

Carnegie Mellon University
MELLON COLLEGE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN THE FIELD OF PHYSICS

TITLE: "Large-Scale Simulations of Hydrogen and Helium Reionization"

PRESENTED BY: Paul La Plante

ACCEPTED BY THE DEPARTMENT OF PHYSICS

HY TRAC 9/28/16

HY TRAC, CHAIR PROFESSOR DATE

STEPHEN GAROFF 9/28/16

STEPHEN GAROFF, DEPT HEAD DATE

APPROVED BY THE COLLEGE COUNCIL

REBECCA DOERGE 9/28/16

REBECCA DOERGE, DEAN DATE

Large-Scale Simulations of Hydrogen and Helium Reionization

by

Paul La Plante

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Hy Trac

September 27, 2016

Abstract

I present here discussion and results of large-scale simulations of cosmic reionization, both hydrogen and helium. For hydrogen reionization, I have focused on the impact that the light cone effect has on observables related to the 21 cm hyperfine signal, with specific attention given to how the duration of reionization affects the results. I show that the light cone effect can introduce significant anisotropy in the detected signal, especially for (cosmically) brief reionization scenarios. Following this, I discuss a series of simulations of helium reionization, focusing this time on observables related to the Lyman- α forest. First I present a method by which dark matter halos from simulations can be populated with quasars to match the latest quasar luminosity function and clustering measurements. Then I discuss a suite of simulations of helium reionization where the quasar model is modified to explore the impact on the timing and duration of reionization, as well as related measurements of the Lyman- α forest. I show that many of the statistical features from the H I Lyman- α forest are similar once the global optical depth is the same across simulations. However, differences of up to a factor of two remain, which might facilitate detection in the future. Finally, I show that differences in the He II Lyman- α forest are significant, and provide a promising way forward for determining the helium reionization history of the Universe.

Acknowledgments

I would like to start off by thanking my wife, Emma, without whom none of this would have been possible. She has always been there for me, through good times and bad, and encouraged me to follow my passion for physics. I would not have made it without her, and I am forever indebted to her for all of the love and support she has shown me throughout the years.

I would also like to thank my family, who have sacrificed so much to help me accomplish my dreams. My parents, Paul and Lori, and my brother, Jon, have all played a major role in helping me get to this point. I would also like to thank my grandparents, who also provided encouragement and reassurance at every turn. Thank you all so much.

Next I would like to thank those who helped start me on the path to research. Mary Lowe, Michael Bonitz, and Lincoln Greenhill all instilled in me an early appreciation for physics research, and helped me decide on pursuing graduate school. I would like to thank in particular Mike Clark, who taught me about GPUs and correlators, but more importantly about what it means to do research.

After arriving at CMU, and being hit with the first year of classes, I would not have survived without the help of Dave Menasche, Sergio de la Barrera, and Bai-cian Ke. After all of the late nights working on problem sets and studying for exams, I'm glad to be on the other side of the branch cut. I'd also like to thank Nick Battaglia and Aravind Natarajan, for all of their guidance while transitioning to research. Their sometimes after-hours help was invaluable for getting started on the right track. I'd also like to thank Michelle Ntampaka, Ying Zu, and Alex Geringer-Sameth for helping me cross the finish line.

I would like to extend my gratitude to of my collaborators, who provided valuable insight and helped shape the direction of my research. I'd like to thank my committee members, Rupert Croft, Tiziana Di Matteo, and Nick Gnedin, for making suggestions to better this work as well as showing me gaps in my understanding. Working to improve my research as well as my approach to it has been incredibly instructive. I'd also like to thank Renyue Cen and Avi Loeb for their guidance, and for graciously hosting me during my visits.

Finally, I would like to extend particular thanks my advisor, Hy Trac, for all of the work and effort put in to get me to this point. He has always graciously given of his time, code, and ideas to put me on the right track, but also to help me think for myself. His encouragement, motivation, and insight were invaluable in my growth as a researcher, and I strive to challenge myself and hold myself to ever-higher standards as I move forward to the next chapter of my professional career.

Contents

1	Introduction	1
1.1	Physics of Reionization	4
1.1.1	Hydrogen Reionization	5
1.1.2	Helium Reionization	6
1.2	Numerical Simulations of Reionization	7
1.3	Outline	10
2	Observations of Reionization	11
2.1	21 cm Radiation	11
2.2	Lyman- α forest	14
3	Predictions for the 21 cm Signal Incorporating the Light Cone Effect	18
3.1	Introduction	18
3.2	Methodology	21
3.3	Analysis	23
3.3.1	21 cm Theory	23
3.3.2	3D Power Spectrum	24
3.3.3	Bias Parameter and Average Bias	27
3.4	Light Cone Effect	28
3.4.1	3D Power Spectrum with the Light Cone Effect	30
3.4.2	Cross-correlation Coefficient with the Light Cone Effect	34
3.4.3	Anisotropic Power Spectrum	34
3.4.4	Power Wedges	35
3.4.5	Comparison to Previous Work	37
3.5	Observational Comparison	39
3.6	Discussion	40
3.7	Conclusions	41
4	Modeling Quasars as Radiation Sources	43
4.1	Introduction	43
4.2	Modeling Quasars as Radiation Sources	46
4.2.1	Radiation-hydrodynamic Simulations	46
4.2.2	Quasar Light Curves	47

4.2.3	Triggering Rate	48
4.2.4	Abundance Matching	49
4.2.5	The Quasar Luminosity Function	53
4.3	Clustering Measurements	56
4.3.1	Two-point Correlation Function	56
4.3.2	Calculating χ^2 Values	57
4.3.3	Characteristic Luminosity and Lifetime	60
4.4	Discussion	63
4.4.1	Mass-to-Light Ratio	63
4.4.2	Mass Function and Duty Cycle of Halo Hosts	65
4.5	Predictions for Helium Reionization	67
4.6	Conclusion	71
5	Signatures of Quasar Activity on the IGM	72
5.1	Introduction	72
5.2	Radiation-Hydrodynamic Simulations	74
5.2.1	Populating Simulations with Quasars	74
5.2.2	Quasar Properties	76
5.2.3	Simulation Features	77
5.2.4	Details of the Simulation Suite	79
5.3	He III Ionization Fraction	83
5.3.1	Ionization Fraction Evolution	84
5.4	The Temperature History of the IGM	89
5.4.1	Temperature-density relation	89
5.4.2	Temperature at mean density	93
5.5	Lyman- α Forest Measurements	94
5.5.1	Effective Optical Depth	95
5.5.2	Flux PDF	97
5.5.3	One-dimensional flux power spectra	99
5.5.4	Three-dimensional flux power spectra	101
5.6	Conclusion	102
6	The Helium Lyman-α Forest	104
6.1	Introduction	104
6.2	Radiation-Hydrodynamic Simulations	106
6.3	He II Lyman- α Forest Measurements	107
6.3.1	Effective Optical Depth	107
6.3.2	Flux PDF	108
6.3.3	One-d flux power spectra	110
6.4	Discussion	111
6.5	Conclusion	112

7	Conclusions	113
7.1	Future Observations of Reionization	113
7.1.1	Hydrogen Reionization	113
7.1.2	Helium Reionization	115
7.2	Future Research Directions	116
	Appendices	119
A	Additional Material for the Light Cone Effect	120
A.1	Exclusion of $k_{\perp} = 0$ Modes	120
A.2	Additional Figures	121
B	Additional Material for Quasars as Radiation Sources	126
B.1	Fitting the parameters of the quasar luminosity function	126
B.2	Bias as a function of redshift	129
B.3	Bias as a function of luminosity	130
C	Additional Material for Signatures of Quasar Activity on the IGM	133
C.1	Renormalizing τ_{eff}	133
C.2	Renormalizing the Lyman- α flux PDF	133
C.3	The Quasar Luminosity Function	134
C.3.1	Model Q1	136
C.3.2	Model Q2	138

List of Tables

4.1	QLF parameters of the datasets incorporated	52
4.2	Parameters for the composite QLF	55
4.3	Parameters of some of the quasar models considered	57
5.1	Parameters for helium simulation suite	80
B.1	Best-fit parameters for our QLF	129
C.1	QLF parameters for helium reionization suite	137

List of Figures

2.1	The hyperfine structure of the hydrogen atom	12
3.1	H I neutral fraction and global 21 cm signal evolution with redshift	23
3.2	Comoving 3D 21 cm brightness temperature power spectrum	25
3.3	Scale-dependent and average bias factors	26
3.4	Visualization of the light cone effect on 21 cm radiation	29
3.5	Light cone 3D power spectrum as a function of reionization history	31
3.6	Light cone 3D power spectrum as a function of neutral fraction	32
3.7	Effect of box size on the power spectrum	33
3.8	Cross-correlation coefficient for coeval and light cone cases	36
3.9	Two-dimensional $(k_{\perp}, k_{\parallel})$ power spectra	37
3.10	Power wedges measurement	38
3.11	Comparison with observational data	39
4.1	Comparison between the observed and method-generated QLF	54
4.2	The correlation matrix for one of our models	58
4.3	The two-point correlation function	59
4.4	Constraints on the model in t_0 - γ space	61
4.5	Allowed quasar lifetimes as a function of γ	62
4.6	The mass-to-light ratio predicted by our model	64
4.7	The mass range of halo hosts predicted by our model	65
4.8	The duty cycle predicted by our model	66
4.9	The volume-filling fraction Q_i	68
5.1	Number of photons produced in each simulation of helium suite	83
5.3	Visualization of reionization, single simulation	85
5.4	Visualization of reionization, multiple simulations	86
5.2	Evolution of x_{HeIII} as a function of redshift	87
5.5	Temperature-density relation	90
5.6	Evolution of T_0 and γ with redshift	92
5.7	IGM temperature at mean density	93
5.8	$\tau_{\text{eff,HI}}$ as a function of redshift	96
5.9	H I flux PDF	98
5.10	One-dimensional H I Lyman- α forest flux power spectrum	100

5.11	One-dimensional H I Lyman- α forest flux power spectrum	101
6.1	$\tau_{\text{eff,HeII}}$ as a function of redshift	108
6.2	He II flux PDF	109
6.3	One-dimensional He II Lyman- α forest flux power spectrum	111
A.1	Light cone power spectrum at $x_i = 0.75$	122
A.2	The power spectrum with $k_{\perp} = 0$ included	123
A.3	Light cone power spectrum at $x_i = 0.25$	124
A.4	Anisotropic power spectrum as a function of box size	125
B.1	Evolution of QLF parameters with redshift	127
B.2	t_0 - γ parameter space evolution with redshift	130
B.3	Evolution of quasar clustering as a function of γ and light curve	131
C.1	The flux PDF after renormalization of τ_{eff}	134
C.2	QLF parameter evolution of helium simulation suite	136

In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.

Douglas Adams, *The Restaurant at the End of the Universe*, 1980.

Chapter 1

Introduction

From time immemorial, astronomy has concerned itself with the bright points in the sky. Starting in antiquity with stars and constellations to the development of telescopes capable of observing galaxies beyond our own, we are naturally drawn to investigate the patches of light in the endless sea of darkness. However, this does quite a disservice to those very oceans of darkness. These voids between galaxies, which comprise more than 99% of the Universe by volume, can provide just as much meaningful information as the galaxies themselves. In our quest to understand more about the history and trajectory of our Universe, we would do well to heed the knowledge coming from the space of outer space.

Before delving too deeply into specifics, it bears looking at the “big-picture” of cosmology, and what the first steps in the evolution of the Universe are. In the seconds and minutes following the big bang and inflation, the constituent parts of the Universe were in thermal equilibrium. As the Universe continued to adiabatically expand, the constituent components began to cool. Among these constituents was baryonic matter, which has undergone several key transitions during the evolution of the Universe. Initially, baryonic matter formed a “soup” of particles, in a phase known as the quark-gluon plasma. Within a second of the big bang, this plasma had cooled enough to form protons and neutrons, the building blocks of all matter. After about ten minutes, big bang nucleosynthesis had occurred, fixing the primordial abundances of hydrogen, helium, and heavier elements. By mass, baryonic matter was composed of about three-quarters hydrogen nuclei (unbound protons) and one-quarter helium-4 nuclei (alpha particles), with negligible amounts of other elements or isotopes. At this point, the Universe was still hot enough that protons, alpha particles, and electrons formed a plasma, and were not bound in neutral atoms. In other words, during this era baryons were ionized, instead of forming neutral atoms. During this plasma phase of baryons, the mean free path of photons was incredibly short, making this era opaque to observation by photons.

Following additional expansion and cooling of the Universe, baryons were eventually able to form neutral atoms. This era is referred to as the Epoch of Recombination;

however, this is something of a misnomer, since neutral matter had not been formed earlier (although “Epoch of Combination” admittedly does not have the same ring to it). Recombination occurred about 380,000 years after the big bang, at a redshift of $z \approx 1100$. Shortly after recombination, the Universe became transparent to photons, which lead to the emission of the cosmic microwave background (CMB). The CMB is sometimes referred to as the “surface of last-scattering”, because observation of the CMB requires photons travelling to Earth without further scattering off of ions. Exquisite measurements of the CMB by the *WMAP* and Planck satellites has provided a wealth of information, including the most precise determination of many of the cosmological parameters that define our Universe. As will be discussed further below, the CMB also serves as an important reference point for a key observational technique of hydrogen reionization.

After recombination and last-scattering, neutral baryons evolve primarily under the influence of gravitational attraction. The gravitational potential, sourced primarily by dark matter particles, amplifies of initial inhomogeneities in the matter distribution laid down by inflation: regions with above-average density accrete matter and become denser, while regions at below-average density become less dense over time. Immediately following recombination, the inhomogeneities are roughly one part in 10^5 . Over the next 500 million years, though, the inhomogeneities continue to grow. Eventually the density of gas is sufficient for the first stars in the Universe to form, emitting radiation as they burn. These early proto-galaxies are relatively isolated, and surrounded by low-density neutral gas. The radiation from these first stars propagates outward into the space between galaxies, typically referred to as the intergalactic medium (IGM).

As the radiation propagates outwards from stars composing proto-galaxies into the IGM, photons with sufficient energy can ionize the neutral hydrogen atoms encountered. Specifically, photons with an energy $h\nu \geq 13.6$ eV interacting with neutral hydrogen will reionize the gas, returning the hydrogen atom to its constituent proton and electron. Eventually, enough radiation is produced by stars to completely reionize neutral hydrogen of the IGM, marking the end of hydrogen reionization. All told, most models predict that this epoch lasts for 500–700 million years, a mere fraction of the total lifetime of the Universe. Nevertheless, this era represents one of the most exciting eras of development, and sets the stage for the subsequent evolution of the Universe.

Throughout this epoch of the first stars forming, neutral helium resides in the IGM alongside of hydrogen. However, the binding energy of helium is significantly greater than that of hydrogen: whereas ionizing hydrogen requires a mere 13.6 eV of energy, the first ionization of helium requires 24.6 eV of energy. The second ionization of helium has an even steeper energy requirement, of 54.4 eV. This factor of four may seem insignificant on first blush, but it actually means that helium reionization happens at a dramatically different point in the Universe’s evolution compared to hydrogen reionization.

The first ionization of helium likely happens simultaneously with hydrogen reionization. The energy requirement is greater than that of hydrogen by almost a factor of two. The stars that are responsible for ionizing hydrogen do not produce many photons energetic enough to ionize helium relative to the number of hydrogen atoms. Counteracting this, there is an order of magnitude fewer helium atoms in the Universe. The fact that the mass fraction of helium is about one-quarter of the total baryon mass means that by number density, helium composes only about 7% of the baryons. There are therefore competing effects for the first ionization of helium: stars produce fewer photons of energy $h\nu \geq 24.6$ eV compared to the number of 13.6 eV photons, but there is far less helium in the Universe to reionize. It turns out that the latter effect wins out, and based on the stellar models for sources of hydrogen-ionizing radiation, there is a sufficient number of helium-ionizing photons to singly ionize helium during the reionization of hydrogen.

The second ionization of helium represents a chasm in energy space that is too vast to cross. The radiation output from the first stars simply is not energetic enough to doubly ionize helium. The ionization of helium is temporarily paused until sources of higher energy photons emerge. These sources take the form of quasars, which are the most luminous objects in the galaxy. Thought to be accretion disks surrounding supermassive black holes, quasars can outshine their host galaxies when at their peak luminosities. Conversely, quasars are relatively short-lived, with lifetimes lasting only a few tens of millions of years. Nevertheless, the radiation from quasars is sufficient to dramatically alter the gas of the IGM.

Due to their increased mass and structure requirements, quasars do not appear until later in the timeline of the Universe compared to the stars that power hydrogen reionization. Based on observations, it seems that a significant number density of quasars does not appear until about 1 billion years after the big bang, following the end of hydrogen reionization. The general picture of helium reionization also looks different from hydrogen reionization. Singly ionized helium has a significantly lower cross-section to photons compared to neutral hydrogen, which leads to a longer mean free path of energetic photons from quasars. Additionally, quasars are highly biased sources, which means that they are typically found clustered together rather than homogeneously spread throughout a given volume. These features mean that helium reionization was much more anisotropic than hydrogen reionization, a property that will be discussed more below.

One additional feature of helium reionization is due to the finite lifetime of quasars, as alluded to earlier. Rather than acting as slowly burning candles that gradually reionize the IGM, quasars explode like fireworks out into the IGM, rapidly eating away at the singly ionized helium. The resulting imprint on the IGM is one of a sudden growth of doubly ionized helium, followed by the source switching off; consequently, there may no longer be a source of radiation within this region. As a result, the doubly ionized helium may begin to recombine with free electrons back into singly ionized helium. The gas will then undergo subsequent ionization, which can have interesting

implications for the thermal state of the gas in the IGM. Overall, helium reionization contains richer interplay between the gas of the IGM and radiation sources than hydrogen reionization, and provides a fascinating glimpse into the later evolution of the Universe.

Now that the stage has been set regarding the main players of reionization, it can be helpful to take a finer look at some of the physics governing the interplay of gas and radiation. This will allow for a more detailed description of the broad brush-strokes of reionization, allowing for a deeper treatment in further discussion.

1.1 Physics of Reionization

Surprisingly, the rich and complex world of reionization can be described using very simple physics. At its core, reionization is a reaction between photons and neutral atoms. In general, these physical processes fall under the category of radiative transfer, which follows the absorption and emission of photons. Combined with simple conservation laws, the process of reionization can be described incredibly accurately. As already mentioned, the broad umbrella term of “reionization” actually encompasses two primary flavors: hydrogen and helium. In the discussion of the microscopic physics that follows, I will discuss hydrogen reionization primarily, specifically. Nevertheless, the relevant physics is analogous for the case of helium.

Consider a hydrogen atom in its ground state in free space. An incoming photon with an energy of $h\nu \geq 13.6$ eV will interact with the hydrogen atom and ionize the neutral atom, leaving behind a proton and an electron:



This reaction is referred to photoionization. To a good approximation, any energy remaining after overcoming the 13.6 eV ionization energy is imparted to the much less massive electron as kinetic energy. In this ionized state, the protons and electrons are subsequently transparent and will no longer react with additional incoming photons.

The reverse reaction is also possible, wherein a proton and an electron combine to form a neutral hydrogen atom and a photon with an energy of 13.6 eV. This process is known as recombination and will readily occur when protons and electrons are found in sufficient density. The combination of photoionization and recombination will determine the local ionization state of the gas: for highly dense regions with little photon flux, gas will tend to be neutral. Alternatively, a large photon flux through low density gas will lead to a highly ionized state. The degree of ionization can be quantified with the ionization fraction x_i , which is defined as the number density of an ionized species compared to the total number. For instance, the ionization fraction of ionized hydrogen H II can be written as:

$$x_{\text{HII}} = \frac{n_{\text{HII}}}{n_{\text{H}}}. \quad (1.2)$$

Broadly speaking, reionization science is concerned with determining the ionization fields as a function of redshift for all points in space.

The process of helium reionization is quite similar to that of hydrogen described in Equation 1.1, except that there are two such equations: a transition from He I to He II, and from He II to He III. Despite commonality in the microscopic physics, the macroscopic picture of hydrogen and helium reionization are very different. This difference is due in large part to the different epochs of reionization, and differences in the sources of reionization. Before proceeding further, it helps to have a general picture in mind for the different flavors of reionization.

1.1.1 Hydrogen Reionization

As already discussed in the introductory portion of the chapter, hydrogen reionization was precipitated by the radiation from the first stars, early on in the history of the Universe. Quantitatively, most models suggest that hydrogen reionization took place during a redshift range of $20 \gtrsim z \gtrsim 6$ (*e.g.*, Barkana & Loeb 2001), featuring a relatively gradual transition between the neutral and ionized phases. Further, this transition proceeds in an “inside-out” fashion, in which the radiation from stellar sources first reionizes the high-density gas near galaxies. As the Universe becomes increasingly ionized, the radiation fronts propagate further outward toward low-density gas, which is ionized later than the high-density regions. This general picture is in large part due to the relatively large cross-section of neutral hydrogen, combined with a sharp drop-off in the number of high-energy photons produced by the stellar population in proto-galaxies. Hence, the photons from stars have a very short mean free path, and virtually all of the photons that escape into the IGM are absorbed immediately upon encountering neutral atoms. This leads to generally well-defined ionization fronts. The initial “bubbles” of ionized gas surrounding sources eventually grows larger, leading to ever-growing portions of ionized gas surrounded by neutral atoms.

Thus far, most of the discussion has focused on photoionization of hydrogen. However, this process only tells part of the story. Just as important is photoheating, and the effect that hydrogen reionization has on the thermal state of the IGM. As mentioned in Sec. 1.1, following photoionization, excess energy from the photon is transferred to kinetic energy in the electron. If the kinetic energy of the electron is greater than 13.6 eV, then secondary ionizations between these energetic electrons and neutral hydrogen is possible (Shull, 1979). Given the spectra of radiation from the stellar sources, though, secondary ionizations are thought to be uncommon. Accordingly, the excess kinetic energy in the electrons is eventually deposited into the gas as heat. For low-density gas in the IGM, photoheating raises the temperature of the gas from $\mathcal{O}(10^2 \text{ K})$ to $\mathcal{O}(10^4 \text{ K})$. Thus, the general picture of hydrogen reionization is one of photons from the first stellar sources propagating outward into the IGM, encountering initially cool gas, then ionizing and heating the gas, leaving a

warm plasma in their wake. Eventually, these regions of ionized gas overlap, soon after which the entire Universe has been reionized.

The modeling of hydrogen reionization is generally straightforward. For semi-analytic calculations, one can define simple relationships between halo properties and the number of hydrogen-ionizing photons that can be expected to escape. Such semi-analytic tools have been developed (*e.g.*, 21CMFAST, Mesinger et al. 2011), and generally provide a sufficient level of accuracy for forecasting the results from different ionization histories. Nevertheless, a full numerical treatment is necessary to capture the effect of anisotropies of the reionization process, as well as capturing the feedback between the radiation and the gas of the IGM.

1.1.2 Helium Reionization

In general, helium reionization proceeds in a manner similar to hydrogen reionization: radiation from quasars propagates outward into the IGM, ionizing the gas as it travels. As the radiation ionizes the gas, a large amount of heat is deposited due to photoheating. The regions of doubly ionized gas grow larger until eventually they completely overlap, signaling the end of reionization. So far, so similar.

However, there are several key differences between hydrogen and helium reionization that lead to helium reionization presenting a much more challenging problem to treat correctly. First, due to the fact that quasars are the primary sources of reionization, a proper understanding of quasars being hosted in dark matter halos is required. In general, the relationship between halo mass and quasars is less well-understood than that of halo mass and stellar content, and so this presents a difficulty in modeling the sources of helium reionization.

Related to the fact that the sources in helium reionization are quasars rather than galaxies, the finite lifetime features prominently in helium reionization. The regions of doubly ionized gas grow significantly during the lifetime of quasars, growing to be up to tens of Mpc in diameter for the most luminous sources. However, following the end of the quasar's lifetime, the ionized region ceases to grow, and the gas is no longer necessarily exposed to an external radiation field. Accordingly, the gas in the denser regions begins to recombine, allowing for additional reionization events. When the gas is ionized again, after being exposed to radiation from a new quasar, it undergoes further photoheating. This process has interesting ramifications on the thermal state of the IGM, and leads to an interesting relationship between temperature and density.

As a final characteristic of reionization precipitated by quasars rather than galaxies, the highly biased nature of the sources implies that quasars are typically found clustered together. This clustering leads to an earlier overlap of ionized regions compared to a more homogeneous distribution of sources, and in particular compared to hydrogen reionization. When the ionized regions begin to overlap, the propagation of radiation from sources becomes highly anisotropic, which makes accurate semi-analytical modeling difficult.

Another important difference between helium and hydrogen reionization is the longer mean free path of photons in the case of helium reionization. As discussed earlier, the cross-section of helium to photons at the edge of its ionization energy is lower than that of hydrogen's, and so there is less absorption of photons as they propagate. Accordingly, the ionization fronts in the case of helium reionization are not well-defined to the same extent they are for hydrogen. There is a gradual transition between singly ionized and doubly ionized gas in the IGM, especially far from sources due to spectral filtering: the photons closer to the ionization energy of helium have a higher cross-section, and so are preferentially absorbed by gas near the sources.

In summary, the process of helium reionization is tricky to properly treat using semi-analytic methods: the early overlap of ionized regions, the gradual transition between singly and doubly ionized regions, and the recombination of gas in dense doubly ionized regions all argue for a more precise treatment than is possible using semi-analytic methods. Therefore, numerical simulations of helium reionization are necessary to properly understand the implications that reionization has on the IGM. Further, the numerical simulations that properly examine helium reionization must be sufficiently sophisticated to capture all of relevant features. This is no small feat, and so it is helpful to discuss some details of numerical methods to discuss how to implement these simulations best.

1.2 Numerical Simulations of Reionization

An important tool for better understanding reionization is numerical simulation of the Universe on cosmological scales. By attempting to create physically accurate models of the matter and radiation on large scales, we are able to make predictions for what observational efforts to detect reionization should see. In general, this process requires a tradeoff between faithfully capturing the small-scale physics relevant to the source of radiation inside of galaxies and probing large-scale distributions of matter for the purposes of making accurate statistical summaries of scales where the cosmological principle applies. On the small-scale, simulations in principle must resolve feedback processes generated by accretion disks surrounding blackholes, which have a physical extent of 10^{-4} – 10^{-3} parsecs. On the large-scale, we would like to generate a simulation with a volume comparable to the observable universe in order to generate realistic statistics, which has an extent of 1–10 gigaparsecs (Gpc). These requirements span a dynamic range of 12–15 orders of magnitude per dimension, which is flatly computationally infeasible. Thus, approximations must be made, either by limiting the resolution element of the simulation, or by resolving a smaller volume of the simulated universe, or both.

In practice, simulations of reionization typically opt for sacrificing the resolution on small scales in favor of generating a representative volume of the mock universe. This approach employs so-called sub-grid models for physical phenomena that occur below the resolution limit of the numerical simulation. In some incarnations of this

approach, entire galaxies are abstracted away to individual point-sources of radiation, and the focus of the simulation becomes the low-density gas of the IGM between point-sources. Further computational demands are made by requiring the simultaneous solution of relevant physical properties, such as hydrodynamics for baryonic physics and radiative transfer for solving interactions with radiation. The compromises made in the design and execution of numerical experiments are ultimately informed by the purpose and focus of the current investigation. These considerations lead to decisions about which physics and scales are most relevant for the application at hand.

For the reionization simulations discussed in the body of this work, there are several different sets of physics involved. We will now turn to each of them, and further discuss their relevance to the broader picture.

1. ***N*-body methods.** The large-scale structure of the universe can be explained exceedingly well using purely gravitational interactions of collisionless dark matter particles. *N*-body methods refer to techniques for determining the positions and momenta of a collection of *N* particles subjected to self-gravitation. One approach is to use Newton’s Law of Universal Gravitation directly to find the force acting on each particle due to the influence of every other particle. However, this approach scales poorly as the number of particles grows large, and so a more sophisticated scheme must be developed. More typically in practice, the gravitational potential is calculated at all points in a cosmological volume, and then the resulting accelerations, positions, and momenta are found. The large-scale structure lays down the scaffolding of sheets, filaments, and dark matter halos, the latter of which can host high densities of baryonic matter and ultimately stars and galaxies.
2. **Hydrodynamics.** In addition to the gravitational potential sourced primarily by dark matter, baryonic matter is subject to self-interaction terms such as pressure. The fundamental equations of hydrodynamics are the Euler Equations, which are applied to fluids and derived by assuming conservation of mass, conservation of linear momentum, and conservation of energy. The local fluid density ρ , velocities \vec{v} , and total internal energy E are tracked for all particles composing the fluid (Lagrangian scheme) or all volumes in space (Eulerian scheme). Computational fluid dynamics techniques have been applied to cosmology to incorporate baryonic effects and feedback that is important for galaxy evolution and IGM physics.
3. **Radiative transfer.** The final main ingredient for simulations of the IGM is radiative transfer. This branch of physics concerns the interactions between matter and radiation. As has been discussed in Sec. 1.1, the processes of photoionization and photoheating are crucial for studying the impact of reionization on the IGM. Radiative transfer determines the ionization state of different gases in the IGM, as well as the number of free electrons.

In the most basic simulations, only N -body methods are included. These types of simulations are the crudest representation of the observable Universe, but are relatively straightforward to implement. For studies concerned merely with the large-scale structure, such as the generation of mock catalogs for comparison with galaxy surveys, N -body-only simulations are sufficient.

The inclusion of hydrodynamics in cosmological simulations has been quite fruitful for our understanding of galaxy formation and baryonic feedback effects. Specific implementations of hydrodynamics in a cosmological context have been developed using Lagrangian schemes (*e.g.*, GADGET, Springel 2005), Eulerian ones with optional adaptive mesh refinement (*e.g.*, Enzo, O’Shea et al. 2005), or a hybrid scheme tracking particles and employing a moving mesh (*e.g.*, AREPO, Springel 2010). Although much more computationally intensive than pure N -body schemes, these approaches allow the user to track the gas properties during the evolution of the Universe.

Finally, radiative transfer has been applied to simulations in a variety of ways. The most common approach is to perform what are known as post-processing simulations, in which the radiative transfer equations are applied to a snapshot taken from a hydrodynamic simulation. In this approach, the feedback effects on the gas are not incorporated into the solution of the hydrodynamics equations of the simulation, which can lead to inaccuracies. Most vitally for reionization, post-processing simulations do not incorporate the effect of photoheating on the IGM. As the gas heats up, its density and pressure can change dramatically, affecting how it interacts with nearby gas. This outcome is clearly disfavored. On the other hand, post-processing simulations typically implement an inhomogeneous radiation field, which is an essential feature of the reionization process.

