and thermal history. The redshift evolution of the gas temperature is illustrated in Figure 3.10 which is generated from a slice through a high-resolution simulation ($L = 50\,h^{-1}$Mpc, $N = 2048^3$ cells and particles) using our best-fit photoionization and photoheating rates. The figure shows the monotonic increase in the temperature of the IGM during hydrogen reionization at $z \gtrsim 6$. After hydrogen reionization completes by $z \sim 6$, the input of energy into the IGM falls dramatically, and the gas then cools primarily through adiabatic expansion. This first epoch of IGM cooling lasts until the onset of helium reionization ($z \sim 4.5$) when extreme UV radiation from AGNs ionizes He II atoms and drives a second epoch of reheating that completes by $z \sim 3$ and is followed by a second epoch of adiabatic cooling.

The thermal state of diffuse IGM gas is often modeled with the power-law relation (Hui & Gnedin 1997; Puchwein et al. 2015; McQuinn 2016)

$$T(\Delta) = T_0 \Delta^{\gamma-1}. \tag{3.7}$$

We fit the power law relation to the gas density-temperature distribution in each of the simulations from the CHIPS grid and at multiple epochs, $2 \leq z \leq 9$, following the procedure presented in Villasenor et al. (2021a). We restrict the fit to the overdensity range $0 \leq \log_{10}\Delta \leq 1$, as we find that in our simulations a single power law does not accurately describe the wider range $-1 \leq \log_{10}\Delta \leq 1$ (see Appendix 3.9).

Figure 3.11 shows the redshift evolution of the parameters T_0 and γ from our best-fit model and the 95% confidence interval that results from our MCMC marginalization over the posterior distribution of the photoionization and photoheating rates. For comparison, we also depict the data points for these parameters inferred from the properties of the Lyman-α forest by Bolton et al. (2014), Hiss et al. (2018), Boera et al. (2019), Walther et al. (2019), Gaikwad et al. (2020), and Gaikwad et al. (2021).

The inference from Boera et al. (2019) and Walther et al. (2019) follow similar methodologies. They generate flux power spectra from simulations run with different thermal histories, resulting in multiples trajectories for the evolution of T_0 and γ. For each redshift bin they determine the best-fit T_0, γ, and mean transmitted flux $\langle F \rangle$by performing Bayesian inference and comparing the simulated flux power spectra to observations of the Lyman-α forest $P(k)$. Bolton et al. (2014) and Hiss et al. (2018) measure a set of values for the Doppler parameter b and H I column density $N_{\rm HI}$ directly from the forest by decomposing the absorption spectra into a collection of Voigt profiles. They infer the parameters T_0 and γ by comparing simulations with different $b - N_{\rm HI}$ distributions to the observed one. Gaikwad et al. (2020) follow a similar approach by comparing simulated Lyman-α forest spectra to Voigt profiles fitted to the observed transmission spikes in the inverse transmitted flux $1 - F$ at $z > 5$. Gaikwad et al. (2021) report more precise determinations by inferring T_0 and γ from the combined constraints obtained through a comparison of simulated Lyman-α forest absorption with the observed flux power spectra, $b - N_{\rm HI}$ distributions, wavelet statistics, and curvature statistics.

As shown in Figure 3.11, the temperature evolution from our best-fit model

presents a first peak ($T_0 \simeq 1.3 \times 10^4\,\mathrm{K}$) at the end of hydrogen reionization ($z \sim 6.0$) followed by an epoch of adiabatic cooling from cosmic expansion. Our results agree well with the high redshift measurements of T_0 and γ at $5.4 \leq z \leq 5.8$ from Gaikwad et al. (2020). We note that their estimates also suggest a period of cooling at these epochs, and from their result it is possible to infer a peak in T_0 from H reionization sometime at redshift $z \gtrsim 5.8$.

In our model, the IGM continues to cool until the onset of helium reionization, and the temperature reaches a local minimum of $T_0(z \sim 4.5) \simeq 9.5 \times 10^3\,\mathrm{K}$. Evidence of this transition can also be seen in the measurements from Boera et al. (2019), where T_0 shows little evolution from $z = 5.0$ to $z = 4.6$ and then a slight increase to $z = 4.2$. Nevertheless, there are significant differences between T_0 from the model at $4 \lesssim z \lesssim 5$ and the measurements from Boera et al. (2019), as the temperature predicted by our model is higher than their inferred values of $T_0 \sim 7.4 \times 10^3\,\mathrm{K}$ and $T_0 \sim 8.1 \times 10^3\,\mathrm{K}$ at $z = 4.6 - 5$ and $z = 4.2$, respectively. The higher temperatures in our model reflect a suppressed power spectrum of the Lyman-α flux on small scales ($0.1 \lesssim k \lesssim 0.2\,\mathrm{s\,km^{-1}}$) compared to the $P(k)$ measurement from Boera et al. (2019) at $4.2 \leq z \leq 5.0$ (see Figure 3.6). Decreasing the photoheating from the UVB during $z \gtrsim 4$ would decrease the temperature of the IGM at this epoch and potentially alleviate this discrepancy.

In Appendix 3.8 we present scenarios were the mid-redshift IGM is set to be colder compared to our model by decreasing the best-fit H I and He I photoheating rates at $4.2 \leq z \leq 6.2$. We find that reducing $\mathcal{H}_{\mathrm{HI}}$ and $\mathcal{H}_{\mathrm{HeI}}$ by $\sim 80\%$ at $z \sim 6$ decreases the IGM temperature T_0 by $\sim 20\%$ making it consistent with the estimates from Boera

et al. (2019) at $4.2 \leq z \leq 5.0$ with minimal impact in T_0 at $z \lesssim 3.5$ (see Figure 3.21). Nevertheless, we find that such colder evolution of T_0 is in conflict with the $z \sim 5.4$ estimate from Gaikwad et al. (2020) (see Figure 3.21). This conflict indicates some degree of tension between the higher $T_0 = 1.10 \pm 0.16 \times 10^4\,\mathrm{K}$ at $z \sim 5.4$ from Gaikwad et al. (2020) and the low $T_0 = 7.37^{+1.13}_{-1.39} \times 10^3\,\mathrm{K}$ at $z \sim 5.0$ from Boera et al. (2019).

After $z \sim 4.5$, radiation from AGN ionizes He II atoms in the Universe and heats the IGM for a second time. Our model predicts that T_0 increases monotonically until He II reionization completes at $z \sim 3$, resulting in a second peak in the temperature ($T_0 \simeq 1.4 \times 10^4\,\mathrm{K}$) followed by a second epoch of cooling due to cosmic expansion. Our results for the evolution of T_0 during $z \lesssim 4.5$ are consistent with the determinations from Gaikwad et al. (2021) and Walther et al. (2019) that show a similar T_0 history within the uncertainties during and after He II reionization, as both show a peak in T_0 at $z \sim 2.8 - 3.0$. Our $T_0(z)$ results are higher yet consistent within the uncertainties from the measurement by Bolton et al. (2014) at $z = 2.4$. The results presented by Hiss et al. (2018) also show the effects of He II reionization on the temperature of the IGM in the form a peak in the temperature at $z \sim 2.8$, but their peak value of $T_0 \sim 2 \times 10^4\,\mathrm{K}$ is significantly higher than our result and the measurements from Gaikwad et al. (2020) and Walther et al. (2019).

The right panel of Figure 3.11 shows our result for the evolution of the density-temperature power-law index γ (black line and shaded 95% confidence interval). At the end of hydrogen reionization, the gas in the IGM is mostly isothermal ($\gamma \sim 1$). As the IGM cools and the low-density gas cools more efficiently, the index γ increases in

the interval $4.5 \lesssim z \lesssim 6$. During the reheating of the IGM from He II reionization, low-density gas heats faster and γ decreases until helium reionization completes. After helium reionization cooling from cosmic expansion causes an increase on γ for a second time.

The evolution of the power-law index in our model is consistent with measurements from Hiss et al. (2018), Boera et al. (2019), Gaikwad et al. (2020), and Gaikwad et al. (2021), and shows deviations only for a few redshift bins after He II reionization completes. The transition in γ after He II reionization in our model is not as pronounced as the determinations from Gaikwad et al. (2021) and Hiss et al. (2018).

The results from Walther et al. (2019) show significantly higher values of γ compared to all the other measurements. We have evaluated the plausibility of a steep density-temperature relation ($\gamma > 1.6$) by simulating the extreme case in which all photoheating and photoionization from the UVB stops after hydrogen reionization completes, i.e. $\Gamma = 0$ and $\mathcal{H} = 0$ for $z > 6$. We find that in the absence of external heating, as the IGM cools by adiabatic expansion, the overdensities cool down at a slower rate from compression by gravitational collapse. Here γ tends to increase with decreasing redshift at a roughly constant rate of $\Delta\gamma/|\Delta z| \sim 0.18$. Starting from an isothermal distribution of the gas in the IGM when H reionization finishes ($\gamma = 1$), it takes a change in redshift $|\Delta z| \sim 3 - 3.5$ for the gas distribution to steepen to $\gamma \sim 1.6$. Hence, we can reproduce values of $\gamma > 1.6$ at $z \sim 5$ only if hydrogen reionization completes very early at $z > 8$.

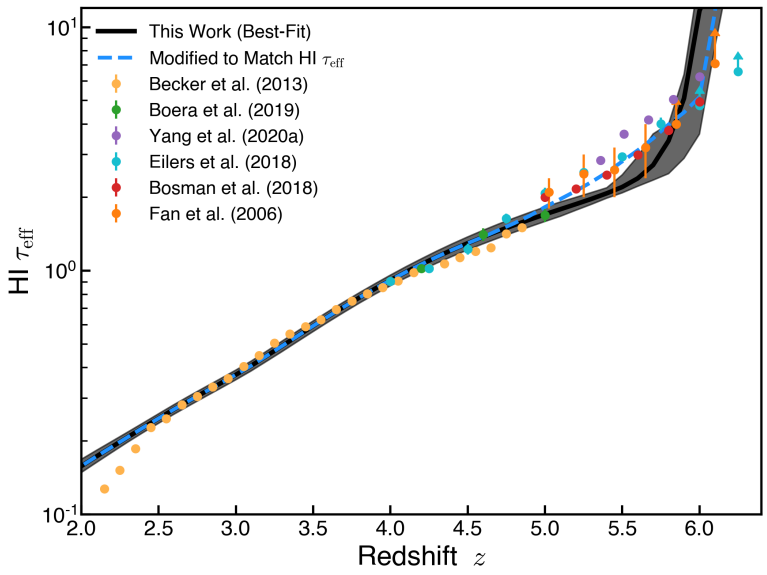

Figure 3.12: Redshift evolution of the hydrogen effective optical depth $\tau_{\rm eff,H}$ from our best-fit determination of the photoheating and photoionization rates (black line) and the corresponding 95% confidence interval. Data points show the observational measurements of $\tau_{\rm eff}$ from Fan et al. (2006), Becker et al. (2013b), Bosman et al. (2018), Eilers et al. (2018), Boera et al. (2019), and Yang et al. (2020a). The model results show consistency with the measurement from Becker et al. (2013b) (yellow) for $2.5 \lesssim z \lesssim 4.2$ and are in good agreement with the determination from Boera et al. (2019) (green) for $4.2 \lesssim z \lesssim 5.0$. At high redshift ($z > 5$) the results from Yang et al. (2020a) lie significantly higher than those from Eilers et al. (2018) and Bosman et al. (2018) by $\sim 10 - 30\%$. In the redshift range $5 \lesssim z \lesssim 5.8$, the model shows lower $\tau_{\rm eff,H}$ compared with the observations. By modifying the best-fit H I photoionization rate $\Gamma_{\rm HI}$ as shown in §3.3.7, we can obtain a high-z evolution of $\tau_{\rm eff,H}$ (dashed blue) consistent with the measurement from Bosman et al. (2018) and Fan et al. (2006).

3.3.4 Evolution of the Hydrogen Effective Optical Depth

The H I effective optical depth $\tau_{\rm eff,H} = -\ln\langle F\rangle$ measured from the Lyman-α forest reflects the overall H I content of the gas in the IGM. Hence, $\tau_{\rm eff,H}$ probes the ionization state of hydrogen in the medium and can be used to constrain the intensity of the ionizing UVB. In our work, constraints obtained for the H I photoionization rate $\Gamma_{\rm HI}$ derive from the power spectrum of the Lyman-α transmitted flux itself as we do not include the observational determinations of $\tau_{\rm eff,H}$ as constraints in our inference procedure.

The power spectrum $P(k)$ of the flux fluctuations [Eq. (3.1)] is itself sensitive to the hydrogen effective optical depth. Because of the non-linear relation $F = \exp(-\tau)$, the normalization of $P(k)$ on most scales relevant to this work ($0.002 \lesssim k \lesssim 0.1$ s km^{-1}) is affected by the value of $\tau_{\rm eff,H}$ obtained from the skewer sample used for the measurement. Thus, including the effective optical depth of the forest does not provide additional independent information for constraining the model. See Appendix 3.6 for a discussion on the impact that H I $\tau_{\rm eff}$ has on the Lyman-α flux power spectrum.

Figure 3.12 shows the redshift dependence of $\tau_{\rm eff,H}$ from our best-fit determination of the photoheating and photoionization rates (black line) and the corresponding 95% confidence interval. Data points in the figure show the observational measurements of $\tau_{\rm eff,H}$ reported by Fan et al. (2006), Becker et al. (2013b), Bosman et al. (2018), Eilers et al. (2018), Boera et al. (2019), and Yang et al. (2020a). Our results are consistent with the evolution of H I $\tau_{\rm eff}$ measured by Becker et al. (2013b) (yellow points) for the redshift range $2.5 \lesssim z \lesssim 4.2$. Our model results in a more opaque IGM compared to

their measurements at lower redshifts $2.2 \lesssim z \lesssim 2.5$ and higher redshifts $4.2 \lesssim z \lesssim 4.8$. Our model agrees well with the determination from Boera et al. (2019) (green points) during the redshift range $4.2 \lesssim z \lesssim 5.0$.

At high redshift ($z > 5$), the measurements of the H I effective optical depth from Bosman et al. (2018) (red points), Eilers et al. (2018) (cyan points), and Fan et al. (2006) (orange points) are similar, with only small differences ($< 12\%$) toward higher $\tau_{\rm eff,H}$ from Eilers et al. (2018) compared with Bosman et al. (2018). The measurements by Yang et al. (2020a) (purple points) suggest a more opaque IGM with a $\tau_{\rm eff,H}$ that is significantly higher ($\sim 20 - 30\%$) compared to the measurements by Bosman et al. (2018).

Shortly after hydrogen reionization completes ($5 \lesssim z \lesssim 5.8$), our best-fit UVB model significantly underestimates $\tau_{\rm eff,H}$ compared with the observational measurements, suggesting that the hydrogen in the IGM is overly ionized in our model at these redshifts. To address this possible discrepancy, we can modify our best-fit result for the H I photoionization rate such that $\Gamma_{\rm HI}$ is reduced only in the redshift range $4.8 < z < 5.8$ and increased for $5.8 < z < 6.1$ (see §3.3.5 and §3.3.7). As shown in Figure 3.12, the high redshift evolution ($z > 5$) of $\tau_{\rm eff,H}$ from the modified model (dashed blue line) is consistent with the measurements from Bosman et al. (2018). The subsequent evolution at redshifts $z < 4.8$ remains virtually unchanged from the best-fit model as hydrogen is in photoionization equilibrium at these times and the ionization fraction is therefore determined by the instantaneous amplitude of the H I photoionization rate $\Gamma_{\rm HI}$. We refer the reader to §3.3.7 for a discussion on the effect that the modified UVB

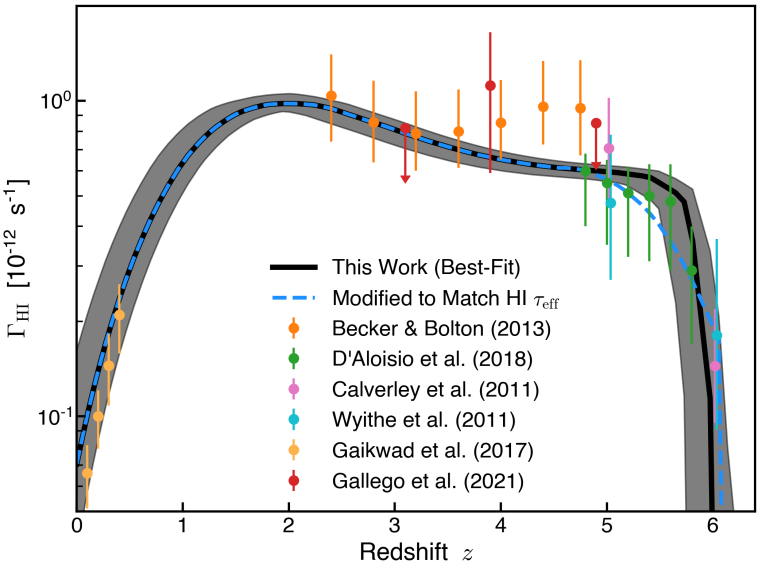

Figure 3.13: Evolution of the hydrogen photoionization rate Γ_{HI} from our best-fit determination and the 95% confidence interval (black line and shaded region). Data show observationally inferred photoionization rates measured by Calverley et al. (2011), Wyithe & Bolton (2011), Becker & Bolton (2013), Gaikwad et al. (2017),D'Aloisio et al. (2018), and Gallego et al. (2021). A modified model for Γ_{HI} designed to match the observational measurements of $\tau_{eff,H}$ from Bosman et al. (2018, see Figure 3.12) is shown as the dashed blue line. Our models agree well with the observationally-inferred results, except for visible differences with the estimate from Becker & Bolton (2013) during $4 \lesssim z \lesssim 5$. These differences in Γ_{HI} reflect small differences between our best-fit model predictions for $\tau_{eff,H}$ and the observational $\tau_{eff,H}$ measurement by Becker et al. (2013b) over this redshift range.

model has on the properties of the gas in the IGM.

By providing a simple modification to our best-fit UVB model that allows to change the high-redshift evolution of the hydrogen effective optical depth to achieve consistency with the observation and with minimal impact on the subsequent evolution of the properties of the IGM for $z \lesssim 5.0$, we show that the high-z discrepancy of the observed $\tau_{eff,H}$ and the model is not a significant challenge to our results and the conclusions of this work.

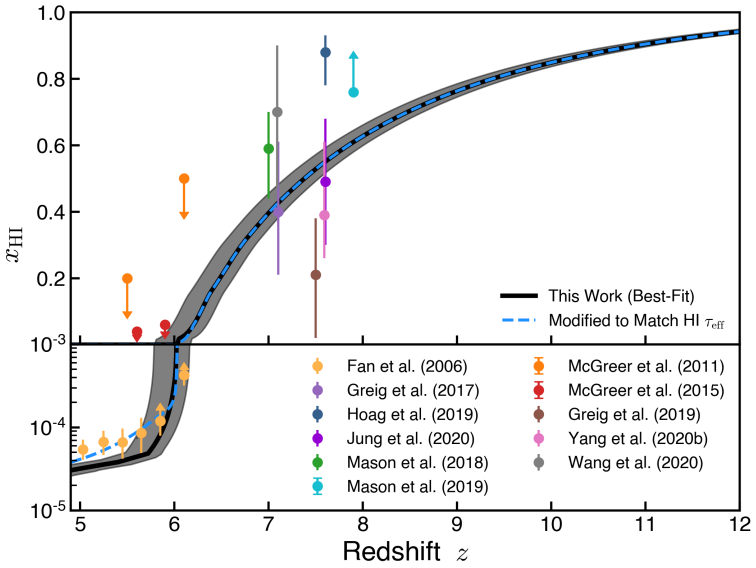

Figure 3.14: Redshift evolution of the volume-weighted neutral fraction of hydrogen for our best-fit model and the corresponding 95% confidence interval (black line and shaded region). Data points show the observational estimates reported in Fan et al. (2006), McGreer et al. (2011, 2015), Greig et al. (2017, 2019), Mason et al. (2018, 2019), Hoag et al. (2019), Jung et al. (2020), Yang et al. (2020b), and Wang et al. (2020). For $z \gtrsim 7$ the observational estimates show a wide range of $x_{\rm HI}$, from $x_{\rm HI} \sim 0.2$ to $x_{\rm HI} \sim 0.8$. Our models result in a $z \sim 7-8$ neutral fraction of $x_{\rm HI} \sim 0.4-0.5$, consistent with the results from Greig et al. (2017), Jung et al. (2020), and Yang et al. (2020b). After hydrogen reionization completes at $z \lesssim 6.0$, our best-fit model shows an evolution of $x_{\rm HI}$ below the measurement by Fan et al. (2006). By modifying our best-fit photoionization rates to better match $\tau_{\rm eff,H}$ (see Figure 3.12), we can also better match the $x_{\rm HI}$ data from Fan et al. (2006) (dashed blue line).