An alternative approach that captures the radiative feedback on the gas is to use a uniform radiation field. In the limit that the gas is optically thin to radiation, this approximation is not necessarily inaccurate. For instance, at late times ($z \lesssim 6$), the IGM is optically thin to UV photons, and so the extragalactic background can be well modeled as a uniform field. However, for the process of reionization, this is clearly a poor approximation. Nevertheless, this approach incorporates the feedback effects of radiation on the gas.

The optimal approach for radiative transfer in simulations is to incorporate an inhomogeneous radiation field on-the-fly so that the radiation feedback is incorporated into the hydrodynamical quantities of the simulation. In most of the analysis that follows, I will be presenting results from the RadHydro code, which features just such a scheme. The RadHydro code includes an Eulerian approach to solving hydrodynamics, and uses ray tracing to solve the radiative transfer equations. It has been previously applied to hydrogen reionization simulations (Trac & Cen, 2007; Trac et al., 2008), and below I will discuss how it has been modified for application to helium reionization.

1.3 Outline

The following chapters discuss work I have done in analyzing hydrogen and helium reionization. The methods used primarily involve numerical simulations, and make predictions for what real-world experiments might observe. Chapter 2 details the relevant physics involved in connecting reionization to observations. I detail the hyperfine transition in neutral hydrogen, the most promising avenue for detecting hydrogen reionization, and the Lyman- α forest, which contains invaluable information about helium reionization. In Chapter 3, I discuss applying the patchy reionization model of Battaglia et al. (2013b) to hydrogen reionization, and incorporating the light cone effect when computing quantities of interest. In Chapter 4, I present a method by which dark matter halos can be populated with quasars to reproduce an observed quasar luminosity function (QLF) and quasar clustering. I match the QLF from the COSMOS survey (Masters et al., 2012), and the Sloan Digital Sky Survey (SDSS; McGreer et al. 2013; Ross et al. 2013), while matching the quasar clustering from the Baryon Oscillation Spectroscopic Survey (BOSS; White et al. 2012). In Chapter 5, I discuss a new suite of simulations which explore helium reionization. These simulations include hydrodynamics and radiative transfer solved simultaneously, and represent a significant step forward in helium reionization studies, improving upon previous studies (McQuinn et al., 2009, 2011; Compostella et al., 2013, 2014) in meaningful ways. In Chapter 6, I discuss some of the key impacts that helium reionization has on the helium Lyman- α forest. This set of results from simulations shows significant differences in certain statistical measures, and can clearly indicate the ionization level of the helium gas. When compared with real-world observations of the helium Lyman- α forest (Syphers et al., 2009b,a), a meaningful measurement of helium reionization history should be made in future work. In Chapter 7, I discuss some general conclusions and mention avenues for future investigation.

Chapter 2

Observations of Reionization

In this chapter, I will briefly discuss some of the methods by which hydrogen and helium reionization may be detected. In both cases, rather than looking directly at the sources of reionization, it is generally more fruitful to examine the IGM, and look for some of the effects that reionization has on it. In both cases, the general technique involves examining a source of radiation, and measuring the amount of absorption or emission that is contributed as the radiation passes through the IGM. By understanding how the different features of reionization manifest in the observables, it becomes easier to make inferences regarding the processes of hydrogen and helium reionization.

2.1 21 cm Radiation

One of the most prominent observables related to hydrogen reionization is the 21 cm signal (*e.g.*, Pritchard & Loeb 2012). In typical parlance, 21 cm radiation is said to come about when the spins of the proton and electron flip relative to each other. However, this explanation is an oversimplification, and does not accurately capture the rich physics involved. In neutral hydrogen, the interaction between the magnetic field created by the proton magnetic moment \vec{I} and electron magnetic moment of the electron spin \vec{S} lead to the so-called “hyperfine structure” of hydrogen (Cohen-Tannoudji et al., 2005). In the ground state $1s$ of hydrogen, the fine structure terms denoted by $1s_{1/2}$ (which includes relativistic corrections, spin-orbit coupling, and the Darwin effect) only produce an overall energy shift, and do not break the degeneracy of the energy. However, including the interaction $\vec{I} \cdot \vec{S}$, the so-called “contact term”, does break this energy degeneracy.

We can write the total angular momentum of the hydrogen atom \vec{F} as $\vec{F} = \vec{L} + \vec{S} + \vec{I} = \vec{J} + \vec{I}$, where \vec{J} is the total angular momentum of orbital angular momentum \vec{L} and spin angular momentum \vec{S} . However, because $\vec{L} = 0$ in the ground state of hydrogen, this reduces to $\vec{F} = \vec{S} + \vec{I}$. Representing the spins of the proton and electron as $|+\rangle$ for spin-up and $|-\rangle$ for spin-down, we note that the state $F = 1$ is a spin-triplet

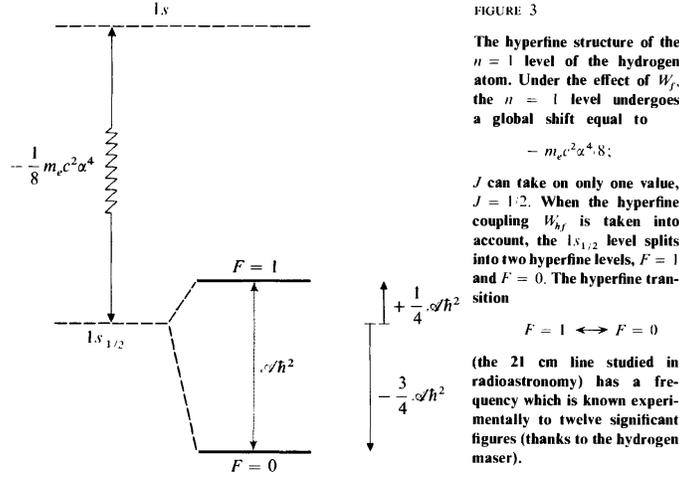


FIGURE 3

The hyperfine structure of the $n = 1$ level of the hydrogen atom. Under the effect of W_{hf} , the $n = 1$ level undergoes a global shift equal to

$$- m_e c^2 \alpha^4 / 8;$$

J can take on only one value, $J = 1/2$. When the hyperfine coupling W_{hf} is taken into account, the $1s_{1/2}$ level splits into two hyperfine levels, $F = 1$ and $F = 0$. The hyperfine transition

$$F = 1 \leftrightarrow F = 0$$

(the 21 cm line studied in radioastronomy) has a frequency which is known experimentally to twelve significant figures (thanks to the hydrogen maser).

Figure 2.1: An energy diagram showing the energy structure of the fine and hyperfine structure of the hydrogen atom. The ground state ($1s$) undergoes an overall shift due to the fine structure ($1s_{1/2}$), but does not have the degeneracy broken. Upon including the hyperfine structure, a transition between states of $F = 0$ and $F = 1$ appears with an energy difference equal to a photon with $\lambda = 21$ cm appears. Figure reproduced from Cohen-Tannoudji et al. (2005).

state. Using the notation of $|F, m_F\rangle$, we can write the basis vectors of $F = 1$ as:

$$\begin{aligned} |1, 1\rangle &= |++\rangle \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle) \\ |1, -1\rangle &= |--\rangle. \end{aligned} \tag{2.1}$$

Meanwhile, we have a spin-singlet state for the $F = 0$ state:

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle). \tag{2.2}$$

Figure 2.1 shows an energy diagram of the relative energy of the hyperfine structure.

In free space, a hydrogen atom that is in the spin triplet state can transition to the spin singlet state, releasing a photon with frequency 1420 MHz. This frequency corresponds to a rest-frame wavelength of 21 cm, leading to the name of the transition. This transition is mediated by the magnetic dipole moment and is thus a “forbidden transition.” Accordingly, the time scale of this transition is exceedingly long, on the order of 10 million years. Nevertheless, in great enough quantities, as in the IGM, there is still an appreciable signal at 21 cm.

Use of the 21 cm signal as a probe of neutral hydrogen has a long history in astrophysics and cosmology. As with any radiative process, the use of this signal

can be understood in terms of the radiative transfer equation. Given radiation of a specific intensity I_ν propagating through free space and interacting with matter, the intensity of radiation can change through both absorption or emission processes. Along a path length ds , the intensity can be calculated as:

$$\frac{dI_\nu}{ds} = j_\nu - \alpha_\nu I_\nu, \quad (2.3)$$

where j_ν represents emission from matter along the path length, and α_ν represents the absorption coefficient at the specific frequency ν .

Formally, with 21 cm radiation, we can set up an expression using the radiative transfer equation of Equation (2.3), where the source of radiation I_ν is the cosmic microwave background (CMB), which propagates toward us on Earth while potentially interacting with matter along the way. As is customary, we express this interaction with matter in terms of an optical depth $\tau \equiv \int ds a_\nu(s)$. The optical depth quantifies the relative opacity of matter, in this case neutral hydrogen, and determines to what extent the radiation from the CMB is enhanced (through emission) or diminished (through absorption).

Field (1958) introduced the concept of “spin temperature” T_S to quantify the relative occupation between spin states of neutral hydrogen (the spin-triplet vs. the spin-singlet). The spin temperature is not actually a thermodynamic temperature, but is instead a measure of the relative occupation of the two hyperfine levels. The occupation of the hyperfine levels is governed by the interaction between neutral hydrogen and ambient radio radiation (characterized by a brightness temperature of radiation of T_γ), collisional coupling to gas with a kinetic temperature T_K , and coupling to Lyman- α photons, the latter of which can affect the spin temperature via the Wouthuysen-Field effect (Wouthuysen, 1952; Field, 1958).

The radiation from the first stars and galaxies responsible for hydrogen reionization propagates outwards into the IGM, photoheating as it moves. As a result of this photoheating, the kinetic temperature of the gas becomes much hotter than the CMB. At this point $T_S \gg T_\gamma$, and so the spin temperature is saturated. The spatial inhomogeneities caused by differences in the local ionization fraction, gas density, and temperature become unimportant, and globally the 21 cm signal is seen in emission compared to the CMB. The excess radio emission compared to the background source (in this case, the CMB) is referred to as the 21 cm brightness temperature δT_b . At this point in the Universe’s history, radiation from galaxies has sufficiently heated the IGM such that this increase is ~ 10 mK, which is much greater than the inherent anisotropy of the CMB sourced by density fluctuations at the time of recombination. This simplification allows for a straightforward expression for the brightness temperature that depends only on local density fluctuations, ionization state of the gas, and cosmological parameters.

As mentioned above in Sec. 1.1, reionization proceeds in an inside-out fashion. As a result, regions of high density are ionized before regions at mean and low density.

Because the 21 cm signal only appears in regions that contain neutral hydrogen, the signal seen during reionization is anti-correlated with density fluctuations. Regions of high density are reionized earlier, and so for these regions $\delta T_b = 0$ at an earlier redshift compared to lower density regions. As reionization proceeds, the 21 cm signal is produced only in increasingly less dense portions of the Universe, until finally at redshift $z \sim 6$ the signal disappears entirely. Following the conclusion of reionization, neutral hydrogen is only appreciably found inside of galaxies. As a result, the 21 cm signal is seen only in very dense regions. In this era of $z \lesssim 6$, it has been proposed to use 21 cm radiation for intensity mapping galaxies at low redshift (Peterson et al., 2009).

From an observational point of view, the 21 cm signal is the most promising probe of the Universe at this epoch. The typical galaxies that are thought to drive reionization are far too faint to be seen by current or next-generation telescopes. The Hubble Space Telescope (HST) has observed $7 \lesssim z \lesssim 9$ galaxies only relatively recently (Robertson et al., 2013), with ones from even higher redshift remaining exceedingly rare occurrences (Oesch et al., 2016). The projected luminosity function limits for the upcoming James Webb Space Telescope have detection limits which are significantly fainter than the HST (Gardner et al., 2006); however, only the equivalent of the HST Frontier Fields (*i.e.*, magnification from gravitational lensing) will provide the necessary sensitivity to detect first light (Windhorst et al., 2006). Additionally, since reionization concerns the ionization state of the IGM, it is more meaningful to make observations related to these regions. Accordingly, observing only galaxies from this epoch gives just a part of the story.

A more complete picture is provided instead by detecting the 21 cm signal using a radio telescope. Current generation endeavors, such as the Murchison Widefield Array (MWA, Bowman et al. 2005), the Low Frequency Array (LOFAR, Harker et al. 2010), and the Precision Array for Probing the Epoch of Reionization (PAPER, Parsons et al. 2010) have placed increasingly better upper-limits on the global power spectrum at various redshifts. Next-generation experiments, such as the Hydrogen Epoch of Reionization Array (HERA, Pober et al. 2014) or the Square Kilometer Array (SKA, Koopmans et al. 2015). These experiments aim to observe hydrogen reionization through observing the 21 cm signal as opposed to the luminous sources that drive reionization. Although the method is fraught with difficulty, most notably the fact that foreground contaminants of the signal are several orders of magnitude larger than the meaningful signal, the potential benefit of using the voluminous IGM to provide information leads to a profound increase in what can be learned about reionization.

2.2 Lyman- α forest

At later times in the Universe's history, the gas of the IGM can be detected through its absorption of background radiation. At redshifts below $z \lesssim 6$, distant quasars

in the early universe emit sufficient radiation as to be visible from Earth. These quasars emit radiation over a broad range of frequencies, which propagate toward detectors on Earth. Along the way, the spectrum of the quasar becomes redshifted due to the expansion of the Universe. As the radiation passes through gas in the IGM, neutral hydrogen (H I) absorbs radiation at frequencies corresponding to certain electron transitions. Microscopically, a hydrogen atom in the ground state $1s$ can absorb a photon with a wavelength of 1216 \AA and transition into the $2p$ state, due to selection rules governed by conservation of angular momentum (*i.e.*, $\Delta l = \pm 1$ when absorbing or emitting photons). This photon is subsequently re-radiated when the atom transitions from this excited state back to the ground state; however, from the point of view of an observer on Earth, the radiation is generally re-radiated in a different direction, and thus the radiation at this frequency appears as a net absorption compared to the emitted spectrum of the distant quasar.

Macroscopically, we can again make use of the radiative transfer equation from Equation (2.3), where we have only an absorption term. Because the absorption coefficient, and by extension the optical depth, depends on the properties of the gas, measuring the optical depth of the radiation along the line of sight allows for making statistical inferences about the average state of the IGM as a function of redshift.

As mentioned above, the distant source emitting radiation is a quasar at redshift z_q , which has a very broad spectrum. As the spectrum propagates toward Earth, it passes through neutral hydrogen from the IGM. Suppose we have a gas cloud at redshift $z_g < z_q$. From the point of view of an observer on Earth, the intrinsic quasar spectrum with a redshift of z_q would show an absorption feature due to the Lyman- α transition centered at a wavelength of $1216(1 + z_g) \text{ \AA}$, which will be blueward of $1216(1 + z_q) \text{ \AA}$. The extent of the absorption feature in frequency space is determined by the cross section $\sigma(\nu)$. Because the cross section is sharply peaked about some particular frequency ν_0 , the shape of the cross section can be written as

$$\sigma(\nu) = \sigma_0 \phi(\nu), \quad (2.4)$$

where $\phi(\nu)$ is the line profile of the cross section and is suitably normalized such that $\int_0^\infty \phi(\nu) d\nu = 1$.

In a naïve approach, the line profile can be written as a simple Dirac delta function:

$$\phi(\nu) = \delta(\nu - \nu_0), \quad (2.5)$$

where the atom only transitions when interacting with a photon of frequency ν_0 . However, due to the Heisenberg uncertainty principle $\Delta E \Delta t \sim \hbar$ and the random thermal motions of the gas, the gas at a particular frequency will interact with all photons $\nu \approx \nu_0$. The exact width of the cross section in frequency space can be determined by analyzing the process in greater depth.

When analyzing the interaction between photons and bound electrons, it suffices to treat the atom quantum mechanically but the radiation field classically. Using this

approach, the cross section of neutral atoms can be modeled as a driven harmonically bound particle. The line profile $\phi(\nu)$ is sharply peaked about the natural frequency of the system $\omega_0 = 2\pi\nu_0$. The total cross section can be written as (Rybicki & Lightman, 1986):

$$\phi(\nu) = \frac{\gamma/2\pi}{(\nu - \nu_0)^2 + (\gamma/2)^2}, \quad (2.6)$$

where γ is the rate of spontaneous decay of the atom to the target energy level. In the case of an atom with a purely Coulomb potential (*i.e.*, single-electron atoms such as neutral hydrogen H I, ionized helium He II, doubly ionized lithium Li III, etc.), an alternative approach to finding the line profile is to write down the wavefunction for each quantum state, and then calculate the matrix element between the two states assuming an electronic dipole transition. In this approach, the finite width of the cross section is attributed to the Heisenberg uncertainty principle. The form of the profile in Equation (2.6) is called a Lorentz profile, and is an intrinsic property of the atom.

In addition to the natural broadening, the random thermal motions of the gas within a cloud further broadens the width of the line profile. Microscopically, some of the gas with velocity toward the observer will appear to absorb at a slightly higher frequency, whereas gas moving away from the observer will appear to absorb at a lower frequency. The shift in frequency therefore depends on the component of velocity along the line of sight v_z (taken to be the z -axis here). To lowest order in v/c , the shift in frequency can be written as:

$$\nu - \nu_0 = \frac{\nu_0 v_z}{c}. \quad (2.7)$$

If the particles of gas in the cloud have a temperature T and individual mass m_g , the number of particles dN_g found within a differential velocity range dv_z is given by a Maxwellian distribution:

$$dN_g = \exp\left(-\frac{m_g v_z^2}{2kT}\right) dv_z, \quad (2.8)$$

where k is the Boltzmann constant. Rearranging Equation (2.7) to isolate for the velocity leads to the relations:

$$v_z = \frac{c(\nu - \nu_0)}{\nu_0} \quad (2.9)$$

$$dv_z = \frac{c d\nu}{\nu_0}, \quad (2.10)$$

which means that Equation (2.8) can be written in terms of frequency as:

$$dN_g \propto \exp\left(-\frac{m_g c^2 (\nu - \nu_0)^2}{2\nu_0^2 kT}\right) d\nu. \quad (2.11)$$

The line profile is proportional to this number, and should be properly normalized. After normalization, the line profile can be written as:

$$\phi(\nu) = \frac{1}{\Delta v_D \sqrt{\pi}} \exp\left(-\frac{(\nu - \nu_0)^2}{(\Delta v_D)^2}\right), \quad (2.12)$$

where Δv_D is the Doppler width:

$$\Delta v_D \equiv \frac{v_0}{c} \sqrt{\frac{2kT}{m_g}}. \quad (2.13)$$

In practice, the line profile of hydrogen in the IGM will be a convolution of the Lorentz (Equation (2.6)) and the Doppler (Equation (2.13)) profiles. Essentially, each atom with velocity along the line of sight v_z will have its own Lorentz profile interacting with the radiation field. The result is an integral over all possible velocities:

$$\phi(\nu) = \int_{-\infty}^{\infty} dv_z \frac{1}{\sqrt{2\pi kT/m_g}} \exp\left(\frac{-v_z^2}{(2kT/m_g)}\right) \frac{\gamma/2\pi}{(\nu - \nu_0 - \nu_0 v_z/c)^2 + (\gamma/2)^2}. \quad (2.14)$$

One can define the quantities a and u such that:

$$a \equiv \frac{\gamma}{4\pi\Delta v_D}; \quad (2.15)$$

$$u \equiv \frac{\nu - \nu_0}{\Delta v_D}. \quad (2.16)$$

Including these definitions, Equation (2.14) becomes:

$$\phi(\nu) = \frac{H(a, u)}{\Delta v_D \sqrt{\pi}}, \quad (2.17)$$

where $H(a, u)$ is the Voigt-Hjerting integral (Hjerting, 1938):

$$H(a, u) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{\exp(-y^2)}{a^2 + (u - y)^2} dy. \quad (2.18)$$

The optical depth of the IGM to incoming radiation can then be calculated using the properties of the gas, such as the ionization level, temperature, and peculiar velocities. Essentially, using the Voigt profile of the gas from Equation 2.17, each parcel along the line of a mock Lyman- α sightline contributes to every pixel along the line of sight, in principle. In practice, the gas only contributes to nearby pixels in redshift space, and so the proper calculation of the optical depth can be sped up by only considering pixels for which the contribution is large. Further discussion of calculating the optical depth from simulations can be found in Chapter 5

Chapter 3

Predictions for the 21 cm Signal Incorporating the Light Cone Effect

3.1 Introduction

During the cosmological dark ages, the massive components of the universe were largely cold dark matter and neutral hydrogen. As the first stars and galaxies began to form, the UV photons emitted into the surrounding intergalactic medium (IGM) reionized the hydrogen. This phase transition is known as the Epoch of Reionization (EoR, Loeb & Furlanetto 2012). During the reionization process, it is expected that ionized hydrogen formed bubbles in the IGM surrounding stars, creating patches of reionized gas. As the photons travelled further out into the IGM, the ionized bubbles grew larger, until they eventually joined together. Subsequently, most of the remaining neutral hydrogen was localized to the inside of galaxies, with the rest of the IGM being highly ionized. For reviews of the EoR, see Furlanetto et al. (2006), Morales & Wyithe (2010), Loeb & Furlanetto (2012), and Pritchard & Loeb (2012).

This currently accepted description is overly simplistic because the precise details of reionization are still largely unknown. From observing the Gunn-Peterson absorption trough (Gunn & Peterson 1965) in the Ly α forest, we can infer that the global neutral hydrogen fraction f_{HI} was greater than 10^{-3} until $z \sim 6$ (Fan et al. 2006). Recent probes of the cosmic microwave background radiation (CMB) such as the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) and *Planck* have measured the Thomson optical depth of the IGM, which is a measure of the integrated electron density (Hinshaw et al., 2013; Planck Collaboration et al., 2013). *WMAP*-9 reports a value of $\tau = 0.089 \pm 0.014$, which assuming an instantaneous reionization gives $z_{\text{reion}} = 10.6 \pm 1.1$. Another experimental constraint comes from using the Hubble Space Telescope Ultra Deep Field observations of the very first galaxies, which contains information about the UV luminosity of star-forming galaxies at early times (Robertson et al., 2013).

One of the most promising tools for further probing this epoch comes from the

hyperfine transition of neutral hydrogen. The rest-frame wavelength of this transition is $\lambda \approx 21$ cm. The precise nature of the 21 cm signal depends on several factors, including when the midpoint of reionization occurred, the duration of reionization, and the dominant method by which hydrogen is reionized (*e.g.*, ionization via UV vs. x-ray photons). When making a measurement using the 21 cm brightness temperature, one can observe the *global signal* or the *power spectrum*. The former is the brightness temperature average over the entire sky, which during reionization is $\mathcal{O}(10)$ mK. The power spectrum is a statistical measure of the fluctuations in the field as a function of k -space. More information about the importance of the 21 cm signal can be found in, for example, Loeb & Zaldarriaga (2004), Cooray (2004), and Bharadwaj & Ali (2004).

With the advent of large radio-telescope and dipole arrays constructed specifically to observe the EoR, there have recently been several exciting advances regarding 21 cm observations. Some of the observational probes that are currently taking EoR data (or will be in the near future) are, for example, the Low Frequency Array (LOFAR¹; Harker et al., 2010), the Precision Array for Probing the Epoch of Reionization (PAPER²; Parsons et al., 2010), the Giant Metrewave Radio Telescope (GMRT³; Pen et al., 2009), the Murchison Widefield Array (MWA⁴; Bowman et al., 2005), and the Experiment to Detect the Global EoR Step (EDGES⁵; Bowman & Rogers, 2010). These arrays are designed to extract the 21 cm signal over a relatively narrow frequency band, targeting a particular redshift. An upcoming telescope, such as the Square Kilometer Array (SKA⁶; Mellema et al., 2013), will be designed to take full tomographic data of the EoR, and map the 21 cm signal as a function of frequency.

When performing a three dimensional measurement of the 21 cm signal, there are several important caveats to bear in mind. Two of the major effects are the light cone effect and redshift space distortions (RSD). The light cone effect comes purely from the time delay of propagation of the signal to the observer. In general, different comoving distances from an observer correspond to different points in redshift space. For sufficiently large scales, the comoving distance spanned by the observed volume corresponds to a large duration in redshift space. The neutral hydrogen fraction can change significantly if the length of the observed redshift interval is comparable to or larger than the duration of reionization. This evolution of the neutral fraction also introduces anisotropy along the line of sight in the 3D power spectrum. The light cone effect has been explored with respect to 21 cm observations semi-analytically by Barkana & Loeb (2004) and numerically by Datta et al. (2012). In previous works, the light cone was deemed to have a significant effect on the 21 cm brightness

¹www.lofar.org

²eor.berkeley.edu

³gmrt.ncra.tifr.res.in

⁴www.mwatelescope.org

⁵www.haystack.mit.edu/ast/arrays/Edges

⁶www.skatelescope.org

temperature two-point correlation function or power spectrum, respectively. We show in this work that the light cone can have a similar effect for sufficiently large volumes. Furthermore, we show that the light cone is most important around the midpoint of reionization, where $0.4 \lesssim f_{\text{HI}} \lesssim 0.6$.

RSD are the result of peculiar velocities of the signal sources. Since the simplest computation of the 21 cm signal assumes that the only source of velocity is the Hubble flow, peculiar velocities lead to a correction of the predicted signal. The effect of RSD has already been applied to 21 cm cosmology (*e.g.*, Barkana & Loeb, 2005; Bharadwaj & Ali, 2005; Mao et al., 2012; Jensen et al., 2013; Shapiro et al., 2013; Majumdar et al., 2013). In general, RSD have a significant effect on the 3D 21 cm brightness temperature power spectrum at the largest scales. The effects of RSD are thought to be most prominent early in reionization. For example, Jensen et al. (2013) show that RSD are most important for $0.7 \lesssim f_{\text{HI}} \lesssim 1.0$, peak at $f_{\text{HI}} \sim 0.9$, and have little impact after the midpoint of reionization.

The light cone effect also has important implications for measurements that use the baryon acoustic oscillation (BAO) method. The BAO method is important for understanding the accelerating expansion of the universe, and is used to make measurements of fundamental parameters such as $H(z)$. The BAO scale is large, typically 150 comoving Mpc. As we show, the light cone effect also becomes important on these scales. The BAO method can be subjected to the Alcock-Paczyński test (Alcock & Paczyński, 1979), which uses spherical features and relates their angular diameter distance to their extent in redshift space to determine cosmological parameters. Proper application of this test requires an accurate understanding of any anisotropies between perpendicular and parallel behavior of these features. As is discussed more in the body of this chapter, the light cone effect can introduce anisotropy in the 21 cm signal in the parallel direction. Therefore, if the 21 cm signal is to be used in BAO methods, the light cone effect must be properly understood and included in calculations. For application of the BAO method to the 21 cm signal, see Nusser (2005) and Barkana (2006); for discussion of the BAO theory and current implementations, see Weinberg et al. (2012).

Our approach combines numerical simulations with semi-analytic tools. We first perform a reionization simulation including hydrodynamics and radiative transfer on a relatively small volume. Once a statistical measure has been devised for how the matter overdensity field is related to the redshift of reionization, this statistical measure is used on a matter-only simulation in a larger volume that still accurately predicts reionization observables. In addition, different reionization histories can be explored rapidly without rerunning computationally expensive simulations. For a more thorough explanation of the general method outlined here, see Battaglia et al. (2013b). For applications of this method to EoR observables related to the CMB, see Natarajan et al. (2013) and Battaglia et al. (2013a).

The main purpose of this work is to quantify how the 21 cm power spectrum signal changes with the inclusion of the light cone effect. In Sec. 3.2, we discuss the

methodology behind the analysis and briefly describe the numerical techniques being applied. In Sec. 3.3, we discuss the basic science of the 21 cm brightness temperature power spectrum, and the types of statistical tests we perform on the data. Also in this section, we examine the application of these tests to data which comes from performing the analysis on a simulation box at a single redshift snapshot. (Hereafter, we refer to this type of data as “coeval cubes.”) In Sec. 3.4, we discuss the light cone effect on the 3D power spectrum. Then, in Sec. 3.5, we talk about specific applications to various observational endeavors, and how this signal might appear in real-world measurements. In Sec. 3.6, we discuss other effects and potential difficulties related to the 21 cm signal. To conclude, in Sec. 3.7, we talk about future prospects and outlooks. We assume a Λ CDM cosmology with $\Omega_\Lambda = 0.73$, $\Omega_m = 0.27$, $\Omega_b = 0.045$, $h = 0.70$, and $\sigma_8 = 0.80$. These values are consistent with the *WMAP-9* results (Hinshaw et al., 2013).

3.2 Methodology

In Battaglia et al. (2013b), a semi-analytic model was developed for relating the matter content in a computational simulation cell with the redshift at which the cell becomes 90% ionized. This approach exploits the fact that the matter overdensity field, defined as

$$\delta_m(\vec{x}) \equiv \frac{\rho_m(\vec{x}) - \bar{\rho}_m}{\bar{\rho}_m}, \quad (3.1)$$

is highly correlated with fluctuations in the redshift of reionization field ($z_{\text{re}}(\vec{x})$) defined as

$$\delta_z(\vec{x}) \equiv \frac{[z_{\text{re}}(\vec{x}) + 1] - [\bar{z} + 1]}{\bar{z} + 1}, \quad (3.2)$$

on large scales ($\gtrsim 1 \text{ Mpc}/h$) (Battaglia et al., 2013b). To motivate this observation, note that in an “inside-out” reionization scenario, the densest regions are the ones which form stars and galaxies capable of producing reionizing photons the earliest. The difference in amplitude between the two fields can be quantified using the bias parameter $b_{zm}(k)$ which is applied to the two fields in Fourier space. The bias parameter can be written as:

$$b_{zm}^2(k) \equiv \frac{\langle \delta_z^* \delta_z \rangle_k}{\langle \delta_m^* \delta_m \rangle_k} = \frac{P_{zz}(k)}{P_{mm}(k)}, \quad (3.3)$$

where $P_{xx}(k)$ is the auto-power spectrum of a field δ_x . In order to quantify how similar two fields are, the cross-correlation coefficient r can be used. This quantity can be defined as:

$$r_{zm}(k) \equiv \frac{\langle \delta_z^* \delta_m \rangle_k}{\sqrt{\langle \delta_z^2 \rangle_k \langle \delta_m^2 \rangle_k}} = \frac{P_{zm}(k)}{\sqrt{P_{zz}(k) P_{mm}(k)}}, \quad (3.4)$$

where $P_{xy}(k)$ is the 3D cross-power spectrum of the fields δ_x and δ_y . The normalization ensures that $r \in [-1, 1]$. For values where the cross-correlation coefficient becomes 1, the fields are highly correlated, and the amplitudes of the fields differ only by their bias factor. This is true for the matter and reionization fields during the EoR on large scales (Battaglia et al., 2013b).

Since the matter and reionization fields are highly correlated on large scales, the bias parameter can be used to relate their amplitude difference. In general, the bias will change as a function of k . We have chosen a functional form of the bias defined in Equation (3.3) in such a way to reproduce the relationship observed in simulations. We define this bias b_{zm} to be:

$$b_{zm} = \frac{b_0}{\left(1 + \frac{k}{k_0}\right)^\alpha}. \quad (3.5)$$

There are essentially three free parameters in this model: b_0 , k_0 , and α . The value of b_0 can be predicted using excursion set formalism in the limit that $k \rightarrow 0$ (Barkana & Loeb, 2004). We have chosen b_0 to be 0.593.

In order to determine best-fit values for the parameters k_0 and α , we compare the matter overdensity and reionization-redshift fields using a RadHydro code, which contains radiative transfer + hydrodynamics + N -body simulation (Trac et al., 2008). These particular simulations contain 2048^3 dark matter particles, 2048^3 gas cells, and 17 billion adaptive rays in a 100 Mpc/ h cubical box. We find that the best fits for the values were $\alpha = 0.564$ and $k_0 = 0.185 h \text{ Mpc}^{-1}$. In addition to these physically motivated “fiducial” values, two other sets of values were chosen to represent more extreme reionization scenarios: a long and short reionization history, parameterized in our model with the values of $(\alpha, k_0) = \{(1.8, 0.1), (0.2, 0.9)\}$, respectively. Examining different reionization histories allows for the identification of features in the power spectrum which may indicate how quickly reionization occurred.

Once the values in the bias relationship have been fixed, the matter overdensity field can be used in order to construct the reionization-redshift field. Accordingly, we performed a dark-matter-only simulation with a particle-particle-particle-mesh (P³M) N -body code using 2048^3 dark matter particles in a 2 Gpc/ h box. Then, using a snapshot of the matter overdensity field at the midpoint of reionization \bar{z} , we apply the bias relation in Equation (3.5). For more details on this method, see Battaglia et al. (2013b).

Figure 3.1 shows a plot of the ionization fraction of the simulation volume, both mass- and volume-weighted. All neutral fractions reported in the rest of this work, unless otherwise noted, are mass-weighted. The duration of reionization is measured by finding the redshift range for when the simulation cube is 25% ionized to 75% ionized, which measures the “50% ionization width” Δz_{50} . The 50% reionization duration in redshift for the long, fiducial, and short cases (weighted by mass) are: $\Delta z_{50} = 2.11, 1.10,$ and 0.24 . The reionization model considered here does not allow

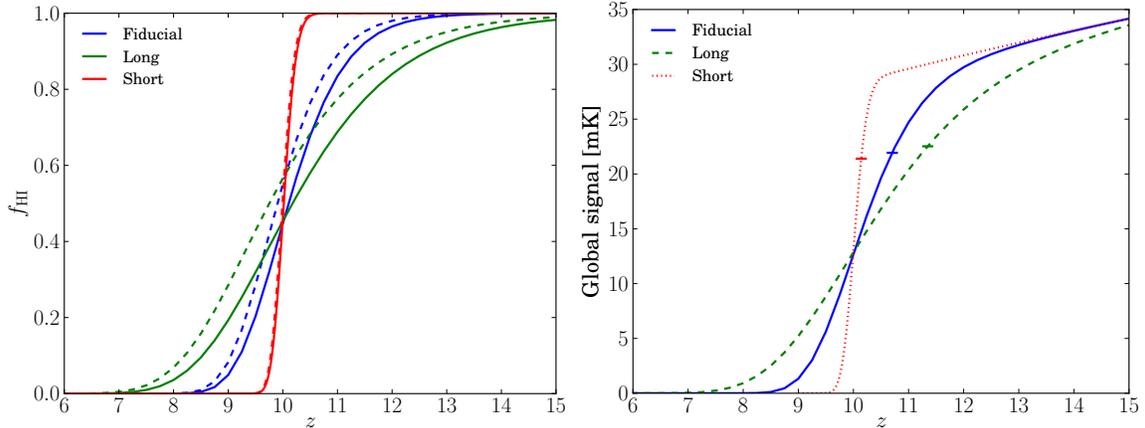


Figure 3.1: Left: the average neutral hydrogen fraction as a function of redshift. Plotted are the mass-weighted average (solid lines) and the volume-weighted average (dashed lines). In an inside-out reionization scenario, the densest regions of the universe are the first ones to reionize, so the mass-weighted neutral fraction is always lower than the volume-weighted one. Right: the global 21 cm signal as a function of redshift for the different reionization histories from Equation (3.6), with the simplifying assumption that $T_S \gg T_{\text{CMB}}$. This approximation is only physically justifiable for mass-weighted neutral fractions of $f_{\text{HI}} \lesssim 0.75$; nevertheless, we plotted the global temperature predicted by Equation (3.6) for higher redshifts since it is still approximately true in this redshift range. We have marked the points where $f_{\text{HI}} \sim 0.75$ by small ticks on the lines. By construction, all of the histories have the same midpoint of reionization of $\bar{z} = 10$, which accounts for the point of intersection.

for “exotic” reionization scenarios, such as extended reionization or recombination before a second ionization.

3.3 Analysis

3.3.1 21 cm Theory

The 21 cm signal tracks regions of neutral hydrogen in the IGM. The application of the radiative transfer equation to CMB photons free-streaming from the surface of last scattering and passing through neutral hydrogen in the intergalactic medium predicts whether the neutral hydrogen will absorb or emit radiation at 21 cm.

The difference between the brightness temperature and the temperature of the

CMB is given as (Madau et al., 1997; Harker et al., 2010):

$$\begin{aligned} \frac{\delta T_b}{\text{mK}} &= 38.6 h(1 + \delta_m)x_{\text{HI}} \left(\frac{T_S - T_{\text{CMB}}}{T_S} \right) \times \left(\frac{\Omega_b}{0.045} \right) \left[\left(\frac{0.27}{\Omega_m} \right) \left(\frac{1+z}{10} \right) \right]^{\frac{1}{2}} \\ &= T_0(z)(1 + \delta_m)x_{\text{HI}}, \end{aligned} \quad (3.6)$$

where x_{HI} is the neutral hydrogen fraction (assumed to be 0 or 1 for an individual gas cell), and T_0 is the redshift-dependent ‘‘average temperature’’ of the signal, which is modulated by the spatial fluctuations of the matter overdensity field and the ionization state. This analysis was performed in a regime where Ω_Λ can be safely ignored. Equation 3.6 gives the difference of the brightness temperature at a frequency corresponding to 21 cm from the CMB as a function of redshift and spatial position.

In the following analysis, it has also been assumed that the spin temperature is large compared to the CMB temperature, $T_S \gg T_{\text{CMB}}$. Following from the results of Santos et al. (2008), this factor is approximately 1 for mass-weighted neutral fractions $f_{\text{HI}} \lesssim 0.75$. Once the neutral fraction has reached this value, the spin temperature is collisionally coupled to the kinetic temperature of the gas, which is typically 2 orders of magnitude larger than the effective CMB temperature. This assumption is applicable to a large range of reionization scenarios, *e.g.*, ones where UV photons from stars photo-ionize and photo-heat the neutral hydrogen, so that hydrogen’s spin temperature couples to the kinetic energy of the gas particles and becomes much hotter than the CMB. Exotic reionization scenarios, *e.g.*, those where reionization is caused by x-ray heating, do not necessarily meet the condition that $T_S \gg T_{\text{CMB}}$. However, these scenarios have not been examined in this analysis, and these considerations have been saved for future work.

Figure 3.1 shows the global 21 cm signal as a function of redshift for the different reionization scenarios. The duration of the reionization history affects the rate at which the global signal diminishes: the long reionization scenario drops gradually, whereas the short reionization scenario drop rapidly. Observationally, the signal from a shorter reionization scenario is easier to measure than a longer one (Bowman & Rogers, 2010), since a shorter reionization scenario would appear as a sharper feature in frequency space.

3.3.2 3D Power Spectrum

We define the 3D power spectrum as $P_{xx}(k) = \langle \delta_x^* \delta_x \rangle_k$, and the dimensionless power spectrum $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$. The features of the coeval matter overdensity field power spectrum have already been extensively explored, so we will only list some common features. As the universe evolves over time, the amplitude of the matter power spectrum increases monotonically. Because the 21 cm brightness temperature (*cf.* Equation (3.6)) is proportional to the matter overdensity field, one might expect the 21 cm brightness temperature power spectrum also to increase monotonically.

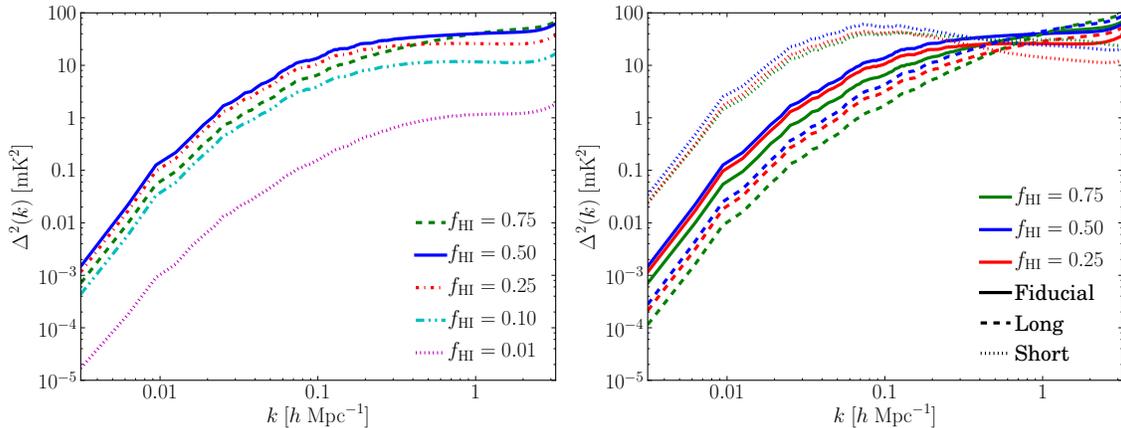


Figure 3.2: Left: a plot of the 3D 21 cm brightness temperature power spectrum, as a function of neutral hydrogen fraction. On large scales, the power peaks at $f_{\text{HI}} \sim 0.5$, but at smaller scales it peaks for a larger neutral fraction, $f_{\text{HI}} \sim 0.75$. The reason for this is that at a larger neutral fraction, only the densest regions are ionized, so the 21 cm power spectrum looks more like the matter power spectrum. At a neutral fraction of less than 50%, the differences in amplitude for different values of f_{HI} on large scales is roughly proportional to the difference in neutral fraction. This phenomenon is due to the fact that the redshift evolution of the 21 cm signal after the midpoint is dominated by the changing neutral fraction. Right: the evolution of the power spectrum for different reionization histories. Across all reionization histories, the power spectrum is larger near 50% ionization. The shape of the power spectrum changes dramatically for different reionization histories, where in general, a shorter duration of reionization implies more large-scale power and less small-scale power.

However, the 21 cm signal also incorporates the neutral hydrogen fraction, and so as the universe becomes increasingly ionized, the signal diminishes. This evolution causes the amplitude of the 21 cm signal to increase as the universe begins to ionize, peak at a particular neutral fraction, and then decrease as the universe ionizes further. The shape of the 21 cm power spectrum in the coeval case has also been examined (*e.g.*, Lidz et al., 2008), and the value corresponding to a peak in large scale power is $\sim 50\%$ ionization fraction.

We calculate the power spectrum as a function of neutral fraction, since equal neutral fractions between reionization scenarios capture the same physics better than equal redshifts. We linearly interpolate between matter overdensity fields from adjacent snapshots in order to create a power spectrum as a function of specific neutral fractions. The 21 cm brightness field was computed from this interpolated matter overdensity field using Equation (3.6) where x_{HI} was determined from $z_{\text{re}}(\vec{x})$.

Figure 3.2 shows the features of the 3D power spectrum from the fiducial reionization scenario. On large scales, the amplitude peaks near the midpoint of reionization,

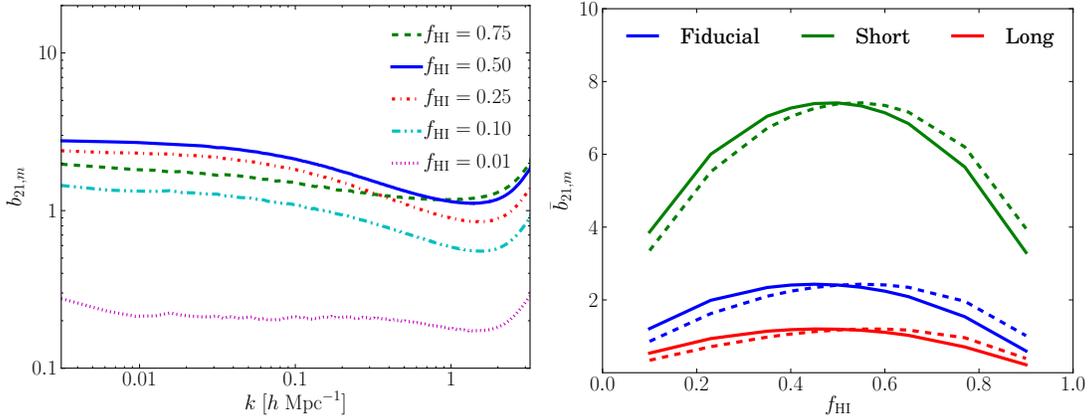


Figure 3.3: Left: the scale-dependent bias between the two fields, defined in Equation (3.3), for the fiducial reionization history. To remove the redshift dependence between different neutral fractions, we divide the 21 cm brightness temperature by $T_0(z)$ defined in Equation (3.6). One can see that the value is fairly constant in the region $k \lesssim 0.1 h \text{ Mpc}^{-1}$, leading to the choosing of this value for the large-scale bias parameter. The small-scale structure for large k -values changes noticeably as the universe becomes more ionized. Right: A plot of the large scale bias relationship between the 21 cm power spectrum and matter power spectrum at different neutral fractions. Shown are the mass-weighted neutral fraction (solid lines) and the volume-weighted neutral fraction (dashed lines). The bias is calculated according to Equation (3.7), which only takes into account the largest scales ($k < 0.1 h \text{ Mpc}^{-1}$). When the bias is largest, there is the most 21 cm signal relative to the underlying matter overdensity field. Note that as the reionization history becomes shorter, the bias becomes larger at all neutral fractions.

$f_{\text{HI}} \sim 0.5$. On small scales, the power is largest for the highest neutral fraction, $f_{\text{HI}} \sim 0.75$. Early on in reionization, only the densest regions have become ionized, which means the 21 cm brightness temperature power spectrum looks very similar to the matter power spectrum. As the universe becomes more ionized, this small-scale power is lost due to the ionized regions growing larger.

Figure 3.2 also shows the 3D power spectrum across the different reionization scenarios of our model. The general shape of the spectra changes dramatically as a function of reionization history: as the duration of reionization decreases, more power is transferred from small scales to large ones. For our model, although the underlying matter overdensity field is identical across the simulations, the reionization history dramatically changes the predicted shape of the 21 cm power spectrum.

3.3.3 Bias Parameter and Average Bias

Figure 3.3 shows a plot of the bias parameter (Equation (3.3)) between the 21 cm brightness temperature field and matter overdensity field. (The cross-correlation coefficient is discussed further in Sec. 3.4.2.) As already mentioned, the bias parameter can be used to quantify the relative amplitudes between the different power spectra. This quantity has already been applied in a number of settings (*e.g.*, Fry & Gaztañaga 1993, Heavens et al. 1998, Croft et al. 2002, etc.). The application at hand is the bias factor between the matter overdensity field and the 21 cm brightness temperature field. Note that for the calculation of the bias parameter, the average temperature $T_0(z)$ is divided out in Equation (3.6) (resulting in $\widehat{\delta T}_b = (1 + \delta_m)x_{\text{HI}}$), in order to remove dependence on redshift. For high values of the neutral fraction, the bias is flatter, meaning that the 21 cm brightness temperature field is more similar to the matter overdensity field. As the universe becomes more ionized, the bias changes more dramatically as a function of k . As in the case of the 3D power spectrum, the amplitude of the bias on large scales peaks at $f_{\text{HI}} \sim 0.5$. The evolution of the matter overdensity field is small compared to the change in the 21 cm brightness temperature.

In regions where the bias is roughly constant, an “average bias” can be defined as:

$$\bar{b}_{21,m} = \left\langle \sqrt{\frac{\delta_{21}^* \delta_{21}}{\delta_m^* \delta_m}} \right\rangle_{k < k_*} \quad (3.7)$$

where k_* ($0.1 h \text{ Mpc}^{-1}$) is a predefined cutoff value to ensure that the selected regime is relatively constant. For a given reionization history, the power spectrum of both the 21 cm field and the matter overdensity fields is calculated. The average of the ratio of the two power spectra is computed for all k -values out to k_* at several different values of the neutral fraction. The large scale bias is important because it predicts the amplitude of the 21 cm power spectrum compared to the matter power spectrum, especially at large scales. As in the case of the scale-dependent bias, the average temperature of the 21 cm brightness temperature field $T_0(z)$ has been divided out.

Figure 3.3 shows the average bias for the three different reionization scenarios. As seen in the figure, the bias peaks at an ionization fraction of roughly 50% by mass. As already discussed in Sec. 3.3.2, this coincides with the peak in the power of the 21 cm power spectrum. A large value of the bias implies that the sources of reionization are themselves “highly biased,” in the sense that they are larger and rarer for larger values of the bias parameter. Figure 3.4, which shows the 21 cm brightness field in the coeval and light cone cases, demonstrates this visually. In the coeval column on the left, the short reionization scenario has larger but fewer ionized regions, which implies that the sources are massive and rare. Indeed, the difference between the large voids in the case of short reionization and the small pockets of ionized gas in the long reionization is striking. Thus, the large scale bias parameter is important not only because it yields valuable information about the relation between the 21 cm

brightness temperature and the matter overdensity field, but also because it is related to the sources of reionization.

3.4 Light Cone Effect

The light cone effect on 21 cm power spectra has been examined semi-analytically in Barkana & Loeb (2006) and numerically in Datta et al. (2012). The previous numerical work was concerned only with relatively small volumes, and found that the light cone effect is a significant effect on their largest scales. We examined the impact of the light cone effect on volumes larger than those used by Datta et al. (2012), and we conclude that this effect is an essential consideration for 21 cm measurements.

In essence, the light cone effect is due to evolution of the signal along the line of sight. Although the coeval power spectrum is easy to compute in a simulation volume, it is not representative of a 3D power spectrum that would be observed. Given a flat Λ CDM cosmology, the comoving distance from an observer today can be calculated as a function of redshift:

$$r(z) = \int_0^z \frac{c}{H(z')} dz'. \quad (3.8)$$

As an example, if the center of the 2 Gpc/ h box is placed at a comoving distance corresponding to a redshift of $z = 10$ for our particular cosmology (*i.e.*, the 21 cm signal at the center of the box has a redshift of 10 relative to an observer), then a signal from the far side of the box (from the perspective of the observer) has a redshift of $z \sim 21$, whereas the near side of the box has a redshift of $z \sim 6$. The duration in redshift space spanned by the box is much larger than the Δz_{50} for all of the reionization histories of our model. This means that, even for very extended reionization scenarios, the far side of the box would correspond to a totally neutral universe, and the near side would be completely ionized. The matter overdensity field also evolves from $z \sim 21$ to $z \sim 6$. Intuitively, one would expect that such a radical change could affect the power spectrum of 21 cm, because the signal is dependent upon the presence of neutral hydrogen. In other words, the evolution along the line of sight is non-negligible for these large volumes.

To produce the light cone effect, we divided the full simulation volume into a series of cubes with smaller dimensions, since 2 Gpc/ h spans a redshift range that always exceeds the duration of reionization in our model. We treated these different sub-boxes as fully independent, because the matter overdensity field, which generates the reionization field, has the same statistical values (*e.g.*, mean value, standard deviation, σ_8 , etc.) in each sub-volume, with some acceptable fluctuation. Specifically, we cut the 2 Gpc/ h box into sub-volumes of 500 Mpc/ h , 250 Mpc/ h , and 125 Mpc/ h . This yields 64, 512, and 4096 independent cubes, respectively. We placed the center of the sub-boxes at the redshift corresponding to 25%, 50%, or 75% neutral hydrogen fraction by mass. For each cell in the simulation volume, the comoving distance r from the observer is calculated along with the redshift corresponding to that distance

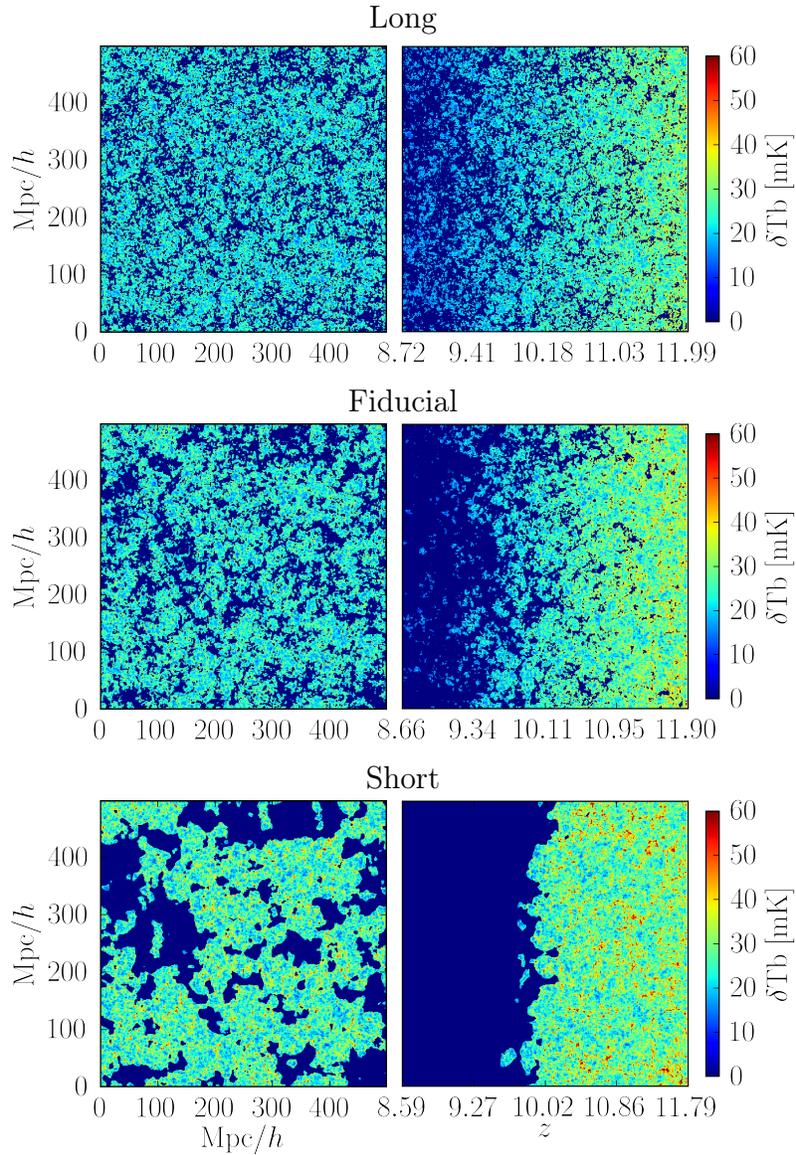


Figure 3.4: A visualization of the evolution of the 21 cm brightness of the simulation cube. Left: a coeval sub-box at 50% ionization fraction with side length of 500 Mpc/h for long (top), fiducial (middle), and short (bottom) reionization scenarios. Right: the corresponding light cone cube, which includes evolution of the ionization field. The x -axis on the right shows the redshift instead of the comoving distance, with the center of the box placed at the redshift equalling 50% ionization by mass. We notice that the 21 cm signal initially follows the underlying matter fluctuations at the side of the box farther from the observer where it is almost entirely neutral, then gradually fades to zero brightness as the IGM becomes increasingly ionized. For the long reionization scenario, the coeval case has smaller bubble sizes at 50% reionization, and the light cone effect is not as pronounced. For the short reionization scenario, the coeval case has very pronounced bubbles of ionized gas at 50% reionization, and the light cone effect is quite dramatic.

$z(r)$, the inverse of Equation (3.8). Then, the mass of the cell is linearly interpolated from the snapshots of coeval mass density arrays from the bracketing redshifts, just as is done for the coeval case. Finally, the 21 cm brightness temperature is computed as in Sec. 3.3.2.

Figure 3.4 shows the evolution of the 21 cm signal in the simulation volume as a function of redshift. One can see that the late-time portion (left side of the box) contributes almost nothing to the signal, and the earlier times (right side) has variation in the temperature proportional to the fluctuations in the matter overdensity field.

3.4.1 3D Power Spectrum with the Light Cone Effect

To determine the impact the light cone effect has on the 3D power spectrum, we find the power spectrum of each individual sub-box, take the average, and then compute the standard deviation to get the corresponding 1σ values. Because the simulation volumes were constructed in this way, periodicity was explicitly broken which altered the power on large scales. However, we found that this does not greatly affect the computation of the power spectrum. Furthermore, many of these results involve ratios between power spectra that are both affected by the problem of broken periodicity, so the problems introduced do not significantly change the predictions. Also note that when computing the 3D power spectrum with the light cone effect, Fourier modes where $k_{\perp} = 0$ relative to the line of sight have been removed. The inclusion of these modes leads to significantly more power on large scales, but they cannot be observed by radio interferometers. (See Appendix A.1 for more discussion.)

In Figures 3.5 and 3.6 (along with Figures A.1, A.2 and A.3 in Appendix A.2), we present the 3D power spectra with and without the light cone effect. Figure 3.5 compares the power spectra across box sizes and reionization histories, but with constant $f_{\text{HI}} = 0.5$. A general feature is that the power is suppressed at all scales. Including the light cone effect is somewhat analogous to averaging over the duration in redshift range spanned by the volume. For the large sub-volume size (500 Mpc/h), this leads to an effective averaging over a significant portion of the reionization history. This explains why there is less power on all scales: the neutral fractions where the large- and small-scale power peak ($f_{\text{HI}} = 0.50$ and $f_{\text{HI}} = 0.75$, respectively) are being averaged with other neutral fractions that contain less power. Thus, the averaging tends to decrease power on all scales for our reionization scenarios. Note also that in the limit where the redshift space duration is relatively small (*i.e.*, the 125 Mpc/h volume), there is little deviation from the coeval case.

One feature to point out is the 1σ spread of the power spectrum, represented by a shaded region surrounding the light cone line. As discussed in Sec. 3.4, the sub-boxes are treated as independent and identically distributed sub-samples of the larger volume. Specifically, we treat the power spectrum from each sub-volume as the random variable of an underlying cosmological distribution. The standard deviation calculated here is that of the power spectra themselves, computed over the 64, 512, or

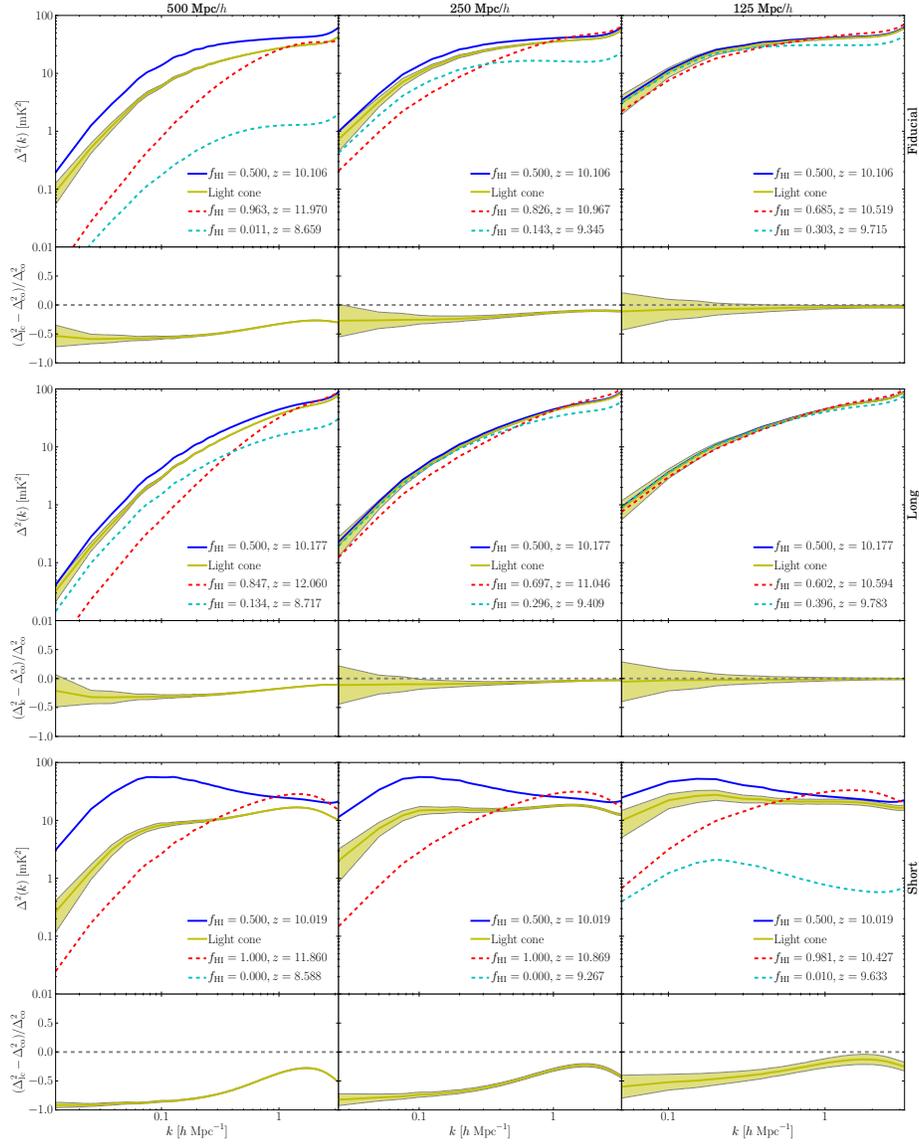


Figure 3.5: The light cone effect for a sub-box with side length 500 Mpc/h (left column), 250 Mpc/h (center column), and 125 Mpc/h (right column), for the fiducial (top row), long (middle row), and short (bottom row) reionization scenarios. The coeval power spectrum (solid blue line) is computed at the midpoint of reionization. The light cone effect (yellow line) has 1σ error regions shaded in. Note that these spectra do not include modes where $k_x = k_y = 0$ (see Appendix A.1). Also shown are coeval power spectra corresponding to the bracketing redshifts of the light cone cube for the far side from the observer (red dashed line) and the near side (cyan dashed line). The percent difference between the coeval and light cone lines is shown in the bottom panel, with the same 1σ error regions shaded in. The light cone effect is most pronounced at the largest scales. The light cone effect can also change the shape of the power spectrum, where a shorter reionization scenario leads to more deviation from the coeval case.