3.3.5 Hydrogen Photoionization Rate

Our best-fit model results for the hydrogen photoionization rate $\Gamma_{\rm HI}$ provide several opportunities for comparisons with observations, even though observationally inferred $\Gamma_{\rm HI}$ measurements are not used to constrain our model. There are observational determinations of $\Gamma_{\rm HI}$ informed by simulations where the photoionization rate is rescaled to match the observational $\langle F \rangle$(Becker & Bolton 2013; D'Aloisio et al. 2018). Our results can also be compared to estimates of $\Gamma_{\rm HI}$ from the quasar proximity effect and the size of the near-zone of high Lyman-α transmission around quasars (Calverley et al. 2011; Wyithe & Bolton 2011). Observations have measured $\Gamma_{\rm HI}$ by detecting the florescent Lyman-α emission produced by the Lyman limit systems (LLS) illuminated by background radiation (Gallego et al. 2021). Finally, there are $\Gamma_{\rm HI}$ determinations from combining the PDF and power spectrum of the Lyman-α transmitted flux from observations with simulations that apply different photoionization rates $\Gamma_{\rm HI}$ (Gaikwad et al. 2017).

Figure 3.13 shows our result for the HI photoionization rate with the corresponding 95% confidence limits (black line and shaded band) along with the observational inferences of $\Gamma_{\rm HI}$ mentioned above. Our result is consistent with the previous observational determinations that show a rapid evolution in $\Gamma_{\rm HI}$ for $z \gtrsim 5.6$, followed by a gradual increase during $2 \lesssim z \lesssim 5.6$ and a rapid decrease at $z < 2$. The only visible differences with Becker & Bolton (2013) occur in the redshift range $4 \lesssim z \lesssim 4.8$. Their measurement was obtained by tuning the photoionization rate $\Gamma_{\rm HI}$ in simulations such that the Lyα effective optical depth $\tau_{\rm eff,H}$ was consistent with the observational

measurement from Becker et al. (2013b). The higher estimate of $\Gamma_{\rm HI}$ from their result reflects the lower $\tau_{\rm eff,H}$ from Becker et al. (2013b) compared with the evolution of $\tau_{\rm eff,H}$ from our model for the redshift range $4.2 \lesssim z \lesssim 4.8$, as shown in Figure 3.12.

As described in §3.3.4, shortly after hydrogen reionization completes our best-fit model significantly underestimates the Lyman-α effective optical depth $\tau_{\rm eff,H}$ compared with the observations in the redshift range $5 \lesssim z \lesssim 5.8$. To address this discrepancy, we presented an alternative model where the sharp transition in $\Gamma_{\rm HI}$ at $z \sim 5.6$ from the original best-fit model is replaced by a softer increase that extends over the redshift range $4.8 < z < 5.8$ (dashed blue line in Figure 3.13). Decreasing $\Gamma_{\rm HI}$ during this epoch increases the neutral fraction of hydrogen in the IGM in photoionization equilibrium, thereby increasing $\tau_{\rm eff}$. Our modified model for $\Gamma_{\rm HI}$ was chosen such that the resulting evolution of $\tau_{\rm eff,H}$ is consistent with the observational measurement presented by Bosman et al. (2018)(dashed blue line in Figure 3.12), and the altered transition of $\Gamma_{\rm HI}$ from our modified model is still within the uncertainties of the observational inference by D'Aloisio et al. (2018) in the redshift interval $4.8 \lesssim z \lesssim 5.8$.

3.3.6 Ionization History

We present the redshift evolution of the volume-weighted neutral fraction of hydrogen $x_{\rm HI}$ resulting from our best-fit determination of the UVB model and the corresponding 95% confidence limits (black line and shaded band) in Figure 3.14. For comparison we show several observational estimates. We show constraints from the optical depth of the Lyman-α, Lyman-β, and Lyman-γ transitions in the forest (Fan et al. 2006). We also show constraints on the IGM neutrality from properties of Lyman-α

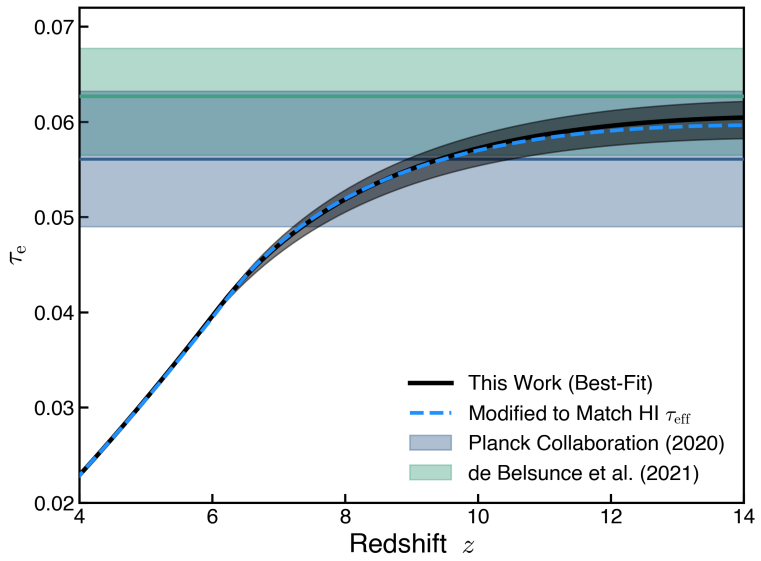

Figure 3.15: Thomson optical depth from electron-scattering of the CMB τ_e from the best-fit model and the 95% confidence limit (black line and shaded bar) and our modified model to match the $z > 5$ $\tau_{eff,H}$ (dashed blue line). Also shown are the observational measurements from the Planck satellite presented in Planck Collaboration et al. (2020) and the constraint from de Belsunce et al. (2021). Our model results for τ_e lie within the Planck limits.

emission from galaxies at high redshift (Hoag et al. 2019; Mason et al. 2018, 2019) and the damping wing absorption in the spectra of $z \gtrsim 7$ quasars (Greig et al. 2017, 2019; Jung et al. 2020; Yang et al. 2020b; Wang et al. 2020). Finally, we show constraints from the covering fraction of dark pixels in the Lyα/β forest of high-z quasars (McGreer et al. 2011, 2015).

Our model results in a prolonged hydrogen reionization history, extending from $x_{\rm HI} \sim 0.9$ at $z \sim 11$ to $x_{\rm HI} \sim 0.1$ at $z \sim 6.5$. The duration results in part from the gradually increasing ionization rate $\Gamma_{\rm HI} < 1 \times 10^{-15}\,{\rm s}^{-1}$ at $z > 6.5$ associated with radiation emitted by early star-forming galaxies.

For $7 \lesssim z \lesssim 8$, the observational estimates display a wide range of $x_{\rm HI}$, from a highly ionized ($x_{\rm HI} \sim 0.8$) to a mostly neutral ($x_{\rm HI} \sim 0.2$) IGM. Our model lies within this range, and at $z = 7$ our result is in agreement with the $x_{\rm HI} \sim 0.4$ estimates from Greig et al. (2017) and Yang et al. (2020b) as well as with the $x_{\rm HI} \sim 0.5$ estimate from Jung et al. (2020) at $z \sim 7.6$.

The redshift at which hydrogen reionization completes $z_{\rm R}$, defined as the redshift at which $x_{\rm HI} \leq 1 \times 10^{-3}$ for the first time, is $z \sim 6.0$ for our best-fit model. After hydrogen reionization completes, our best-fit model results in an ionization fraction that falls below the estimate from Fan et al. (2006) (reflected by the lower optical depth $\tau_{\rm eff}$ in Figure 3.12). Nevertheless, our modified model (dashed blue line) shows better consistency with their estimate.

Later in cosmic history, high energy radiation emitted by AGNs leads to the ionization of singly ionized helium (He II). For our best-fit model He II reionization starts

at $z \sim 5$ and completes at $z \sim 3.0$ when the He II fraction reaches $x_{\rm HeII} \leq 1 \times 10^{-3}$ for the first time. As the He II effective optical depth from our model is consistent with the observation from Worseck et al. (2019) for $2.4 \lesssim z \lesssim 2.9$, we argue that the end of He II reionization by $z \sim 2.9$ is suggested by their measurement.

Thomson scattering of the CMB by the free electrons in the IGM provides another diagnostic of the reionization history of the IGM. From the evolution of the electron density $n_{\rm e}$ given by the ionization state of hydrogen and helium from our models, we can compute the electron scattering optical depth $\tau_{\rm e}$ as

$$\tau_{\rm e}(z) = \int_0^z \frac{c\sigma_{\rm T} n_{\rm e}(z)}{(1+z)H(z)} {\rm d}z \tag{3.8}$$

where $\sigma_{\rm T}$ represents the Thomson scattering cross section. Figure 3.15 shows the electron scattering optical depth $\tau_{\rm e}$ from our best-fit model (black line and shaded region shows the 95% confidence limit). Also shown are constraints from the Planck satellite (Planck Collaboration et al. 2020) and the recent constraint from de Belsunce et al. (2021). Our result for $\tau_{\rm e} = 0.60$ lies within the upper limit of the $\tau_{\rm e} = 0.0540 \pm 0.0074$ constraint from Planck Collaboration et al. (2020) and in good agreement with the determination of $\tau_{\rm e} = 0.0627^{+0.0050}_{-0.0058}$ from de Belsunce et al. (2021).

3.3.7 Modified UVB Rates for Matching the Observed High-Redshift Hydrogen Effective Optical Depth

In §3.3.4 and §3.3.5 we discuss how the IGM from our best-fit model is possibly too highly ionized after hydrogen reionization completes. The hydrogen effective optical

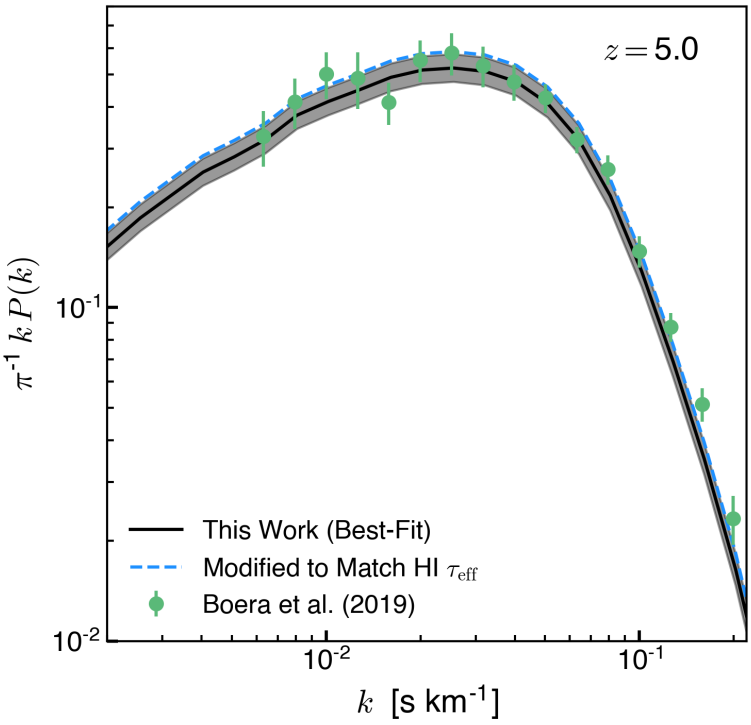

Figure 3.16: Power spectrum of the Lyman-α transmitted flux $P(k)$ at $z = 5$ from our best-fit model (black) and from our alternative model of the UVB (dashed blue) where the H I and He I photoionization and photoheating rates are modified in the redshift range $4.8 \lesssim z \lesssim 6.1$ such that $\tau_{\rm eff,H}$ is consistent with the observation from Bosman et al. (2018). The effect on the power spectrum from the modified model is to increase $P(k)$ by a roughly constant factor of $\sim 12\%$ compared with the best-fit model due to the $\sim 6\%$ increase in the H I opacity at $z = 5$. Both models are consistent with the observation from Boera et al. (2019) for $k \lesssim 0.1$ s km^{-1}.

depth $\tau_{\rm eff,H}$ from the model is significantly lower compared with observations in the redshift range $5 \lesssim z \lesssim 5.8$ (see Figure 3.12). We can address this issue by decreasing the H I photoionization rate $\Gamma_{\rm HI}$ such that the sharp transition at $z \sim 5.8$ from the best-fit model is replaced by a more gradual increase of $\Gamma_{\rm HI}$ during the redshift range $4.8 \lesssim z \lesssim 6.0$ (dashed blue line in Figure 3.13). This alternative transition in $\Gamma_{\rm HI}$ was chosen such that the resulting evolution of H I $\tau_{\rm eff,H}$ is consistent with the observations from Bosman et al. (2018).

Assuming that changes made to the photoionization rate $\Gamma_{\rm HI}$ correspond to a change of the mean-free-path of ionizing photons $\lambda_{\rm mfp}$, then the He Iphotoionization rate $\Gamma_{\rm HeI}$ should also reflect the modification applied to $\Gamma_{\rm HI}$. Correspondingly, we rescale the helium photoionization rate $\Gamma_{\rm HeI}$ such that the ratio $\Gamma_{\rm HI}(z)/\Gamma_{\rm HeI}(z)$ from the modified model matches the best-fit model.

Changing $\lambda_{\rm mfp}$ would also affect the photoheating rates $\mathcal{H}_{\rm HI}$ and $\mathcal{H}_{\rm HeI}$. Assuming that the average energy of the ionizing photons remains the same in the modified model, we rescale the photoheating rates such that the ratios $\mathcal{H}_{\rm HI}(z)/\Gamma_{\rm HI}(z)$ and $\mathcal{H}_{\rm HeI}(z)/\Gamma_{\rm HeI}(z)$ match the best-fit model. Results from our modified model for photoheating and photoionization rates are shown in Figure 3.9 as dashed blue lines.

After hydrogen reionization completes at $z \lesssim 6.0$, hydrogen in the IGM is in photoionization equilibrium. During this epoch, decreasing the H I and He I photoionization rates effectively increases the the neutral fraction of hydrogen and helium. Consequently the opacity of the IGM, quantified as the optical depth $\tau_{\rm eff}$, also increases during redshift range. The temperature of the gas in the IGM is not strongly affected

by the modified photoionization and photoheating rates, because, in equilibrium, the gas temperature $T(z) \propto \mathcal{H}(z)/\Gamma(z)$ and this ratio is unchanged from the best-fit model.

The modified model only changes the photoionization and photoheating rates during the redshift range $4.8 \leq z \leq 6.1$. These changes result in an increase of $\tau_{\rm eff,H}$ during $4.8 \leq z \leq 5.8$ and a decrease during $5.8 < z \leq 6.1$ but do not strongly affect the evolution of the gas temperature. For redshifts $z < 4.8$, the ionization fraction of hydrogen in the IGM in photoionization equilibrium is determined by the ratio of the photoionization rate to the recombination rate $x_{\rm HII}(z) \propto \Gamma_{\rm HI}(z)/\alpha_{\rm HII}(z,T)$. Since the thermal evolution resulting from the modified and best-fit models are very similar and the rates Γ and $\mathcal{H}$ at $z < 4.8$ are the same. Thereby, the evolution of the neutral fraction $x_{\rm HI}$, the effective optical depth $\tau_{\rm eff,H}$, and the Lyman-α power spectrum $P(k)$ resulting from the modified model is nearly unchanged from the best-fit model at redshifts $z < 4.8$.

The increase in the hydrogen effective optical depth $\tau_{\rm eff,H}$ during the redshift range $4.8 \leq z \leq 5.8$ in the modified model influences the Lyman-α power spectrum at this epoch. Given the available data, this modification only affects comparisons with the observed $P(k)$ at $z = 5.0$. Figure 3.16 shows $P(k)$ from the modified (dashed blue) and best-fit model (black) at $z = 5$. Relative to the best-fit model, using the modified model results in a small increase ($\sim 12\%$) in $P(k)$ owing to the small increase ($\sim 6\%$) in $\tau_{\rm eff,H}$. Either model shows consistency with the observational $P(k)$ measurement from Boera et al. (2019).

3.3.8 Limitations of the Model

For this work, we have modeled the evolution of the properties of the IGM using a spatially homogeneous ionizing background. Simulations of a more realistic, spatially inhomogeneous hydrogen reionization process show that spatial fluctuations in the temperature–density relation of the post-reionization IGM have a minor effect on the flux power spectrum (Keating et al. 2018) at $z \leq 5$ while the inhomogeneous UVB allows large islands of neutral hydrogen to persist up to redshift $z \leq 5.5$ and can reproduce the observed distribution of Lyman-α opacity (Kulkarni et al. 2019). Similarly, radiative transfers simulations of He II reionization show that the fluctuations in the ionization state of helium have a minor effect on observations of the hydrogen Lyman-α forest (La Plante et al. 2017; Upton Sanderbeck & Bird 2020). Not including the impact of galactic winds or AGN-feedback on the forest is a conservative approach for simulations aimed at constraining effects that suppress small-scale power. AGN feedback in the form of heating or mass redistribution from small to large scales is also expected to suppress the 1D power spectrum on large scales, and to have an increased effect at low redshifts (Viel et al. 2013b). Ignoring the impact of AGN feedback may lead to a few percent bias in the determination of cosmological and astrophysical parameters (Chabanier et al. 2020). This model uncertainty is comparable to the statistical uncertainties of the eBOSS data used in this work.

Another limitation of our method results from the UVB photoionization and photoheating rates used for our simulation grid being constructed from simple transformations of a template set of rates. We therefore do not probe the full range of ionization

and thermal histories that could be allowed by the observations of the Lyman-α forest. However, our model produces statistical properties of the Lyman-α forest that agree with a wide range of observations and a thermal evolution of the IGM consistent with previous inferences. These features of our work represent a significant achievement enabled by the ability to explore a wide range of models for the UVB from self-consistently evolved simulations. We emphasize that with our computational capabilities, performing a very large number of simulations (e.g., thousands) is now a possibility. We therefore defer more flexible explorations of models for the heating and ionization from the UVB to future work.

In the approach used for this work, we modify the photoionization and photoheating jointly. This joint variation results in another important limitation of our study. The large scales of the power spectrum of the forest are sensitive to the ionization state of H I which, in equilibrium, is set by the balance between photoionization and recombination. The large-scales of $P(k)$ depend on the temperature of the gas through the recombination coefficient $\alpha(T) \propto T^{-0.72}$ but are mostly determined by the intensity of the photoionization rate $\Gamma_{\rm HI}$. Since a large fraction of the dataset used for our inference probes the large-scale $P(k)$ the best-fit photoheating rates are influenced by the determination of the best-fit photoionization rates. We have shown that the photoheating from our best-fit model is consistent with other estimates of the thermal state of the IGM determined independently. Nevertheless, the relatively small uncertainty in the thermal state parameters T_0 and γ from this work is in part a consequence of the of the well-constrained determination of the photoionization rate from the large-scale

$P(k)$. In future work we will explore a more flexible approach in which the photoheating has some degree of freedom with respect to the photoionization rate, such as using density-dependent UVB rates to better model a inhomogeneous reionization.

3.4 Summary

With the objective of finding a photoionization and photoheating history that results in properties of the IGM consistent with observations of the hydrogen and helium Lyman-α forest, we have used the GPU-native *Cholla* code to perform an unprecedented grid of more than 400 cosmological simulations spanning a variety of ionization and thermal histories of the IGM. These calculations extend our CHIPS suite of hydrodynamical simulations initially presented in Villasenor et al. (2021a). We compare the properties of the Lyman-α forest from our simulations to several observational measurements to determine via a likelihood analysis the best-fit model for the photoionization and photoheating rates. From our best-fit model we have inferred the thermal history of the IGM, and demonstrate consistency with recent estimates obtained from the properties of the Lyman-α forest. A summary of the efforts and conclusions from this work follows.

- We present a direct extension of the CHIPS suite (Villasenor et al. 2021a) consisting of a grid of 400 simulations ($L = 50\,h^{-1}$Mpc, $N = 1024^3$) that vary the spatially-uniform photoionization and photoheating rates from the metagalactic UVB. The UVB rates applied for our grid use the Puchwein et al. (2019) model as a template, and use four parameters that control a rescaling of the amplitude and redshift-timing of the hydrogen and helium photoionization and photoheating

rates.

- The CHIPS simulations self-consistently evolve a wide range of ionization and thermal histories of the IGM. We compare the properties of the Lyman-α forest in the form of the power spectrum $P(k)$ of the hydrogen Lyman-α transmitted flux and the helium (He II) effective optical depth $\tau_{\rm eff,HeII}$ from our simulations to several observational measurements covering the redshift range $2.2 \leq z \leq 5.0$ for $P(k)$ (Iršič et al. 2017b; Boera et al. 2019; Chabanier et al. 2019c) and $2.4 \lesssim z \lesssim 2.9$ for $\tau_{\rm eff,HeII}$ (Worseck et al. 2016).

- We perform a Bayesian MCMC marginalization to determine the best-fit UVB model. The performance of each model in reproducing the observations is evaluated over the entire redshift evolution instead of comparing for each redshift bin independently. Additionally, our simulation grid naturally probes a large range of ionization histories that we match directly to evolution of the ionization state of hydrogen encoded in the power spectrum of the Lyman-α forest. We thereby avoid any need to rescale the optical depth from the simulations in post-processing to match the observed mean transmission of the forest, which is a common shortcoming of previous analyses.

- Our approach does not require an assumption of a power-law relation for the density-temperature distribution of the gas, as the Lyman-α spectra is constructed from our self-consistently evolved simulations. We find that a single power law does not accurately describe the $\rho_{\rm gas} - T$ distribution of the gas in the density range relevant to generating the signal of the Lyman-α forest.