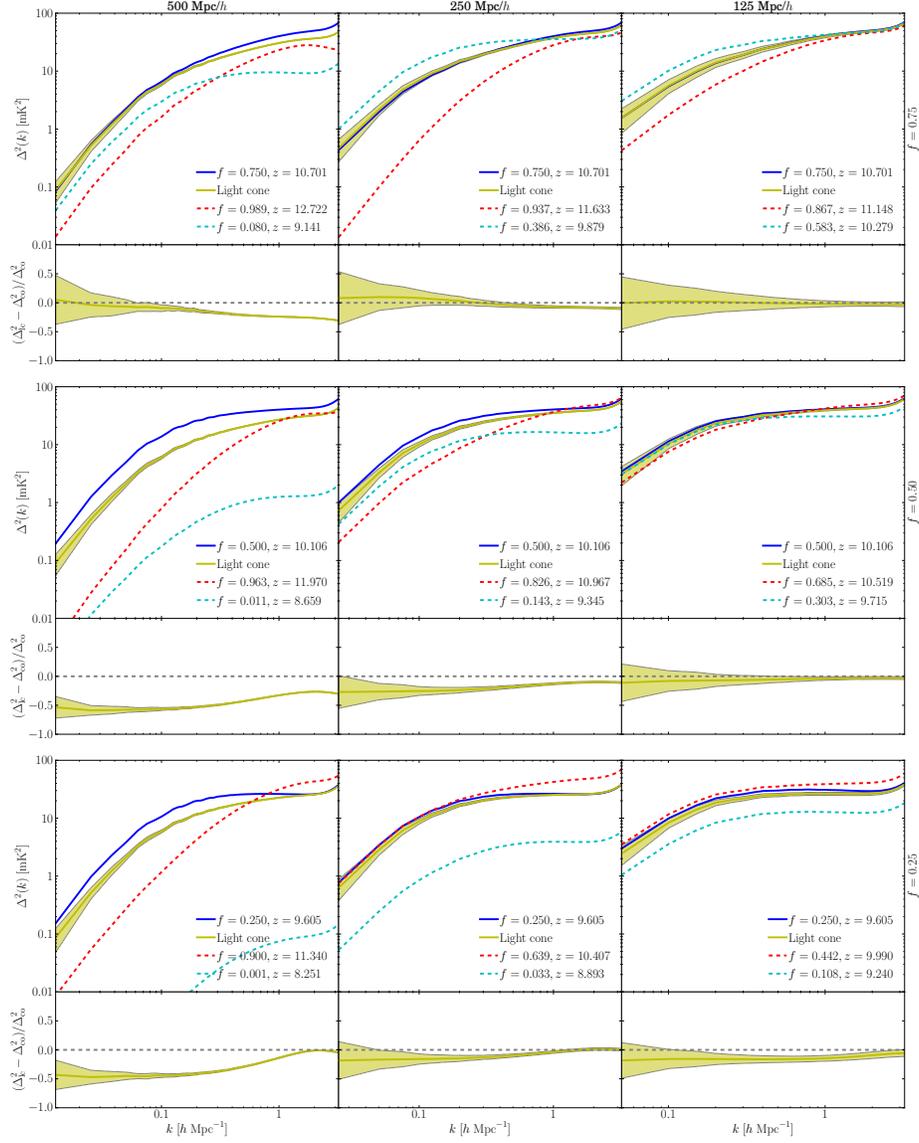


Figure 3.6: A plot similar to Figure 3.5, but showing the power spectrum as a function of neutral fraction. All plots are for the fiducial reionization history, with rows corresponding to $f_{\text{HI}} = 0.75$ (top), 0.50 (center), and 0.25 (bottom). The columns have their same ordering as in Figure 3.5. We can see that only the small-scale power changes appreciably between different neutral fractions. Thus, on large scales, only the coeval power spectrum changes shape appreciably. Compare the coeval shape change to Figure 3.2. This implies that the shape of the light cone power spectrum might not change as dramatically as in the coeval case, especially for large sample volumes.

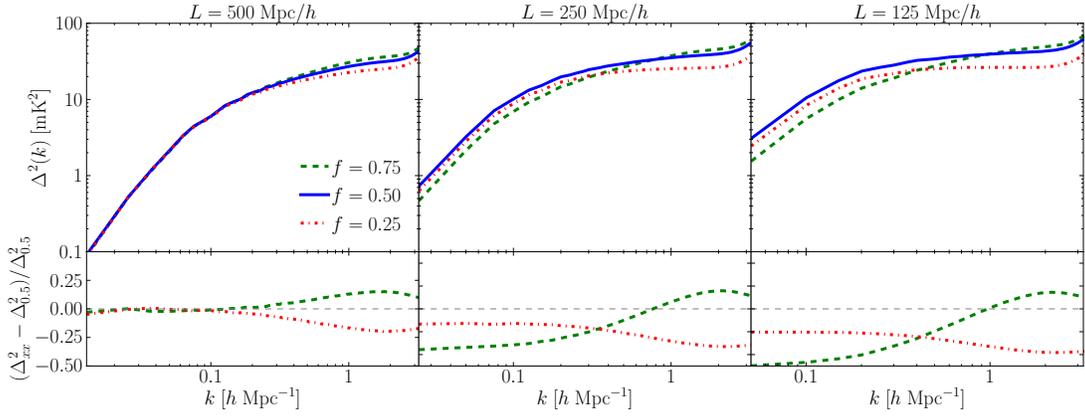


Figure 3.7: Top: light cone power spectra, plotted at different neutral fractions across all sub-box sizes for the fiducial reionization scenario. These spectra are the same as in Figure 3.6, but reproduced here for more straight-forward comparison. Bottom: percent difference of the $f = 0.75, 0.25$ spectra from $f = 0.50$ spectra. As the extent in redshift space becomes larger with respect to the duration of reionization, the large scale power becomes increasingly similar across neutral fractions. This effect is most apparent in the short scenario, but also partially seen in the long scenario.

4096 sub-boxes for a particular sub-volume size. The relatively larger spread for the smaller sub-volumes demonstrates there is more fluctuation when examining smaller scales.

Figure 3.6 compares the 3D power spectrum across different box sizes and different neutral fractions, but only for the fiducial reionization scenario. Similar to Figure 3.5, the light cone enhances power at large scales and diminishes power at small scales. Another similarity is that the deviation from the coeval case is greater for large sub-volumes than for small ones. An interesting feature of these plots is that the shape of the light cone power spectrum does not change as drastically for different ionization fractions as it does for different reionization histories in our model. This implies that differences in the shape of the power spectrum are most sensitive to the duration of reionization, and are not as dependent on the midpoint of reionization.

Figure 3.7 shows the power spectra for the fiducial reionization scenario at different neutral fractions. As the sub-box size becomes larger and the extent in redshift space becomes large compared to the duration of reionization, the large-scale power of the different neutral fractions becomes increasingly similar. This is due to how the region of maximal contrast near $f_{\text{HI}} \sim 0.5$ relates to where the box is centered in redshift space. Since the light cone cube is centered on the redshift corresponding to a particular neutral fractions, longer reionization scenarios will have a greater change in where the cubes are centered. Sub-box sizes where the region of maximal contrast is adequately spanned for all neutral hydrogen fractions will have similar amounts of large scale power. The Figure shows this is true for the largest sub-box

size in the fiducial reionization scenario. As an observational implication, our model predicts that future measurements will not be able to easily distinguish different neutral fractions for briefer reionization scenarios.

Related to this phenomenon, the average bias (*cf.*, Sec. 3.3.3) also behaves differently when the light cone effect is included. In the coeval case, the average bias is initially relatively small early in reionization, rises with increased ionization, and then falls following the midpoint of reionization (see Figure 3.3). The inclusion of the light cone effect flattens out this curve, so that the average bias does not change significantly as a function of neutral fraction. Again, this phenomenon is related to the duration of reionization compared to sub-box size, with briefer reionization scenarios being flatter. This result further demonstrates that when the light cone effect is included, it becomes difficult to determine the change in neutral fraction as a function of redshift. One alternative to the 3D power spectrum would be to measure the 2D angular power spectrum as a function of frequency, where the large scale bias would likely rise and fall as a function of neutral fraction in a manner similar to the coeval case.

3.4.2 Cross-correlation Coefficient with the Light Cone Effect

We examined the cross-correlation coefficient between the 21 cm brightness temperature and the matter overdensity fields for the light cone effect. We computed the cross-correlation between the two fields using Equation (3.4). Figure 3.8 shows the cross-correlation coefficient for the light cone. In general, on large scales the fields show less statistical correlation than in the coeval case. We can motivate this by noting that when the box is completely neutral, there is perfect correlation between the two fields. Conversely, once the box becomes totally ionized, there is no longer any correlation between the matter overdensity and 21 cm fields, because the 21 cm signal is zero everywhere. Because this effect is more pronounced in the short reionization scenario (*cf.* Figure 3.4), the short histories (the dotted lines in Figure 3.8) deviate the most from perfect anti-correlation. In fact, the combination of zero correlation in ionized regions and almost perfect correlation in neutral regions accounts for why the short reionization scenario exhibits a large degree of positive correlation on small scales. The amount of anti-correlation grows larger for longer reionization scenarios, and the fields tend toward perfect anti-correlation on large scales for the fiducial and long reionization scenarios.

3.4.3 Anisotropic Power Spectrum

We are interested in quantifying any anisotropy in the power spectrum because the light cone effect inherently alters the signal along the observer’s line of sight of the volume, but does not affect the signal perpendicular to the line of sight. The computation of an anisotropic power spectrum proceeds in a fashion similar to that of

the 3D power spectrum (as in Sec. 3.3.2); however, instead of binning in terms of a single spherical magnitude $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$, the binning is done in terms of two quantities $k_{\parallel} \equiv k_z$ and $k_{\perp} \equiv \sqrt{k_x^2 + k_y^2}$. When decomposing the power spectrum in this manner, we use a “flat-sky” approximation which neglects the curvature of the sky. In our calculations, the distance to the observer is large enough that the effects of the flat-sky approximation are negligible. Additionally, we noticed that on small scales, there is a significant amount of anisotropy in the Figure even in the coeval matter power spectrum. The density field is constructed by assigning particles to a Cartesian grid using an anisotropic cubical top hat filter. Deconvolution with this filter is not perfect, and does not completely remove the anisotropy. The deviation from isotropy becomes increasingly important on scales that are close to the size of a grid cell. Accordingly, we only trust this statistic for which $k \lesssim 1 h \text{ Mpc}^{-1}$.

Figure 3.9 shows a pseudo-color plot in which the different k -modes k_{\perp} and k_{\parallel} are on the x - and y -axes, respectively. The power spectrum $P(k)$ is plotted as a function of these two modes on a linear scale, so that the isotropy (or anisotropy) is apparent in the plot. An interesting feature to point out is that the light cone introduces a subtle deviation from the isotropy seen in the coeval case. There is slightly less power in modes where $k_{\perp} \sim k_{\parallel}$ compared to modes where $k_{\perp} \gg k_{\parallel}$ or vice versa. Compare this to the coeval case, where the contours are almost perfectly circular with little deviation from isotropy. The anisotropy indicates there is more power for volumes with small extent in redshift space or small extent in the plane of the sky, compared to ones where the extent is almost equal.

In the case of an isotropic box with no preferred direction (*e.g.*, a coeval cube containing the matter overdensity field), one would expect the contours of equal power to be roughly circular, because there should be equal contributions in all directions without a preferred orientation. When the light cone effect is included, we find that there is generally less power at all scales $k \lesssim 1 h \text{ Mpc}^{-1}$, which is consistent with Figure 3.5. Figure 3.5 demonstrates that including the light cone effect leads to a similar spectrum but with less power at all scales, and Figure 3.9 shows that there is little anisotropy introduced by the effect.

3.4.4 Power Wedges

We quantify the anisotropy produced by the light cone effect using a tool we name “power wedges,” in analogy to the “clustering wedges” tool recently introduced in BAO analysis for the two-point correlation function (*e.g.*, Kazin et al., 2013; Sánchez et al., 2013). To perform the power wedges analysis, the plane of k_{\parallel} and k_{\perp} is bisected along the line $k_{\parallel} = k_{\perp}$. Then, the power corresponding to these combinations is binned as a function of k . This process produces decompositions P_{\parallel} and P_{\perp} , where $k_{\parallel} > k_{\perp}$ or vice versa. Finally, the ratio of the power spectra ($\chi(k)$) is taken:

$$\chi(k) \equiv \frac{\langle P_{\parallel}(k) \rangle}{\langle P_{\perp}(k) \rangle}. \quad (3.9)$$

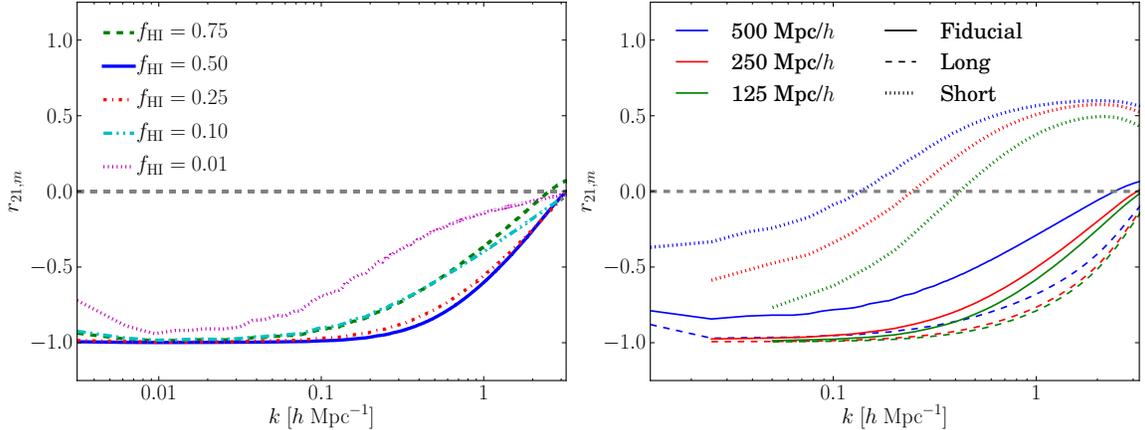


Figure 3.8: Left: the cross-correlation coefficient between the 21 cm brightness temperature and the matter overdensity field, defined in Equation (3.4), for the fiducial reionization history plotted at different neutral fractions. On very large scales, there is almost perfect anti-correlation between the two fields. Right: cross-correlation coefficient including the light cone effect, across different sub-box sized and reionization scenarios at constant $f_{\text{HI}} = 0.5$. For the fiducial and long reionization scenarios, there is generally a tendency toward -1 on large scales, though the anti-correlation is not as pronounced as in the coeval case. However, this correlation does not exist to the same extent for the short reionization scenario.

In the case that the k -values are equal, the contribution to the power is added to both spectra. In a perfectly isotropic case, this parameter should be equal to 1 (with some fluctuation). If the parameter is greater than 1, then there is more power coming from the modes along the line of sight of the simulation box, and vice versa.

Figure 3.10 shows the results of using the power wedges analysis. One can see that χ changes noticeably as a function of reionization history. For the fiducial and long histories, the value is very close to 1, meaning that the signal is isotropic. However, the short history demonstrates a moderate degree of anisotropy on large scales (almost $\sim 40\%$). The deviation from 1 becomes less as the sub-box size becomes smaller. Physically, the shorter reionization scenario displays a greater change in the variance of the 21 cm signal along the line of sight compared to the long reionization scenario.

Another interesting result evident in Figure 3.10 is how the evolution within the volume affects $\chi(k)$. There is a much larger deviation from unity for the case of the short reionization scenario compared to the fiducial and long ones. For a larger sub-box size, there is more evolution of the neutral hydrogen fraction, especially for the short reionization scenario. Because the anisotropy induced depends on this evolution, the larger anisotropy for larger box sizes makes sense.

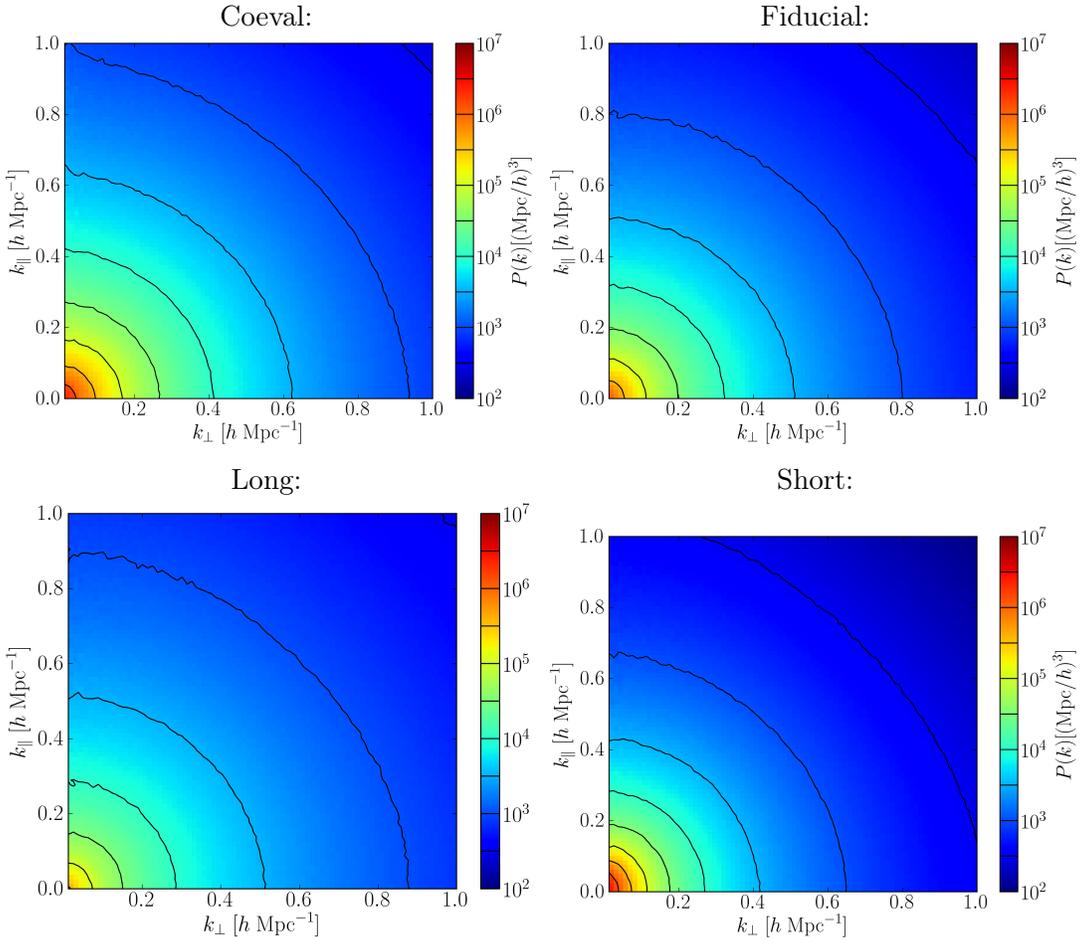


Figure 3.9: A plot of the anisotropic power spectrum, broken down into parallel and perpendicular Fourier modes. Top left: the anisotropic power spectrum of the coeval 21 cm field, $f_{\text{HI}} = 0.5$, fiducial reionization scenario. As with the 3D power spectrum, the result has been averaged over the 64 independent 500 Mpc/ h sub-boxes. Top right: same plot, but including the light cone effect and all Fourier modes for the fiducial reionization scenario. Here, k_{\parallel} is taken to be along the line of sight and coincident with the direction of the light cone effect. Both cases appear similarly isotropic. The anisotropy changes slightly based on the reionization history, especially in the short case. The effect is also more pronounced for larger scales. Bottom: the long and short reionization scenarios, respectively.

3.4.5 Comparison to Previous Work

As mentioned previously, the light cone effect has been investigated in Datta et al. (2012). The work presented here differs from the previous one in several key aspects. First, some of the volumes considered here are significantly larger. The simulation volume in the previous work was 163 Mpc (≈ 114 Mpc/ h) on a side, compared to the

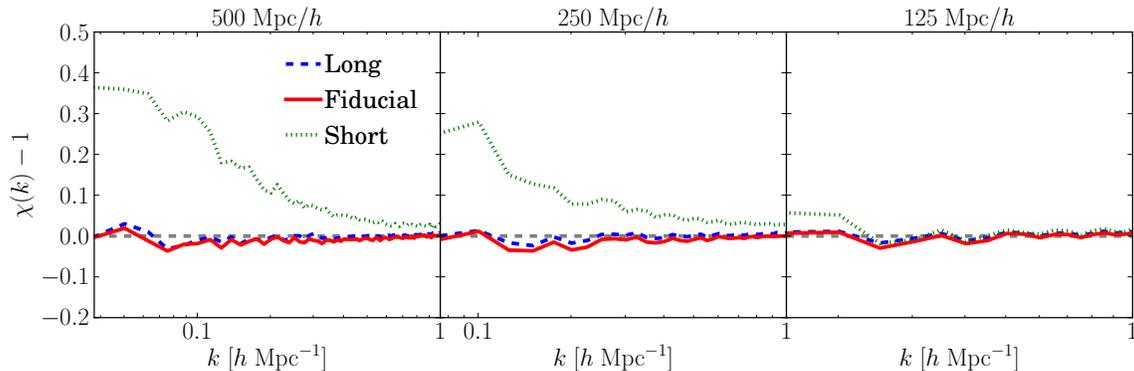


Figure 3.10: The power wedges measurement, across different sub-box size and reionization histories. The fiducial and long reionization scenarios show little anisotropy. However, for larger sub-box sizes and on large scales, the short reionization scenario does display some anisotropy. This is due to the fact that for the short reionization scenarios, there is greater variance in the 21 cm signal along the line of sight compared to the longer reionization scenarios. Note that as the scales considered become smaller, all of the scenarios are nearly isotropic.

light cone sub-box volumes of 500, 250, and 125 Mpc/h. In the previous work, the light cone was predicted to deviate from the coeval signal by $\sim 30\text{-}40\%$. Additionally, the previous work also found for the early- and mid-points of reionization, there was an increase in power when compared to the coeval case on large scales, and a decrease at small scales. When looking at the results for the 125 Mpc/h sub-box, we find that the predictions presented here match the ones presented previously, but only in the long reionization scenario early in reionization. Since our long reionization duration is comparable to their fiducial case, there is good agreement. We also note that the light cone effect becomes increasingly important as the scales get larger. As Figure 3.5 shows, for the 500 Mpc/h volumes, the light cone effect can deviate by more than 50% for the fiducial scenario and up to an order of magnitude for the short scenario. To see the full effect of the light cone, larger volumes must be used.

Another difference is that the light cone cubes presented here are constructed from a sub-volume of the entire simulation volume available. In the previous work, the light cone volume was the same size as the total simulation volume. This leads to pseudo-periodic boundary conditions in the perpendicular directions. Breaking the periodicity of the FFT can have important implications on the predicted power spectrum, especially for large-scale modes. These considerations are especially important for real-world data acquisition, where in general periodic boundary conditions do not apply. So, by explicitly breaking periodicity with the light cone cubes, we present predictions that will more readily conform to practical data processing.

The use of sub-volumes in the light cone calculation also means we are able to eliminate much of the cosmic variance for large scales. By averaging the power spectra

over many independent sub-volumes of the total simulation volume, we reduce the scatter inherent in the large scale modes. Accordingly, we are able to make progress toward a smooth power spectrum, creating an improved statistical measure of the 21 cm brightness temperature field.

3.5 Observational Comparison

Recently, upper limits on the 21 cm signal were derived based on data from the Precision Array for Probing the Epoch of Reionization (PAPER) (Parsons et al., 2010; Pober et al., 2013). Specifically, we are interested in recent results presented in Parsons et al. (2014), which reported an observational upper-limit on the 21 cm power spectrum of 2700 (mK)^2 at a redshift of $z = 7.7$ in the neighborhood of $k \sim 0.1 h \text{ Mpc}^{-1}$. We computed a predicted observation using the bias model discussed in Sec. 3.2. We considered here a prediction for the light cone power spectrum, with a midpoint of reionization to be $\bar{z} = 8$ for a more apt comparison, using the $500 \text{ Mpc}/h$ sub-box size, measured at $f_{\text{HI}} = 0.5$ by volume (which corresponds to $z = 7.9$).

The Giant Metrewave Radio Telescope (GMRT) also has derived upper limits on the 21 cm signal from measurements (Paciga et al., 2011, 2013). In this result, GMRT has upper limits on the power spectrum amplitude at a redshift of $z = 8.6$ in the neighborhood of $k \approx 0.50 h \text{ Mpc}^{-1}$. The most restrictive measurement at 2σ is $(248 \text{ mK})^2$ at $k = 0.50 h \text{ Mpc}^{-1}$, with 4 singular value decomposition (SVD) modes removed to correct for foreground contamination. (See Paciga et al. 2013 for further explanation.) One aspect to note is that the foreground removal techniques of PAPER and GMRT are different, and the measurements are reported for different redshift values. A direct comparison should not be made between the two, but instead compared directly to the theoretical prediction (solid line).

Figure 3.11 presents the 21 cm power spectrum upper-limits from PAPER (Par-

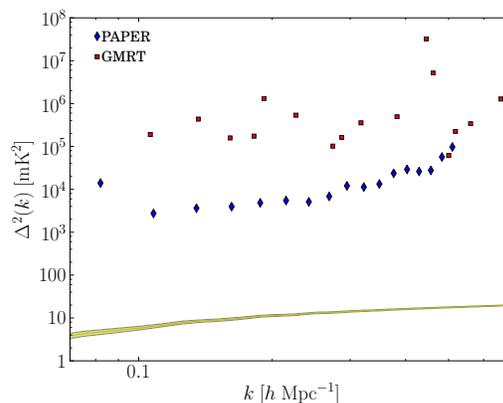


Figure 3.11: A comparison of experimental results from PAPER (Parsons et al., 2014) and GMRT (Paciga et al., 2013) with theoretical predictions incorporating the light cone. The data from PAPER represents 2σ upper limits. The data from GMRT also represents 2σ upper limits of the power spectrum. The solid curves are the predicted power spectrum of the $500 \text{ Mpc}/h$ sub-box for the fiducial reionization scenario at 50% ionization with a midpoint of reionization at $\bar{z} = 8$. The difference between the predictions and the data is several orders of magnitude.

sons et al., 2014) and GMRT (Paciga et al., 2013), compared to the reionization model at 50% reionization for $\bar{z} = 8$ with the light cone effect. For the plot of GMRT data, we selected the most restrictive point among the different number of SVD modes removed. The predicted amplitude is $\sim 10 - 100 \text{ mK}^2$, which is at least two orders of magnitude smaller than the upper limits reported by PAPER and GMRT. However, other theoretical predictions that do not include exotic reionization scenarios have similar order of magnitude differences (*e.g.*, Zahn et al., 2007; Iliev et al., 2008). Varying the reionization history did not raise the signal to the same order of magnitude of the upper limits.

Another important observational constraint comes from the Experiment to Detect the Global EoR Step (EDGES) experiment (Bowman & Rogers, 2010). In this result, the authors reported a lower limit to the duration of reionization, stating that the total duration of reionization is $\Delta z_{50} \gtrsim 0.07$ with 95% confidence. We have converted the EDGES definition of Δz , which assumes a functional form of a hyperbolic tangent, to the definition of Δz_{50} discussed in Sec. 3.3.1. The short reionization scenario has a 50% reionization duration of $\Delta z_{50} = 0.24$. Thus, the EDGES observations do not yet rule out any of the theoretical models presented here.

3.6 Discussion

An important observational consideration when measuring the 21 cm signal is the process of foreground removal. 21 cm brightness temperature fluctuations are typically 3-5 orders of magnitude smaller than signals coming from foreground contamination, such as galactic synchrotron radiation and extragalactic point sources. Typical schemes for removing these contaminants are to look at their spectra in frequency space. The 21 cm signal is expected to vary rapidly as a function of frequency, whereas these contaminants are expected to vary smoothly (Zaldarriaga et al., 2004; McQuinn et al., 2006; Liu et al., 2009). By removing these smoothly varying components from the spectrum, the true 21 cm signal emerges from the foregrounds. Unfortunately, this technique may also remove some of the long-frequency modes of the power spectrum, which is also the region of interest for the light cone effect. Care must be taken to ensure that 21 cm signal is not being discarded along with the foregrounds.

The implications of the light cone effect can also be compared to the effect of redshift space distortions (RSD). Recent work by Jensen et al. (2013) showed that RSD are most important at early stages of reionization ($0.7 \lesssim f_{\text{HI}} \lesssim 1.0$). At these stages, RSD contribute to a significant enhancement of the 3D power spectrum on large scales $k \lesssim 0.3 h \text{ Mpc}^{-1}$. For later stages in reionization, RSD have a less important effect, and by $f_{\text{HI}} \sim 0.5$ onwards, RSD induce only a percent-level change on the 3D power spectrum. As Figure 3.6 demonstrates, the light cone can have a decrement of up to $\sim 50\%$ for scales $k \lesssim 0.02 h \text{ Mpc}^{-1}$. Figure 3.5 shows that the light cone effect has an important effect on large scales for the midpoint of reionization, $0.25 \lesssim f_{\text{HI}} \lesssim 0.75$. RSD also introduce an anisotropy to the 3D power spectrum,

though the formalism presented in Jensen et al. (2013) expresses the anisotropy as an expansion in terms of $\mu \equiv \cos \theta$, the angle between the line of sight and the direction in k -space. We plan to further investigate the different implications of the light cone effect and RSD in future work.

3.7 Conclusions

We accomplished the following in this work:

- Using a parametrized bias factor between the redshift of reionization and the matter overdensity field, we created a reionization field for a large (~ 2 Gpc/ h) simulation volume.
- We made predictions about the global 21 cm brightness signal using this large volume.
- We calculated the 3D power spectrum and cross-correlation coefficient for both the coeval and light cone cases.
- We showed that including the light cone effect makes a moderate difference in the amplitude (up to 50% for small k -modes), and can change the shape of the spectrum at all scales.
- Using “power wedges” analysis, we showed that the anisotropy introduced by the light cone is only present for our short reionization scenario. We also showed this anisotropy is most sensitive to large changes in the neutral fraction of the contained volume. Thus, the light cone effect likely will not induce significant anisotropy in upcoming experiments.
- We compared predictions from our model to the recent results from the PAPER and GMRT surveys, and showed that our predictions are an order of magnitude smaller than their upper-limits on the 3D power spectrum of the 21 cm brightness temperature signal.

As mentioned in Sec. 3.1, the light cone effect has important implications for measurements that use the BAO method. The BAO scale, ~ 150 comoving Mpc ($k \sim 0.06 h \text{ Mpc}^{-1}$), approaches the scale where the light cone effect becomes non-negligible. The light cone effect can have up to a $\sim 50\%$ effect on the predicted signal on these scales. We have also shown that the light cone effect can introduce an anisotropy along the line of sight for short reionization scenarios. This complicates using the Alcock-Paczyński test to determine the proper cosmological parameters of the universe. Future applications of the BAO method to the 21 cm signal will have to account for the light cone effect in their analyses.

In future work, we would like to include redshift space distortions with the light cone effect. RSD have been investigated with respect to the 21 cm signal (*e.g.*, Bharadwaj & Ali, 2004; Barkana & Loeb, 2005; Bharadwaj & Ali, 2005; Mao et al., 2012; Jensen et al., 2013; Shapiro et al., 2013; Majumdar et al., 2013). However, these previous explorations did not include the light cone effect in their analysis. We would like to examine both simultaneously, and determine which scales are important for the effects, and how measurements are affected by each. As the data thus far suggests, smaller volumes are affected less by the light cone effect; the logical conclusion of this observation would be to analyze the 2D power spectrum, where we only examine the signal in the plane of the sky for a very narrow redshift range. By performing the analysis in this fashion, we are no longer plagued by the problem of disproportionate power from along the line of sight, but we potentially lose out on valuable three dimensional information. Thus, we hope to make predictions at different points in redshift/frequency space and then combine the results to reconstruct the 3D signal.

Chapter 4

Modeling Quasars as Radiation Sources

4.1 Introduction

Helium reionization is an important epoch in the Universe's history, and the most recent large-scale transition of the intergalactic medium (IGM). During the epoch of hydrogen reionization, the first stars and galaxies emitted photons capable of ionizing hydrogen and singly ionizing helium (whose ionization energies are 13.6 and 24.6 eV, respectively). However, the spectra of these first sources did not contain a sufficient number of high-energy photons capable of doubly ionizing helium, which requires a much larger ionization energy (54.4 eV). Consequently, helium was predominantly singly ionized following hydrogen reionization until a burst of quasar activity at redshifts $6 \gtrsim z \gtrsim 2$. Quasars are thought to be the first objects to emit an appreciable number of photons capable of doubly ionizing helium. However, because the birth of quasars requires additional time for structure to form and sufficient mass to assemble inside dark matter halos, this period of evolution occurs later in the Universe's history.

Recent and upcoming efforts to look for quasars include the Baryon Oscillation Spectroscopic Survey (BOSS) of SDSS-III (Dawson et al., 2013), the Hyper Suprime Cam of the Subaru telescope (Kashikawa et al., 2015), and DESI (Schlegel et al., 2011). There are currently about 420,000 unique quasar objects (Flesch, 2015), with this number projected to increase by an order of magnitude after the conclusion of the next generation of experiments. This rich set of observations allows us to characterize quasars to an unprecedented level of accuracy, and better characterize their properties. This is especially true at high redshift ($z \gtrsim 6$), where there are currently few observations. Determining quasar properties at high redshifts is helpful for understanding the growth of structure, as well as providing observations of reionization through measuring their absorption spectra.

Observations have shown that quasar activity peaks between $2 \lesssim z \lesssim 3$ (Warren

et al., 1994; Schmidt et al., 1995). The Gunn–Peterson trough (Gunn & Peterson, 1965) of helium has been detected at $z > 3$ (Jakobsen et al., 1994; Zheng et al., 2008; Syphers & Shull, 2014), implying that some fraction of helium was still present as He II at these redshifts. Helium absorption then transitions to becoming patchy, with extended regions of absorption and transmission in the He II Lyman- α forest (Reimers et al., 1997), and seems to be completed by $z \sim 2.7$ (Dixon & Furlanetto, 2009; Worseck et al., 2011), which coincides with the peak in quasar activity. However, to observe the Gunn–Peterson trough of He II, the sight line must be free of any intervening Lyman-limit systems. This means that the number of observations for these measurements is rather small (of $\mathcal{O}(10)$).

When discussing helium reionization, it is important to understand the properties of the ionization sources, such as quasars’ lifetimes and light curves. On the theoretical side of the problem, there are some predictions for quasar properties, but also a fair degree of uncertainty. By treating quasars as accretion disks around super-massive black holes (SMBHs), one can show that the maximal conversion efficiency ϵ for converting mass to luminosity is $\epsilon \sim 0.3$ (Thorne, 1974). Further, for matter accreting onto an SMBH at the Eddington limit (Eddington, 1926), one obtains an exponential increase in mass and luminosity with a characteristic time scale (called the Salpeter e -folding time) of $\tau = 45$ Myr for $\epsilon = 0.1$ (Salpeter, 1964; Wyithe & Loeb, 2003). Cosmological simulations that seek to capture the relationship between quasars and their galaxy hosts have treated quasar activity as being the result of a major-merger event between two galaxies (Springel et al., 2005; Hopkins et al., 2006, 2008), or a cold-flow accretion of gas onto the central SMBH (Di Matteo et al., 2012). However, there is no definitive evidence that quasars accrete exclusively at the Eddington limit, or are limited to a single episode of highly luminous activity.

Observations can also help us understand the physics of quasars, though typically at larger scales than theory or simulation. Since the entire rise and fall of quasar number density spans a time of roughly 10^9 years, the quasar lifetime must be shorter than this (Osmer, 2004, p. 324). At the other extreme, observations of the quasar proximity zone show that quasar lifetimes should be at least 10^5 years (Martini, 2004, p. 169). This time scale corresponds to the photoionization timescale of relatively high-density neutral hydrogen systems observed to be ionized in the IGM, and so the lifetime of the quasar must be at least this long in order to maintain the highly ionized level of these systems observed in the Lyman- α forest. Further constraints are difficult to obtain, and usually rely on indirect methods such as quasar clustering measurements (*e.g.*, Porciani et al. 2004; Porciani & Norberg 2006; White et al. 2012). Estimates made using these methods yield values for the quasar lifetime that are 10–100 Myr, with most values being ~ 30 Myr, which is comparable to the Salpeter e -folding time. Further, there are few definitive constraints on quasar light curves (though see Hopkins & Hernquist 2009).

For the universal populations of quasars, the major pieces of data are their number density as a function of luminosity and redshift (*i.e.*, the quasar luminosity function

(QLF) $\phi(L, z)$, *e.g.*, Schmidt & Green 1983; Boyle et al. 2000; Ross et al. 2013), and their spatial clustering (Outram et al., 2003; Porciani et al., 2004; White et al., 2012). These observations can constrain scaling relations between quasars and their hosts (*e.g.*, Conroy & White 2013), or used to calibrate subgrid models for simulations (*e.g.*, Feng et al. 2014). However, as mentioned above, the properties of individual quasars are difficult to extract from these observations, due to degeneracies. The imposed constraints are typically weak, and only provide order-of-magnitude precision.

Cosmological simulations are an ideal tool for furthering our knowledge about this portion of the universe’s history. Helium reionization leaves a lasting impression on the thermal history of the IGM: the relative hardness of quasar spectra means that there is a large degree of photoheating of the IGM while reionization is occurring. Thus, it is important to include hydrodynamics in simulations, in order to include the effects of baryonic physics. Additionally, due to the relatively long mean free path of far-UV and soft X-ray photons when looking at helium reionization, it becomes important to include radiative transfer calculations in simulations. Thus, semi-analytic calculations that assume a sharp reionization front are typically poor approximations of the physical situation. Even 1D radiative transfer codes are not realistic enough to calculate the inhomogeneous reionization process, especially when reionized regions begin to overlap. Due to the highly biased nature of quasar sources, this is typically early in the reionization process. Therefore, 3D radiative transfer calculations are essential for capturing the complicated physics of helium reionization. As mentioned earlier, the large degree of thermal heating argues for simulations in which the hydrodynamics calculations are coupled to the radiative transfer ones. This work builds on and extends previous investigations of helium reionization, which either were semi-numerical (Furlanetto & Oh, 2008b; Dixon et al., 2014) or applied radiative transfer in post-processing (McQuinn et al., 2009, 2011; Compostella et al., 2013, 2014).

Our approach to helium reionization uses simulations, with N -body, hydrodynamics, and radiative transfer solved simultaneously. An essential first step of this calculation is to understand the sources of reionization, and ensure that their properties match the observations as nearly as possible. To this end, we use the observed QLF from the SDSS and the COSMOS survey across various redshift epochs (Masters et al. 2012; McGreer et al. 2013; Ross et al. 2013, hereafter M12, M13, and R13) and the clustering measurements from BOSS (White et al., 2012) to inform the properties of individual quasars for our simulation input. By using these two constraints, as well as a formalism for populating dark matter halos with quasars that we will outline below, we are able to select simulated quasar hosts that agree well with the latest observational constraints. Specifically, matching the QLF means that we have an observationally accurate number of ionization sources, and matching the clustering measurements means our topology of reionization (*e.g.*, the size and overlap of reionized regions) will be similar to the actual reionization process. The clustering can also have an effect on the spatial correlations present in the radiation field, which can affect the BAO measurement from the Lyman- α forest.

This first part of the series discusses the way in which we create sources for our simulations of helium reionization. In Sec. 4.2, we describe our simulation strategy, and how we construct a quasar catalog from an N -body halo catalog. In Sec. 4.3, we explain how we modify our quasar properties in order to match recent observations. In Sec. 4.4, we explore implications of our findings for quasar populations. In Sec. 4.5, we discuss implications for helium reionization. Finally, in Sec. 4.6, we summarize our presentation and lay out future directions. Throughout this work, we assume a Λ CDM cosmology with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.045$, $h = 0.7$, $\sigma_8 = 0.8$, and $Y_{\text{He}} = 0.24$. These values are consistent with the *WMAP*-9 year results (Hinshaw et al., 2013).

4.2 Modeling Quasars as Radiation Sources

4.2.1 Radiation-hydrodynamic Simulations

When modeling helium reionization, we employ the RadHydro code, which includes N -body, hydrodynamics, and radiative transfer calculations simultaneously. The code includes a particle mesh (PM) solver for gravity calculations, a fixed-grid Eulerian code for solving hydrodynamics, and radiative transfer solved by performing ray-tracing. For more details on the hydrodynamics portion of the simulation code, see Trac & Pen (2004). For more details regarding the RadHydro code and its application to hydrogen reionization, see Trac & Cen (2007) or Trac et al. (2008).

The simulation strategy we employ for our exploration of helium reionization consists of two steps. First, a high-resolution N -body simulation is run for a given set of initial conditions. Halos are found on-the-fly using the friend-of-friends algorithm, and a corresponding catalog of spherical overdensity halos are saved at even steps in cosmological time (Trac et al., 2015). Then, using the same initial conditions, a medium-resolution simulation using the RadHydro code is run. In order to provide accurate sources of ionizing photons for the radiative transfer calculations, it is necessary to convert the halo catalogs produced from the first simulation into quasar catalogs for the second simulation. Since the resolution of the RadHydro simulations is comparatively low (typically a hydro grid unit is 10-100 h^{-1} kpc), the simulations are not able to accurately capture the subgrid, galaxy-level physics to include quasars directly. Thus, either a halo-level scaling relation or observational constraint is needed in order to create a physically reliable sample. Rather than having to rely on scaling relations that require several steps to convert between halo mass and quasar luminosity, we use abundance matching to calculate luminosity as a function of mass, and then use observations to create a population with the proper characteristics.

In order to calibrate the proper quasar properties to use, a suite of 10 N -body P^3M simulations with $L = 1 h^{-1}$ Gpc and 2048^3 dark matter particles were run, which corresponds to a particle mass of $m_p = 8.72 \times 10^9 h^{-1} M_\odot$. The total volume is thus $10 (h^{-1} \text{ Gpc})^3$; the BOSS measurement of the two-point correlation function

in White et al. (2012) has an effective volume of $9.8 (h^{-1} \text{ Gpc})^3$, so the volumes are comparable. Then halo finding was performed which produced the associated halo catalog snapshot every 20 Myr between $2 \leq z \leq 10$. Since only comparatively massive halos serve as hosts for the bright quasars of interest, the simulations have a sufficient resolution to capture the required number of halos.

4.2.2 Quasar Light Curves

The first step in our model construction is to define the properties of individual quasars. The two most important of these are the light curve (*i.e.*, $L(t)$) and the quasar lifetime. The most common model found in the literature for the light curve of quasars is the so-called lightbulb model, in which a quasar emits radiation at a constant luminosity for a lifetime t_q before turning off. Though largely unphysical, this model has the convenience of being simple to implement in calculations. A further simplification is typically made in which it is assumed that t_q is independent of luminosity, so that this quantity becomes a universal property.

A more realistic model of the light curve is to assume an exponential form. This type of model can be motivated physically by noting that it corresponds to Eddington accretion onto the central SMBH. Several variations on this version include an exponential ramp-up to some peak luminosity followed by abrupt turn off, a symmetric exponential about some peak luminosity, or an exponential ramp-up with a power-law fall off in luminosity (Hopkins & Hernquist, 2009; McQuinn et al., 2009). While these models are more physically motivated, they are slightly more complicated. The approach we outline below is able to reproduce a given luminosity function for quasar light curves of this form.

Specifically, we consider here two classes of quasar light curves: the “lightbulb” model and “exponential” model, defined as:

$$L_{\text{lb}}(t) = L_{\text{peak}} \Theta(t + t_q/2 - t_0) \Theta(t_q/2 - t + t_0), \quad (4.1)$$

$$L_{\text{exp}}(t) = L_{\text{peak}} \exp(-|t_0 - t|/\tau), \quad (4.2)$$

where $\Theta(t)$ is the Heaviside theta function and t_0 is the time when the quasar reaches its peak luminosity L_{peak} . In the exponential case, the parameter τ can be treated as a free parameter in a manner analogous to t_q in the lightbulb case. Nevertheless, we relate τ to t_q , which we will describe in more detail in Sec. 4.2.4.

Another consideration is the quasar lifetime itself, which in general need not be a universal property of all quasars. We have parameterized quasar lifetime as a function of luminosity using a power-law form:

$$t_q(L) = t_0 \left(\frac{L}{10^{10} L_{\odot}} \right)^{\gamma}, \quad (4.3)$$

where we vary the values of t_0 and γ . We explore models in which $10^7 \leq t_0 \leq 10^9$ yr, and $-0.25 \leq \gamma \leq 0.10$. Positive values of γ imply that brighter quasars have longer

lifetimes compared to dimmer ones, and $\gamma = 0$ is the case of a universal lifetime for all quasars.

4.2.3 Triggering Rate

We have discussed considerations for the individual quasars (*i.e.*, light curves and lifetimes), and we wish to connect them to the universal quasar population (*i.e.*, the QLF). In order to do so, we use the concept of a triggering rate $\dot{n}(L_{\text{peak}}, z)$, which dictates the differential number density of quasars that reach their peak luminosity L_{peak} as a function of luminosity and redshift per unit logarithmic luminosity. Using the formalism outlined in Hopkins et al. (2006), we distinguish between the peak luminosity of a quasar L_{peak} and the instantaneous luminosity at which it is measured for the construction of the QLF L , and relate the two with the triggering rate \dot{n} . Essentially, the triggering rate must be convolved with the light curve of the quasars, since the measured luminosity function reflects a given quasar's current luminosity L rather than its intrinsic peak luminosity L_{peak} . The result of this convolution is the observed QLF from the intrinsic triggering rate:

$$\phi(L, z) = \int \frac{dt(L, L_{\text{peak}})}{d \log L} \dot{n}(L_{\text{peak}}, z) d \log L_{\text{peak}}. \quad (4.4)$$

As explained in Hopkins et al. (2006), $\phi(L)$ is the QLF (*i.e.*, the comoving number density of quasars per logarithmic bin in luminosity), and the quantity $dt(L, L_{\text{peak}})/d \log L$ is the amount of time that a quasar spends in a logarithmic luminosity bin. Essentially, the triggering rate can be thought of as analogous to the halo mass function, though with the light curve convolution to account for changes in quasar brightness. In simple cases of the light curve the triggering rate can be solved for analytically: for a lightbulb model, $dt(L, L_{\text{peak}})/d \log L$ is a delta function at $L = L_{\text{peak}}$, and so the triggering rate is proportional to the quasar luminosity function:

$$\dot{n}_{\text{lightbulb}}(L, z) = \frac{1}{t_q} \phi(L, z). \quad (4.5)$$

In the case of an exponential light curve as defined in Equation (4.2), we have

$$\dot{n}_{\text{exp}}(L, z) = \frac{1}{2\tau} \left. \frac{d\phi(L, z)}{d \log L} \right|_{L=L_{\text{peak}}}, \quad (4.6)$$

where the factor of 2 arises because a quasar will be observed at a luminosity L while its luminosity is increasing and then decreasing. In practice, the QLF is typically reported in magnitude units rather than luminosity. One common convention is to report the quasar's absolute i -band magnitude at $z = 2$, $M_i(z = 2)$. This quantity is then converted to the specific luminosity at 2500 Å, L_{2500} , in cgs units ($\text{erg s}^{-1} \text{Hz}^{-1}$) by using Equation (4) of Richards et al. (2006):

$$\log_{10} \left(\frac{L_{2500}}{4\pi d^2} \right) = -0.4[M_i(z = 2) + 48.60 + 2.5 \log_{10}(1 + z)], \quad (4.7)$$

where $d = 10 \text{ pc} = 3.08 \times 10^{19} \text{ cm}$. To find the approximate bolometric luminosity, the relation of Shen et al. (2009) can be used to convert $M_i(z = 2)$ to luminosity in erg/s:

$$M_i(z = 2) = 90 - 2.5 \log_{10}(L). \quad (4.8)$$

One should note that this relation is approximate, and depends on the assumed spectral energy distribution (SED) of the quasar. Equation (4.4) is soluble for a few classes of light curves, such as the ones explored here.

4.2.4 Abundance Matching

The technique of abundance matching has already been applied to populations of galaxies with great success (*e.g.*, Simha et al. 2012; Hearin et al. 2013), and has also been discussed in the context of quasars (*e.g.*, Martini & Weinberg 2001; Porciani et al. 2004; Croton 2009). However, we wish to extend the techniques mentioned above to include different quasar light curves and lifetimes. The methods we outline below are also fairly general, and can be extended to include semi-analytic models as well. We start with the *Ansatz* for abundance matching of galaxies, namely that the most luminous galaxies are found in the most massive halos. This makes intuitive sense: more massive halos have more dark matter and baryonic matter to eventually convert to stars. Specifically, halo mass is highly correlated with the luminosity in the red bands, which shows the percentage of older stellar mass.

For quasars, we have a similar situation where the most luminous quasars are found in the most massive halos. However, in this case the situation is slightly more complicated because quasars have a lifetime which is much shorter than the period from the halo's formation to the activation of the quasar. Thus, we need to introduce a factor to account for the fact that not all halos host an active quasar. If we assume that the fraction of halos hosting an active quasar is universal (*i.e.*, independent of halo mass or quasar luminosity), we can express abundance matching for quasars, assuming a lightbulb light curve, as:

$$\phi(> L) = f_{\text{on}} n_{\text{halo}}(> M). \quad (4.9)$$

Expressed this way, f_{on} is simply the fraction of halos of a mass M that host an active quasar. Alternatively, we could define this fraction in terms of the quasar lifetime:

$$f_{\text{cov}} \times f_{\text{on}}(L, z) = \frac{t_q(L)}{t_H(z)}, \quad (4.10)$$

where in some models $t_H(z)$ is formulated as the halo lifetime (Martini & Weinberg, 2001), or the Hubble time (Conroy & White, 2013). We follow Conroy & White (2013) and use the Hubble time. As we shall see, though, the exact choice for $t_H(z)$ does not strongly affect the results. We have also introduced the quantity f_{cov} , which represents the solid angle fraction that the average quasar covers (where 4π steradians represents

$f_{\text{cov}} = 1$). Since in general, quasars likely have a finite opening angle (Borisova et al., 2015), what is actually constrained is the product of $f_{\text{cov}} \times f_{\text{on}}$. However, in the following analysis, we assume that $f_{\text{cov}} = 1$, though in future work we can explore the impact of setting f_{cov} to a smaller value. For the redshifts of interest, for a uniform value of $t_q = 30$ Myr and $f_{\text{cov}} = 1$, this implies that $f_{\text{on}} \sim 0.1 - 1\%$.

We can generalize the procedure of abundance matching to different light curves by using the triggering rate. In integral form, we can write abundance matching as equating the cumulative number of quasars above a particular peak luminosity given by the triggering rate with the cumulative number of halos given by the halo mass function. The total number of halos which should host quasars within a time interval Δt is:

$$\begin{aligned} \int_{\Delta t} \int_L^\infty \dot{n}(L^*) d \log L^* dt &= \int_{\Delta t} \int_L^\infty \frac{dn_{\text{halo}}(L^*)}{d \log M^*} \frac{d \log M^*}{d \log L^*} \frac{dP}{dt} d \log L^* dt \\ &= \frac{\Delta t}{t_H} \int_M^\infty \frac{dn_{\text{halo}}(M^*)}{d \log M^*} d \log M^* . \end{aligned} \quad (4.11)$$

This form of our abundance matching equation becomes the central mechanism by which we are able to equate quasar luminosity with host halo mass. In this construction, we have implicitly used the mass-to-light ratio $d \log M / d \log L$ to convert halo mass to quasar luminosity. Additionally, we have introduced the factor dP/dt to represent the probability that an individual halo will host a quasar. We have set this quantity to be equal to $1/t_H$. Thus, for the case of a lightbulb light curve and a universal quasar lifetime, this formalism reduces to Equation (4.9). Formally, this expression is an expansion of $\dot{n}(L^*, z)$ about z that is first-order accurate to $\Delta t/t_H$ (Hopkins et al., 2006). Thus, so long as the time-steps between determining the triggering rate are small compared to t_H (defined either as the Hubble time or the halo lifetime, both several orders of magnitude longer than the typical quasar lifetime), this expression should reproduce the target QLF.

In the exponential case, we are free to choose the parameter τ in any way that we like, as long as it is constant with respect to L (though it may vary with L_{peak}). We have chosen τ such that $\dot{n}(L_{\text{peak}})$ is the same between the lightbulb and exponential cases for all luminosities. We accomplish this by equating Equation (4.5) and Equation (4.6), and solving for τ in terms of t_q . The expression involves the ratio of the QLF and its derivative. This means that when we perform abundance matching, the same implicit mass-to-light ratio is used in the two cases. Since the halo mass function is the same between the two cases (due to the same population of halos being used), and the functional form of \dot{n} is the same, we must have the same form of $d \log M / d \log L$. This has the advantage of allowing us to apply certain intuition from the lightbulb case to the less straightforward exponential case. The downside to this approach is that when exploring the parameter space of quasar lifetimes t_q in the lightbulb case, it is not immediately obvious how this translates to the exponential time constant τ , since we effectively have different values of t_q for different luminosities.

ties. For instance, even in cases where t_q is independent of luminosity, τ still changes as a function of L . However, the benefits of being able to interpret the results of the exponential case using the intuition provided by the lightbulb case outweigh the downsides of not exploring parameters in τ directly.

The general procedure is as follows.

1. The halo mass found from the halo catalog at redshift z_{cat} is read and converted to an expected number density in a particular cosmology using the universal mass function described in Tinker et al. (2008). The fitted form of the mass function is used rather than the empirical one from the catalog in order to decrease the variation in number density at the high-mass end, since these quasars are disproportionately important for the reionization process.
2. Using Equation (4.11), the halo number density is converted to an expected quasar number density using a specified QLF.
3. The quasar magnitudes are binned into equal intervals in magnitude ΔM , such that the expected triggering rate $\dot{n}(M, z_{\text{cat}}) \dots \dot{n}(M + \Delta M, z_{\text{cat}})$ is found, which is converted from a number density to a total number $\dot{N}(M)$ using the volume of the simulations.
4. Within each magnitude bin, each quasar is assumed to have an equal probability of becoming active. Each quasar candidate is randomly turned on with probability $1/\dot{N}_{\text{bin}}(M)$.
5. To ensure that the volume self-consistently follows the merging of the underlying host halos, the quasars are propagated forward using a halo merger tree. By design, the halo catalog snapshots are made at times that are shorter than the expected lifetimes of the quasars. This approach allows for halos hosting quasars to be tracked throughout the simulation. In most cases, an active quasar from time step $i - 1$ in a progenitor halo passes to the single descendent halo at time step i . Additionally, this halo hosting an active quasar is not eligible to host a new quasar. This approach covers the majority of halos for the majority of time steps. However, there are several special cases related to merger events worth discussing. Specifically, when two progenitor halos merge into a single descendent and one of them is hosting an active quasar, the descendent halo inherits the active quasar. If a single active progenitor halo splits to form two descendent halos, the larger halo retains the quasar. In the case of a merger between two active quasar halos, only the larger quasar survives. These cases represent a comparatively few number of instances of our total population evolution, and do not strongly influence our conclusions.

Dataset	z	$\log_{10}(\phi^*)$	M_0^*	k_1	k_2	α	β
R13	2.2-3.5	$-5.93^{+0.02}_{-0.01}$	$-26.57^{+0.04}_{-0.02}$	$-0.689^{+0.021}_{-0.027}$	$-0.809^{+0.033}_{-0.166}$	$-1.29^{+0.15}_{-0.03}$	$-3.51^{+0.09}_{-0.18}$
M12	3.2	$-6.58^{+0.26}_{-0.79}$	-27.03 ± 0.68	-1.73 ± 0.11	-2.98 ± 0.21
M12	4	$-7.12^{+0.62}$	-27.13 ± 2.99	-1.72 ± 0.28	-2.6 ± 0.63
M13	5	$-8.47^{+0.20}_{-0.24}$	$-28.70^{+0.27}_{-0.33}$	-0.47	...	$-2.03^{+0.15}_{-0.14}$	-4.00
M13	5	$-7.63^{+0.30}_{-0.25}$	$-27.34^{+0.60}_{-0.49}$	-0.47	...	-1.50	$-3.12^{+0.28}_{-0.41}$
M13	5	$-7.93^{+0.03}_{-0.03}$	-27.88	-0.47	...	-1.80	-3.26

Table 4.1: A list of the QLF parameters of the datasets incorporated. These data are from the QLF as reported in Masters et al. (2012), McGreer et al. (2013), and Ross et al. (2013). In the table, the parameter ϕ^* has units of $\text{Mpc}^{-3} \text{mag}^{-1}$. As discussed in the text, the value of M_0^* is defined such that $M_0^* = M_i(z=2) = M_{1450} - 1.486$. The parameters k_1 and k_2 are defined for models with redshift evolution in Equations (4.13)–(4.15). Note that the authors of M12 provide a value for ϕ_0^* at $z \sim 4$ where the reported error is greater than the value itself. Since this value must be positive, the resulting lower-bound is unphysical. We reproduce the value and upper-bound here for completeness, but do not include this value directly when determining the values of the QLF. See Appendix B.1 for further details. Additionally, in M13, the authors provide three fits, each with at least one parameter held constant. Values without error ranges indicated correspond to the parameters held fixed for a particular fit.

4.2.5 The Quasar Luminosity Function

Throughout this work, we use a series of QLFs as determined at different epochs. For relatively low-redshift ($2 \lesssim z \lesssim 3$), we use the QLFs as determined by R13 from the BOSS survey, specifically the high- z stripe 82 sample (S82) form which includes luminosity evolution and density evolution (LEDE). Above a redshift of 3, the QLF has been measured at $z \sim 3.2$ and $z \sim 4$ by M12 using data from COSMOS.¹ At $z \sim 5$, the QLF has been measured by M13 using data from the SDSS.² Although these works use slightly different values for cosmological parameters from the ones assumed here, the impact on the reported quantities is minimal.

In order to span the different epochs over which the luminosity function has been measured, it is necessary to combine the different data sets. All of the data sets fit to a double power law form of the QLF, written as:

$$\Phi(M) = \frac{\phi^*}{10^{0.4(1+\alpha)(M-M^*)} + 10^{0.4(1+\beta)(M-M^*)}}, \quad (4.12)$$

where Φ is the comoving number density of quasars of magnitude M per unit magnitude, ϕ^* is the normalization of the QLF, α is the faint-end slope of the luminosity function, β is the steep-end slope (which is reversed from the parameterizations of M12), and M^* is the so-called break magnitude where the luminosity function transitions from the faint-end to the steep-end. In most formulations at high-redshift, redshift evolution is incorporated by a change in ϕ^* , M^* , or both, that is linear in redshift. For the data from R13, the evolution is given by the equations:

$$\log_{10} \phi^*(z) = \log_{10} \phi_0^* + k_1(z - 2.2), \quad (4.13)$$

$$M_i^*(z) = M_0^* + k_2(z - 2.2). \quad (4.14)$$

For the data in M13, there is linear evolution in $\log_{10} \phi^*$ as well, given as:

$$\log_{10} \phi^*(z) = \log_{10} \phi_0^* + k_1(z - 6). \quad (4.15)$$

Table 4.1 lists the parameters that we include from the measurements of R13, M12, and M13. The values from M12 are not included in the fitting procedure directly, and serve primarily as a consistency check due to their comparatively large error bars. The parameters from M13 are determined at $z \sim 5$, and the ones from M12 are determined at $z \sim 4$ and $z \sim 3.2$. Note that the authors of M13 provide 3 independent fits to their data, which are all incorporated into the final QLF parameterization. (See

¹Additionally, the QLF at $z \sim 4$ has also been measured by Glikman et al. (2011) and Ikeda et al. (2011). As noted in M12, the normalization of the QLF of Ikeda et al. (2011) is comparable, whereas the normalization of Glikman et al. (2011) is larger than the others by a factor of ~ 4 . M12 notes that the difference can be caused by contamination of the faintest-magnitude bins from dwarf stars and high-redshift galaxies. In the following analysis, we use the results from M12.

²An upper limit for the QLF at $z \sim 5$ was found by Ikeda et al. (2012), which is consistent with the results of M13.

Appendix B.1 for more details.) For the measurements from R13, whose fiducial LEDE model include redshift evolution in ϕ^* and M^* , the model is valid over a range of redshift, from $2.2 \leq z \leq 3.5$. For the purposes of generating our quasar catalogs, we are interested in exploring the QLF until $z = 2$. For the sake of simplicity, we simply extend the LEDE model from R13 to this redshift. Although the LEDE fit is ostensibly not valid below $z = 2.2$, we expect helium reionization to be largely finished by this redshift, and so the precise form of the QLF at $z \sim 2$ is not of fundamental importance to our study.

Also, for the value of M^* , it is necessary to convert to a single magnitude system. As explained in Sec. 4.2.3, we use $M_i(z = 2)$, the absolute i -band magnitude at $z = 2$. The QLF of M12 and M13 use M_{1450} , which is related to $M_i(z = 2)$ by $M_i(z = 2) = M_{1450} - 1.486$ (Richards et al. 2006; Ross et al. 2013, Appendix B). Note that this conversion assumes that the quasar SED follows a power-law with an effective spectral index of $\alpha = 0.5$ (using the convention that $f_\nu(\nu) \propto \nu^{-\alpha}$). Modifying the spectral index α changes the magnitude conversion, so care must be taken when converting between magnitude systems. See Appendix B.1 for further discussion.