- From our analysis, we infer the evolution of the thermal state of the IGM. The temperature history of the IGM shows a first temperature peak ($T_0 \simeq 1.3 \times 10^4$K) due to hydrogen reionization at $z \simeq 6$. This peak is followed by an epoch of cooling due to adiabatic expansion of the Universe until the onset of helium reionization from radiation emitted by AGNs. The ionization of helium leads to a second increase of the temperature until He II is fully ionized ($z \simeq 3$), resulting in a second peak of $T_0 \simeq 1.4 \times 10^4$K. The second peak is followed by a second period of cooling from cosmic expansion. Our result is consistent with previous estimates from Gaikwad et al. (2020) and Gaikwad et al. (2021). We note that the method employed in this work where we modify the UVB photoionization and photoheating rates by rescaling and shifting the model from Puchwein et al. (2019) limits the variation on the evolution of the thermal history of the IGM in our simulations. In future work we will allow for more flexibility in the photoheating history which will result in a more complete sample of the IGM density-temperature distribution. The improved flexibility of the models may permit a better inference of the thermal history of the IGM, as for now our low-redshift ($z < 4$) constraints are largely informed by the ionization state of hydrogen which likely results in a underestimated uncertainty in our $T_0 - \gamma$ evolution.

- We compare the evolution of the hydrogen effective optical depth $\tau_{\rm eff,H}$ from our best-fit model to several observational determinations. We find that after hydrogen reionization completes ($5 \lesssim z \lesssim 6$), the H I effective optical depth resulting from the model may underestimate the observations. We provide a modification to

our best-fit model where the photoionization and photoheating rates are reduced during this epoch such that the evolution of $\tau_{\rm eff,H}$ is consistent with measurements by Bosman et al. (2018). Additionally, the neutral fraction of hydrogen from the modified model shows consistency with the measurements by Fan et al. (2006) during this redshift interval.

- The model for the photoionization and photoheating rates from the UVB obtained from our analysis shows consistency with the observations of the Lyman-α power spectrum and the effective optical depth from both hydrogen and helium (He II), the optical depth from the CMB probed by Planck Collaboration et al. (2020), and previous inferences of the thermal state of the IGM. This model can be applied in future cosmological simulations that aim to reproduce properties of the IGM consistent with the observed Lyman-α forest.

Our work shows that an exploration of the IGM properties from hundreds of self-consistently evolved models for the astrophysical processes that impact the gas in the medium is now possible by exploiting modern computational techniques on the world's largest supercomputers. Using our efficient GPU-based code *Cholla* with *Summit*, we are able to run hundreds of cosmological simulations in just a few days using a small fraction of the system. We anticipate that when combined with the exquisite picture of the Lyman-α forest that experiments like DESI Collaboration et al. (2016a) will provide, this capability will revolutionize future studies of the properties of the IGM. We can leverage next-generation exascale systems and simulate large volumes ($L \sim 50h{-}1$Mpc) at high resolution ($N = 2048^3$) for thousands of models describing the various the

astrophysical processes that affect the IGM with a range of cosmological parameters, and study different models for the nature of dark matter and the mass hierarchy of neutrinos based on their impact on the small-scale power spectrum of the Lyman-α forest.

3.5 Resolution Convergence Analysis

To assess the possible impact of the simulation spatial resolution on our results, we compare the Lyman-α transmitted flux power spectrum measured from simulations with different resolutions. Each run was performed using the same box size ($L = 50h^{-1}$cMpc) for identical cosmological parameters (Planck Collaboration et al. 2020) and our best-fit determination for the photoionization and photoheating rates, and differ only in their grid resolution. Our comparison is made between three runs with resolutions $N = 512^3$, $N = 1024^3$, and $N = 2048^3$ cells and dark matter particles, with comoving spatial resolutions of $\Delta x \simeq 98$, 49, and 24 h^{-1}kpc, respectively. The initial conditions for the runs were generated to preserve common large-scale modes, such that the results from the simulations could be compared directly over shared spatial scales.

Figure 3.17 shows the power spectrum of the Lyman-α flux measured for our three simulations at redshifts z =2, 3, 4, and 5. As shown, the structure of the Lyman-α forest becomes better resolved as the number of cells increase. The lower panels present the fractional difference $\Delta P(k)/P(k)$ of the power spectrum measured from the $N = 512^3$ and $N = 1024^3$ simulations compared with the $N = 2048^3$ simulation on overlapping spatial scales. Our comparison shows that the effect of the decreased

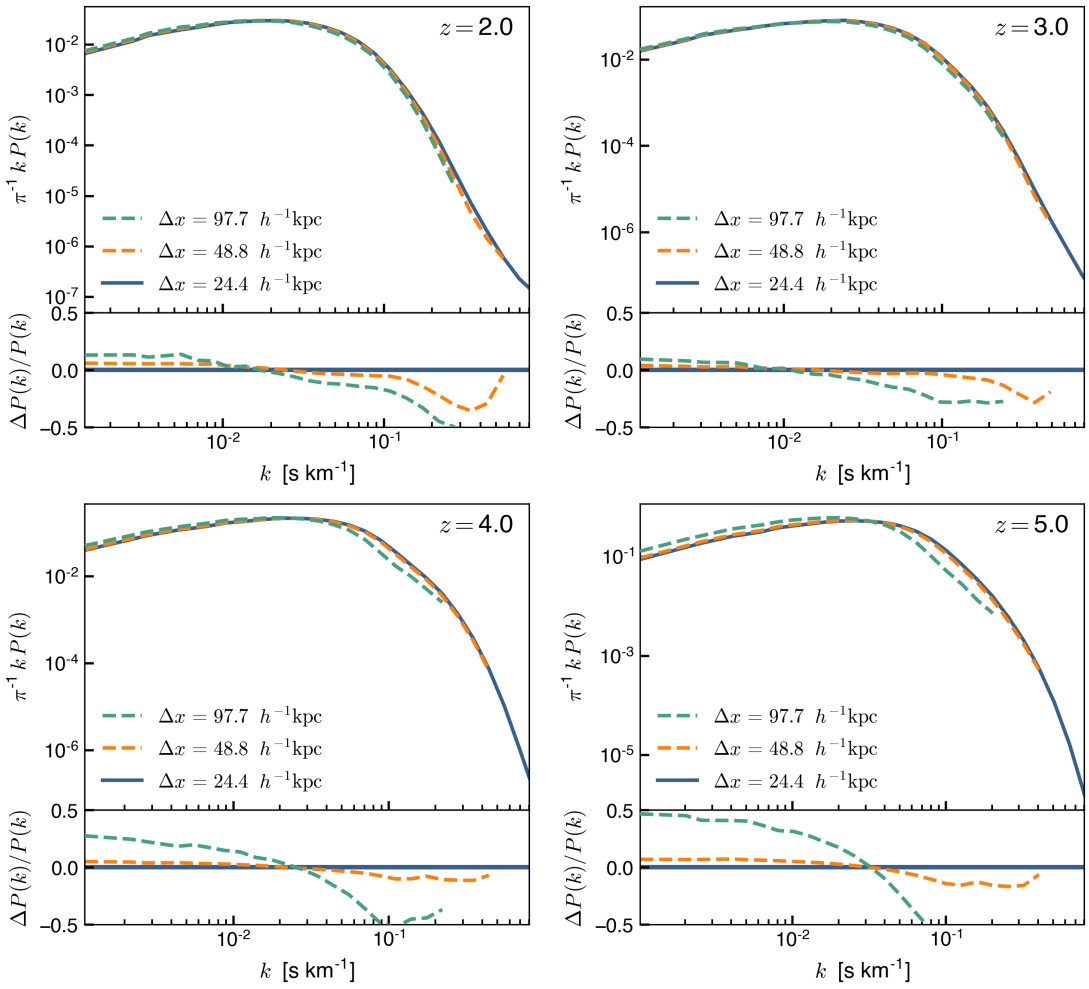

Figure 3.17: Power spectrum of the Lyman-α transmitted flux $P(k)$ measured from simulations with different comoving spatial resolutions of $\Delta x \simeq 98$, 49, and 24 h^{-1}kpc. The three simulations model a $L = 50h{-}1$Mpc box with the Planck Collaboration et al. (2020) cosmology and apply our best-fit determination for the photoionization and photoheating rates. The bottom panels show the fractional difference in the power spectrum $\Delta P(k)/P(k)$ between the $N = 512^3$ and $N = 1024^3$ runs and the $N = 2048^3$ simulation. Low-resolution simulations show increased power on large scales ($k \lesssim 0.03$ s km^{-1}) and suppressed structure in the small scales relative to higher resolution simulations. For the intermediate-resolution simulation $N = 1024^3$, which corresponds to our fiducial CHIPS grid resolution, the differences in $P(k)$ with respect to the $N = 2048^3$ simulation are $\lesssim 7\%$ on the large scales and $\lesssim 10 - 25\%$ on the small scales. We account for this resolution effect during our inference procedure by adding a systematic error to the observational measurements of $P(k)$ in the form of $\sigma_{\rm res} = \Delta P(k, z)$, where $\Delta P(k, z)$ is the redshift- and scale-dependent difference in the power spectrum measured from the $N = 1024^3$ run compared with the $N = 2048^3$ simulation.

resolution is to increase the power on large scales ($k \lesssim 0.02\ \mathrm{s\,km^{-1}}$) while the small-scale power is suppressed. For the low-resolution simulation ($N = 512^3$) the differences are significant, and on large scales the power spectrum is overestimated by $\sim 50\%$ at redshift $z = 5$. As the redshift decreases the differences also decrease to $\sim 13\%$ by $z = 2$. On small scales, the power spectrum is suppressed by $20 - 60\%$.

Our fiducial resolution for the CHIPS simulations was $N = 1024^3$. At this resolution we measure only small differences in the Lyman-α structure compared with the $N = 2048^3$ simulation, as on large spatial scales the power spectrum is overestimated by $\lesssim 7\%$, and for small scales ($0.03 \lesssim k \lesssim 0.2\ \mathrm{s\,km^{-1}}$) we measure a suppression on $P(k)$ of $\lesssim 10 - 25\%$. To account for the effect of resolution on simulations used to constrain the UVB model, we include a systematic uncertainty of the form $\sigma_{\rm res} = \Delta P(k, z)$, where $\Delta P(k, z)$ is the redshift- and scale-dependent difference in the power spectrum measured between the $N = 1024^3$ and $N = 2048^3$ simulations.

3.6 Effect of Rescaling the H I Effective Optical Depth on the Lyman-α Flux Power Spectrum

The power spectrum of the Lyman-α transmitted flux $P(k)$ is computed from flux fluctuations $\delta_F = (F - \langle F \rangle)/\langle F \rangle$. The power spectrum is sensitive to changes on the ionization state of hydrogen in the IGM, which in turn changes the effective optical depth $\tau_{\rm eff,H}$ and the mean transmitted flux $\langle F \rangle = \exp(-\tau_{\rm eff})$. To estimate how changes in the overall ionization state of the IGM affect the power spectrum of the Lyman-α flux, we can rescale the optical depth of the simulated skewers and re-measure $P(k)$. We

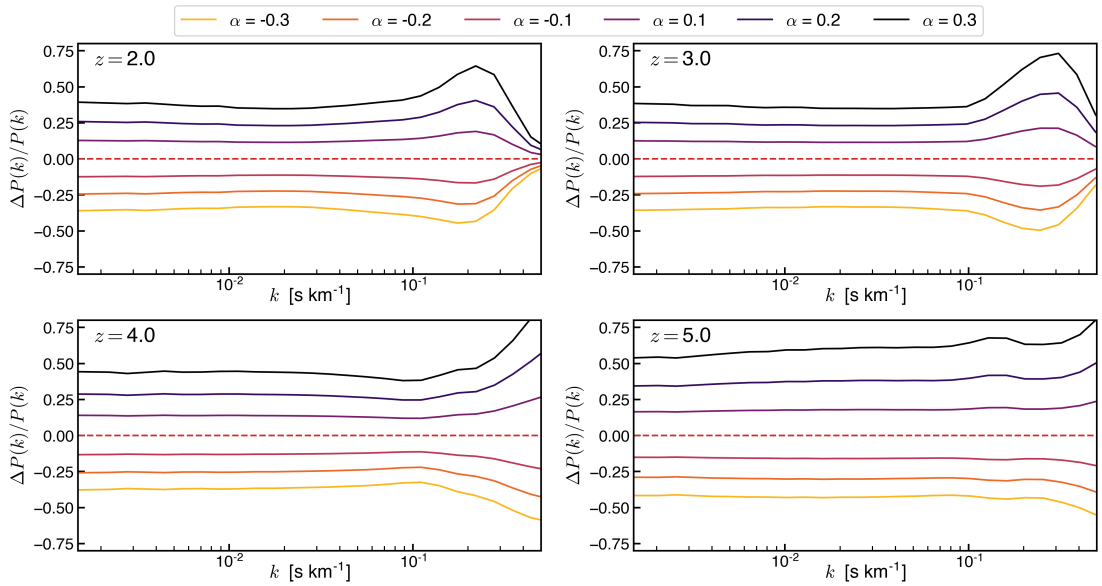

Figure 3.18: Consequences of rescaling the effective optical depth for the power spectrum of the Lyman-α transmitted flux at redshifts z =2, 3, 4, and 5. Shown is the fractional difference $\Delta P(k)/P(k)$ after rescaling the optical depth along the skewer sample from our simulations by a constant factor such that $\widetilde{\tau_{\rm eff,H}} = (1+\alpha)\tau_{\rm eff}$ for α in the range [-3, 3]. Rescaling the optical depth along the skewers such that $\tau_{\rm eff,H}$ increases (decreases) has the effect of increasing (decreasing) $P(k)$. On scales in the range $0.002 \lesssim k \lesssim 0.1\ {\rm s\,km^{-1}}$ the change induced on $P(k)$ is almost uniform, while for the smallest scales $k \gtrsim 0.1\ {\rm s\,km^{-1}}$ the effect is redshift- and scale-dependent.

rescale by a constant factor tuned such that the effective optical depth measured from the rescaled skewers follows $\widetilde{\tau_{\rm eff,H}} = (1+\alpha)\tau_{\rm eff,H}$, where $\tau_{\rm eff,H}$ is the original effective optical depth obtained from the simulated skewers. From the rescaled skewers, we compute the corresponding fluctuations of the transmitted flux $\widetilde{\delta_F} = (\widetilde{F} - \langle\widetilde{F}\rangle)/\langle\widetilde{F}\rangle$, where $\langle\widetilde{F}\rangle = \exp(-\widetilde{\tau_{\rm eff,H}})$. Finally, from $\widetilde{\delta_F}$ we compute the mean flux power spectrum $\widetilde{P}(k)$ for the rescaled sample.

Figure 3.18 shows the fractional difference of the flux power spectrum $\Delta P(k)/P(k) = \widetilde{P}(k)/P(k) - 1$ measured between the rescaled skewers and the original sample for several values in the range $\alpha \in [-0.3, 0.3]$. Because of the non-linear relation

between the optical depth τ and the transmitted flux $F = \exp(-\tau)$, rescaling the effective optical depth $\tau_{\rm eff,H}$ in the skewer sample to higher values $\alpha > 0$ has the effect of increasing the overall normalization of $P(k)$ on most of the scales relevant for this work, namely $0.002 \lesssim k \lesssim 0.1$ s km^{-1}. On a similar way, decreasing $\tau_{\rm eff,H}$ decreases the normalization of $P(k)$ at these scales. For smaller scales $k > 0.1$ s km^{-1} the effects are redshift dependent and we find that increasing (decreasing) $\tau_{\rm eff,H}$ tends to also increase (decrease) $P(k)$ for $z \gtrsim 3.5$, while it has the opposite effect for $z \lesssim 3.5$ as $P(k)$ decreases (increases) when $\tau_{\rm eff,H}$ is increased (decreased).

This study shows that the Lyman-α power spectrum itself is sensitive to the hydrogen effective optical depth, and for this reason we do not include the observational measurements of $\tau_{\rm eff,H}$ for our inference of the UVB model presented in this work.

3.7 Covariance Matrices of the Transmitted Flux Power Spectrum from the Simulations

In Section 3.2.9 we present the likelihood function employed for our MCMC analysis (Eq. 4.15). When comparing the power spectrum of the Lyman-α transmitted flux from the simulations to the observational measurements we employ the covariance matrices of $P(k)$ reported by the observational works (Chabanier et al. 2019c; Iršič et al. 2017b; Boera et al. 2019). In this section, we quantify the effect on the covariance of the simulated $P(k)$ from variations in our model parameters.

Figure 3.19 shows the normalized covariance of $P(k)$ at $z = 4.6$ for simulations with different values for the parameters $\beta_{\rm H}$ (top panels) and $\Delta z_{\rm H}$ (bottom panels).

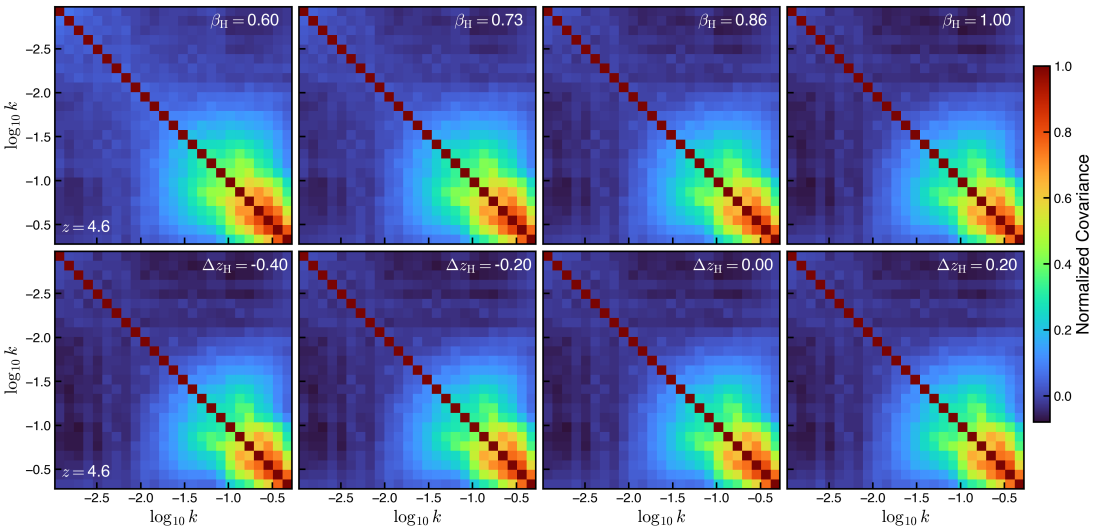

Figure 3.19: Normalized covariance matrix of the Lyman-α transmitted flux power spectrum at $z = 4.6$ measured from simulations that vary the parameters $\beta_{\rm H}$ (top panels) and $\Delta z_{\rm H}$ (bottom panels) independently. The structure of the covariance is maintained across the simulations. Decreasing the parameter $\beta_{\rm H}$ increases the normalization of $P(k)$ and its covariance on roughly all scales. We measure small elementwise differences < 0.1 in the normalized covariance across simulations with different $\beta_{\rm H}$. The effect of changing $\Delta z_{\rm H}$ is minimal with elementwise differences < 0.03.

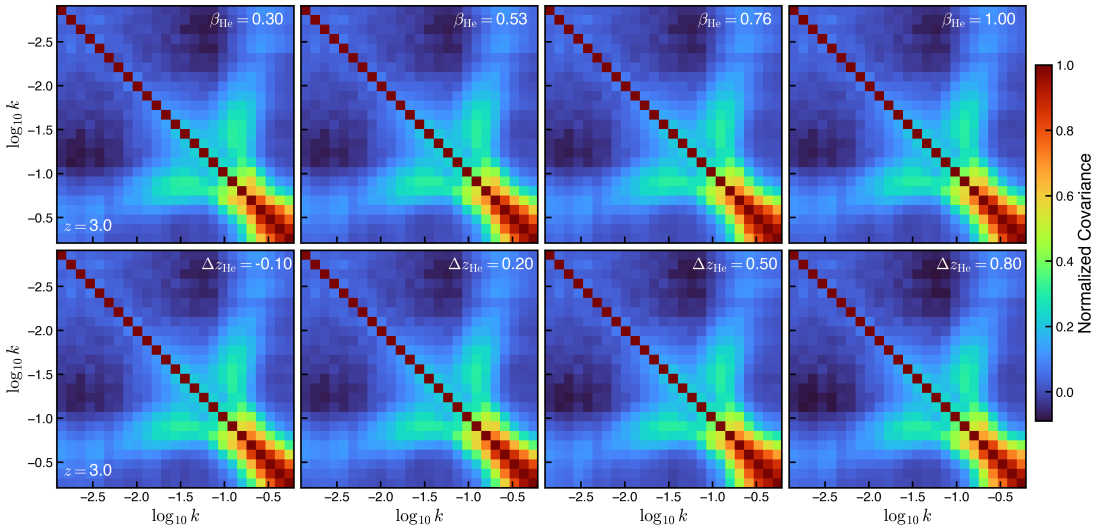

Figure 3.20: Normalized covariance matrix of the Lyman-α transmitted flux power spectrum at $z = 3.0$ measured from simulations that vary the parameters $\beta_{\rm He}$ (top panels) and $\Delta z_{\rm He}$ (bottom panels) independently. The structure of the covariance is maintained across the simulations. Changes in $\beta_{\rm He}$ and $\Delta z_{\rm He}$ cause small variation in the normalized covariance matrix, we measure only small elementwise differences < 0.05 over these simulations.