To combine the R13, M12, and M13 data sets into a single set of quantities, we first assume that the results from R13 are accurate for redshifts $z \leq 3.5$. This is the nominal limit of the LEDE fits, and though there are small differences between the fit QLF and the binned data, overall the fits are excellent. To incorporate the results at higher redshifts, we cast the four parameters of the QLF (ϕ^* , M^* , α , and β) as quantities that have lin-

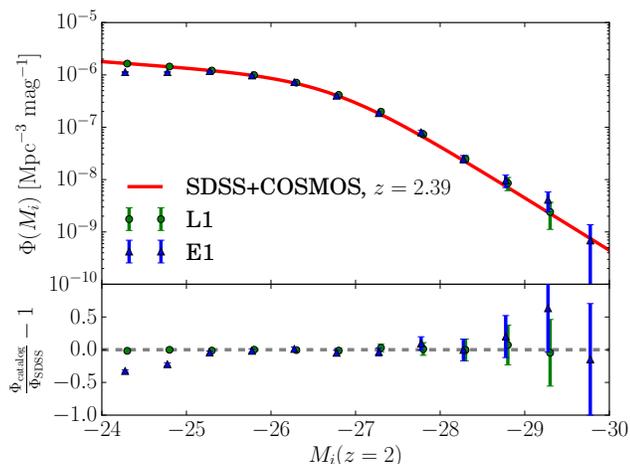


Figure 4.1: A comparison of the composite quasar luminosity functions from the SDSS+COSMOS measurements (Ross et al., 2013; Masters et al., 2012; McGreer et al., 2013) to our abundance matching method, plotted with Poisson error bars. The two different quasar models (defined in Table 4.3) are offset from each other for visual clarity. The agreement is excellent for comparatively dim quasars which are more common, but there is some discrepancy for bright objects. The reason for this disagreement is primarily due to Poisson noise, since these objects are rare even for the large ($1 (h^{-1}\text{Gpc})^3$) simulation volume. At low luminosity in the exponential case, the completion limits of dark matter halo hosts at this mass become noticeable. See Sec. 4.2.4 for further discussion.

Parameter	Fiducial value	Parameter range
$\log_{10} \phi_0^*$	-6.82	-
c_1	-0.790	$[-1.10, -0.536]$
M_0^*	-27.6	-
c_2	-0.238	$[-0.716, 0.170]$
α_0	-1.29	-
c_3	-0.324	$[-0.493, -0.140]$
β_0	-3.51	-
c_4	0.0333	$[-0.327, 0.260]$

Table 4.2: A list of the parameters used in Equations (4.16a-4.16d) based on the data listed in Table 4.1.

ear evolution in redshift. We define these parameters as:

$$\log_{10} \phi^*(z) = \log_{10} \phi_0^* + c_1(z - 3.5), \quad (4.16a)$$

$$M^*(z) = M_0^* + c_2(z - 3.5), \quad (4.16b)$$

$$\alpha(z) = \alpha_0 + c_3(z - 3.5), \quad (4.16c)$$

$$\beta(z) = \beta_0 + c_4(z - 3.5). \quad (4.16d)$$

These parameterizations are applied to redshifts where $z > 3.5$. The constant values are defined to be equal to the values of R13 at $z = 3.5$, and the values for the slopes (c_1 - c_4) are allowed to take on a range of values. The range is generally chosen such that the values for the different parameters brackets the range of best-fit values provided by the highest redshift (M13) data. The fiducial values for the slopes are taken to be ones that reasonably reproduce the high-redshift measurements. Table 4.2 shows the fiducial values for the slopes, as well as the range of values for the parameters at $z \sim 5$ used in the parameter space exploration in Sec. 4.5. For a complete discussion on selecting the parameters for the QLF, see Appendix B.1.

Figure 4.1 shows the combined QLF from R13, M12, and M13 (which at this epoch is essentially that of R13), as well as two different quasar models at $z \sim 2.4$. We can see that there is generally very good agreement between the constructed quasar catalog and the target luminosity function, as should be expected. The differences between the constructed catalogs and target luminosity function are typically on average $\lesssim 5\%$, which is comparable to or smaller than the uncertainties in the luminosity function itself at these redshifts. At high luminosities ($M_i \lesssim -28$), though, there are some comparatively large differences that can arise between the predicted and empirical luminosity functions. This deviation is largely due to Poisson shot-noise introduced by the rarity of the objects. For objects in this luminosity range, there are typically only a few objects ($\mathcal{O}(10)$) in the entire $1 (h^{-1}\text{Gpc})^3$ volume. At the dim end of the QLF, there can be insufficient halos of a particular mass given the mass resolution of our simulation. The minimum halo mass is $M_{\text{halo},\text{min}} = 4.36 \times 10^{11} h^{-1} M_\odot$.

Since quasars with $M_i \leq -25$ are most important for this study, this does not affect our results significantly.

Throughout most of the following analysis, we focus our attention on several models in particular, parameterized in terms of t_0 and γ as in Equation (4.3). The first four of these models have particularly good agreement with the BOSS measurements. The last two are included to demonstrate how the clustering signal changes as a function of t_0 for a fixed value of γ : models L1, L3, and L4 all have the same γ value. We summarize these models in Table 4.3.

4.3 Clustering Measurements

4.3.1 Two-point Correlation Function

By construction, our method matches the input QLF at all redshifts, regardless of the individual properties of the underlying quasar population. However, we are not guaranteed to match the observed clustering of quasars. Changing the implicit mass-to-light ratio of Equation (4.11) through changing the quasar lifetimes will affect how halos are populated with quasars. In general, longer quasar lifetimes lead to quasars of the same luminosity being matched into hosts of larger masses. Since their hosts are more biased, this leads to quasars of the same luminosity showing a larger clustering signal. This is true at all luminosities. We want to match the clustering because it can affect the topology of reionization. There can also be spatial correlations present in the radiation field as a result of reionization, which are important for making measurements of the baryon acoustic oscillation (BAO) from the Lyman- α forest (*e.g.*, White et al. 2010; Slosar et al. 2013).

Here, we explore how to include clustering measurements from the two-point correlation function in our quasar catalog. Recent results from the BOSS survey for the clustering of quasars in the redshift range of interest are presented in White et al. (2012). The above work examines the clustering signal of quasars in both 2D-projected and 3D-redshift-space correlation functions at intermediate scales ($3 \lesssim s \lesssim 25 h^{-1}\text{Mpc}$). The authors also introduce luminosity cuts to make the results more robust. For the purposes of this comparison, we consider their selection for which they imposed luminosity cuts on both the bright and faint ends, so that only objects with $-25 \geq M_i \geq -27$ were considered across the entire redshift range (Sample 4 as defined by the authors). For a fair comparison, we impose similar cuts on our object selection. We also examine the redshift evolution of the results, and compare against the high- z /low- z samples (Samples 5 and 6) as well. See Appendix B.2 for further discussion of these different redshift samples.

We explore the parameter space of available quasar models by examining the lightbulb and exponential light curves defined in Equations (4.1) and (4.2), as well

³ t_0 and γ as defined in Equation (4.3).

Model Name	Light curve	$\log_{10}(t_0/\text{yr})^3$	γ
L1	Lightbulb	7.75	0
L2	Lightbulb	8.25	-0.125
E1	Exponential	7.25	0
E2	Exponential	7.75	-0.15
L3	Lightbulb	7	0
L4	Lightbulb	8.5	0

Table 4.3: A list of the parameters of some quasar models considered.

as luminosity-dependent quasar lifetimes defined in Equation (4.3), parameterized by t_0 and γ . For each combination of parameters, we construct a quasar catalog in the manner described above.⁴ Then, we extract from this catalog all objects that satisfy the magnitude constraints at the central redshift of the survey $z = 2.39$. This redshift represents the average redshift of quasars chosen in the BOSS sample; the actual quasar objects span in redshift from $2.2 < z < 2.8$. However, as is noted in White et al. (2012), the redshift evolution of the signal is weak. Thus, extracting objects from our quasar catalogs at a single redshift rather than a range should have little effect on our overall conclusions. We measure the monopole of the two point correlation function using the “natural estimator” ξ :

$$\xi(s) = \frac{\langle DD(s) \rangle}{\langle RR(s) \rangle} - 1, \quad (4.17)$$

where $\langle DD(s) \rangle$ is the average number of quasar pairs from the quasar catalog separated by a real-space distance of $[s - \Delta s/2, s + \Delta s/2]$, and $\langle RR(s) \rangle$ is the number of pairs of points at the same separation drawn from a distribution with Poisson noise.

4.3.2 Calculating χ^2 Values

In order to quantify the statistical uncertainty in our catalog, we ran a suite of 10 N -body simulations with different initial conditions. We then performed our abundance matching procedure on each of the different simulations, including several realizations for each volume. Since our abundance matching procedure stochastically determines which halos should be hosting active quasars at a given time step, we create several

⁴There are several extreme models where the number of objects is significantly fewer than the number predicted by the quasar luminosity function. This is not a failure of our methodology, but rather instances of there being too few halo objects of a given mass to host quasar objects. In essence, f_{on} is so small that we reach the resolution limits of the simulation. In these cases, we add particles from a second-order Lagrangian perturbation theory (2LPT) simulation of the same initial conditions at the same redshift in order to define a set of “random” particles that are still representative of the underlying matter distribution. We randomly sample from these particles in order to fill out the catalog to the expected number. This ensures that we do not measure a statistically significant clustering measurement when the catalog is clearly unphysical.

quasar catalogs for each individual halo catalog, using a different initial random seed (three realizations per volume for these results). Additionally, we have augmented the effective number of samples by including redshift space distortions along the different principal axes of the simulation. This strategy gives us a total of 90 samples for which to measure the clustering signal. The best estimate for the correlation function $\xi(s)$ for a given radial bin s_i is given by averaging over all of the individual estimates ξ_k :

$$\bar{\xi}(s_i) = \frac{1}{N} \sum_{k=1}^N \xi_k(s_i) \quad (4.18)$$

We then estimate the covariance between the radial bins by computing the entries of the covariance matrix C_{ij} . We compute the entries of the covariance matrix as (Zehavi et al., 2005):

$$C_{ij} = \frac{1}{N} \sum_{k=1}^N (\xi_k(s_i) - \bar{\xi}(s_i)) (\xi_k(s_j) - \bar{\xi}(s_j)). \quad (4.19)$$

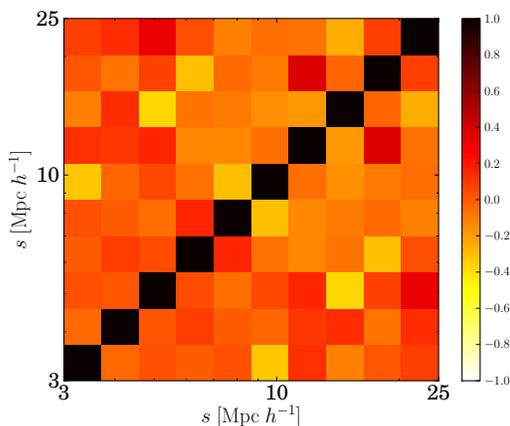


Figure 4.2: The correlation matrix for the L1 model. Note how the matrix is dominated by the diagonal entries, which is to be expected for shot-noise dominated measurements. The small off-diagonal terms suggest that the covariance matrix has converged numerically, and should be stable when inverting. This type of structure is seen in all models considered.

The correlation matrix entries for our model L1 is plotted in Figure 4.2. Notice that the diagonal entries dominate, which means that the bins are mostly independent of each other and dominated by shot-noise (Valageas et al., 2011; White et al., 2012). Implicitly, the samples have been treated as being independent, and this is almost surely not the case. Although the 10 volumes as a whole can be treated as being statistically independent, the different realizations based on the same halo catalog are likely correlated. Further, the projections of peculiar velocities along different axes for the same realization are also likely to produce correlated results. However, producing a sufficient number of independent realizations to decrease the noise in the covariance matrix is computationally infeasible. Further, the variance in the clustering signal among quasar catalog realizations for a given

(t_0, γ) pair is comparable to small displacements in the t_0 - γ parameter space, so it is necessary to include this source of uncertainty. Since we are interested only in finding models that are consistent with the BOSS measurements which have their

own set of observational uncertainties, we feel that this approach produces sufficiently accurate results.

Once the entries of the covariance matrix have been computed, the difference vector $\delta(s_i) \equiv \xi_{\text{model}}(s_i) - \xi_{\text{BOSS}}(s_i)$ is calculated. The correlation function ξ_{BOSS} is fit to a power law: $\xi_{\text{BOSS}}(s) = (s/s_0)^\beta$, where the authors have fixed the value of $\beta = -2$. In order to investigate the impact this choice has on the conclusions, we performed fits on the correlation function measured from our quasar catalogs using two different parameterizations: one where the best-fit value of s_0 was found when fixing $\beta = -2$, and another where the value of s_0 and β were both fit. In the length scales used for our analysis ($3 \leq s \leq 25$), the deviation of β from the fiducial value of -2 was small, typically less than 5%. Furthermore, the values for s_0 were also largely similar between a fixed slope or a varying one, with deviations typically less than 1%. Thus, the choice to set $\beta = -2$ does not strongly bias the results presented here, or the values reported in ξ_{BOSS} .

When comparing one of the quasar models with the BOSS results, the χ^2 value of the model is then given by:

$$\chi^2 = \delta^T C^{-1} \delta. \quad (4.20)$$

To define the model that fits the BOSS observations best, we want to minimize the χ^2 value of the model. A two-dimensional space in t_0 and γ is constructed for both of the light curves, this space is explored using regular grid points. Following the analysis of White et al. (2012), a χ^2 distribution with 9 degrees of freedom is assumed. Using this distribution, the χ^2 value for a particular model is converted to a confidence interval. An equivalent $n\sigma$ value is computed based on the confidence interval (1σ if the enclosed probability is 0.683, 2σ if it is 0.955, etc.). This statistic demonstrates how “consistent” a particular model is with the BOSS observations.

Figure 4.3 shows the clustering measurements for several of our well-fitting models compared to the BOSS measurements. The values of these models are given in Table 4.3. In general, as t_0 increases at a fixed value of γ , the clustering signal increases

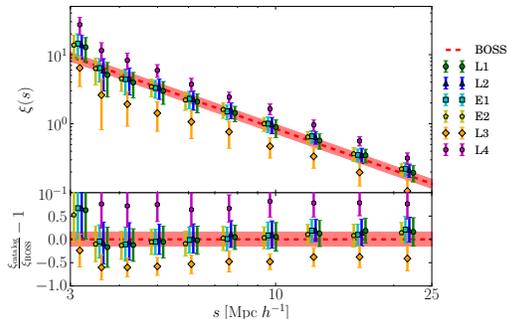


Figure 4.3: The quasar two-point correlation function from White et al. (2012), compared to several models whose parameters are described in Table 4.3. All measurements were made at the same values of s , but are offset from each other for visual clarity. The shaded error regions on the measurements from BOSS are the reported 1σ error bars, and the error bars on the models are the square root of the diagonal elements of the covariance matrix. Note that for the same value of γ , increasing t_0 leads to a larger clustering signal (compare L3, L1, and L4 in order of increasing t_0). See the text for additional details.

as well. Compare specifically the L3, L1, and L4 models, which have the same value of γ but have respectively increasing values of t_0 . Mathematically, this behavior can be seen from the form of Equation (4.11): for the same luminosity and mass functions but a larger value of $f_{\text{on}} \propto t_q$, quasars of the same luminosity will shift to more massive host halos. Since the clustering signal increases with the mass, it follows that increasing t_0 will increase the clustering signal. For similar reasons, increasing values of γ for constant values of t_0 are also associated with a stronger clustering signal, since this also effectively increases the quasar lifetime t_q .

4.3.3 Characteristic Luminosity and Lifetime

Figure 4.4 shows the χ^2 values in the two-dimensional parameter space t_0 and γ , as defined by Equation (4.3), for the different light curves. The region of good agreement between the BOSS measurements and our models takes on a linear relationship between $\log_{10}(t_0)$ and γ . Such a relationship can be parameterized as:

$$\log_{10}(t_0/\text{yr}) = \log_{10}(t_{\text{eff}}/\text{yr}) + L_0\gamma. \quad (4.21)$$

The parameters t_{eff} and L_0 can be thought of as a characteristic timescale and a characteristic luminosity, respectively. From the functional form of our power-law for quasar lifetime in Equation (4.3), L_0 can be interpreted as changing the normalization luminosity. This is the luminosity at which all models have the same lifetime, regardless of the value of γ . In other words, the characteristic luminosity of the power law becomes:

$$\log_{10}(L_{\text{eff}}/L_{\odot}) = 10 - L_0, \quad (4.22)$$

where L_0 is defined in Equation (4.21). The parameter t_{eff} is the characteristic time because all models have this same lifetime at the luminosity L_{eff} .

For the lightbulb model, the best-fit values for t_{eff} and L_{eff} are $\log_{10}(t_{\text{eff}}/\text{yr}) = 7.76$ and $\log_{10}(L_{\text{eff}}/L_{\odot}) = 13.29$. (See Table B.1 in Appendix B.2 for evolution of these parameters with redshift.) The characteristic luminosity inferred from this value is $L_{\text{eff}} = 10^{13.29} L_{\odot}$, which has a corresponding magnitude of $M_i = -27.2$. This value is not surprising, given that quasars were selected for the clustering measurements near this magnitude range. More interesting is the value of $\log_{10}(t_{\text{eff}}/\text{yr}) = 7.77$, which gives a characteristic lifetime of $10^{7.77} = 59$ Myr. This is a quasar lifetime that is slightly longer than those typically quoted in the literature (Yu & Tremaine, 2002; Yu & Lu, 2004; Porciani et al., 2004; Conroy & White, 2013), which are closer to the Salpeter e -folding time scale or shorter (~ 45 Myr for a quasar accreting at Eddington luminosity and a mass conversion efficiency of $\epsilon = 0.1$). Although t_{eff} is slightly higher than these values, it is within a factor of 2.

In the exponential model, the best fit values for t_{eff} and L_{eff} defined in Eqn (4.21) are $\log_{10}(t_{\text{eff}}/\text{yr}) = 7.18$ and $\log_{10}(L_{\text{eff}}/L_{\odot}) = 13.05$. This luminosity implies a slightly dimmer characteristic luminosity ($M_i = -26.6$). As discussed in Sec. 4.2.4, there is not a single τ for all quasars for a given value of t_{eff} : $L \approx L^*$ quasars have $\tau \approx t_{\text{eff}}$, with brighter quasars having $\tau > t_{\text{eff}}$. However, the difference between τ and t_{eff} does not differ by more than a factor of 2 in either direction, and so to a good approximation $\tau \sim t_{\text{eff}}$, especially for the luminosity range used to match the clustering measurements. Compared to the lightbulb case, the quasars with an exponential light curve have a shorter characteristic lifetime of 15.1 Myr. The characteristic lifetime is smaller for the exponential than in the lightbulb case because quasars do not shut off entirely after a single lifetime, so the time that a quasar is “bright enough” to be included within the luminosity cuts is longer than its lifetime t_{eff} . This lifetime is about a third of the Salpeter e -folding time scale, which implies that if quasar light curves are roughly exponential, the combination of the measured QLF and the clustering measurements favors quasars that either radiate at luminosities dimmer than their Eddington ratio ($L/L_{\text{edd}} < 1$), have a mass-conversion efficiency is less than the fiducial value ($\epsilon < 0.1$), or both. Unfortunately, since our model does not track the underlying physics present, we are not able to distinguish between these two cases.

The reason for the different best-fit values between the two models can be understood as follows. By construction, we have fixed the lifetime of the exponential quasars such that their peak luminosity to mass ratio is the same as in the case of the lightbulb for a given choice of t_0 and γ . (See Sec. 4.2.4 for more details.) However, the mass-to-light ratio for the two light curves are significantly different. This is due to the fact that the observed luminosity for an exponential quasar can be much smaller

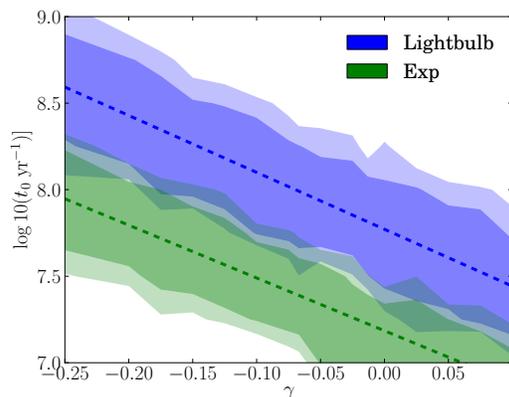


Figure 4.4: A comparison of the parameter space exploration in terms of the parameters t_0 and γ from Equation (4.3). Both the parameter space for the lightbulb model (Equation (4.1)) and exponential model (Equation (4.2)) are shown. The dashed lines represent the best linear fits to the data for a particular light curve. The class of models that are consistent with the BOSS measurements at 1σ and 2σ correspond to the darkly and lightly shaded regions. In general, we find that for the exponential model, shorter lifetimes are preferred (smaller values of t_0 for the same γ). Since we abundance match against the quasar’s peak luminosity, and the quasar spends comparatively little time at or near the peak luminosity, we effectively increase the clustering signal for lower luminosity quasars.

than its peak luminosity. A particular luminosity range is selected for the clustering measurements, but the clustering of these quasars is tied to their peak luminosity rather than the observed one. Thus, quasars will tend to have higher clustering at a given luminosity in the exponential case compared to the lightbulb, since they spend comparatively little time at or near their peak luminosity. This luminosity selection includes quasars with a higher peak luminosity than the chosen range (and thus a higher clustering signal), so we must also include quasars that have lower mass hosts to match the average clustering signal. This means that there is a larger spread in host mass compared to the lightbulb case. This behavior explains why the characteristic luminosity is slightly smaller for the exponential model compared to the lightbulb: there is an increased number of low-luminosity quasars occupying high-mass hosts.

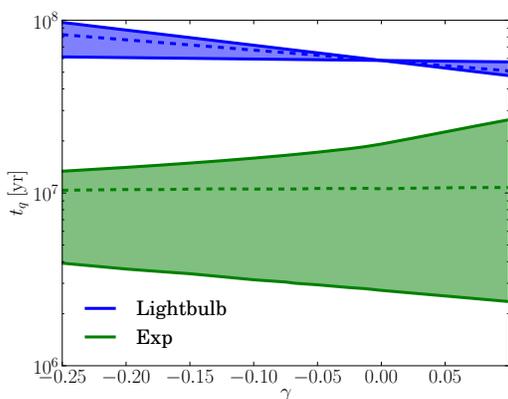


Figure 4.5: The range of quasar lifetimes for models with the best-fit values of t_0 as a function of γ , based on the linear relationship extracted in Equation (4.21). The solid regions show the span in quasar lifetimes, while the dashed line shows the median value within a given model. Note that the median value is fairly constant across all quasar lifetimes. Thus we are able to characterize a quasar model reasonably well using the characteristic lifetime. The comparative large spread in quasar lifetime in the exponential case is due to our method of selecting t_q , rather than reflecting a truly large spread in the data. See the text for additional details.

Figure 4.5 shows the range of quasar lifetimes as a function of model parameter γ . The quasar lifetime is broadly similar across different model choices. The exponential model has a lower overall value due to the effect discussed above, *i.e.*, that quasars from a higher peak luminosity will be included in the sample, bringing along a higher clustering signal. Since this is true for nearly all the quasars in the sample, there is an overall decrease in the selected lifetime of quasars. The large difference in the span of quasar lifetimes is due to the way that we have defined the quasar lifetime in the exponential model. As discussed in Sec. 4.2.4, the exponential lifetime τ is selected such that the same relationship between host mass and quasar peak luminosity exists in the exponential case as in the lightbulb case. Even in a model where for the lightbulb t_q is independent of L (*i.e.*, when $\gamma = 0$), the exponential model parameter τ does have luminosity dependence. In general, quasars with luminosities above L_* will have a lifetime longer than an equivalent luminosity in the lightbulb case for the same choice of t_0 and γ in Equation (4.3), and those with low luminosities will have a shorter lifetime. This choice for our model leads to the spread in lifetimes of a factor of ~ 5 , as seen in the case of $\gamma = 0$. For

models in which $\gamma > 0$, there is a widening in the range of values. This is due to the fact that brighter quasars live longer than dimmer ones. Since our choice of quasar lifetime already enforces this difference, these models see an increased effect. Conversely, for $\gamma < 0$, there are competing effects between brighter quasars being less long-lived due to the choice of γ , but still simultaneously living longer than their lightbulb counterparts due to the choice for t_q . The latter effect wins out, and these quasars end up having a significantly larger spread than in the lightbulb case. Note that the contours in this figure are smooth compared to Fig. 4.4 because these are results lying along the best-fit line, and the figure shows the range in values rather than a single number (*i.e.*, the χ^2 value) that fluctuates as a function of position in t_0 and γ .

4.4 Discussion

4.4.1 Mass-to-Light Ratio

The combination of the QLF and clustering measurements produces an important set of constraints on the space of potential quasar models. Here we investigate the implications of these models. One important implication is the mass of a typical halo for a given quasar luminosity. It is trivial to predict this for the case of a lightbulb model, but less straightforward for the case of the exponential model. Here the peak luminosity L_{peak} is used to define the mass-to-light ratio, since this is the quantity used in our abundance matching approach. This ratio defines a typical mass for quasars, which can be compared with results of previous analysis (*e.g.*, Martini & Weinberg 2001; Shen et al. 2007; White et al. 2012).

Figure 4.6 shows the luminosity of quasars as a function of the mass of the host halo for the lightbulb and exponential models for all combinations of t_0 and γ . This plot shows the mass to light ratio of the entire catalog. The weight assigned to the luminosity as a function of mass $L_{\text{peak}}(M)$ for a particular model i is given by a χ^2 likelihood. We also find the ± 1 and 2σ values that enclose 68 and 95% of the likelihood.

Figure 4.7 shows the mass range as a function of the model parameter γ . Note that the range is essentially constant with respect to γ . By averaging the median mass across all values of γ , a characteristic mass for the two models can be defined. This characteristic mass is $2.5 \times 10^{12} h^{-1} M_{\odot}$ for the lightbulb model, and $2.3 \times 10^{12} h^{-1} M_{\odot}$ for the exponential model. These values are broadly consistent with previous studies of quasar clustering measurements (*e.g.*, Porciani et al. 2004; Croom et al. 2005; Porciani & Norberg 2006; Lidz et al. 2006; White et al. 2012). Since in all models the same clustering signal of the quasars is being selected, there is an implicit requirement for the hosts to lie within a certain mass range.

Additionally, the mass range in the exponential case is significantly larger than that of the lightbulb model. This is again related to the fact that due to the light curve, quasars with a higher clustering signal are included within the luminosity sample, and so there must also be lower-mass hosts included as well to balance the average clustering strength. There is a significantly larger spread above the median mass than below. The reason for this asymmetry is due to the difference in number density: since the high mass objects are rarer, a comparatively smaller range in low-mass halos is necessary to make the clustering signal equivalent to the light bulb case. See Sec. 4.3.3 for further discussion.

In the case of the lightbulb model, the halo mass that correspond to the selected luminosity range of quasars is relatively tightly constrained. For the models that agree with the BOSS measurements at 1σ , the average halo mass ranges from $1.35 \times 10^{12} h^{-1} M_{\odot}$ to $4.93 \times 10^{12} h^{-1} M_{\odot}$ for hosts of quasars within the magnitude cutoff. For the exponential model, there is a much larger range in halo mass: for the collection of models that agree at 1σ , the mass ranges between $6.69 \times 10^{12} h^{-1} M_{\odot}$ and $5.85 \times 10^{13} h^{-1} M_{\odot}$, almost an entire order of magnitude (compared to about half an order of magnitude for the lightbulb model). Also note that the mass range is much larger than in the lightbulb case. This fact can be explained by noting that there is evolution in the mass-to-luminosity ratio during the lifetime of the quasar. Further, the e -folding time for these models is comparatively long, with typical values being $\tau \approx 40$ Myr. This means that there are high-mass hosts included in the sample of quasars chosen for the clustering measurements whose quasars are not at their peak luminosity. Since these hosts have a bias larger than the value preferred by the BOSS measurements, this sample must necessarily include hosts which have a smaller clustering signal, so that on average, the total bias agrees with BOSS.

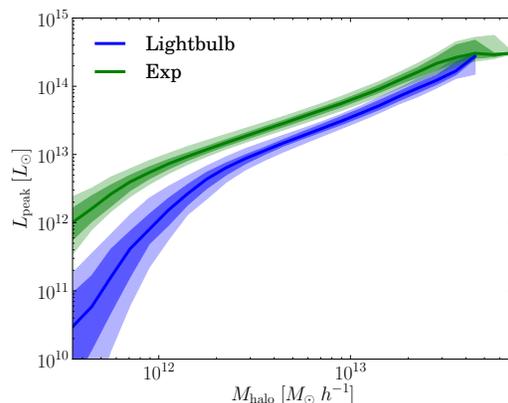


Figure 4.6: A comparison of the mass-to-light ratio between the different quasar models. We have computed this ratio for each model in our parameter space, and weighted their contribution by their corresponding χ^2 value. The lines show the median value by weight and the shaded regions show $\pm 1\sigma$ and 2σ . Note that for a given luminosity, a quasar in the exponential model is found in a halo with a smaller mass. This is due to the fact that quasars with a peak luminosity significantly greater than the observed one are included in the luminosity range selected for the clustering measurements. These hosts have a higher clustering signal than quasars with a peak luminosity in the luminosity selection. This means we must also select lower-mass objects as well. See the text for additional details.

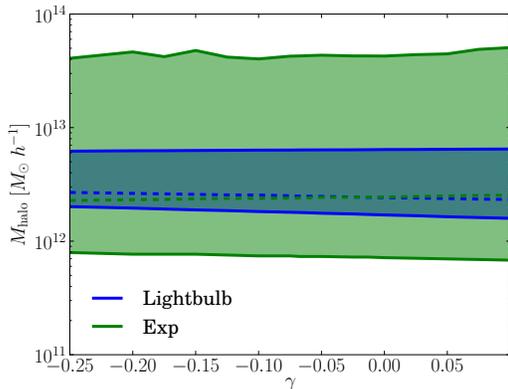


Figure 4.7: The selected mass range for our different models. For the lightbulb case, there is a relatively narrow range in mass. Since the clustering of quasars is fixed as a function of luminosity, there is a very tight relationship between observed luminosity and underlying host mass. In the exponential case, the range is much more extended. Since the mass to light ratio is fixed to be the same for the peak luminosity, the evolution within the model means that there will be quasars with higher peak luminosity (and higher mass hosts) selected by the evolution.

related to evolution of quasar luminosity.