Decreasing the parameter $\beta_{\rm H}$ increases the Lyman-α opacity of the IGM, which increases the normalization of $P(k)$ (see §3.6). The increase of $P(k)$ also increases its covariance on roughly all scales. We measure small elementwise differences < 0.1 in the normalized covariance matrices across simulations that vary $\beta_{\rm H}$, while for simulations with different $\Delta z_{\rm H}$ the impact is minimal and results in only < 0.03 elementwise differences. Figure 3.20 presents the covariance matrix of $P(k)$ at $z = 3.0$ for simulations that vary the parameters $\beta_{\rm He}$ (top panels) and $\Delta z_{\rm He}$ (bottom panels). Here we also measure the impact to be small with differences < 0.05.

3.8 Colder Mid-redshift IGM From Reduced Photoheating

In §3.3.3 we discuss how our best-fit model results in a warmer IGM compared to the estimates from Boera et al. (2019) during the interval $4.2 \lesssim z \lesssim 5.0$ as the temperature T_0 from our model is $\sim 1\sigma$ higher compared to their result. We explore scenarios where the mid-redshift IGM is cooled relative to our best-fit model by decreasing the H I and He I photoheating rates in the redshift range $4.2 \lesssim z \lesssim 6.2$. The modified photoheating rates are shown in Figure 3.21 (middle and center panels) along with the fractional differences relative to the best-fit model shown in the respective bottom panels. To compute the history of T_0 for the reduced photoheating models, we integrate the evolution of the temperature of a single-cell at $\rho_{\rm gas} = \bar{\rho}$ following the method from Hui & Gnedin (1997) (see Section 2 of their work for a detailed description). The resulting evolution of T_0 for the different models is presented in the left panel of Figure 3.21. We show that reducing the H I and He I photoheating rates by $\sim 80\%$ at $z \sim 6$ results in a colder IGM where T_0 is reduced by $\sim 20\%$ at $z \sim 5$ such that $T_0 \sim 8 \times 10^3\,{\rm K}$ for $4.2 \lesssim z \lesssim 5.0$ agrees well with the estimate from Boera et al. (2019). However, we find that for such a scenario T_0 at $z \sim 5.4$ is lower than the inference from Gaikwad et al. (2020). This conflict exhibits some degree of tension between the estimates at $z \sim 5.0$ and $z \sim 5.4$ from Boera et al. (2019) and Gaikwad et al. (2020) respectively.

The photoheating $\mathcal{H}$ and the photoionization Γ rate from the UVB are given by the intensity of the background radiation $J(\nu, z)$ as

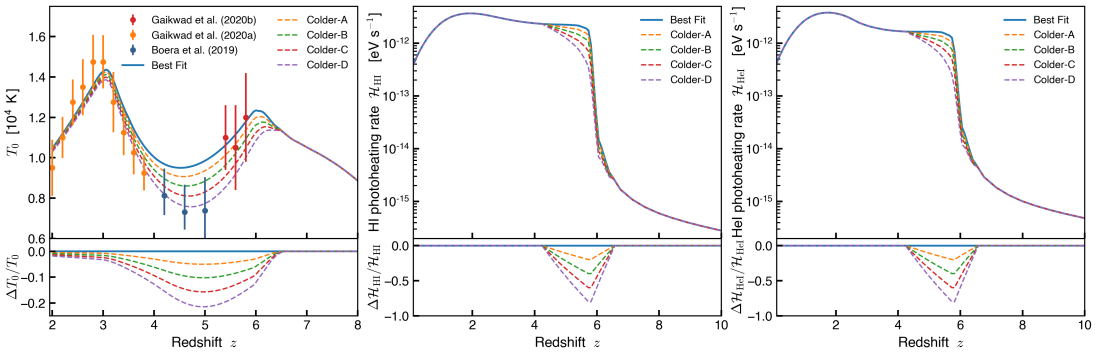

Figure 3.21: Evolution of the IGM temperature T_0 (left panel) from models of the UVB where the H I and He I photoheating rates have been reduced in the interval $4.2 \leq z \leq 6.2$ relative to our best-fit model (center and right panel). The fractional differences of T_0 and the heating rates $\mathcal{H}_{\rm HI}$ and $\mathcal{H}_{\rm HeI}$ with respect to the best-fit model are shown in the bottom panels. The reduced photoheating rates decrease T_0 for $z < 6.2$ but the change is most significant for $3.5 \lesssim z \lesssim 6.0$. At $z \lesssim 3.5$ the impact on T_0 is minimal as heating from He II reionization dominates. A reduction of $\sim 80\%$ in the photoheating rates at $z \sim 6.0$ causes a decrease in T_0 of $\sim 20\%$ at $z \sim 5.0$. For $z \lesssim 3.5$ the reduced photoheating has a minimal impact on T_0 of $\lesssim 5\%$.

$$\Gamma(z) = \int_{\nu_0}^{\infty} \frac{4\pi J(\nu, z)}{h\nu} \sigma(\nu) d\nu, \qquad \mathcal{H}(z) = \int_{\nu_0}^{\infty} \frac{4\pi J(\nu, z)}{h\nu} (h\nu - h\nu_0)\, \sigma(\nu) d\nu \tag{3.9}$$

where ν_0 and $\sigma(\nu)$ are the threshold frequency and photoionization cross-section, respectively. Consider power-law models for the cross-section and the intensity of the radiation at wavelengths $\lambda > 912\,\text{Å}$, which can be written as $\sigma(\nu) = \sigma_0(\nu/\nu_0)^{\phi}$ and $J(\nu) = (\nu/\nu_0)^{\alpha}$, with indices $\phi < 0$ and $\alpha < 0$. Physically, reducing the photoheating rate relative to the photoionization rate can be achieved by changing the spectral index of the ionizing radiation α. By solving the integrals in Eqs 4.2 assuming these power-law models and evaluating the fractional change in the photoionization $\Delta\Gamma/\Gamma$ and photoheating $\Delta\mathcal{H}/\mathcal{H}$ for a change in the spectral index $\Delta\alpha$, we find the following

relation is satisfied

$$\Delta\alpha = (1 + \alpha + \phi)\frac{\frac{\Delta\Gamma}{\Gamma} - \frac{\Delta\mathcal{H}}{\mathcal{H}}}{1 + \frac{\Delta\mathcal{H}}{\mathcal{H}}} \tag{3.10}$$

Equation 3.10 relates the change of the spectral index of the radiation necessary to produce some variation of the photoionization and photoheating from a given UVB model. By applying Eq. 3.10, we can modify the photoheating relative to the photoionization of a UVB model within a physically-plausible range for the index α. In future work, we will explore which variations in the IGM temperature T_0 from changes of the photoheating rate match the observed hydrogen effective optical depth at $z > 5$ while using physically-plausible source populations.

3.9 Accuracy of the Power-Law fit to the Density-Temperature Distribution of the Gas in Our Simulations.

A common method to infer the thermal state of the IGM from observations of the Lyman-α forest involves marginalizing over the thermal properties T_0 and γ in the approximate power-law density-temperature relation $T = T_0\,(\rho_{\rm gas}/\bar{\rho})^{\gamma-1}$ (Bolton et al. 2014; Nasir et al. 2016b; Hiss et al. 2018; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021). The density of the IGM gas that contributes to the majority of the Lyman-α forest signal lies in the range $-1 \leq \log_{10}(\rho_{\rm gas}/\bar{\rho}) \leq 1$. From our simulations we find that a single power law fails to reproduce the density-temperature distribution

of the gas over this density interval. The left panels of Figure 3.22 show the density-temperature distribution of the gas in one of our simulations and the corresponding power-law fit to the distribution over the density range $-1 \leq \log_{10}(\rho_{\rm gas}/\bar{\rho}) \leq 1$ at redshift $z = 3$ (top) and $z = 4$ (bottom). The deviations of the gas temperature in the simulation relative to the power-law fits are presented in the right panels, showing that the fractional differences $\Delta T/T$ from the density-temperature distribution in the simulation with respect to the power-law fit can be as large as $\sim 15\%$.

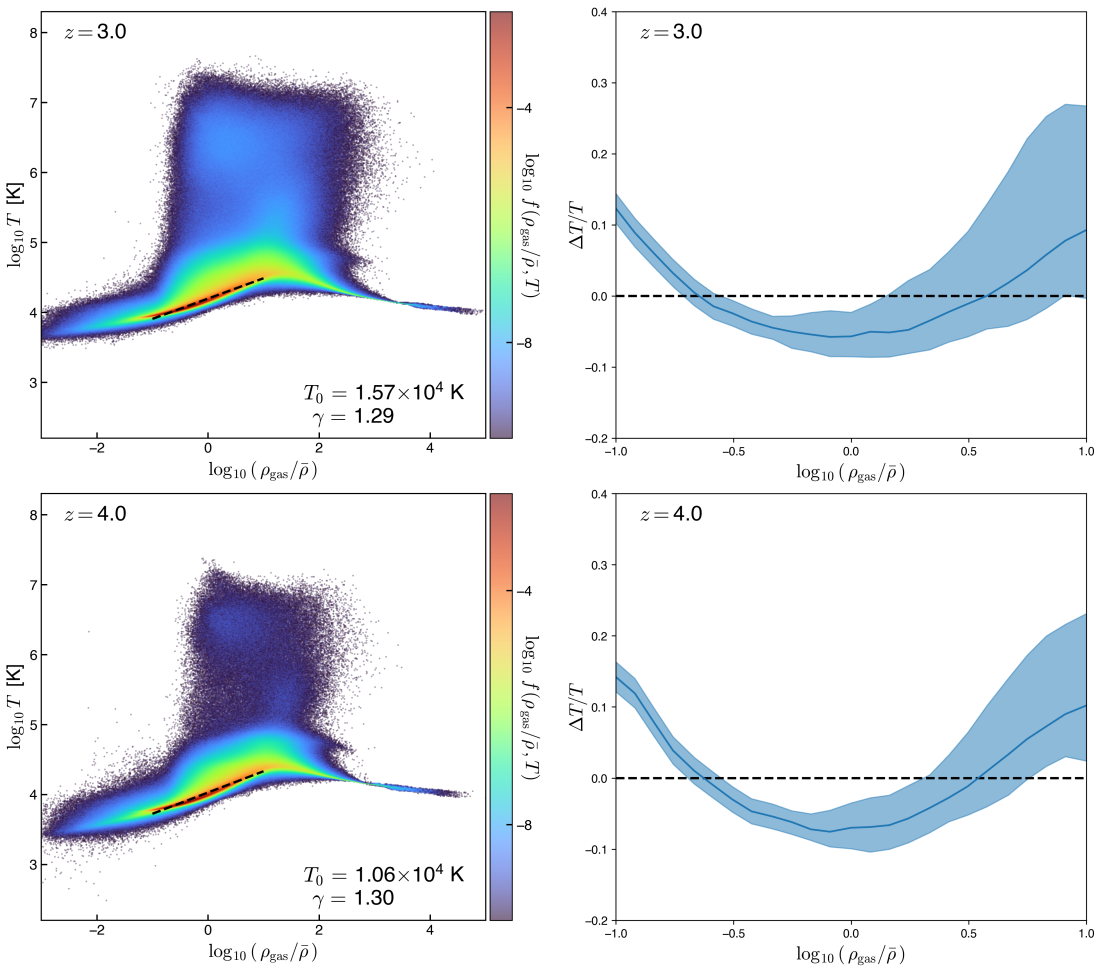

Figure 3.22: Density-temperature distribution of the IGM gas (left column) from one of our simulations at redshift $z = 3$ (top) and $z = 4$ (bottom). A power law fit of the form $T = T_0\,(\rho_{\rm gas}/\bar{\rho})^{\gamma-1}$ over the range $-1 \leq \log_{10}(\rho_{\rm gas}/\bar{\rho}) \leq 1$ is shown (black dashed lines). The right columns shows deviations of the density-temperature distribution with respect to the power-law fit over the fitted range. The blue region corresponds to the 68% highest probability interval for the temperature as function of the overdensity $\rho_{\rm gas}/\bar{\rho}$. The differences between the distribution of the gas relative to the power-law fit can be as large as $\sim 15\%$.

Chapter 4

New Constraints on Warm Dark Matter from the Lyman-α Forest Power Spectrum

4.1 Introduction

The Λ-cold dark matter (ΛCDM) cosmological paradigm has been immensely successful at matching across cosmic time observations spanning physical scales from the horizon length (Planck Collaboration et al. 2020) all the way down to galaxy scales (e.g., Chabanier et al. 2019a), and a vast menagerie of hypothetical non-baryonic elementary particles has been proposed to explain the astrophysical data (Feng 2010). Cold dark matter particles have negligible thermal velocities in the matter-dominated era and therefore clump gravitationally even on the smallest sub-galactic scales, a property

that has caused persistent challenges with observations of the abundances and density profiles of dwarf galaxies in the local Universe (e.g., Bullock & Boylan-Kolchin 2017). Warm dark matter (WDM) is a simple modification of CDM that has been proposed to suppress small-scale power and alleviate some of these problems (Bode et al. 2001). WDM particles of a few keV have significant intrinsic velocities from having decoupled as thermal relics or been produced by non-equilibrium processes, and one of the effects of their Mpc-range *free-streaming length* is to interfere with the gravitational collapse of structures and produce a cut-off in the matter power spectrum.

Intergalactic hydrogen at redshift $2 \lesssim z \lesssim 5$ scatters Lyman-α radiation and produces absorption features in the spectra of distant quasars. This "Lyman-α forest" is a powerful cosmological probe as it traces density fluctuations, the underlying dark matter web-like distribution (the "cosmic web"), and the ionization state and temperature of the intergalactic medium (IGM) on scales and redshifts that cannot be probed by any other observable (Hernquist et al. 1996; Croft et al. 1998b; Meiksin 2009b; McQuinn 2016). The primary statistic derived from spectroscopic data is the 1D power spectrum of the flux distribution in the forest – the Fourier transform of the fractional flux autocorrelation function in velocity space – which arguably provides the best tool for distinguishing between CDM and WDM models (Seljak et al. 2006; Viel et al. 2013a; Baur et al. 2016; Garzilli et al. 2019, 2021). Interpreting such observations requires expensive hydrodynamical simulations of the IGM that cover an extensive range of uncertain IGM photoionization and photoheating histories and consistently maintain high resolution throughout a statistically representative sub-volume of the Universe, a tradi-

tional limiting factor in previous analyses (e.g., Viel et al. 2013a; Lukić et al. 2015; Iršič et al. 2017a; Bolton et al. 2017; Walther et al. 2021). At present, the tightest lower limit to the mass of a thermal WDM relics, $m_{\rm WDM} > 1.9$ keV (95 percent CL), is obtained for highly conservative thermal histories of intergalactic gas (Garzilli et al. 2021).

In this Paper, we revisit these constraints in light of recent measurements with the *Keck* and *VLT* telescopes of the Lyman-α FPS down to the smallest scales ever probed at redshifts $4 \lesssim z \lesssim 5$ (Boera et al. 2019), using a massive suite of 1080 high-resolution cosmological hydrodynamics simulations that are part of the Cholla IGM Photoheating Simulations (CHIPS; Villasenor et al. 2021a,b). High-redshift observations provide better limits on the mass of a WDM particle, as free-streaming becomes more prominent in velocity space and the effect of the non-linear evolution of the matter density field – resulting in a transfer of power from large scales to small scales – is weaker at these epochs. There is also observational evidence for a local minimum of the temperature of the IGM – and therefore a corresponding minimum in the thermal cut-off length scale – in the redshift range $4 < z < 5$ (Villasenor et al. 2021b; Gaikwad et al. 2020), making this era an optimal epoch for deriving bounds on dark matter properties.

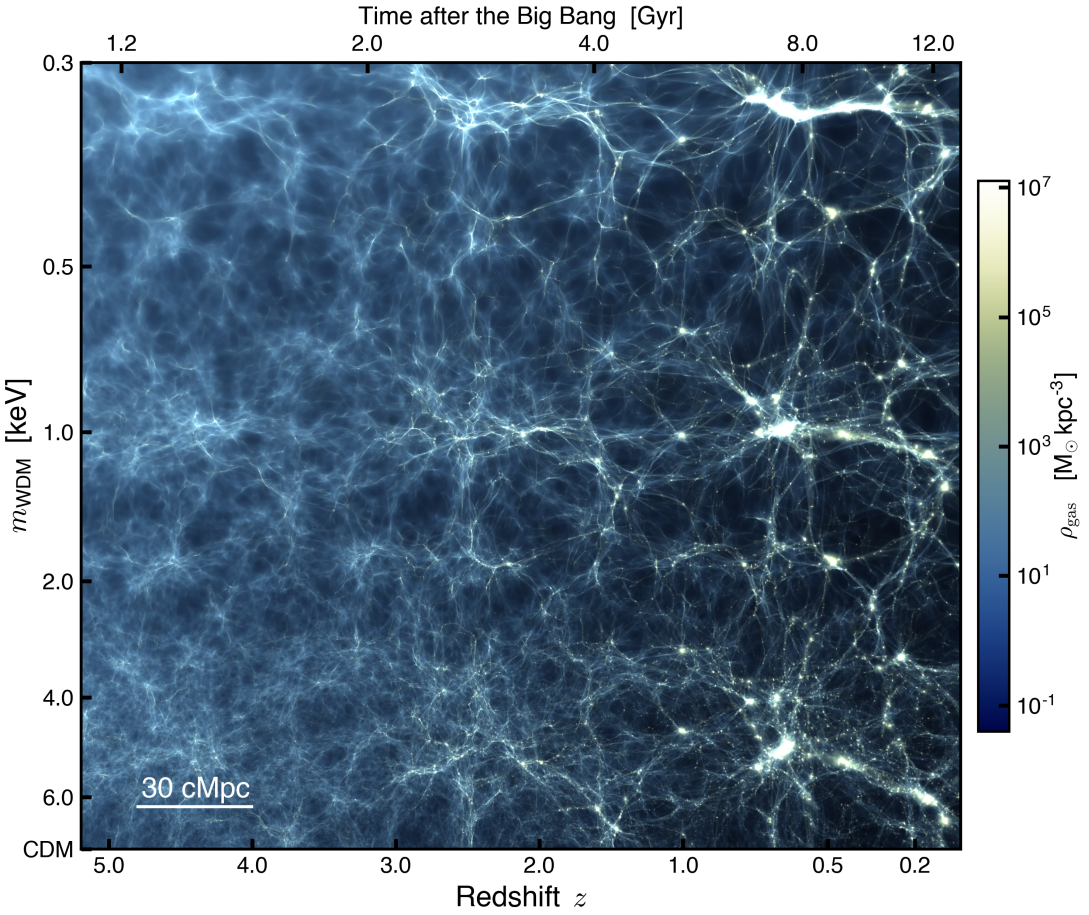

Figure 4.1: Impact of particle free-streaming on baryonic structures at $0 < z < 5.2$. The evolution of the gas density from a slice through the IGM was obtained from a subset of 8 CHIPS simulations where the mass of the warm dark matter particle $m_{\rm WDM}$ was increased from 0.3 keV to ∞. All simulations assume the same gas thermal history – the best-fit model presented in Villasenor et al. (2021b). Because of thermal pressure, the gas distribution is smoothed compared to the dark matter.

4.2 Basic Equations and Characteristic Scales

The hydrogen and helium photoheating H_i and photoionization Γ_i rates per atom i depend on the intensity of a uniform UV background radiation field $J(\nu, z)$ as

$$\Gamma_i(z) = 4\pi \int_{\nu_i}^{\infty} \frac{J(\nu, z)}{h\nu} \sigma_i(\nu) d\nu, \tag{4.1}$$

$$\mathcal{H}_i(z) = 4\pi \int_{\nu_i}^{\infty} \frac{J(\nu, z)}{h\nu} (h\nu - h\nu_i)\, \sigma_i(\nu) d\nu, \tag{4.2}$$

where ν_i and $\sigma_i(\nu)$ are the threshold frequency and photoionization cross-section, respectively. The total photoheating rate $\mathcal{H}$ is summed over the species $i =$ H I, He I, and He II each of proper number density n_i, $\mathcal{H} = \sum_i n_i \mathcal{H}_i$. These rates, together with radiative recombinations, adiabatic compression, Compton and expansion cooling, and the gas density and peculiar velocity fields shaped by gravity and by pressure forces, fully determine the hydrogen Lyman-α absorption optical depth as a function of velocity Hui et al. (1997),

$$\tau_v = \sigma_\alpha \int \frac{c\, n_{\rm HI}}{\sqrt{\pi} b (1+z)} \left| \frac{du}{dx} \right|^{-1} e^{-(v-u)^2/b^2} du. \tag{4.3}$$

Here, $\sigma_\alpha = \pi e^2 f_{12}/(m_e c\, \nu_\alpha)$, ν_α and f_{12} are the frequency and upward oscillator strength of the Lyman-α transition, b is the Doppler parameter, x is the comoving radial coordinate along the line of sight, and we have assumed a Gaussian profile. Denoting with $\bar{z}$ some mean redshift of interest, e.g. the redshift of a simulation output or the average redshift of any given data subset, one can write

$$u(x) = \frac{H(\bar{z})}{(1+\bar{z})} (x - \bar{x}) + v_p(x), \tag{4.4}$$

where v_p is the peculiar velocity along the line of sight at position x, $\bar{x}$ is the position at which the redshift from Hubble expansion is exactly $\bar{z}$, and $H(\bar{z})$ is the Hubble expansion rate. The velocity coordinate u is related to the observed frequency ν_0 by

$$\nu = \nu_0(1+\bar{z})\left(1+\frac{u}{c}\right), \tag{4.5}$$

where ν is the frequency of the photon with respect to the rest frame of intervening hydrogen at position x. Line center, $\nu = \nu_\alpha$, occurs at velocity $u = v$. Because of peculiar velocities, a photon observed at ν_0 can have the same rest-frame frequency ν at more than one place in its trajectory from the quasar to the observer.

In practice, only a limited range of u values contributes to τ_v, and one can replace in Equation (4.3) the redshift z with $\bar{z}$. We can then define the flux contrast $\delta_F(v) = F(v)/\langle F\rangle - 1$, where $F(v) = \exp(-\tau_v)$ is the flux at velocity v and $\langle F\rangle$ is the mean transmitted flux at a given redshift, and decompose each absorption spectrum into Fourier modes $\tilde{\delta}_F(k)$. Their variance as a function of the Fourier wavenumber $k = 2\pi/v$ is the FPS over some velocity interval Δv,

$$P_F(k) = \Delta v\langle\tilde{\delta}_F(k)^2\rangle, \tag{4.6}$$

which is commonly expressed in terms of the dimensionless quantity $\Delta_F^2(k) = kP_F(k)/\pi$.

Four physical effects act to erase small-scale power in the linear FPS:

1. Thermal Doppler broadening, which causes a sharp feature in $F(v)$ to be smoothed over a velocity width $b = (2k_BT/m_{\rm H})^{1/2}$, where T is the gas temperature and $m_{\rm H}$

is the mass of the hydrogen atom. A cut-off in the linear FPS will arise at the proper wavenumber Garzilli et al. (2019)

$$k_b = \frac{\sqrt{2}}{b} = 0.11\, T_4^{-1/2}\,\mathrm{km}^{-1}\,\mathrm{s}, \tag{4.7}$$

where $T_4 \equiv T/10^4\,\mathrm{K}$.

2. Pressure smoothing, which reduces power below a comoving "filtering scale" Gnedin & Hui (1998)

$$\lambda_F = g\,\lambda_J = g\,\frac{c_s}{a}\sqrt{\frac{\pi}{G\bar{\rho}}}, \tag{4.8}$$

where λ_J is the Jeans length, $c_s = [5k_BT/(3\mu m_{\mathrm{H}})]^{1/2}$ is the sound speed, μ is the mean molecular weight ($\mu = 0.61$ for an admixture of ionized hydrogen and singly ionized helium), $\bar{\rho}$ is the mean total (dark matter plus baryons) mass density of the Universe, and a is the scale factor. The g factor accounts for the fact that the filtering scale depends on the prior thermal history of the IGM, not just the instantaneous Jeans length. Typically, $g < 1$ after reionization. The corresponding velocity broadening, $\Delta v_F = Ha\lambda_F/(2\pi)$, produces a feature in the linear FPS at proper wavenumber

$$k_F = \frac{\sqrt{2}}{\Delta v_F} = \frac{\sqrt{3}}{gc_s} = \frac{0.12}{g}\,T_4^{-1/2}\,\mathrm{km}^{-1}\,\mathrm{s}. \tag{4.9}$$

Under the physically-motivated assumptions that the gas temperature decays as $1/a$ after reionization completes around redshift $z \sim 6$, the factor g takes numer-

ical values in the range $0.1 - 0.25$ at the redshifts $4 \lesssim z \lesssim 5.2$ of interest here Gnedin & Hui (1998). Pressure filtering, always lagging behind the growth of the Jeans length, is therefore subdominant compared to the thermal broadening of the absorption spectrum Peeples et al. (2010); Kulkarni et al. (2015b); Nasir et al. (2016a); Boera et al. (2019).

3. Peculiar velocities in the gas, another line-of-sight effect that smoothes baryonic structures in redshift space and depresses the FPS at wavenumbers greater than

$$k_p = \frac{\sqrt{2}}{\Delta v_p} = 0.02\,\mathrm{km}^{-1}\,\mathrm{s}. \tag{4.10}$$

4. Dark matter free-streaming, an effect that can be described by a relative "transfer function" $T(k)$ – the square root of the ratio of the matter power spectrum in the presence of WDM relative to that in CDM, which is well approximated for $k < 3.5\ \mathrm{Mpc}^{-1}$ by the fitting formula Viel et al. (2005)

$$T(k) = [1 + (k/k_{\mathrm{FS}})^{2.24}]^{4.65}. \tag{4.11}$$

In the adopted cosmology and in the case of thermal relics, the comoving scale of the break is

$$k_{\mathrm{FS}} = 13.68 \left(\frac{m_{\mathrm{WDM}}}{1\,\mathrm{keV}}\right)^{1.11}\ \mathrm{Mpc}^{-1}. \tag{4.12}$$

The conversion to a proper wavenumber in inverse km s^{-1}, $k_{\mathrm{DM}} = k_{\mathrm{FS}}(1+z)H^{-1}$,

finally yields a cut-off in the FPS at

$$k_{\rm DM} = 0.15 \left(\frac{m_{\rm WDM}}{1\,{\rm keV}}\right)^{1.11} \left(\frac{6}{1+z}\right)^{1/2} {\rm km}^{-1}\,{\rm s}. \tag{4.13}$$

The numerical values of k_b, k_p, and $k_{\rm DM}$ above suggest that free-streaming, thermal and velocity broadening effects all set in at similar values of k.

4.3 Simulations

The simulations used in this work were performed using the GPU-native MPI-parallelized code Cholla (Schneider & Robertson 2015; Villasenor et al. 2021a,b) and evolve the equations of hydrodynamics on a uniform Cartesian grid while simultaneously tracking the non-equilibrium ionization states of hydrogen and helium using the GRACKLE library (Smith et al. 2017). A spatially uniform, time-dependent UV radiation background was assumed in the form of redshift-dependent photoionization and photoheating rates per ion, Γ_i and H_i, for the species H I, He I, and He II. The initial conditions at $z = 100$ were generated using the MUSIC code (Hahn & Abel 2011b) for a flat cosmology with parameters $h = 0.6766$, $\Omega_M = 0.3111$, $\Omega_\Lambda = 0.6889$, $\Omega_b = 0.0497$, $\sigma_8 = 0.8102$, and $n_s = 0.9665$, consistent with constraints from Planck data (Planck Collaboration et al. 2020). The initial conditions for all runs were generated from identical random number seeds to preserve the same amplitude and phase for all initial Fourier modes across the simulation suite. The volume and numerical size of the simulations correspond to $L = 25\ h^{-1}$Mpc and $N = 2{\times}1024^3$ cells and particles. All simulations

were run on the *Summit* computing system at the Oak Ridge National Laboratory. The effect of a thermally produced WDM particle was introduced by modifying the input transfer function following the fitting formula of Viel et al. (2005), and no attempt was made to explicitly incorporate particle thermal velocities in the initial conditions (Macciò et al. 2013).

To compare the properties of the IGM in our simulations to observations, we extracted synthetic Lyman-α forest spectra measuring H I densities, temperatures, and peculiar velocities of the gas using 4096 skewers along each axis of the box. The optical depth τ_v along each discretized line of sight was estimated as described in Villasenor et al. (2021a). From the set of lines of sight, we computed the transmitted flux and the mean one-dimensional power spectrum of the fractional transmission (see §2.5.2 and §2.5.4 for a detailed description).

4.3.1 WDM Effects on the IGM

The impact of particle free-streaming on gas the density structure of the IGM at fixed gas thermal history is shown in Figure 4.1. The image displays a slice through the gaseous cosmic web at $0 \leq z \leq 5.2$ generated from a set of 8 simulations that vary the mass of the dark matter particle $m_{\rm WDM}$ increasing from 0.3 keV to ∞. The simulations follow the thermal history produced by the baseline photoheating and photoionization history presented in Villasenor et al. (2021b).

The principal effect of free-streaming from WDM on the matter density distribution is to suppress the formation of small-scale structures, reducing the clumpiness of the cosmic web. As free-streaming washes out small-scale inhomogeneities and decreases

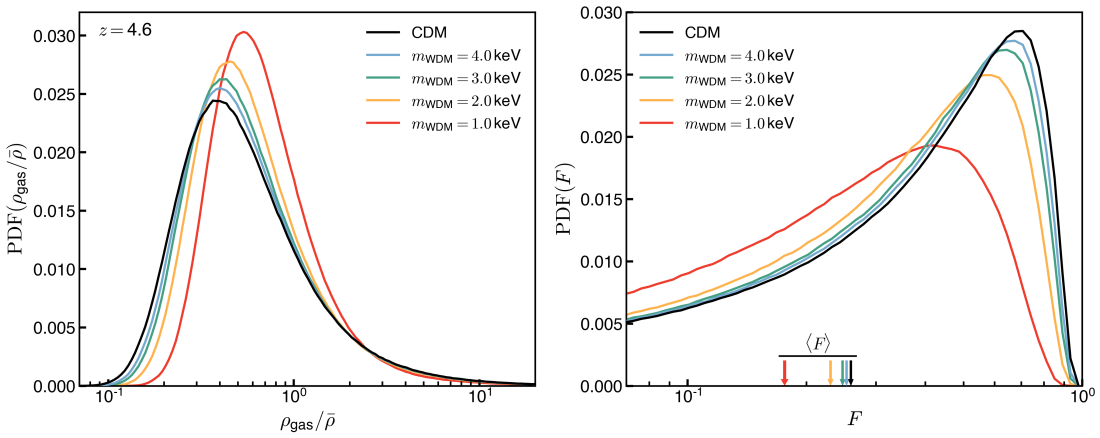

Figure 4.2: *Left:* Volume-weighted probability distribution function of the gas density at $z = 4.6$ in WDM cosmologies compared to ΛCDM. Due to free-streaming, gas that otherwise would collapse into small-scale structures, remains on larger scales. This suppressed collapse shifts the distribution of the gas density to values closer to the mean gas density $\bar{\rho}$. *Right:* The higher density of intergalactic gas for WDM cosmologies results in an increase of the Lyman-α scattering optical depth of the IGM and a decrease of the transmitted flux F. Arrows display the decrease of the mean transmission $\langle F \rangle$ as $m_{\rm WDM}$decreases.

the concentration of overdense regions, an excess of diffuse gas is left to permeate the IGM. Therefore, the volume-weighted distribution of baryonic density shifts to higher values, closer to the gas mean density. Figure 4.2 (left) shows the change in the volume-weighted probability distribution function of the gas density at $z = 4.6$ for different $m_{\rm WDM}$ cosmologies compared to an analogous CDM simulation with fixed thermal history across the simulations.

The enhanced density of the intergalactic gas due to WDM free-streaming increases the opacity of the IGM to Lyman-α transmission. Figure 4.2 (right) shows the probability distribution function of the Lyman-α transmitted flux F along the skewer set for the WDM simulations compared to CDM at $z = 4.6$. As free-streaming increases, the distribution of transmitted flux shifts to lower values, which reflects the increased

opacity of the IGM. The arrows at the bottom of the figure display the decrease of the mean transmission $\langle F \rangle$ of the forest as $m_{\rm WDM}$ decreases.

From our simulations we find that free-streaming from WDM (in the regime $m_{\rm WDM} \geq 1$keV) has almost no effect on the density-temperature distribution of the IGM, as the redshift evolution of the parameters T_0 and γ from the power-law relation $T = T_0(\rho/\overline{\rho})^{\gamma-1}$, are virtually unchanged from the CDM case for fixed photoionization and photoheating history.

4.3.2 WDM Effects on the Flux Power Spectrum

The Lyman-α forest probes the density distribution of intergalactic gas, which itself follows the distribution of dark matter in the IGM. Thereby, the forest is sensitive to the effects of free-streaming from WDM cosmologies on the matter distribution. Suppression of small-scale density fluctuations of gas in the IGM due to free-streaming translates into less small-scale fluctuations in the forest, which results in a decrease of the small-scale flux power spectrum $P(k)$ compared to CDM.

Figure 4.3 shows the dimensionless power spectrum of the Lyman-α transmitted flux $\pi^{-1}kP(k)$ at $z = 4.6$ measured from simulations with different warm dark matter particle mass $m_{\rm WDM}$, compared the CDM. The ratio of $P(k)$ with respect to the CDM simulation is shown in the bottom panel. Free-streaming from the WDM particles affects the power spectrum differently on small and large scales. On small scales ($k \gtrsim 0.02\,{\rm s\,km^{-1}}$) the suppression of density fluctuations decreases the small-scale $P(k)$. On the other hand, the large scale $P(k)$ ($k \lesssim 0.02\,{\rm s\,km^{-1}}$) increases with respect to CDM due to the higher overall normalization of $P(k)$ attributed to the in-

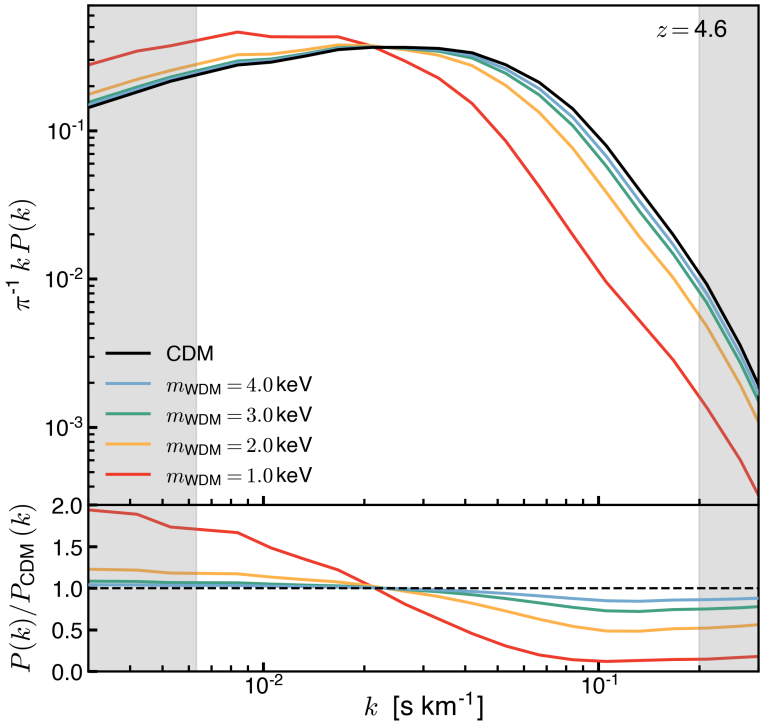

Figure 4.3: One-dimensional transmitted flux power spectrum $P(k)$ at $z = 4.6$ from simulations that vary the warm dark matter particle mass $m_{\rm WDM}$ compared to CDM for fixed thermal history. The bottom panel displays the fractional difference of $P(k)$ from the WDM simulations with respect to CDM. Suppression of small-scale structure due to free-streaming results in a decrease of $P(k)$ compared to CDM for $k \gtrsim 0.02$ s km^{-1}. The increased effective optical depth due to the excess diffuse gas in the IGM results on a higher normalization of $P(k)$ with respect to the CDM measurement. This causes the an increase of large-scale $P(k)$ from the WDM simulations compared to CDM for $k \lesssim 0.02$ s km^{-1}. The gray bands denote the regions that fall outside of the observational measurement from Boera et al. (2019).

crease of the IGM effective optical depth $\tau_{\rm eff}$.

4.3.3 Simulation Grid

To constrain possible WDM cosmologies, we have performed an unprecedented grid of models consisting of 1080 high resolution simulations ($L = 25\ h^{-1}$Mpc and $N = 2{\times}1024^3$) varying the WDM particle mass $m_{\rm WDM}$ as well as the main astrophysical

processes that shape the small-scale density distribution of the gas in the IGM and the properties of the Lyman-α forest.

Besides free-streaming, the main processes that affect the small-scale structure of the Lyman-α forest during $4 \lesssim z \lesssim 5$ are Doppler broadening of the absorption lines due to the instantaneous thermal velocities of the gas, and pressure smoothing of the gas overdensities determined by the integrated past thermal history of the IGM. Our approach for generating models that vary these two processes is to simulate a wide variety of reionization and thermal histories of the IGM by modifying the photoionization and photoheating rates applied in the simulations.

For this work, we use the fiducial model for the hydrogen and helium photoionization and photoheating rates presented in Chapter 3 (V22 hereafter) as a template to generate different IGM reionization and thermal histories. The V22 model is a modification of the rates from Puchwein et al. (2019) and it produces a post-reionization history of the IGM that is overall consistent with observations of the power spectrum of the hydrogen Lyman-α forest as well as the effective optical depth of the He II Lyman-α forest when applied in simulations with a ΛCDM cosmology (see Villasenor et al. (2021b) for more details).

To generate models of the UVB that result in different properties of the Lyman-α forest, we apply three different transformations to the V22 fiducial model.

1. We rescale the hydrogen (H I) and singly-ionized helium (He I) photoionization rate Γ_i by a constant factor β which changes the ionization fraction of the intergalactic hydrogen, thus changing the opacity of the IGM $\tau_{\rm eff}$ and therefore

rescaling the overall normalization of $P(k)$.

2. We modify the hydrogen and singly-ionized helium photoheating rate $\mathcal{H}_i$ by rescaling the average energy of the background radiation $E(z) = \mathcal{H}_i(z)/\Gamma_i(z)$ by a constant factor $\alpha_{\rm E}$. This rescaling changes the temperature of the IGM during and after reionization, which mainly affects the Doppler broadening of the Lyman-α forest absorption lines.

3. We change the time at which hydrogen reionization completes by jointly shifting the redshift dependence of the hydrogen and singly-ionized helium photoionization and photoheating rates by an offset Δz. Changing the epoch of reionization affects the time available for the IGM to cool due to adiabatic expansion. Additionally, changing the timing of ionization impacts the time available for the gas to respond to thermally induced pressure gradients, changing the pressure smoothing of the post-reionization gas distribution. A value $\Delta z > 0$ shifts reionization to an earlier time (higher redshift), while $\Delta z < 0$ delays reionization.

The transformations applied to the H I and He I photoionization and photoheating rates described above are expressed as

$$\begin{aligned} \Gamma_i(z) &\rightarrow \beta\, \Gamma_i^{\rm V22}\,(z - \Delta z), \\ \mathcal{H}_i(z) &\rightarrow \beta\, \alpha_{\rm E}\, \mathcal{H}_i^{\rm V22}(z - \Delta z). \end{aligned} \tag{4.14}$$

For this work, we compare our models to observational determinations of the Lyman-α forest power spectrum in the redshift range $4.2 \leq z \leq 5.0$. For the V22 model, AGN radiation begins to ionize He II and subsequently heat the intergalactic gas at $z \sim$

Table 4.1: WDM-CHIPS Simulation Grid

Parameter	Parameter Values
$m_{\rm WDM}$ [keV]	1, 2, 3, 4, 5, 6, 8, 12, 20, 40, 80, CDM
β	0.6, 0.8, 1.0, 1.2, 1.4, 1.8
α_E	0.1, 0.5, 0.9, 1.3, 1.7
Δz	-0.5, 0.0, 0.5

Note. — The grid of models consist of the 1080 possible combinations of the parameters $\theta = \{m_{\rm WDM}, \beta, \alpha_{\rm E}, \Delta z\}$ Each simulation evolves an $L = 25h^{-1}$Mpc and $N = 1024^3$ box.

4.5, resulting in a local minima of the IGM temperature at this epoch. Nevertheless, the effect of He II reionization reheating the IGM only becomes significant at $z < 4$. Therefore, He II reionization does not significantly impact $P(k)$ for $z \geq 4.2$. For this reason we do not marginalize over He II reionization, and we fix the He II photoionization and photoheating rates to those from the V22 model.

Each model in our simulation suite is characterized by the four-dimensional parameter vector $\theta = \{m_{\rm WDM}, \beta, \alpha_{\rm E}, \Delta z\}$, where $m_{\rm WDM}$ denotes the mass of the WDM particle in keV, and the remaining three parameters vary the IGM reionization history by changing the H I and He I photoionization and photoheating according to Equation (4.14). The values of each of our four model parameters used to generate our model grid are shown in 4.1. All possible combinations of the four parameters listed in 4.1 construct the 1080 models used for this work.

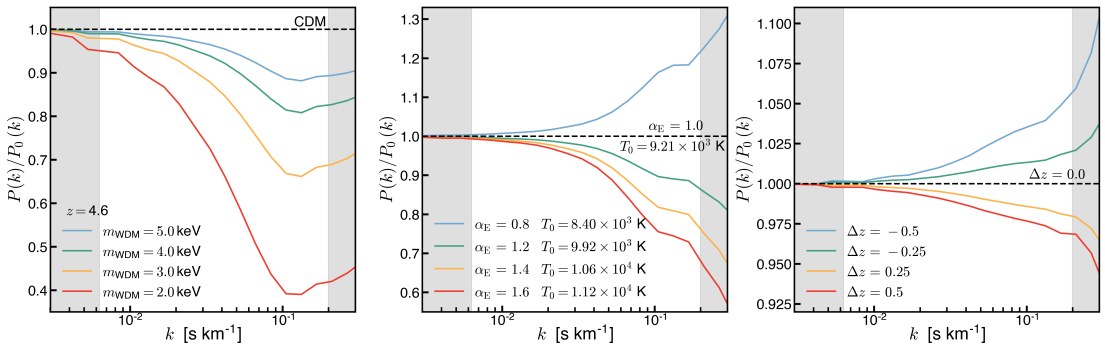

Figure 4.4: Effects of free-streaming (left), Doppler broadening (center) and pressure smoothing (right) on $P(k)$ at $z = 4.6$, compared with the $P_0(k)$ measured from a simulation that adopts the V22 fiducial model (ΛCDM, $\beta = 1$, $\alpha_{\rm E} = 1$, and $\Delta z = 0$). Gray bands show the regions outside the observational measurements from Boera et al. (2019). Contrary to the decrease of $P(k)$ due to Doppler broadening and pressure smoothing, suppression from free-streaming saturates for $k \geq 0.1$ s km^{-1} due to peculiar velocity effects.