4.4.2 Mass Function and Duty Cycle of Halo Hosts

To observe the effect that different points in parameter space have on the halo host properties, the mass function of halos hosting an active quasar has been calculated for the fiducial redshift selection. (For the high- and low-redshift selections, see Appendix B.2.) Figure 4.8 shows the total halo mass function as well as the mass function of halos hosting quasars within the luminosity range $-25 \geq M_i \geq -27$. From this analysis, the duty cycle of halo hosts can be extracted, *i.e.*, the fraction of active halos divided by the total number of halos. As discussed in Sec. 4.2.2, in the lightbulb model the duty cycle can be directly related to the quasar lifetime at that luminosity. However, here the duty cycle is defined simply as the active fraction of halos.

Figure 4.8 compares the case of the lightbulb and the exponential light curves.

There is a systematic shift upward in the mass of the exponential case compared to the lightbulb. This shift is related to the difference in parameter space discussed in Sec. 4.3.1. Due to their exponential change in luminosity, the quasars are not typically found near their peak luminosities. Thus, even though by construction the peak quasar luminosity as a function of halo mass is the same for the two models, the effective luminosity for a given mass is reduced in the exponential case due to the light curve evolution. In other words, quasars of the same luminosity in the two different models are found in more massive hosts in the exponential case. This leads to a systematic shift in the preferred mass range for clustering measurements. To sum up: the increased spread in halo host mass for the exponential model compared to the lightbulb is due to inclusion of highly-biased hosts in the measurement being balanced out by lower-mass ones, and the systematic shift toward higher mass is due to the effective increase in the mass-to-luminosity ratio

When the halo mass function of active halos is examined in the two different cases, it can be seen that the mass range of hosts spanned by an individual model is quite different. In the lightbulb case, there is a very small range in host mass compared to the exponential case. This difference can be explained in terms of which hosts are included in the clustering measurements. In the lightbulb case, since the luminosity is constant as a function of quasar lifetime, the only evolution in the relationship between mass and luminosity comes from mass accretion, which makes up a small fraction of total halo mass over the time scales for which quasars are active. As such, with an essentially static relationship, there is a very strong correlation of mass to light. For a specific model, there is only about a factor of 2 in halo mass included for the quasars in the selected magnitude range. When looking at the duty cycle of quasar hosts, one can see that the fraction is typically 0.5-1%, with little evolution with mass within a model.

Conversely, in the exponential model, there is evolution for individual halos in the mass to light relationship. Most importantly, this implies that massive halos will be included when selecting quasars at a specific luminosity. Since they are more highly clustered (and more biased), smaller, less biased halos must also be included in order to create an average bias consistent with the BOSS measurements. This has the effect of extending the mass range of halos included in the mass function. Note that within a single model, there is a much larger span in halo mass: in some cases, the span is more than an order of magnitude in halo mass. Additionally, the duty cycle is comparable in magnitude to the lightbulb case, though slightly smaller: the ratio of active halos to total halos ranges from 0.05-1%. There also seems to be a trend in the evolution of the duty cycle: there is a central “typical mass” for a given model, and the duty cycle decreases in both directions. A similar trend was found by White et al. (2012).

Note that one result of this measurement is the fact that the mass range of host halos is significantly more extended in the exponential case than the lightbulb case. Thus, one way to break the degeneracy

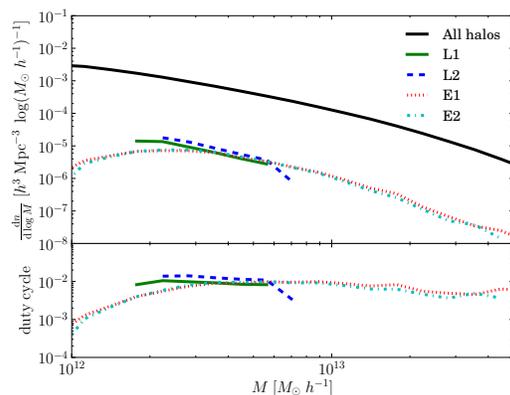


Figure 4.8: The halo mass function for halos hosting quasars versus the total halo population for certain models as defined in Table 4.3. In the case of the lightbulb model, there is a roughly constant value of $\sim 1\%$ of halos hosting quasars as a function of halo mass. For the exponential, there is a constant active fraction at the high mass end, but then the fraction falls off below the median halo mass. This is again related to the fact that there are more halos included at low mass in order to reproduce the average clustering signal.

between the lightbulb and exponential models would be to measure the mass range of underlying host halos, perhaps through using gravitational lensing to independently find the mass of the dark matter halo (Courbin et al., 2012). If the range of masses for quasar hosts is extended, then there would be observational evidence favoring an exponential model (or a model with evolution in the quasar light curve) as opposed to the lightbulb model.

4.5 Predictions for Helium Reionization

One very important prediction that we can make from our quasar catalog is the redshift of helium reionization. In order to understand in detail the implications for helium reionization, we need to run full hydrodynamic plus radiative transfer numerical simulations. However, we can perform a semi-analytic calculation in order to find a rough estimate of the redshift of reionization by computing the fraction of the universe’s volume that has been reionized Q_i (also called the volume filling fraction), where $Q_i = 1$ represents a totally reionized universe (*e.g.*, Madau et al. 1999; Furlanetto & Oh 2008b):

$$\frac{dQ_i}{dt} = \int dL \frac{\dot{N}_\gamma}{\bar{n}_{\text{He}}} \frac{d\phi}{dL} - \bar{C} \alpha_A \bar{n}_e Q_i, \quad (4.23)$$

where \bar{n}_{He} is the number density of neutral helium, \bar{n}_e is the number density of electrons, \dot{N}_γ is the production rate of ionizing photons for an individual quasar, $\alpha_A(T)$ is the recombination coefficient, and $\bar{C} \equiv \langle n_e^2 \rangle / \langle n_e \rangle^2$ is the clumping factor of the ionized IGM. The minimum luminosity of the integral decreases as redshift decreases, in keeping with modeling and observations (Richardson et al., 2012; Shen & Kelly, 2012; Cen & Safarzadeh, 2015; Sijacki et al., 2015). The clumping factor measures the effective distribution of gas inside the scale of volume being averaged (or resolution in the case of simulations). Note that these calculations assume a primordial helium mass fraction of $Y_{\text{He}} = 0.24$. Following the arguments in the appendix of Kaurov & Gnedin (2014), we choose the case A recombination coefficient, which assumes that photons emitted from recombination are not reabsorbed by a neutral atom, increasing the recombination rate.⁵ It is assumed that initially, all of the hydrogen in the IGM has been ionized, and all of the helium is singly ionized. To compute the photoionization rate of an individual quasar \dot{N}_γ , the SED of Lusso et al. (2015) is used to convert the specific luminosity at 2500 Å to the specific luminosity at 912 Å. It is then assumed that quasars have a spectral energy distribution (SED)

⁵Although the arguments presented in the cited work are in the context of hydrogen reionization, the same arguments can be applied equally well to helium reionization. Essentially, the authors argue that the photons redshift out of resonance with the thermally broadened spectral line before they encounter the edge of the ionized region or a Lyman-limit system. Although the ionization fraction of helium might be slightly lower inside an “ionized region” than a comparable hydrogen one, the difference is not significant enough to change the overall conclusion.

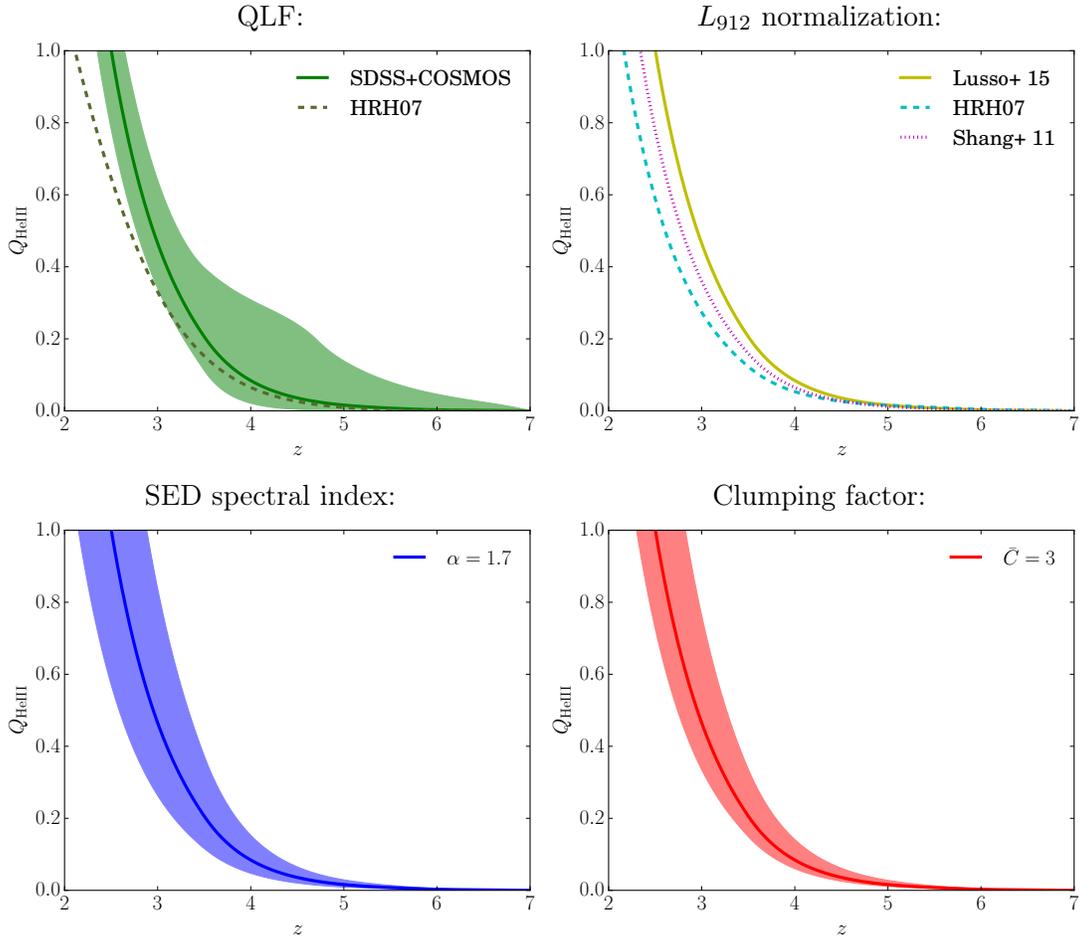


Figure 4.9: The volume filling fraction Q_i of doubly ionized helium defined in Equation (4.23). In each plot, we show the fiducial values we have for volume filling fraction Q_i as a function of redshift, which has the parameters $\alpha = 1.7$, $\bar{C} = 3$, and normalizing the luminosity at 912 \AA using Lusso et al. (2015). This leads to a redshift of reionization of $z \sim 2.5$, comparable to the redshift of $z \sim 2.7$ suggested by observation of the helium Lyman- α forest. We show the change in Q_i as a function of varying these parameters. Top left: we compare the difference between using the composite QLF of SDSS+COSMOS (see Sec. 4.2.5 for more details) and the one in Hopkins et al. (2007). The shaded region reflects differences in ionization level due to jointly varying the parameters over the ranges specified in Table 4.2. Top right: we change the UV SED of the quasar, which affects the normalization at 912 \AA . In addition to the SEDs from Lusso et al. (2015) and HRH07, we show the radio-quiet template from Shang et al. (2011). Bottom left: we allow the EUV SED spectral index for $\lambda < 912 \text{ \AA}$ to vary from $1.4 \leq \alpha \leq 2.0$. Bottom right: we vary the clumping factor of the IGM, from $1 \leq \bar{C} \leq 5$. See the text for additional details.

that follows a power law $L_\nu(\nu) \propto \nu^{-\alpha}$ for values of $\lambda < 912 \text{ \AA}$. The fiducial value chosen is $\alpha = 1.7$, also based on observations of the rest-frame UV spectra of quasars from Lusso et al. (2015). This calculation includes all photons with frequencies in the range $54.4 \text{ eV} \leq h\nu \leq 1 \text{ keV}$. Photons above this energy have a mean free path of helium ionization comparable to the Hubble distance.

Although the two different quasar light curves explored above have different individual properties, both are constrained by the global properties fixed by the QLF. We find that if instead of the statistical calculation outlined above, we use the number of ionizing photons computed directly from the quasar catalogs, the result differs only by a few percent. Therefore, it is much more straightforward to use Equation (4.23). This approach also permits the use of other QLFs in the calculation, so it is possible to explore what effect this has on the results.

Figure 4.9 shows the ionization fraction as a function of redshift computed from Equation (4.23). In the first panel, there is a comparison of the choice of QLF used in the calculation. Included are the QLF used in the main body of this work, the composite QLF composed of the ones from R13, M12, and M13 (the SDSS+COSMOS) as explained in Sec. 4.2.5, and the QLF from Hopkins et al. (2007) (hereafter referred to as HRH07). All other calculations presented use the composite QLF, but then change various other parameters. Note that for the prediction of reionization time using HRH07, both the QLF and the SED are different from the fiducial comparison case. In the figure, the shaded region shows the range of predicted values for the volume filling fraction Q_i at a given redshift z by jointly varying the parameters of the QLF over the range specified by Table 4.2. Interestingly, the late-time ionization level is less sensitive to the variation in parameters at early redshift, due to the interplay between the source and recombination terms present in Equation (4.23). At redshifts $z \leq 3.5$, the source term becomes the same for all histories, since the QLF transitions to that of Ross et al. Further, the recombination rate is proportional to the ionized fraction, so histories that had higher ionization levels at $z \geq 3.5$ will have higher levels of recombination. Since the recombination time is much shorter than the total timescale of the reionization calculation, all histories converge on a similar redshift of total reionization ($Q_i = 1$). Nevertheless, the variation in ionization fraction at early times can have important implications on the topology of ionized regions and the thermal history of the IGM, so such differences may in principle be detectable.

In the second panel of the plot, the specific luminosity of individual objects at 912 \AA L_{912} is varied. One way to achieve this variation is the change the UV SED template used for quasars. Once the specific luminosity L_{2500} is calculated from the observed magnitude according to Equation (4.7), the quasar SED can be used to find L_{912} . In the fiducial approach, we use the SED template from Lusso et al. (2015), which assumes a UV spectral index of $\alpha = 0.61$ for $2500 \geq \lambda \geq 912$. An alternative choice for an SED is one from Shang et al. (2011), which provides a composite quasar SED template by combining observations in different frequency ranges to create a single spectrum. The authors of Shang et al. (2011) divide the sample into radio-loud and

radio-quiet quasars. However, radio-quiet quasars compose $\sim 90\%$ of high-redshift quasars found in SDSS (Shen et al., 2009). Thus, we only include the results of the calculation using the radio-quiet template. This template provides the relative specific luminosity at each frequency, and so can be used to convert L_{2500} to L_{912} . The effective spectral index for this wavelength range for the radio-quiet quasar template is $\alpha = 0.867$. In addition, we show the impact of using the SED from HRH07 (with the QLF from Sec. 4.2.5). Note that the SED from HRH07 is outdated, and used only as a point of comparison. More recent studies (*e.g.*, Stevans et al. 2014; Lusso et al. 2015) are largely inconsistent with the SED of HRH07, and so it is presented here merely to emphasize the importance that using the proper SED has on helium reionization. Given this same specific luminosity L_{2500} , the predicted value of L_{912} from the SED of Lusso et al. (2015) is higher than that of HRH07 by about a factor of 1.7, leading to the earlier reionization time. The second panel of the plot includes these to demonstrate the difference from using different quasar templates.

In the third panel of the plot, the spectral indices are varied, ranging from $1.4 \leq \alpha \leq 2.0$. Recent measurements from Lusso et al. (2015) suggest that at high redshift and bright magnitudes, the spectral index has a value of $\alpha = 1.7 \pm 0.6$. This is slightly softer than the average value of $\alpha = 1.6$ from Telfer et al. (2002). In order to explore some of the implications of changing the spectral index, we vary its value as indicated.

The final panel explores a range of clumping values, from $1 \leq \bar{C} \leq 5$. The precise value for the clumping factor for helium reionization is very uncertain, as most studies on the clumping factor are related to hydrogen reionization (see, *e.g.*, Raičević & Theuns 2011; Kaurov & Gnedin 2014). In Furlanetto & Oh (2008b), the authors explored clumping factors of $0 \leq \bar{C} \leq 3$. More recent results from numerical simulations were calculated by Jeesson-Daniel et al. (2014), who found that the clumping factor of helium range from $3 \leq \bar{C} \leq 8$ for the redshift range of interest, depending on the ionization level of the helium gas.

In Figure 4.9, each panel shows the fiducial evolution of Q_i , which is characterized by the values of $\alpha = 1.7$, $\bar{C} = 3$, the SED of Lusso et al. (2015), and the composite SDSS+COSMOS QLF. In this situation, the redshift of reionization (*i.e.*, when $Q_i = 1$) is $z \sim 2.5$. This value is comparable to, though slightly later than, the redshift suggested by recent observations of $z \sim 2.7$ (Dixon & Furlanetto, 2009; Worseck et al., 2011). However, a smaller volume-averaged clumping factor \bar{C} or a larger amplitude in either the measured QLF or the specific luminosity L_{912} could give an earlier redshift of reionization. Specifically, assuming the fiducial model, changing the clumping factor to $\bar{C} = 1.7$ would give $z \sim 2.7$ as the redshift of reionization. It should be noted that this calculation is not wholly accurate for reionization, since it assumes a single clumping factor for the entire IGM, which is almost certainly not accurate for helium reionization, due to its very inhomogeneous nature. Furthermore, this calculation does not include secondary ionizations from energetic electrons (*e.g.*, Shull 1979; Furlanetto & Stoever 2010), though these interactions are likely unimportant for helium reionization (McQuinn et al., 2009).

When comparing to the results of Furlanetto & Oh (2008b), we notice that the authors' value for the redshift of reionization is significantly earlier than the one that we have found. This is largely due to a different quasar luminosity function used, as well as a different method for calculating a quasar's EUV SED. The referenced paper uses the QLF from HRH07, and assumes an SED that gives more EUV radiation. This luminosity function has a significantly larger amplitude compared to the results from R13, up to an order of magnitude larger for low-luminosity quasars at high redshift. (See Fig. 16 of Ross et al. 2013.) Thus, accurate measurements and a proper understanding of the systematics of the high-redshift QLF, as well as the accompanying quasar SED, are essential for a proper treatment of helium reionization.

4.6 Conclusion

We have provided a technique for populating dark matter halos with quasars that matches a quasar luminosity function by construction for various light curve models of quasars. By using the triggering rate of Hopkins et al. (2006) with the technique of abundance matching, we are able to match the observed quasar luminosity function of SDSS DR9 (R13), COSMOS (M12), and high-redshift SDSS data (M13). After applying this method to dark matter halo catalogs generated from N -body simulations, we have constrained a class of quasar models that reproduce the clustering amplitude measured from the two-point auto-correlation function of the BOSS survey (White et al., 2012) at a redshift of $z = 2.39$. The characteristic mass of the quasar hosts is $2.5 \times 10^{12} h^{-1} M_{\odot}$ for the lightbulb model and $2.3 \times 10^{12} h^{-1} M_{\odot}$ for the exponential model. The effective lifetime as defined in Equation (4.21) of quasars is $t_{\text{eff}} = 59$ Myr for the lightbulb model of quasars and $t_{\text{eff}} = 15$ Myr for the symmetric exponential model.

One of the limitations of this approach is that we have constrained the class of quasar models using a comparatively narrow span in quasar luminosity. By matching the bias of quasars with a different magnitude range, we would have a different effective luminosity range for the bias calculation. This would lead to a different slope in the parameter L_{eff} , which would allow us to break the degeneracy observed in Fig. 4.4. Having the ability to break the sample down into different luminosity intervals would allow us to make tighter constraints on the class of allowed models.

In future work, we plan to use the quasar models explored here as sources of ionizing photons for studying helium reionization using simulations containing hydrodynamics and radiative transfer. These types of simulations will allow us to accurately capture important physical characteristics related to the IGM. Specifically, we are interested in capturing the thermal history of the IGM as it relates to observations. In upcoming simulations, we plan to compute the IGM equation of state and produce synthetic Lyman- α forest fluxes. This will allow us to tap into the wealth of observations available for the Lyman- α forest, such as those currently available from BOSS (*e.g.*, Lee et al. 2013), and from upcoming future surveys such as DESI.

Chapter 5

Signatures of Quasar Activity on the IGM

5.1 Introduction

Helium reionization is a fascinating portion of the Universe's history, and is the last major phase change of the intergalactic medium (IGM). After hydrogen reionization at high redshift ($z \gtrsim 6$) from the first stars and galaxies, helium was singly ionized. However, the second ionization of helium requires significantly more energy (54.4 eV vs. 24.6 eV for the first ionization). The stars providing photons for hydrogen reionization did not emit a significant number of these high energy photons. Thus, helium was not doubly ionized until later in the Universe's evolution, when quasars produced enough high-energy photons to significantly change the ionization level of helium. Following the formation of quasars at redshifts $6 \geq z \geq 2$, the helium of the IGM became totally ionized, leaving an imprint on the IGM.

The process of helium reionization leaves important observational signatures on the Lyman- α forest, which is a measure of the relative amount of photon absorption due to gas in the IGM. The Lyman- α forest can be observed most readily for neutral hydrogen, and has been observed at medium resolution (*e.g.*, the Baryon Oscillation Spectroscopic Survey, BOSS, McDonald et al. 2006; Lee et al. 2015) and high resolution (*e.g.*, Keck-HIRES and Magellan-MIKE, Lu et al. 1996; Viel et al. 2013). To date, there have been more than 150,000 Lyman- α forest spectra measured from BOSS alone (Dawson et al., 2013), and the number of systems is expected to increase by almost an order of magnitude after the deployment of the next generation of telescopes (Myers et al., 2015). This rich observational data set contains much information about the IGM, most notably the abundance of neutral hydrogen and its temperature.

A related measurement to the hydrogen Lyman- α forest is the analogous feature for He II. However, to date, there have been only about 50 systems for which the He II measurement has been made (Syphers et al., 2009b,a, 2012). The reason for

the comparative lack of He II measurements is due to the presence of Lyman-limit systems (LLS), which are optically thick and lead to large absorption features. This absorption contaminates the signal, and makes detection of He II signatures difficult (Møller & Jakobsen, 1990; Zheng et al., 2005). Nevertheless, the detection of the helium analogue of the Gunn-Peterson trough (Gunn & Peterson, 1965) offers an indication of when helium reionization ended. Recent observations have shown a Gunn-Peterson trough for helium at redshifts $z > 3$ (Jakobsen et al., 1994; Zheng et al., 2008; Syphers & Shull, 2014), which shows the He II volume fraction must have been greater than $f_{\text{HeII}} \gtrsim 10^{-3}$ along these sightlines. Helium absorption then becomes patchy, with extended regions of absorption and transmission in the He II Lyman- α forest (Reimers et al., 1997), and seems to be completed by $z \sim 2.7$ (Dixon & Furlanetto, 2009; Worseck et al., 2011). However, the comparatively low number of sightlines that show the Lyman- α forest signature for He II leaves much statistical uncertainty about the exact timing and nature of the reionization process.

In order to better explore some of the signatures that helium reionization leaves on the IGM, we have run a new suite of simulations that include hydrodynamics and radiative transfer solved simultaneously. These simulations represent the first efforts to incorporate all of the relevant physics together using a spatially varying radiation field sourced by quasars, in order to better predict the impact on observations. Previous studies have incorporated various schemes. In some cases, radiative transfer is solved in post-processing of N -body or hydrodynamic simulations (*e.g.*, McQuinn et al. 2009, 2011; Compostella et al. 2013, 2014), which does not incorporate the effect of photoheating on the IGM that accompanies reionization. Alternatively, radiative transfer has been included through the use of a uniform ionization background (*e.g.*, Theuns et al. 1998; Jena et al. 2005; Viel et al. 2013; Puchwein et al. 2015; Bolton et al. 2016), an approach which does not capture the large-scale inhomogeneities of the radiation field. Thus, the simulations presented here represent a step forward in accurately modeling the reionization process, and capture the effects of heating from sources and the inhomogeneous and anisotropic aspects of sources.

This work represents the second endeavor in a series on helium reionization simulations. The work La Plante & Trac 2015 (discussed in Chapter 4) outlines a method whereby dark matter halos from N -body simulations are populated with quasars such that the quasar luminosity function (QLF) from the SDSS and COSMOS surveys (Ross et al., 2013; Masters et al., 2012; McGreer et al., 2013) and the two-point autocorrelation function from BOSS (White et al., 2012) are reproduced. This ensures that our radiation sources match the latest observational constraints in terms of their number density and topology.

We organize the rest of this chapter as follows. In Sec. 5.2, we discuss our simulation technique and describe the method by which we include sources of ionization. In Sec. 5.3, we discuss in more detail the individual models explored here, and the differences apparent in the helium ionization fraction. In Sec. 5.4, we explore impacts of reionization on the thermal history of the IGM. In Sec. 5.5, we discuss generat-

ing synthetic Lyman- α sightlines from the simulations, and comparing with recent observations. In Sec. 5.6, we summarize and explore avenues for future research. Throughout this work, we assume a Λ CDM cosmology with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.045$, $h = 0.7$, $\sigma_8 = 0.8$, and $Y_{\text{He}} = 0.24$. These values are consistent with the *WMAP*-9 year results (Hinshaw et al., 2013).

5.2 Radiation-Hydrodynamic Simulations

To faithfully capture helium reionization, the ideal simulations should include dark matter, baryonic matter, and radiation coupled together. The dark matter is necessary for establishing the large-scale structure of the Universe, and the baryonic matter captures the distribution of neutral and ionized gas in the IGM. By coupling radiation to this gas as the simulation is proceeding, a more accurate state of the IGM is calculated. As mentioned above, due to the large degree of photoheating of the IGM induced by the energetic photons from quasars, the thermal state of the mean-density IGM is dominated by quasars and the reionization of helium. Further, the clustered nature of quasars argues for simulations in which the radiation sources are tracked explicitly, rather than incorporated as a uniform background. Thus, these simulations are able to capture many of the features important to helium reionization, and generate predictions which can be readily compared with observations.

5.2.1 Populating Simulations with Quasars

The simulations presented here have been run using the RadHydro code, which includes N -body, hydrodynamics, and radiative transfer calculations. The code employs a particle mesh (PM) solver for gravity calculations, a fixed-grid Eulerian code for solving hydrodynamics, and a ray-tracing scheme for computing radiative transfer. The radiative transfer calculations use a non-equilibrium solver for the photoionization balance equations, and use many time steps per hydro step to ensure accurate calculation of the thermal state. The code has been used studying hydrogen reionization (Trac & Cen, 2007; Trac et al., 2008; Battaglia et al., 2013b), and has been modified extensively for the current application to helium reionization.

Our simulation strategy is as follows. Due to the requirement of a large box size to capture relatively rare objects, the simulation does not resolve the galaxy-scale physics (and by extension, quasar-scale physics). As such, it is necessary to populate the volume with sources using an alternative method. To this end, we perform the simulation in two steps: a first pass to generate a catalog of quasar sources, and a second pass that uses the sources to perform full reionization simulations. We first run a P^3M N -body simulation including only dark matter (Trac et al., 2015). Initial conditions for these simulations are generated at $z = 150$ using transfer functions generated by CAMB (Lewis et al., 2000). These N -body simulations are run at high resolution, where for our fiducial simulations we use a simulation volume of size

$L = 200 h^{-1}\text{Mpc}$ with 2048^3 particles. This yields a particle mass of $m_p = 6.98 \times 10^7 h^{-1}M_\odot$. Halo-finding is done on-the-fly using a friend-of-friends (FoF) algorithm with mean inter-particle spacing of $b = 0.08$ to find halo members. This value avoids the over-bridging problem in standard FoF with $b = 0.2$. The halo finder is used to locate all halos with 50 or more members. Once the FoF halos are found, a spherical overdensity (SO) algorithm is used to create a corresponding halo catalog. These halo catalogs are produced every 20 Myr in cosmological time while the simulation is running. The halos from the catalogs are then treated as candidate hosts for the quasars to be used in our simulations.

With the halo catalogs from the high resolution simulation in hand, the halos can be populated with quasars, the sources of helium-ionizing radiation. Following the procedure outlined in Chapter 4, we populate these halos with quasars that reproduce the observed quasar luminosity function and clustering measurements. Briefly, the model uses the technique of abundance matching in order to populate potential quasar hosts (*i.e.*, dark matter halos) with quasars in order to reproduce a specified quasar luminosity function (QLF). The method allows the user to specify the QLF to use, and either a lightbulb or exponential model for the quasar light curve. By construction, the method will reproduce the desired QLF at all redshifts (starting at $z \sim 6$, the earliest redshift at which we include quasar sources), provided the quasar lifetime (and time between halo catalog snapshots) is small compared to the Hubble time. The fiducial QLF used in the work presented here combines the results of several different luminosity functions at different redshifts: at high redshift ($z \gtrsim 5$), the QLF reproduces the observations of McGreer et al. (2013). At intermediate redshift ($z \sim 4$), the QLF reproduces the observations of Masters et al. (2012). At lower redshift ($z \lesssim 3.5$), the QLF parameters used are those from Ross et al. (2013). Combining the measurements of the QLF at multiple epochs ensures that the number density sources of helium-ionizing radiation found in the simulations are observationally accurate. Since the timing of reionization is determined, by a large part, by the abundance of sources, having an observationally accurate quasar number density is of the utmost importance. The simulations run here use two slightly different methods for combining the different measurements, which we call Q1 and Q2. See Appendix C.3 for further discussion on the details of the QLF used in these simulations.

In addition to matching the number density of quasar sources, the method of Chapter 4 also matches the observed clustering of quasars. Using the abundance matching technique leaves the lifetime of quasars unconstrained, which affects the bias of quasars. Reproducing the bias of quasars ensures that simulations reproduce the topology of reionization: although the number of sources is fixed by the QLF, the clustering of quasars will affect the size and shape of ionized regions. In general, since quasars are known to be highly biased (White et al., 2012), they are found strongly clustered, which leads to early overlap of doubly ionized regions (McQuinn et al., 2009). In Chapter 4, we use a suite of N -body simulations to study how the lifetime of quasars affects their clustering. We identify a set of parameters that reproduce

the clustering as measured in White et al. (2012) at redshift $z \sim 2.4$. The model developed in Chapter 4 allows for the lifetime of quasars t_q to change as a function of luminosity following a power-law relation, parameterized as $t_q(L) = t_0(L/L_0)^\gamma$, where L is the peak luminosity of the quasar, and t_0 and γ are two parameters allowed to vary. Unless otherwise noted, the models discussed in these simulations used an exponential light curve, with $\gamma = -0.1$. As discussed below, in instances where the QLF is modified to explore a different reionization history, the quasar lifetime t_0 is modified to match the clustering measurements.