4.3.4 Model Variation Effect on the Flux Power Spectrum

Suppression of small-scale fluctuations in the Lyman-α forest can be attributed to decreased small-scale inhomogeneities in the intergalactic gas due to free-streaming from the WDM particle velocities, but also to Doppler broadening of the absorption lines and pressure smoothing of the gas overdensities. Providing an accurate constraint on the WDM particle mass from observations of the flux power spectrum requires differentiating the impact that the three processes have on $P(k)$ and sampling over a dense set of models that simultaneously vary the effect from each mechanism.

Figure 4.4 illustrates the impact on the flux power spectrum due to free-streaming (left), Doppler broadening (center) and pressure smoothing (right) at $z = 4.6$. Shown is the ratio of the $P(k)$ measured on simulations that independently vary the model parameters $m_{\rm WDM}$ (left), $\alpha_{\rm E}$ (center) and Δz (right) to the $P_0(k)$ measured from a simulation that adopts the fiducial V22 model corresponding to a ΛCDM cosmology

and parameters $\beta = 1$, $\alpha_{\rm E} = 1$, and $\Delta z = 0$. To isolate the effects on the small-scale $P(k)$, the optical depth of the skewers from each simulation was rescaled to reflect $\tau_{\rm eff}$ from the fiducial model as this equalizes $P(k)$ on large scales. Additionally, for the simulations that vary Δz (right), we insulate the impact on $P(k)$ due to pressure smoothing by also rescaling the instantaneous temperature of the gas along the skewer set to have the same temperature of the gas at mean density T_0 as the fiducial model.

Variations of the model parameter β mainly impact the ionization fraction of intergalactic hydrogen, consequently changing the IGM effective optical and thereby rescaling the overall normalization of $P(k)$ on the scales relevant for this work. For a detailed analysis on this effect we refer the reader to Appendix B from Villasenor et al. (2021b).

As shown in Figure 4.4 (center), the effect from increasing/decreasing the IGM temperature by increasing/decreasing the photoheating rates is to increase/decrease $P(k)$ for $k \gtrsim 0.01$ s km^{-1}. The $P(k)$ on these scales is similarly impacted by pressure smoothing (right). An earlier reionization ($\Delta z > 0$) suppresses $P(k)$ as overdensities have more time to expand due to the thermally induced pressure gradients, and the opposite effect occurs when reionization is shifted to a later epoch ($\Delta z < 0$).

Figure 4.4 shows that $P(k)$ at $k \gtrsim 0.01$ s km^{-1} can be decreased by the three mechanisms described above, but a significant distinction can be made for the decrease of small-scale $P(k)$ due to WDM cosmologies (left). Contrary to both Doppler broadening (center) and pressure smoothing (right) for which $P(k)$ is increasingly suppressed for increasing k, the suppression of $P(k)$ due to free-streaming with respect to CDM

saturates at $k \geq 0.1\,\mathrm{s\,km^{-1}}$. We argue that this important feature could be key for constraining the nature of dark matter.

Interestingly, we find that if $P(k)$ is measured in real space instead of redshift space (ignoring the Doppler shift of absorption induced by the gas peculiar velocities) then the suppression of $P(k)$ continues to increase to $k \sim 0.2\,\mathrm{s\,km^{-1}}$. Figure 4.5 displays the ratio $P(k)/P_{\mathrm{CDM}}(k)$ for simulations with WDM particle mass $m_{\mathrm{WDM}} = 5\,\mathrm{keV}$ (blue) and $m_{\mathrm{WDM}} = 4\,\mathrm{keV}$ (green). Dashed lines show to the same ratio $P(k)/P_{\mathrm{CDM}}(k)$ but in this case the power spectrum is computed from synthetic spectra calculated real space. Here the effective opacity of the skewers of the WDM simulations has been rescaled to match CDM on large scales.

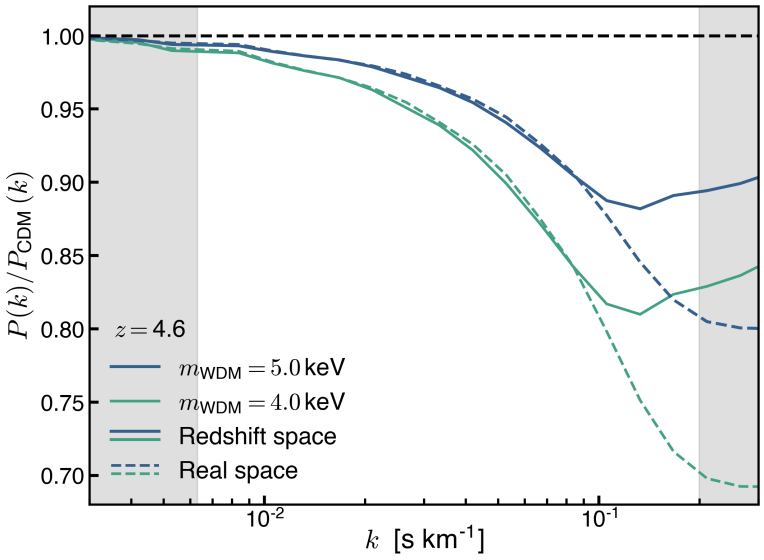

Figure 4.5: Suppression of $P(k)$ due to free-streaming relative to CDM at $z = 4.6$. The full lines show the ratio $P(k)/P_{\mathrm{CDM}}(k)$ where $P(k)$ has been measured from the Lyman-α transmitted flux computed in redshift space. For the dashed lines the transmitted flux from the CDM and WDM case is computed in real space (ignoring peculiar velocities). In real space the suppression of $P(k)$ due to free-streaming continues to increase up to $k \gtrsim 0.2\,\mathrm{s\,km^{-1}}$, while in redshift space the suppression saturates at $k \sim 0.1\,\mathrm{s\,km^{-1}}$.

The comparison presented in Figure 4.5 suggests that peculiar velocities play an important role shaping the structure of the Lyman-α forest. We find that there is an additional suppression of small-scale fluctuations in the forest due to the converging gas flows around density fluctuations. We argue that the Doppler shift of the Lyman-α absorption due to the converging peculiar velocities around an overdensity extend the region of absorption, therefore decreasing small-scale fluctuations in the transmission. We conclude that for $k \gtrsim 0.1\,\mathrm{s\,km^{-1}}$ suppression of $P(k)$ due to peculiar velocities dominates over free-streaming, therefore the decrease in $P(k)$ from WDM relative to CDM saturates at these scales.

4.4 Statistical Comparison

In this section, we present the observations of $P(k)$ employed for this work, and we describe the methodology applied for the Bayesian approach used in this work to constrain the models evolved in our simulation suite.

4.4.1 Observational Flux Power Spectrum

For this work we employ observations of the Lyman-α forest power spectrum from the Keck observatory and the Very Large Telescope, presented in Boera et al. (2019). These measurements represent the highest resolution determination of $P(k)$ to date and probe the structure of the forest in the redshift range $4.2 \leq z \leq 5.0$. Since the impact of free-streaming on $P(k)$ is greater at high redshift, the data set from Boera et al. (2019) is arguably the most constraining determination of $P(k)$ to infer the nature

of dark matter currently available.

The measurements of $P(k)$ from Boera et al. (2019) along our best-fit determination obtained from our analysis is shown in Figure 4.6.

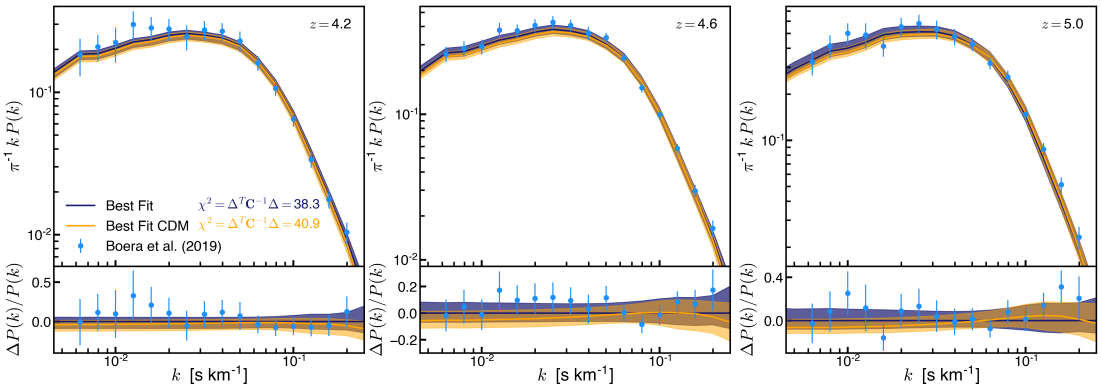

Figure 4.6: Observational determination of $P(k)$ in the redshift range $4.2 \leq z \leq 5.0$ used in this work to constrain WDM cosmologies. The best-fit from the full WDM-CHIPS grid (purple) and the subset CDM-Only grid (orange) are shown as lines and shaded regions which correspond to 95% confidence range of $P(k)$ marginalized over the posterior distribution. From our analysis, $P(k)$ from WDM cosmologies with $m_{\rm WDM} = 4.5^{+50}_{-1.4}$ keV (2σ interval) is preferred over the best-fit CDM model. We note that the WDM best-fit results in lower IGM temperatures compared to the CDM-Only best-fit (see Figure 4.10).

4.4.2 MCMC Inference

To constrain WDM cosmologies from the observations of $P(k)$, we use an MCMC approach to sample over our suite of models and obtain best-fit distributions for our parameters. The likelihood function for the model given by the parameters $\theta = \{m_{\rm WDM}^{-1}, \beta, \alpha_E, \Delta z\}$ is evaluated as

$$\ln \mathcal{L}(\theta) = -\frac{1}{2} \sum_z \left[\mathbf{\Delta}^T \mathbf{C}^{-1} \mathbf{\Delta} + \ln \det(\mathbf{C}) + N \ln 2\pi \right], \tag{4.15}$$

where $\mathbf{\Delta}$ denotes the difference vector between the observational $P(k)$ and the model

$\mathbf{\Delta} = P_{\rm obs}(z,k) - P(z,k|\theta)$, and $\mathbf{C}$ corresponds to the covariance matrix associated with the observation of $P(k)$ as reported by Boera et al. (2019). To compute $P(z,k|\theta)$ for arbitrary values of the parameters θ not directly simulated by our grid, we perform a four-dimensional interpolation of the sixteen neighbouring simulations in parameter space.

The methodology for our inference approach differs from previous works in a number of key respects:

1. Our WDM-CHIPS simulation grid captures a wide range of free-streaming lengths $\lambda_{\rm FS} \propto 1/m_{\rm WDM}$ and UVB models to sample over a variety of cosmic structure formation and gas thermal histories, and thereby produce different statistical properties for the Lyman-α forest. Due to the computational cost of the simulations, evolving large grid of models has not been achieved before. With the advent of efficient numerical codes like Cholla and capable systems like Summit, simulating thousands of models is now possible.

2. We evaluate the performance of a given model in matching the observations of $P(k)$ by marginalizing over the self-consistently evolved hydrogen reionization history in the redshift interval $4.2 \leq z \leq 5.0$. The IGM thermal properties at one redshift cannot be disentangled from its properties at previous epochs, thus the marginalization over the parameter posterior distributions should not be performed independently at each redshift (c.f. Bolton et al. 2014; Nasir et al. 2016a; Hiss et al. 2018; Boera et al. 2019; Walther et al. 2019; Gaikwad et al. 2021)).

3. We do not modify, in post-processing, the mean transmitted flux $\langle F \rangle$ in the forest by rescaling the Lyman-α optical depth, nor do we assume an instantaneous power-law temperature-density relation (cf. Viel et al. (2013a); Iršič et al. (2017a); Boera et al. (2019); Walther et al. (2021)). The latter does not accurately reproduce the thermal state of IGM gas in the range $-1 \leq \log_{10} \Delta \leq 1$. Instead, in our approach simulations self-consistently evolve the ionization and thermal structure of the IGM determined by the wide range of photoionization and photoheating histories applied in our model grid.

4.5 Results

By comparing the high-resolution observation of the flux power spectrum from Boera et al. (2019) to our suite of model that simultaneously vary the impact of free-streaming from WDM cosmologies on the matter distribution and different reionization and thermal histories of the IGM, we infer new constrains on the WDM particle mass $m_{\rm WDM}$. In this section we present the posterior distribution from our MCMC analysis and the marginalized $P(k)$ and thermal evolution of the IGM for our best-fit models. To complement our results, we repeat our analysis modifying $P(k)$ from the simulations to account for a non-uniform UVB and finally we show hypothetical constrains on $m_{\rm WDM}$ from an artificially increased number of quasars.

4.5.1 Distribution of the Model Parameters

The posterior distribution of our four model parameters $\theta = \{m_{\rm WDM}^{-1}, \beta,\, \alpha_E,\, \Delta z\}$ resulting from the Bayesian inference procedure is shown in Figure 4.7. A clear global maximum of the 1D marginalized distributions is shown for $m_{\rm WDM}$ and the parameters β and $\alpha_{\rm E}$ responsible for rescaling the photionization and photoheating rates. The distribution of the parameter Δz is not fully contained by our grid of models, but we argue that this issue does not represent a significant challenge to our conclusions. Values of $\Delta z < -0.5$ would reduce the suppression of small-scale $P(k)$ from pressure smoothing which could be compensated by free-streaming or Doppler broadening, possibly shifting our constrain of $m_{\rm WDM}$ to slightly lower values.

The marginalized distribution for $m_{\rm WDM}^{-1}$ (top left panel, Figure 4.7) is well constrained and it peaks at $m_{\rm WDM}^{-1} = 1/(4.5\ {\rm keV})$. Arguably, the preference for a non-zero free-streaming length is weakly significant as the CDM cosmology is contained within the 3σ interval of the marginalized distribution. The principal result from our analysis is the lower limit $m_{\rm WDM} > 3.0\,{\rm keV}$ (at 2σ level) obtained for the WDM particle mass.

To find the best-fit model to $P(k)$ for a CDM cosmology, we repeat our MCMC approach but restricted to the models that evolve the ΛCDM cosmology varying only the photoionization and photoheating history. Figure 4.8 shows the one- and two-dimensional projections of the posterior distribution for the UVB parameters $\{\beta,\, \alpha_E,\, \Delta z\}$ obtained from the comparison of $P(k)$ from Boera et al. (2019) to the grid of 90 CDM-Only simulations. The posterior distribution of the UVB parameters

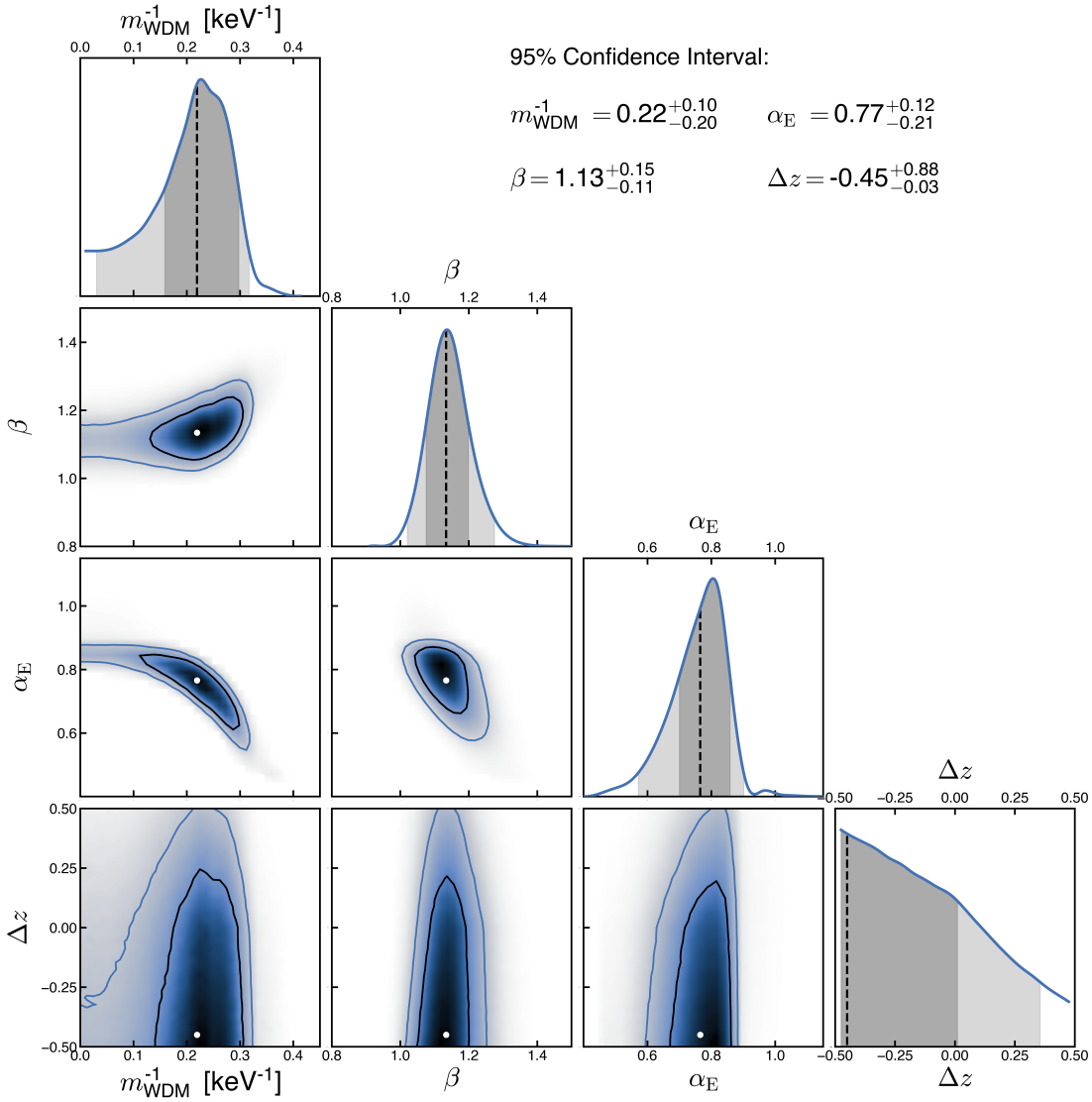

Figure 4.7: Results from the Bayesian inference procedure, showing one- and two-dimensional projections of the posterior distributions for the parameters $\theta = \{m_{\rm WDM}^{-1}, \beta, \alpha_E, \Delta z\}$ recovered by fitting synthetic flux power spectra from our grid of WDM-CHIPS simulations to observations of the Lyman-α forest from Boera et al. (2019). Gray bands in the 1D distributions show the 1σ and 2σ intervals. The marginalized likelihood for $1/m_{\rm WDM} \propto \lambda_{\rm FS}$ peaks at $1/(4.2\,{\rm keV})$. The preference for a non-vanishing free-streaming length is not statistically significant, as the CDM case is contained within the 3σ interval of the distribution. Our main result is a lower bound $m_{\rm WDM} > 3.0$ keV at the 2σ confidence level.

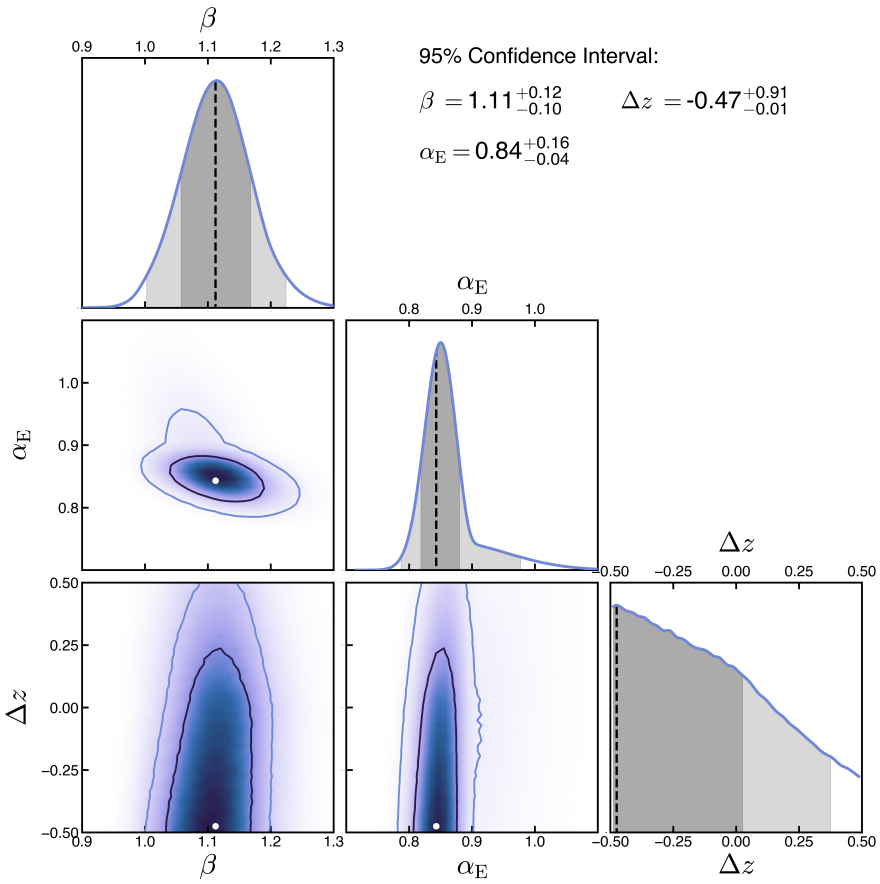

Figure 4.8: Results from our MCMC procedure, showing one- and two-dimensional projections of the posterior distributions for the UVB parameters $\theta = \{\beta, \alpha_E, \Delta z\}$. Here the sampling of models is restricted to the subset of 90 simulations that evolve a CDM cosmology. Gray bands in the 1D distributions show the 1σ and 2σ intervals.

from our CDM-Only fit is slightly different from the distribution of the UVB parameters obtained from the inference that included WDM cosmologies (see Figure 4.7). The main difference between the two distributions is that the parameters β and α_E peak at slightly shifted values. In particular the product $\beta\,\alpha_E$, which rescales the photoheating rates, is higher for the CDM-Only case. Higher photoheating rates directly result in increased IGM temperatures (see section 4.5.3).

4.5.2 Best-Fit Power Spectrum

The marginalized flux power spectrum $P(k)$ over the posterior distribution obtained from our MCMC analysis along the observational data used to constrain the model is shown in Figure 4.6. Lines and shaded regions show the best-fit $P(k)$ and the 95% confidence range. The result from the the fit performed by sampling over the entire WDM model grid (purple) and restricting to the CDM-Only grid (orange) are shown separately. As shown, both results provide a good match to the observed $P(k)$ nevertheless, the relative difference between the model and the data quantified as $\chi^2 = \sum_z \Delta^T \mathbf{C}^{-1} \Delta$ is slightly lower for the best-fit WDM model ($\chi^2 = 38.3$), compared to $\chi^2 = 40.9$ from the CDM-Only best-fit model. Here $\mathbf{C}$ denotes the covariance matrix taken from Boera et al. (2019) and Δ is the difference vector between the observed and model $P(k)$.

We find that the preference for a non-zero free-streaming scale derived from our analysis is driven by the smallest scales probed by the high-resolution measurement of $P(k)$ from Boera et al. (2019). We repeated our inference methodology but excluding the three last data points ($k > 0.1$ s km^{-1}) from the the likelihood calculation. Here we found that excluding the high-k measurements places ΛCDM as the preferred cosmology. The marginalized posterior distribution for $m_{\rm WDM}$ contrasting the fit limited to $k \leq 0.1$ s km^{-1} and the fit to full data set is displayed in Figure 4.9. As shown, the likelihood peaks at CDM and the lower limit constraint at the 2σ level shifts to $m_{\rm WDM} > 3.6$ keV.

In Section 4.3.4 we showed that free-streaming and Doppler broadening impact $P(k)$ differently on small-scales ($k > 0.1$ s km^{-1}). As the temperature of the gas is

increased/decreased, the impact of Doppler broadening is to decrease/increase the small-scale $P(k)$ and the effect monotonically increases with increasing k. On the other hand, free-streaming decreases small-scale density fluctuations but its effect on the flux power spectrum saturates at $k \gtrsim 0.1\,\mathrm{s\,km^{-1}}$ where suppression of $P(k)$ stops increasing with k as the impact from peculiar velocities become dominant on these scales (see Figs. 4.4, 4.5).

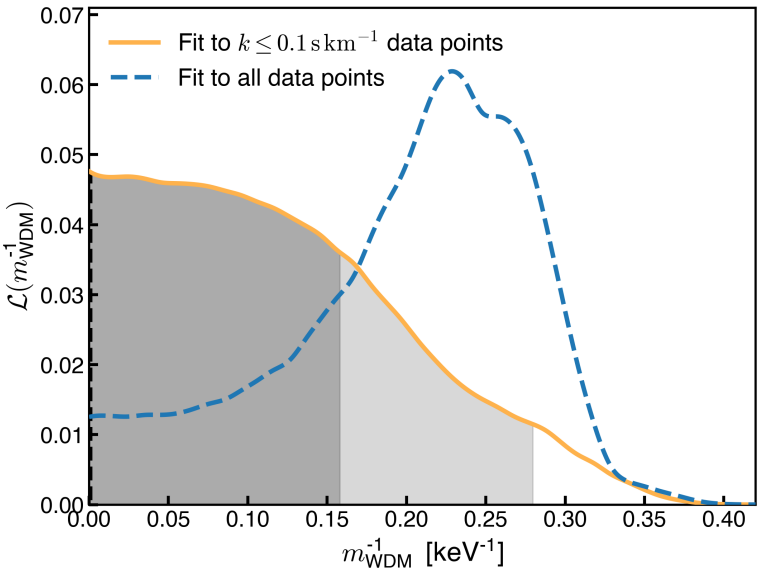

Figure 4.9: One-dimensional posterior likelihood for the WDM particle mass obtained by limiting the observational measurements of $P(k)$ Boera et al. (2019) to $k \lesssim 0.1\,\mathrm{s\,km^{-1}}$ compared to the likelihood distribution from fitting to the complete data set (dashed blue). The limited data set for the observational $P(k)$ is constructed by excluding the last three points of each redshift snapshot. Gray bands show the 1σ and 2σ interval from the fit to the limited data. In this case ΛCDM is the preferred cosmology and the lower limit at 95% confidence shifts to $m_{\mathrm{WDM}} > 3.6\,\mathrm{keV}$ compared to $m_{\mathrm{WDM}} > 3.0\,\mathrm{keV}$ from the fit to the full data set.

We have shown that the preference for a WDM cosmology over ΛCDM originates from the $k > 0.1\,\mathrm{s\,km^{-1}}$ measurement of $P(k)$. Also, we showed that on these scales ($k > 0.1\,\mathrm{s\,km^{-1}}$) the suppression of $P(k)$ due to free-streaming saturates while

the suppression of $P(k)$ from Doppler broadening increases monotonically. Therefore, we conclude that the saturated suppression of $P(k)$ due to a $m_{\rm WDM} \sim 4.5\,{\rm keV}$ combined with a moderately lower IGM temperature provides a slightly better fit to the $k > 0.1\,{\rm s\,km^{-1}}$ observation of $P(k)$ compared to the ΛCDM best-fit model which results in higher IGM temperatures at $4 \leq z \leq 5$ (see Fig. 4.10).

4.5.3 Thermal Evolution of the IGM

One of the advantages of our approach compared to previous works that aim to constrain the WDM free-streaming from observations of the Lyman-α forest power spectrum is that our simulations self-consistently evolve the IGM thermal and ionization history during and after reionization by sampling over a large grid models that vary the UVB photoionization and photoheating rates. Instead, the methodology adopted by previous works Viel et al. (2013a); Iršič et al. (2017a); Garzilli et al. (2019, 2021) was to change the temperature T_0 and the ionization fraction of the IGM in post-processing by rescaling T_0 and the effective optical depth $\tau_{\rm eff}$ of the simulated skewers.

Figure 4.10 shows the redshift evolution of the temperature of the gas at mean density T_0 marginalized over the posterior distribution from our MCMC analysis; shaded bars show the 1σ interval. The best-fit model for WDM cosmologies with $m_{\rm WDM}\ sim 4.5\,keV$ results in moderately lower (5 - 10%) IGM temperatures (purple) due to the slightly reduced photoheating rates compared to the ΛCDM best-fit model (orange). Lower temperatures are expected from the WDM models as they decrease the impact of Doppler broadening in suppressing $P(k)$ which compensates the suppression due to free-streaming.

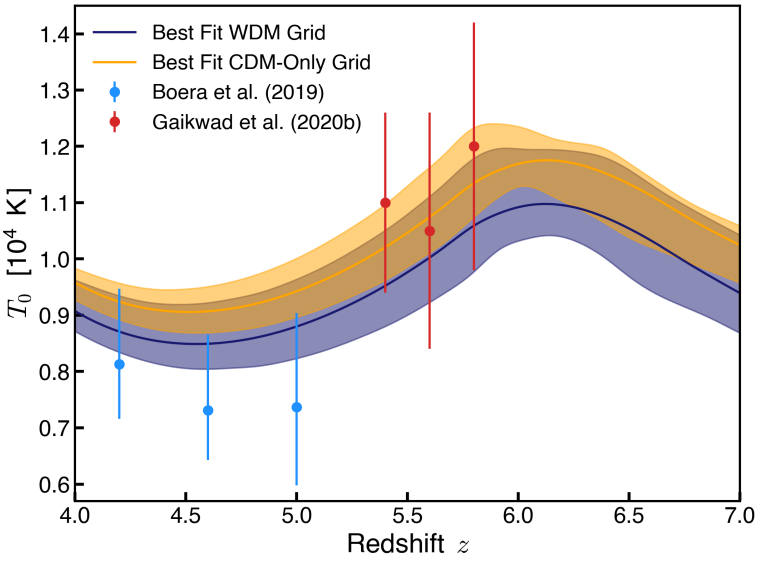

Figure 4.10: Redshift evolution of the temperature T_0 of IGM gas at the mean density from the best-fit ΛCDM and WDM models (solid lines) and 1σ interval (colored bands) obtained from our MCMC analysis. The data points show the values of T_0 inferred from observations of the Lyman-α forest by Boera et al. (2019) and Gaikwad et al. (2020). The best-fit from WDM cosmologies have moderately lower temperatures compared to CDM, which compensates for the suppression of small-scale $P(k)$ due to free-streaming.

In Figure 4.10, we compare the IGM temperature T_0 inferred by this work to other measurements at $z > 4$ Boera et al. (2019); Gaikwad et al. (2020). The high-redshift inference presented by Gaikwad et al. (2020) was obtained by characterising the transmission spikes observed in $z > 5$ spectra. The thermal histories obtained from our analysis are consistent with the results from Gaikwad et al. (2020), suggesting a peak in T_0 due to hydrogen reonization at $z \sim 6$.

The measurement of T_0 presented in Boera et al. (2019) for $4.2 \leq z \leq 5.0$ was made by fitting the observed $P(k)$ to simulated spectra according to a ΛCDM cosmology where the instantaneous density-temperature distribution of the gas is modified in post-processing changing parameters T_0 and γ from the power-law relation $T = T_0(\rho/\bar{\rho})^{\gamma-1}$.

Despite using the same observational determination of $P(k)$ for our inference, the IGM temperatures obtained in this work during $4.2 \leq z \leq 5.0$ are slightly higher than those inferred by Boera et al. (2019). Restricting to a ΛCDM cosmology, we find a best-fit T_0 is 15 - 20% higher at redshift $z = 5.0$ and $z = 4.6$ compared to the result from Boera et al. (2019). At redshift $z = 4.2$ the difference lowers to 10 - 15% and our result agrees with their inference within 1σ. The evolution of T_0 from our best-fit WDM model is also higher (5 - 10%) than the results from Boera et al. (2019), but in this case our results also agree to within 1σ.

4.5.4 Constraining WDM with Increased Quasar Sightlines

Upcoming surveys of the Lyman-α forest (e.g., DESI DESI Collaboration et al. (2016a), WAEVE Pieri et al. (2016), EUCLID Amiaux et al. (2012), LSST) will drastically increase the available observations of quasar sightlines, which will significantly improve the statistics of measurements derived from the forest. The improved statistics will allow tighter constraints on cosmological parameters as well as on WDM and the sum of the neutrino masses from their role in suppressing small-scale structure in the forest.

To assess how better statistics would impact the constraining power of high-redshift and high-resolution observations of the Lyman-α forest for WDM cosmologies, we repeat our analysis but decreasing the uncertainty on the observed $P(k)$ by a factor of one half, which would correspond to an increasing the number of observed quasar spectra from fifteen used by Boera et al. (2019) for their measurement of $P(k)$, to about sixty (a factor of four increase).

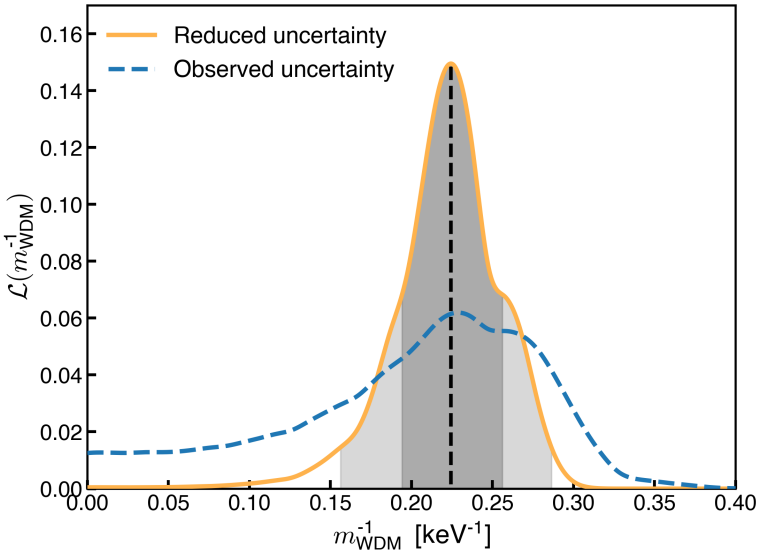

Figure 4.11: One-dimensional posterior likelihood for the WDM particle mass obtained by fitting to the measurements of $P(k)$ from Boera et al. (2019) but with an artificially reduced uncertainty motivated by the improved statistics of Lyman-α spectra from upcoming surveys. Here we rescale the covariance matrix $\mathbf{C}$ of $P(k)$ by a factor of one fourth. In this hypothetical case, the constraint on $m_{\rm WDM}$is tighter, measured as $m_{\rm WDM} = 4.5^{+1.9}_{-1.0}\,$keV at 95% confidence level. The dashed line shows our result from fitting to $P(k)$ with the reported uncertainty from Boera et al. (2019).

In our approach, we approximate a more constraining dataset by rescaling the covariance matrix $\mathbf{C}$ of $P(k)$ by a factor of one fourth in Eq. 4.15. Note that the observational measurement of $P(k)$ is not altered, only the covariance matrix is reduced. Figure 4.11 shows the marginalized likelihood for ${m_{\rm WDM}}^{-1}$ obtained from our analysis using the reduced uncertainty. In this hypothetical case, the improved statistics of the $P(k)$ measurement provide a tighter constrain for the WDM particle mass measured as $m_{\rm WDM} = 4.5^{+1.9}_{-1.0}\,$keV at the 95% confidence level. This exercise demonstrates that increasing the sample of high-z and high-resolution observations of Lyman-α forest would place tight constrains on WDM cosmological models.

4.5.5 Modified $P(k)$ for Inhomogeneous UVB

For this work, the simulations in our WDM-CHIPS grid evolve under a homogeneous UVB in the form of uniform photoionization and photoheating rates. This represents the main limitation of our analysis. We evaluate the effect of this assumption on our conclusions by repeating our analysis but using the modification to $P(k)$ prescribed by Molaro et al. (2022) to account for the impact of a nonuniform UVB on the simulated $P(k)$.

In Molaro et al. (2022), the authors compare a set of simulations that apply a uniform UVB to a set of simulations that follow an hybrid-RT method. For the hybrid-RT approach, spatially-varying maps for the H I photoionization and photoheating rates are computed for the simulated boxes in post-processing and are used as input for a re-run of the base simulation, incorporating the response of the gas to the nonuniform photoionization and photoheating. For the comparison, pairs of uniform and nonuniform UVB simulations are calibrated to have the same average ionization and thermal history.

From their comparison, Molaro et al. (2022) concluded that simulations with a patchy reionization show a suppression (10 - 15%) of $P(k)$ on small scales ($k \sim 0.1\,\mathrm{s\,km^{-1}}$) with respect to the uniform UVB case. This effect is mainly driven by Doppler broadening associated with the high temperatures of the recently ionized regions and the divergent peculiar velocities due to the thermally pressurized gas. On large scales ($k > 0.03\,\mathrm{s\,km^{-1}}$) the variation of the IGM neutral fraction due to large-scale fluctuations of the gas temperature lead to an increase on $P(k)$.

The likelihood distribution from our MCMC inference where we modify the

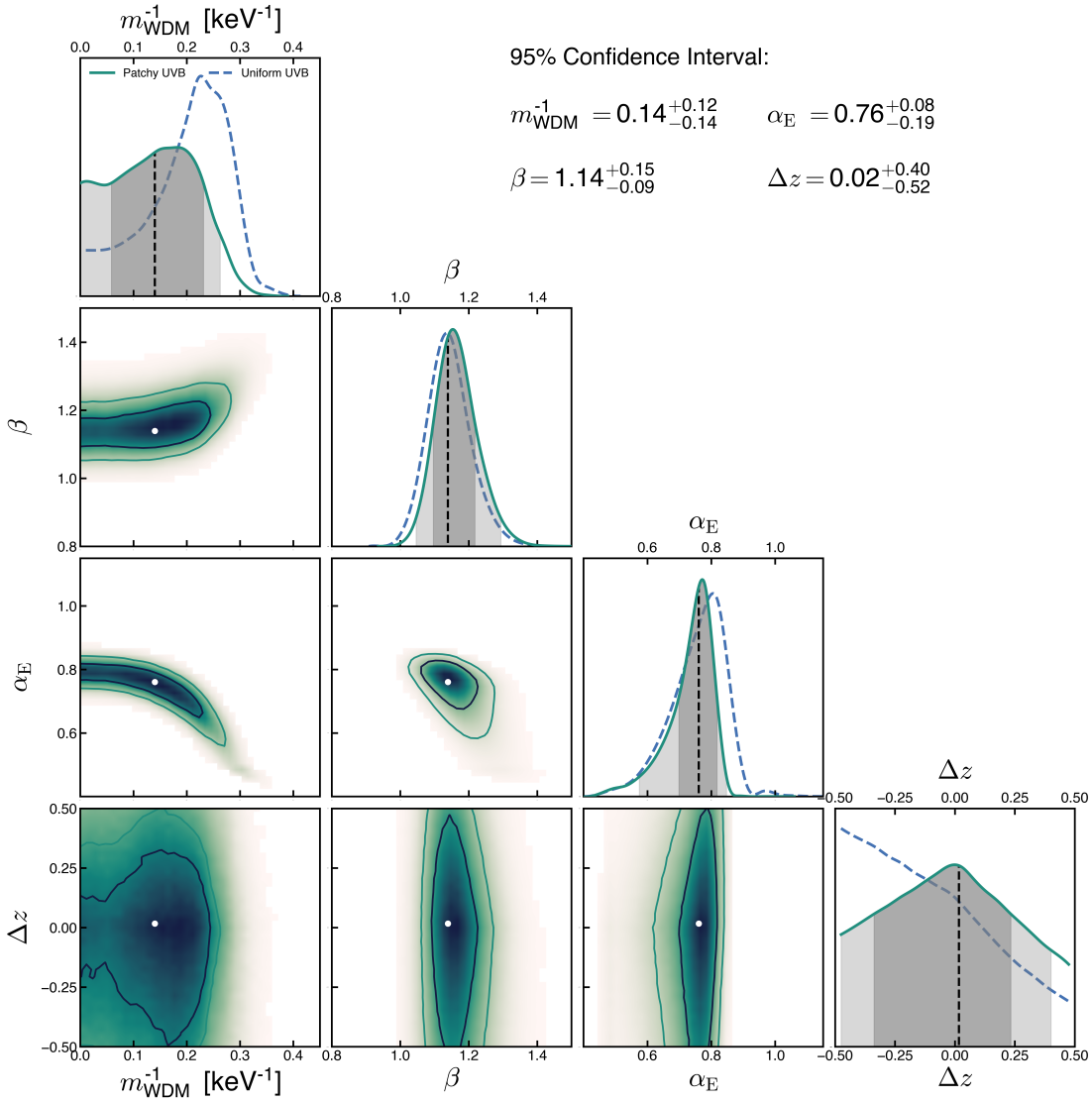

Figure 4.12: Posterior distribution of the parameters $\theta = \{m^{-1}_{\rm WDM}, \beta, \alpha_E, \Delta z\}$ from fitting the observed $P(k)$ from Boera et al. (2019) with models that account for a patchy reionization according to the modification to $P(k)$ presented in Molaro et al. (2022). In this case, the preference for a WDM cosmology persists but the distribution shifts to higher mass. The likelihood peaks at $m_{\rm WDM} = 7.1\,{\rm keV}$ and the lower limit at the 95% levels is $m_{\rm WDM} = 3.8\,{\rm keV}$. In this case the ΛCDM cosmology is contained within the 95% interval of the distribution. The dashed blue lines show the one-dimensional likelihood distributions obtained from fitting $P(k)$ measured directly from the uniform UVB simulations.

$P(k)$ from each one of the simulations in our WDM-CHIPS grid to account for a inhomogeneous reionization according to the transformation presented in Molaro et al. (2022) is shown in Figure 4.12. Dashed blue lines correspond to the one-dimensional marginalized distributions obtained from fitting the base $P(k)$ from the simulations with a uniform reionization.

Figure 4.12 shows that modifying the models $P(k)$ to account for a patchy reionization mainly affects the likelihood of $m_{\rm WDM}$. While WDM cosmologies are still preferred over ΛCDM when fitting the nonuniform UVB models, the likelihood shifts to higher values of $m_{\rm WDM}$ compared to the result from uniform UVB models. An inclination for models with reduced free-streaming compensates for the reduction of small-scale $P(k)$ due to patchy reionization. In this case the maximum likelihood occurs at $m_{\rm WDM} = 7.1\,{\rm keV}$ and the lower limit at 95% level is at $m_{\rm WDM} = 3.8\,{\rm keV}$. Notably, for the patchy UVB models, the ΛCDM cosmology is contained within the 95% interval of the likelihood distribution.

We note that the difference between the observed and model $P(k)$ defined as $\chi^2 = \sum_z \mathbf{\Delta}^T \mathbf{C}^{-1} \mathbf{\Delta}$ (see §4.4.2 is larger for the nonuniform UVB modified best-fit models compared to the best-fit models sampling from the uniform UVB simulated $P(k)$. When fitting with $P(k)$ measured directly from the uniform UVB simulations, we measure $\chi^2 = 38.3$ for the best-fit model with $m_{\rm WDM} = 4.5\,{\rm keV}$ and a slightly higher $\chi^2 = 40.9$ for the ΛCDM best-fit model. On the other hand, when sampling over the modified $P(k)$ to account for patchy reionization, we measure larger values $\chi^2 = 46.4$ and $\chi^2 = 46.6$ for the WDM and CDM-Only best-fit models, respectively.

The higher values of χ^2 obtained for the patchy reionization modified models show that, in the context of this work, $P(k)$ measured directly from the uniform UVB simulations provide a better match to the observation from Boera et al. (2019), therefore we use the likelihood distribution obtained from the sampling over the uniform UVB models tho construct the main results from our analysis.

4.6 Conclusions

With the objective of constraining cosmological models where dark matter free-streaming smooths the matter distribution in the Universe, we have used the GPU-native code Cholla to perform a massive suite of high-resolution hydrodynamical cosmological simulations that simultaneously vary the effect of free-streaming from WDM particles and the IGM reionization history. We compare the power spectrum of the synthetic Lyman-α forest from our simulations to the high-resolution observational measurement presented by Boera et al. (2019) to determine via a likelihood analysis the optimal model for cosmological free-streaming that best matches the observation. A summary of the efforts and conclusions from this work follows.

- We present the WDM-CHIPS suite consisting of a grid of 1080 high-resolution simulations ($L = 25h^{-1}$Mpc, $N = 1024^3$) that vary the free-streaming from WDM cosmologies and the photoionization and photoheating rates from the metagalactic UVB. The UVB rates applied for our grid use the model from Villasenor et al. (2021b) as a template, and use three parameters that control a rescaling amplitude and redshift-timing of the hydrogen photoionization and photoheating

rates. Combined with the WDM particle mass $m_{\rm WDM}$, our four-dimensional grid of models densely sample a wide range of self-consistently evolved reionization and thermal histories of the IGM. This represents a significant improvement over previous studies that aimed to constrain WDM cosmologies from observations of the Lyman-α forest Viel et al. (2005, 2013a); Iršič et al. (2017a); Garzilli et al. (2019, 2021).

- The large range of thermal histories produced by the different UVB models in our grid of simulations results in synthetic measurements of Lyman-α spectra where the impact from Doppler broadening and pressure smoothing on suppressing the small-scale $P(k)$ varies widely. This flexibility is important as these mechanisms have similar effects as free-streaming decreasing small-scale fluctuations in the Lyman-α forest. Additionally, our approach does not require an assumption of a power-law relation for the density-temperature distribution of the gas. Self-consistent evolution of the IGM phase structure proves to be important as we find that a single power law does not accurately describe the $\rho_{\rm gas} - T$ distribution in the density range relevant to generating the signal of the Lyman-α forest (see Appendix E of Villasenor et al. (2021b)).

- We compare our grid of models to the high-redshift ($4.2 \leq z \leq 5.0$) observational measurement of the Lyman-α forest power spectrum from Boera et al. (2019). The observations presented in Boera et al. (2019) provide the highest resolution measurement of $P(k)$ at the time that work was executed. We perform a Bayesian MCMC sampling to determine the best-fit model for free-streaming due to WDM

cosmologies and the IGM photoionization and photoheating history.

- From our MCMC analysis, we find that a cosmological model with non-vanishing free-streaming is preferred over ΛCDM. The best-fit and 95% interval for the WDM particle mass from our result is $m_{\rm WDM} = 4.5^{+50}_{-1.4}$ keV. We find a weak (3σ) preference for WDM over ΛCDM, but both are statistically consistent with the currently available data.

- We repeat our MCMC analysis, restricting to *Lambda*CDM only but with a variable UVB model. We find that best-fit ΛCDM model mainly differs from the WDM optimal model in that the IGM temperatures are slightly lower (5 - 10%) for the WDM case. We find that both thermal histories are in good agreement with the inference from Gaikwad et al. (2020) at $5.4 \lesssim z \lesssim 5.8$ and moderately higher than the temperatures from Boera et al. (2019) at $4.2 \leq z \leq 5.0$.

- We introduce the effect of a patchy reionization in our models by modifying $P(k)$ from our uniform UVB simulations according to the prescription presented in Molaro et al. (2022). The likelihood distribution from our MCMC approach using the modified $P(k)$ shows that the preference for a WDM cosmology is maintained but the distribution shifts to higher values of $m_{\rm WDM}$. The maximum likelihood and the 95% lower limit constraint are $m_{\rm WDM} = 7.1\,{\rm keV}$ and $m_{\rm WDM} = 3.8\,{\rm keV}$, respectively. Additionally, the ΛCDM case is 2σ consistent with the best-fit WDM model when accounting for an inhomogeneous UVB. Notably, the difference χ^2 between the observed and best-fit $P(k)$ model is higher for the modified model to account for a nonuniform UVB, and we therefore use the likelihood distribution

obtained from the uniform UVB models as basis for our results.

Chapter 5

Conclusions

This thesis focused on modeling the properties of the intergalactic medium (IGM) from simulations performed with the hydrodynamics code *Cholla*, which was extended for this purpose as part of this work. Using *Cholla* on Summit (Oak Ridge National Laboratory), we were able to run over 1500 high-resolution cosmological simulations varying the physical models that shape the properties of the IGM, and hence, the structure of the Lyman-α forest. This massive suite of simulations allowed for a study of the properties of the IGM with unprecedented detail.

Chapter 2 describes the methodology applied to extend *Cholla* for cosmological simulations. Additionally, I present our results from comparing the statistical properties of the Lyman-α forest in simulations that used current models for the UV background radiation to the statistics of the observed spectra. Our results indicate that those models fail to reproduce the observations.

Chapter 3 presents our approach to improve the current models of the UV

background. We run a suite of 400 simulations where each simulation evolved a different UVB model. From a Bayesian analysis comparing the simulations to the observed properties of the Lyman-α forest, we find the best-fit model, which in general results in an evolution of the Lyman-α forest power spectrum that is consistent with the observations over a large epoch. From the best-fit model, we infer the thermal and ionization history of the IGM, which is consistent with previous independent determinations.

Chapter 4 introduces an extension of our simulation suite where we include the effects of a possible WDM candidate. Here we run a suite of 1080 simulations that simultaneously vary the mass of the warm dark matter particle and the UVB photoionization and photoheating rates which impact the thermal and ionization history of the IGM. We compare high-resolution observations of the power spectrum of the forest to the simulations. From our Bayesian inference, we conclude that a cosmology with a WDM mass of $\sim 4.5\,\mathrm{keV}$ formally provides a better fit to the power spectrum compared to the ΛCDM cosmology. This interesting result motivates the continued observation of high-resolution, high-redshift spectra to improve the statistics and provide a better constrain on possible WDM cosmologies.

Currently, additional physics is being developed for *Cholla*. A full radiative transfer implementation combined with a machine learning-based method for populating dark matter halos with galaxies will allow for full simulations of the reionization epoch. Combined with spectra from the DESI experiment and observation from JWST, these simulations will be able to provide constraints on cosmological models and the population distribution of the early galaxies responsible for reionization.

Bibliography

Almgren, A. S., Bell, J. B., Lijewski, M. J., Lukić, Z., & Van Andel, E. 2013, ApJ, 765, 39

Amiaux, J., Scaramella, R., Mellier, Y., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. Oschmann, Jacobus M., 84420Z

Arinyo-i-Prats, A., Miralda-Escudé, J., Viel, M., & Cen, R. 2015, Journal of Cosmology and Astro-Particle Physics, 2015, 017

Baur, J., Palanque-Delabrouille, N., Yèche, C., Magneville, C., & Viel, M. 2016, J. Cosmology Astropart. Phys., 2016, 012

Becker, G. D., & Bolton, J. S. 2013, MNRAS, 436, 1023

Becker, G. D., D'Aloisio, A., Christenson, H. M., et al. 2021, arXiv e-prints, arXiv:2103.16610

Becker, G. D., Hewett, P. C., Worseck, G., & Prochaska, J. X. 2013a, MNRAS, 430, 2067

—. 2013b, MNRAS, 430, 2067

Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013, ApJ, 762, 109

Bird, S., Rogers, K. K., Peiris, H. V., et al. 2019, J. Cosmology Astropart. Phys., 2019, 050

Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93

Boera, E., Becker, G. D., Bolton, J. S., & Nasir, F. 2019, ApJ, 872, 101

Bolton, J. S., Becker, G. D., Haehnelt, M. G., & Viel, M. 2014, MNRAS, 438, 2499

Bolton, J. S., Puchwein, E., Sijacki, D., et al. 2017, MNRAS, 464, 897

Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nature, 380, 603

Bosman, S. E. I., Fan, X., Jiang, L., et al. 2018, MNRAS, 479, 1055

Bryan, G. L., Norman, M. L., Stone, J. M., Cen, R., & Ostriker, J. P. 1995, Computer Physics Communications, 89, 149

Bryan, G. L., Norman, M. L., O'Shea, B. W., et al. 2014, The Astrophysical Journal Supplement Series, 211, 19

Bullock, J. S., & Boylan-Kolchin, M. 2017, ARA&A, 55, 343

Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, MNRAS, 412, 2543

Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, ApJ, 437, L9

Chabanier, S., Bournaud, F., Dubois, Y., et al. 2020, MNRAS, 495, 1825

Chabanier, S., Millea, M., & Palanque-Delabrouille, N. 2019a, MNRAS, 489, 2247

Chabanier, S., Palanque-Delabrouille, N., Yèche, C., et al. 2019b, J. Cosmology Astropart. Phys., 2019, 017

—. 2019c, J. Cosmology Astropart. Phys., 2019, 017

Colella, P., & Woodward, P. R. 1984, Journal of Computational Physics, 54, 174

Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998a, ApJ, 495, 44

—. 1998b, ApJ, 495, 44

D'Aloisio, A., McQuinn, M., Davies, F. B., & Furlanetto, S. R. 2018, Monthly Notices of the Royal Astronomical Society, 473, 560

D'Aloisio, A., McQuinn, M., Maupin, O., et al. 2019, Astrophysical Journal, 874, 154

D'Aloisio, A., Upton Sanderbeck, P. R., McQuinn, M., Trac, H., & Shapiro, P. R. 2017, Monthly Notices of the Royal Astronomical Society, 468, 4691

Davies, F. B., Hennawi, J. F., Eilers, A.-C., & Lukić, Z. 2018a, Astrophysical Journal, 855, 106

Davies, F. B., Hennawi, J. F., Bañados, E., et al. 2018b, Astrophysical Journal, 864, 142

de Belsunce, R., Gratton, S., Coulton, W., & Efstathiou, G. 2021, MNRAS, 507, 1072

DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv e-prints, arXiv:1611.00036

—. 2016b, arXiv e-prints, arXiv:1611.00037

Doussot, A., Trac, H., & Cen, R. 2019, Astrophysical Journal, 870, 18

Eilers, A.-C., Davies, F. B., & Hennawi, J. F. 2018, ApJ, 864, 53

Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017, ApJ, 840, 24

Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117

Faucher-Giguère, C.-A. 2020, MNRAS, 493, 1614

Feng, J. L. 2010, ARA&A, 48, 495

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, Rev. Mexicana Astron. Astrofis., 53, 385

Gaikwad, P., Khaire, V., Choudhury, T. R., & Srianand, R. 2017, MNRAS, 466, 838

Gaikwad, P., Srianand, R., Haehnelt, M. G., & Choudhury, T. R. 2021, MNRAS, 506, 4389

Gaikwad, P., Rauch, M., Haehnelt, M. G., et al. 2020, MNRAS, 494, 5091

Gallego, S. G., Cantalupo, S., Sarpas, S., et al. 2021, MNRAS, 504, 16

Garaldi, E., Gnedin, N. Y., & Madau, P. 2019, Astrophysical Journal, 876, 31

Garzilli, A., Magalich, A., Ruchayskiy, O., & Boyarsky, A. 2021, MNRAS, 502, 2356

Garzilli, A., Magalich, A., Theuns, T., et al. 2019, MNRAS, 489, 3456

Gnedin, N. Y. 2016, Astrophysical Journal, 821, 50

Gnedin, N. Y., Becker, G. D., & Fan, X. 2017, ApJ, 841, 26

Gnedin, N. Y., & Hui, L. 1998, MNRAS, 296, 44

Gnedin, N. Y., & Kaurov, A. A. 2014, Astrophysical Journal, 793, 30

Greig, B., Mesinger, A., & Bañados, E. 2019, MNRAS, 484, 5094

Greig, B., Mesinger, A., Haiman, Z., & Simcoe, R. A. 2017, MNRAS, 466, 4239

Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633

Haardt, F., & Madau, P. 2012, ApJ, 746, 125

Hahn, O., & Abel, T. 2011a, MNRAS, 415, 2101

—. 2011b, MNRAS, 415, 2101

Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJ, 457, L51

Hiss, H., Walther, M., Hennawi, J. F., et al. 2018, The Astrophysical Journal, 865, 42

Ho, M.-F., Bird, S., & Shelton, C. R. 2022, MNRAS, 509, 2551

Hoag, A., Bradač, M., Huang, K., et al. 2019, ApJ, 878, 12

Hockney, R. W., & Eastwood, J. W. 1988, Computer simulation using particles

Hui, L., & Gnedin, N. Y. 1997, MNRAS, 292, 27

Hui, L., Gnedin, N. Y., & Zhang, Y. 1997, ApJ, 486, 599

Iršič, V., Viel, M., Haehnelt, M. G., et al. 2017a, Phys. Rev. D, 96, 023522

Iršič, V., Viel, M., Berg, T. A. M., et al. 2017b, MNRAS, 466, 4332

Jung, I., Finkelstein, S. L., Dickinson, M., et al. 2020, ApJ, 904, 144

Kaurov, A. A., & Gnedin, N. Y. 2014, Astrophysical Journal, 787, 146

—. 2015, Astrophysical Journal, 810, 154

Keating, L. C., Puchwein, E., & Haehnelt, M. G. 2018, MNRAS, 477, 5501

Keating, L. C., Weinberger, L. H., Kulkarni, G., et al. 2020, MNRAS, 491, 1736

Khaire, V., Walther, M., Hennawi, J. F., et al. 2019, Monthly Notices of the Royal Astronomical Society, 486, 769

Krolewski, A., Lee, K.-G., White, M., et al. 2018, Astrophysical Journal, 861, 60

Kulkarni, G., Hennawi, J. F., Oñorbe, J., Rorai, A., & Springel, V. 2015a, Astrophysical Journal, 812, 30

—. 2015b, ApJ, 812, 30

Kulkarni, G., Keating, L. C., Haehnelt, M. G., et al. 2019, MNRAS, 485, L24

La Plante, P., Trac, H., Croft, R., & Cen, R. 2017, ApJ, 841, 87

—. 2018, ApJ, 868, 106

Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, MNRAS, 446, 3697

Macciò, A. V., Ruchayskiy, O., Boyarsky, A., & Muñoz-Cuartas, J. C. 2013, MNRAS, 428, 882

Madau, P., & Haardt, F. 2015, ApJ, 813, L8

Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648

Martel, H., & Shapiro, P. R. 1998, MNRAS, 297, 467

Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, ApJ, 856, 2

Mason, C. A., Fontana, A., Treu, T., et al. 2019, MNRAS, 485, 3947

McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, MNRAS, 447, 499

McGreer, I. D., Mesinger, A., & Fan, X. 2011, MNRAS, 415, 3237

McQuinn, M. 2016, ARA&A, 54, 313

Meiksin, A. A. 2009a, Reviews of Modern Physics, 81, 1405

—. 2009b, Reviews of Modern Physics, 81, 1405

Miniati, F., & Colella, P. 2007, Journal of Computational Physics, 227, 400

Miralda-Escudé, J., & Rees, M. J. 1994, MNRAS, 266, 343

Molaro, M., Iršič, V., Bolton, J. S., et al. 2022, MNRAS, 509, 6119

Nasir, F., Bolton, J. S., & Becker, G. D. 2016a, MNRAS, 463, 2335

—. 2016b, MNRAS, 463, 2335

Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2017a, ApJ, 837, 106

Oñorbe, J., Hennawi, J. F., Lukić, Z., & Walther, M. 2017b, ApJ, 847, 63

Osterbrock, D. E. 1989, Astrophysics of gaseous nebulae and active galactic nuclei

Peeples, M. S., Weinberg, D. H., Davé, R., Fardal, M. A., & Katz, N. 2010, MNRAS,

404, 1281

Pieri, M. M., Bonoli, S., Chaves-Montero, J., et al. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé, J. Richard, L. Cambrésy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin, 259–266

Pippig, M. 2013, SIAM Journal on Scientific Computing, 35, C213

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Puchwein, E., Bolton, J. S., Haehnelt, M. G., et al. 2015, MNRAS, 450, 4081

Puchwein, E., Haardt, F., Haehnelt, M. G., & Madau, P. 2019, MNRAS, 485, 47

Qin, Y., Mesinger, A., Bosman, S. E. I., & Viel, M. 2021, MNRAS, 506, 2390

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJ, 802, L19

Rorai, A., Hennawi, J. F., Oñorbe, J., et al. 2017, Science, 356, 418

Schmidt, T. M., Hennawi, J. F., Worseck, G., et al. 2018, ApJ, 861, 122

Schneider, E. E., Ostriker, E. C., Robertson, B. E., & Thompson, T. A. 2020, ApJ, 895, 43

Schneider, E. E., & Robertson, B. E. 2015, ApJS, 217, 24

—. 2017, ApJ, 834, 144

—. 2018, ApJ, 860, 135

Schneider, E. E., Robertson, B. E., & Thompson, T. A. 2018, ApJ, 862, 56

Seljak, U., Makarov, A., McDonald, P., & Trac, H. 2006, Phys. Rev. Lett., 97, 191303

Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119

Slosar, A., Font-Ribera, A., Pieri, M. M., et al. 2011, J. Cosmology Astropart. Phys.,

2011, 001

Smith, B. D., Bryan, G. L., Glover, S. C. O., et al. 2017, MNRAS, 466, 2217

Sorini, D., Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2018, ApJ, 859, 125

Teyssier, R. 2002, A&A, 385, 337

Teyssier, R. 2015, Annual Review of Astronomy and Astrophysics, 53, 325

Tonnesen, S., Smith, B. D., Kollmeier, J. A., & Cen, R. 2017, ApJ, 845, 47

Trac, H., Cen, R., & Mansfield, P. 2015, ApJ, 813, 54

Upton Sanderbeck, P., & Bird, S. 2020, MNRAS, 496, 4372

Upton Sanderbeck, P. R., D'Aloisio, A., & McQuinn, M. J. 2016, MNRAS, 460, 1885

Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013a, Phys. Rev. D, 88, 043502

Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2005, Phys. Rev. D, 71, 063534

Viel, M., Schaye, J., & Booth, C. M. 2013b, MNRAS, 429, 1734

Villasenor, B., Robertson, B., Madau, P., & Schneider, E. 2021a, ApJ, 912, 138

—. 2021b, arXiv e-prints, arXiv:2111.00019

Walther, M., Armengaud, E., Ravoux, C., et al. 2021, J. Cosmology Astropart. Phys., 2021, 059

Walther, M., Hennawi, J. F., Hiss, H., et al. 2018, ApJ, 852, 22

Walther, M., Oñorbe, J., Hennawi, J. F., & Lukić, Z. 2019, ApJ, 872, 13

Wang, F., Davies, F. B., Yang, J., et al. 2020, ApJ, 896, 23

Wolfson, M., Hennawi, J. F., Davies, F. B., et al. 2021, MNRAS, 508, 5493

Worseck, G., Davies, F. B., Hennawi, J. F., & Prochaska, J. X. 2019, ApJ, 875, 111

Worseck, G., Prochaska, J. X., Hennawi, J. F., & McQuinn, M. 2016, ApJ, 825, 144

Wyithe, J. S. B., & Bolton, J. S. 2011, MNRAS, 412, 1926

Yang, J., Wang, F., Fan, X., et al. 2020a, ApJ, 904, 26

—. 2020b, ApJ, 897, L14

Zel'Dovich, Y. B. 1970, A&A, 500, 13