5.2.2 Quasar Properties

For individual quasar objects, there are two components of the spectral energy distribution (SED) that must be specified: the normalization and the spectral index. The QLF is typically reported in terms of magnitude, rather than luminosity. Specifically, the convention used when reporting the QLF in Ross et al. (2013) is to use the absolute i -band magnitude at $z = 2$. In order to determine the energy output of a quasar, we convert from magnitude to luminosity using Equation (4) of Richards et al. (2006):

$$\log_{10} \left(\frac{L_{2500}}{4\pi d^2} \right) = -0.4[M_i(z = 2) + 48.60 + 2.5 \log_{10}(1 + 2)], \quad (5.1)$$

where $d = 10 \text{ pc} = 3.08 \times 10^{19} \text{ cm}$. This formula converts the magnitude of the QLF into a specific luminosity at 2500 Å. Once this specific luminosity has been found, the specific luminosity in the extreme ultraviolet (EUV) region must be calculated to determine the output of radiation relevant to helium reionization. For the purposes of this calculation, we use the quasar SED template of Lusso et al. (2015). This template assumes a power-law form for the SED with a spectral index of $\alpha = 0.61$ ($f_\nu \propto \nu^{-\alpha}$) for $\lambda \geq 912 \text{ Å}$ and $\alpha = 1.7$ for shorter wavelengths. The number of photons is then computed in 7 different frequency bins for the radiative transfer calculation, spanning photon energies from $h\nu = 13.6 \text{ eV}$ to 1 keV. At energies higher than this, the mean free path of photons interacting with singly ionized helium becomes comparable to the Hubble scale, and as a practical matter much larger than the box size of the simulation.

As discussed in Chapter 4, there is a moderate degree of uncertainty in the systematic effects of the quasar population. For instance, reddening of quasars due to dust, obscured quasars, contamination of non-quasar objects in photometric surveys, and poor knowledge of the intrinsic colors of quasars could all systematically shift the normalization of the QLF. In order to marginalize over some of this uncertainty, we have conducted several simulations with the same underlying gas distribution and large-scale structure, but with different quasar populations. Specifically, we modify the normalization of the QLF and the normalization of the SED. These different

simulations allow us to explore some of the effect that these systematic uncertainties generate, and how they might impact different observations of the IGM. We will further discuss all of the models explored below in Sec. 5.2.4.

5.2.3 Simulation Features

Although the main focus of this study is to understand the impact of helium reionization, an accurate treatment of hydrogen reionization is nevertheless important. In some sense, the initial conditions of helium reionization (especially with respect to the temperature of the IGM) is set by the timing of hydrogen reionization, and the inside-out nature of denser regions undergoing reionization earlier than less dense ones.

In order to capture the inhomogeneous effects that hydrogen reionization has on the IGM, the method of “patchy reionization” developed in Battaglia et al. (2013b) is applied to the simulation volume, which predicts a redshift of reionization based on the density field from a dark-matter-only simulation. A mean redshift of reionization $z_{\text{re}} = 8$ was used for these simulations, with the fiducial values for the other parameters in the model which control the duration of reionization. The application of this method better captures the thermal state of the IGM following hydrogen reionization than using a uniform radiation background.

The radiative transfer is calculated using explicit ray tracing of photons from quasars, using the scheme described in Trac et al. (2008). However, tracking rays from galaxies in addition to those from quasars would be prohibitively expensive. The stellar content of galaxies does not produce an appreciable number of photons with $h\nu > 54.4$ eV, and are thus largely unimportant for helium reionization (Furlanetto & Oh, 2008b). However, galaxies do produce photons that contribute to hydrogen ionization. The ionization balance equation for hydrogen can be written as:

$$\frac{dn_{\text{HI}}}{dt} = -\Gamma_{\text{tot}}n_{\text{HI}} + \alpha_{\text{HII}}n_{\text{HII}}n_e, \quad (5.2)$$

where Γ_{tot} is the total photoionization rate per atom in s^{-1} , α_{HII} is the recombination coefficient, and n_i is the comoving number density of species i . For the case of hydrogen, there are contributions from both quasars and galaxies, which can be expressed as $\Gamma_{\text{tot}} = \Gamma_{\text{qso}} + \Gamma_{\text{gal}}$. The computation of Γ_{qso} is computed explicitly via ray tracing, but the value of Γ_{gal} must be specified. For the purposes of running the simulation, the value of Γ_{gal} is assumed to be a uniform value. For late times ($z \lesssim 6$), the hydrogen in the IGM is highly ionized and hence optically thin, and so treating the UV background as uniform is a valid approximation. One approach is to use a value based on a semi-analytic model (*e.g.*, Haardt & Madau 2012, hereafter HM12). However, this approach relies on the specifics of the model chosen, and does not account for other details in the simulation (such as the quasar contribution to hydrogen ionization, patchy hydrogen reionization, etc.)

In order to circumvent some of these issues, we choose to set the value of Γ_{gal} to match the observed effective optical depth τ_{eff} measured by BOSS (Lee et al., 2015). We generate Lyman- α sightlines on-the-fly while the simulation is running, and modify the value of Γ_{gal} in order to match $\tau_{\text{eff}}(z)$. Instead of generating the full number of sightlines available to us (N_{grid}^2), we reduce the number of sightlines drawn by a factor of four in each dimension, for a total of $N_{\text{grid}}^2/16$. In comparisons performed between using the full sample and this reduced subset, we did not find significant differences in the calculated value of τ_{eff} , and therefore inferred the same target value of Γ_{gal} . By matching the value of τ_{eff} by construction, we are better able to compare between simulations and against observation. This also avoids renormalizing the Lyman- α forest in post-processing, which is the usual approach taken in simulations comparing against the Lyman- α forest (*e.g.*, Bolton et al. 2009b). Put another way, Γ_{gal} becomes a free parameter that we adjust at every time step in the simulation in order to match the value of τ_{eff} specified by Lee et al. (2015), such that $\Gamma_{\text{gal}} + \Gamma_{\text{qso}}$ reproduces the proper optical depth.

Below in Sec. 5.2.4, we discuss the simulations performed in our simulation suite. Some of the models have an increased number of photons produced by quasars, above the fiducial values assumed by the quasar properties as discussed in Sec. 5.2.2. For these models with an increased number of photons, the contribution of Γ_{qso} is large enough that even if $\Gamma_{\text{gal}} = 0$, the IGM becomes too highly ionized, and the value of τ_{eff} is lower than that of Lee et al. (2015). Accordingly, it becomes impossible to match the value of τ_{eff} , due to the increased radiation output of quasars.

Given the fact that τ_{eff} from simulations is lower than that of Lee et al. (2015), the value of Γ_{tot} must be decreased in order to match the target value. As stated above, the radiation from quasars is more than sufficient to match the value of τ_{eff} , so the value of Γ_{qso} must be decreased. Therefore, it becomes necessary to choose a minimum value of Γ_{gal} , below which the radiation output of quasars must be decreased to agree with observations. We choose to have a finite value of Γ_{gal} for these simulations, since the stellar output of galaxies still provide a contribution to the hydrogen ionization level at these redshifts. Most models (Haardt & Madau, 1996, 2012) or measurements that infer this value (Bolton & Haehnelt, 2007; Becker et al., 2007; Faucher-Giguère et al., 2008; Becker & Bolton, 2013) of the UV background at these redshifts have a contribution from galaxies of $10^{-13} \text{ s}^{-1} \lesssim \Gamma_{\text{gal}} \lesssim 10^{-12} \text{ s}^{-1}$.

Following the models and measurements, we require for our simulations that $\Gamma_{\text{gal}} \geq 10^{-13} \text{ s}^{-1}$. If τ_{eff} is still too low given this minimum value of Γ_{gal} , the value of Γ_{qso} must be decreased. Because this value is only set indirectly by the number of photons produced by quasars in the ray tracing scheme, the total output of radiation from quasars is decreased to match τ_{eff} . This approach ensures that all of the simulations match the measured value of Lee et al. (2015). As the simulation progresses, if the ionization level needs to be increased to match the desired value, then the photon production of quasars is increased back to its default value before increasing Γ_{gal} .

This approach of modifying the value of Γ_{gal} on-the-fly to match the values of τ_{eff} is,

to our knowledge, unique to the simulations presented here. In addition to facilitating the comparison between the simulations and observations, this approach has several other benefits. For instance, by ensuring that we have the proper thermal state of the IGM, the pressure smoothing of the gas is more accurate. This property has implications for measurements related to the Lyman- α forest, discussed more fully in Sec. 5.5. Observations that depend on the value of τ_{eff} are also true apples-to-apples comparisons, and isolate the effect of the differences in the timing of reionization. Thus, this suite of reionization simulations allows for a straightforward determination of effects directly attributable to quasar activity as it pertains to helium reionization.

Table 5.1 summarizes the properties of the simulations examined in this paper. All simulations are conducted with a box size of $L = 200 h^{-1}\text{Mpc}$. This volume is large enough to include several high-luminosity quasars, which are important for helium reionization.¹ Our default resolution for the gas grid uses $N_g = 2048^3$ resolution elements. For dark matter, we use $N_{\text{dm}} = 2048^3$ particles as well. The grid on which the equations of radiative transfer are solved is coarser by a factor of 4, *i.e.*, $N_{\text{rt}} = N_g/64$. For all of the simulations in the suite, the same initial conditions for the dark matter particles and the gas cells are used, so that the only difference is the helium reionization history sourced by quasars. This allows us to isolate the impact that varying helium reionization has on measurements from our simulations, since the gas and matter distributions are largely the same. Indeed, the power spectra for dark matter in the simulations is effectively identical in all of the simulations, and the gas power spectra only show differences on small scales ($k \gtrsim 10 \text{ Mpc}^{-1} h$).

5.2.4 Details of the Simulation Suite

We will now discuss in detail some of the differences between the various simulations run. All of the simulations use the same set of initial conditions for dark matter and baryons, and as such the halo catalogs from the corresponding N -body simulation are the same. (See Sec. 5.2.1 for more information.) Further, all of the simulations use the patchy hydrogen reionization discussed in Sec. 5.2.3 at high redshift before helium reionization. The one exception to this is the simulation that uses a uniform UV background, Simulation H6, which uses the photoionization and photoheating rates from HM12. Additionally, also as discussed in Sec. 5.2.3, the simulations feature a dynamic renormalization of Γ_{gal} to match the reported value of τ_{HI} as provided by Lee et al. (2015). This renormalization applies to almost all of the simulations,

¹Despite this chosen volume size, the high-luminosity portion of the QLF may not be sufficiently resolved, especially at high redshift ($z \gtrsim 4$). Quantitatively, by integrating the product of the QLF and the number of photons produced per quasar, we are able to see how changing the box size (and hence the most-luminous quasar present in the volume) affects the number of photons included in the simulations. We find that the chosen box size only captures $\sim 75\%$ of the photons produced at $z \gtrsim 4$, with the captured fraction being less at higher redshifts. However, this quantity is highly dependent on the extrapolation of the QLF to high redshift, and steeper values of the bright-end slope mean that the discrepancy is not as large.

Simulation	Box size	N_{grid}	z_{50}	z_{99}	Δz_{50}	Δz_{90}	Quasar Model	t_0	QLF Amplitude	SED Amplitude
H1	200	2048^3	3.34	2.69	0.80	2.31	Q1	30.9	1	1
H2	200	2048^3	3.96	2.73	0.90	2.73	Q1	40	2	1
H3	200	2048^3	2.96	2.23	0.79	2.71	Q1	20	0.5	1
H4	200	2048^3	4.22	2.71	1.83	2.92	Q1	30.9	1	2
H5	200	2048^3	3.65	2.84	1.06	2.25	Q2	30	1.67	1.5
H6	200	2048^3	4.14	3.16	0.58	1.51	UVB

Table 5.1: A list of the parameters of the simulations presented in this work. The box size is in comoving h^{-1} Mpc. The redshifts z_{50} and z_{99} correspond to the redshifts when $x_{\text{HeIII}} = 0.5$ or $x_{\text{HeIII}} = 0.99$ by volume, as defined in Equation (5.3). Δz_{50} is the duration in redshift of the central 50% change in ionization fraction (defined in Equation (5.4)). For more information on the differences between the quasar models, see Appendix C.3 for the differences between quasar models Q1 and Q2. The quantity t_0 is defined in Equation (4.3), measured in Myr.

including H6, where all of the photoheating and photoionization rates are scaled to match τ_{eff} . As a point of comparison, we have run an additional simulation which purposely does not match the functional form of τ_{eff} in order to test for features that may appear as the result of helium reionization. We will discuss this simulation further in Appendix C.1.

1. The simulation H1 is one which uses a QLF that is generated in the manner discussed in Sec. 5.2.1. In general, the amplitude of the QLF is low at early times, but has a relatively steep low-luminosity slope. This leads to a quasar population that features a large number of low-luminosity objects. Since the effective lifetime of quasars is generally proportional to their luminosity, these sources are also relatively short-lived. As the Universe evolves, the amplitude of the QLF becomes greater, and the faint-end slope becomes shallower. This leads to a similar number of objects overall, but with larger, more luminous sources being the primary drivers of reionization. Large objects also tend to have larger regions of doubly ionized helium, since the longer lifetimes lead to larger reionization regions. This evolution becomes clear when visualizing the reionization process (see Fig. 5.4).
2. As noted in Sec. 5.1, there is some uncertainty in the overall amplitude of the QLF. In order to explore this uncertainty, we have run simulations H2 and H3, which use the same input QLF as H1, but with a change to the quasar luminosity function amplitude. In H2 the amplitude of the QLF is increased by a factor of 2 at all redshifts, and in H3, the amplitude is decreased by a factor of 2. In both cases, the lifetime of quasars is modified in order to reproduce the quasar clustering measurements of White et al. (2012), as discussed in Sec. 5.2.1. Although the statistical uncertainty of the QLF is lower than this amount at low redshift (*i.e.*, the data from Ross et al. (2013) has errors that are better than 10%), there are considerable uncertainties at high redshift. Further, there are potential sources of systematic uncertainty (*e.g.*, reddening of objects due to dust, obscured sources, or mischaracterization of potential sources as stars). By exploring changes in the amplitude of the QLF, we are better able to characterize the impact that different redshifts of helium reionization can have on observables.
3. A separate source of uncertainty related to the quasar sources is the normalization of individual quasar objects given a specific luminosity. As explained in Sec. 5.2.1, we use Equation (5.1) to convert from the observed magnitude to the specific luminosity at 2500 \AA L_{2500} , and the SED template of Lusso et al. (2015) to determine the EUV radiation. The statistical uncertainties of Lusso et al. (2015) are very small for the UV portion of the SED (wavelengths where $\lambda > 912 \text{ \AA}$), although differences arise when comparing the spectral indices between different SEDs (*e.g.*, Richards et al. 2006; Hopkins et al. 2007; Shang

et al. 2011). To explore some of the uncertainty associated with the SED, we have run Simulation H4 with a quasar model that has the same QLF amplitude as H1, but in which the photon number count has been increased by a factor of 2. This results in a comparable number of photons being produced as in H2, but with the same number of objects and topology as in H1. As a result, we expect the regions of doubly ionized helium to be larger than those found in H1, which would lead to patchier reionization. We would also expect the timing of reionization to be similar to H2.

4. As mentioned above, an additional uncertainty related to the observed quasar luminosity function involves the method by which observations from different redshift ranges are incorporated into one single QLF that evolves with redshift. We present two alternative methods of performing this combination in Appendix C.3. We call the two models Q1 and Q2. Simulation H5 uses a method slightly different from the fiducial one of Simulation H1. As with the uncertainties explored in Simulations H2 and H3, this comparison underlines the importance of accurately determining the QLF at all redshifts to better understand helium reionization. Note that when creating this QLF, several of the parameters of the QLF were modified in an effort to better reproduce the timing of the reionization found in Simulation H1.
5. Finally, as a point of comparison, we have run a simulation that does not include explicit quasar sources, and instead features a uniform UV background. The photoionization and photoheating rates are given by those in HM12. This allows for a comparison with other studies which employ a uniform UV background (Becker et al., 2011a; Puchwein et al., 2015). However, for a fair comparison with the other simulations presented here, we have renormalized these rates to match τ_{eff} as outlined in Sec. 5.2.3. Although only the value of Γ_{HI} affects the observed τ_{eff} , we apply the same renormalization to all of the photoionization and photoheating rates. Simulation H6 uses this uniform background, and can be thought of as the limiting case of having many, low-luminosity ($\mathcal{O}(10^9 - 10^{10} L_{\odot})$) objects drive helium reionization, rather than comparatively few high-luminosity ($\mathcal{O}(10^{12} - 10^{13} L_{\odot})$) ones.

Figure 5.1 shows the cumulative number of photons capable of ionizing helium ($h\nu \geq 54.4$ eV) as a function of redshift for each of the simulations presented here. The top panel shows as a point of comparison the total number of helium atoms in the volume. Notice that at early times, there are noticeable differences between Simulations H2 and H4, which in principle should both have twice as many photons as Simulation H1. These variations are likely due to shot-noise introduced by the relatively rare quasars, which becomes less extreme at later times. If all of the photons produced by quasars were absorbed by helium atoms, and there were no recombinations, then helium reionization would be completed when equality is reached.

Nevertheless, not all photons are absorbed (especially for the highest-energy frequency bin, due to the very low cross section of helium at these frequencies), and recombination is prevalent, especially in dense regions. Thus, the actual timing of reionization can be significantly different from when photon-helium atom equality is reached.

5.3 He III Ionization Fraction

One of the most basic results from the simulations is the calculation of the He III ionization fraction as a function of redshift. We define the ionization fraction x_{HeIII} as the (volume-weighted) amount of doubly-ionized helium relative to the total amount for all cells i in the volume:

$$x_{\text{HeIII}} \equiv \frac{1}{N} \sum_i \frac{n_{\text{HeIII},i}}{n_{\text{He},i}}. \quad (5.3)$$

Given a particular model for the quasar sources, the ionization fraction reflects the impact of these sources on the IGM. For instance, the duration of reionization gives some information about the important sources: a relatively long reionization argues for more sources that are fainter, and a shorter reionization is driven by a few large sources. When comparing features in observables produced from simulations, it is usually more important to compare results at the same ionization fraction than redshift. We refer to different redshifts related to an ionization fraction with a subscript, such that $z_n \Rightarrow x_{\text{HeIII}} = n\%$. For instance, $z_{50} \Rightarrow x_{\text{HeIII}} = 50\% = 0.5$. In addition to finding the redshift corresponding to different ionization fractions, we are also interested in quantifying the duration of reionization. To this end, we define

$$\Delta z_{50} \equiv z_{25} - z_{75}, \quad (5.4)$$

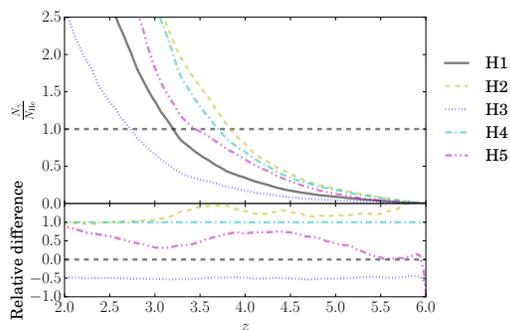


Figure 5.1: A comparison of the number of helium-ionizing photons ($h\nu \geq 54.4$ eV) produced by quasars in each of the simulation models as a function of redshift. Top: the cumulative number of photons in the simulation volume relative to the number of helium atoms. If all photons produced ionized helium with no recombinations, then helium reionization would be completed by the intersection with this line. Bottom: the number of photons produced relative to Simulation H1. The simulations are described in detail in Sec. 5.2.4. Note that Simulation H2 and Simulation H4 in principle produce a comparable number of photons as a function of redshift. Nevertheless, the two simulations have different reionization histories, as well as different reionization topologies.

which corresponds to the duration in redshift of the central 50% change in ionization fraction. We also define a similar quantity Δz_{90} , which represents the difference between $z_5 - z_{95}$. We report the redshifts associated with certain ionization fractions, as well as Δz_{50} and Δz_{90} , in Table 5.1, which summarizes the main results of the simulations. As a reference for converting Δz into time units, the shortest reionization scenario, Simulation H6, has a central duration of $\Delta z_{50} = 0.58 = 252$ Myr, whereas the longest reionization scenario, Simulation H4, has a duration of $\Delta z_{50} = 1.83 = 834$ Myr. These reionization scenarios take place over a relatively extended portion of the Universe’s history, and leave a lasting impression on the IGM.

5.3.1 Ionization Fraction Evolution

Figure 5.2 shows the volume-averaged ionization fraction of the different simulations as a function of redshift. We define the quantities Δz_{50} and Δz_{90} as the duration, in redshift, for the volume to transition from 25-75% ionized (by volume) and 5-95% ionized, respectively. Note that in general, helium reionization is a very extended process, with $\Delta z_{90} \gtrsim 2$ for almost all of the reionization scenarios, with Simulation H4 having very extended reionization times of $\Delta z_{90} \sim 2.9$. However, there is a large variation in the timing of reionization. The earliest simulation to reach 50% ionization is H4, which occurs at $z_{50} \sim 4.22$. The latest simulation is H3, which occurs at $z_{50} \sim 2.96$. The fiducial reionization scenario, H1, is 50% ionized at $z_{50} \sim 3.34$. As pointed out below in Secs. 5.4 and 5.5, in general observations are more sensitive to the end of helium reionization, when the volume becomes 90-95% doubly ionized. The main exception to this result is Simulation H2, which reaches a maximum temperature at $z \sim 3.41$, which corresponds to an ionized fraction of 80%. The reason for the difference is related to the method by which the quasar emission is modified to match τ_{eff} , as outlined in Sec. 5.2.3. Nevertheless, knowing the full reionization history has important implications on the thermal history of the IGM.

Figure 5.3 shows visualizations of Simulation H1. The four columns, from left to right, show the He III ionization fraction x_{HeIII} , the gas temperature, the He II photoionization rate Γ_{HeII} , and the He II photoheating rate Λ_{HeII} . The rows show the same slice of the simulation at increasing values of ionization fraction, which from top to bottom are $x_{\text{HeIII}} = 0.1, 0.25, 0.5, 0.75,$ and 0.99 . The corresponding redshift is shown on the right side of the panels. These slices show a segment of the yz -plane of the simulation, with a thickness of 1 radiative transfer cell in the x -direction. This width corresponds to a comoving distance of $\sim 400 h^{-1}\text{kpc}$. In a loose sense, the first and second columns are integrated quantities corresponding to the third and fourth columns, respectively. In both cases, the figure shows only photoionization and photoheating rates, which in particular does not include collisional ionization and heating prevalent in regions of high density. Nevertheless, the photoionization and photoheating rates are dominated by the contribution of photons from quasars in the volume. Further, note that for the temperature of the IGM (Column 2), the

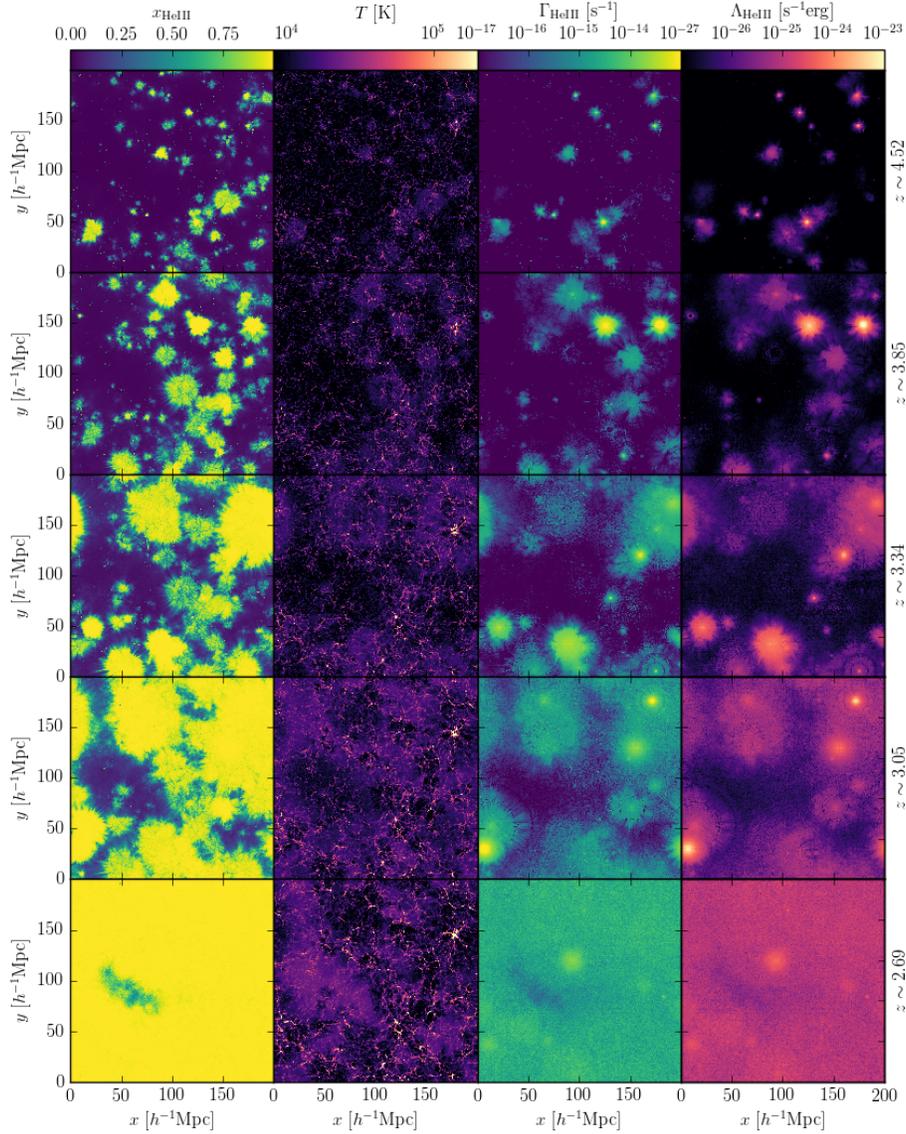


Figure 5.3: A comparison of different properties in Simulation H1. Each panel shows a 2-dimensional slice through the simulation volume with the thickness of a single RT cell ($\sim 400 h^{-1} \text{kpc}$). The columns, from left to right, show the He III ionization fraction, the gas temperature, the He III photoionization rate Γ_{HeIII} , and the He III photoheating rate Λ_{HeIII} . Note that the third and fourth columns only include the contribution to the photoionization and photoheating from the quasar sources, and do not include other sources of ionization and heating (*e.g.*, collisional ionization or heating). The different rows show redshift snapshots corresponding to volume-average ionization fractions of $x_{\text{HeIII}} = 0.1, 0.25, 0.5, 0.75$, and 0.99 , from top to bottom. Note that early on in the reionization process, the average He III bubble size is small ($\sim 5 h^{-1} \text{Mpc}$), but later on in reionization, the size of ionized regions becomes much larger ($\sim 50 h^{-1} \text{Mpc}$ in some cases). This change in bubble size is due to relatively long lifetimes of luminous quasars.

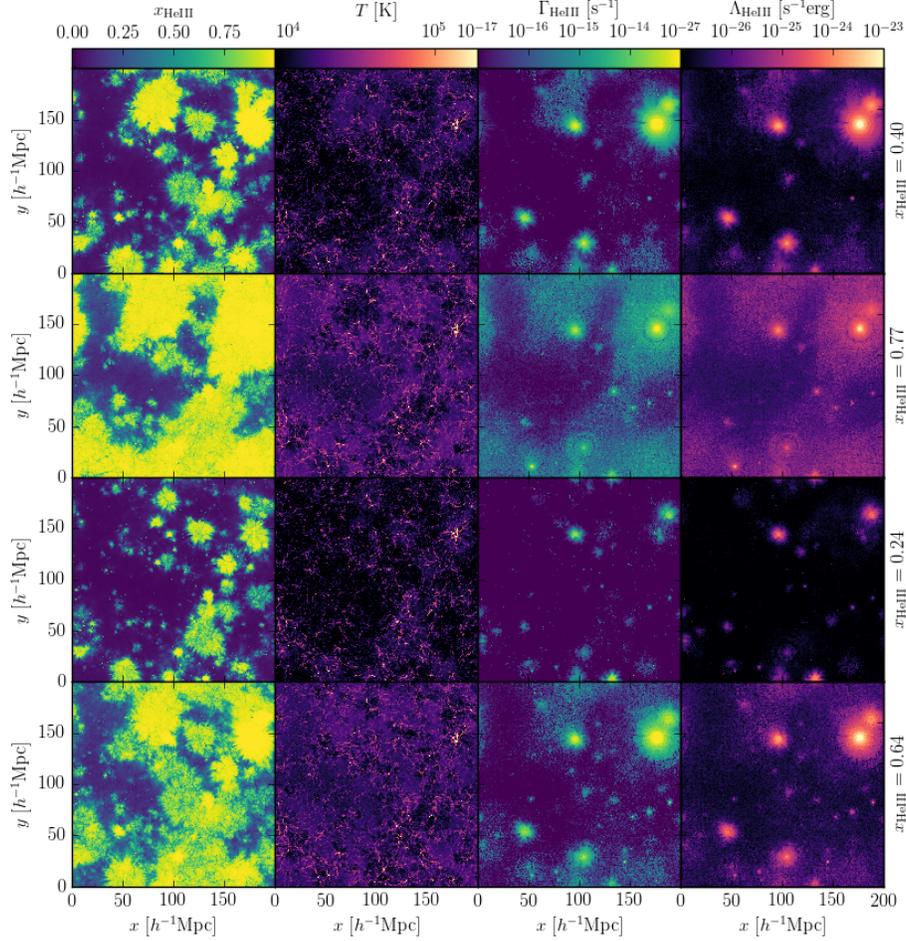


Figure 5.4: A similar plot to Figure 5.3, but comparing different simulations at redshift $z \sim 3.5$. The rows, from top to bottom, show Simulation H1, H2, H3, and H4. The value of x_{HeIII} is shown to the right of each row. Note that in addition to the obvious differences in helium ionization level morphology, the temperature of the IGM (second column) is also very different for the different simulations. There is also an apparent difference between the quasar models used in Simulation H2 (second row) and Simulation H4 (fourth row), which in principle have similar photon counts but are at different ionization levels. The reasons for these differences are discussed in Sec. 5.3.1.

hottest regions are found along filaments and other dense regions of cosmic structure. Despite these regions being the hottest, photons from quasars dramatically heat the low-density IGM by several thousand kelvin. See Sec. 5.4 for further discussion of the IGM temperature.

As discussed in Chapter 4, in our model the clustering of quasars indirectly affects their lifetimes. Because the lifetimes of quasars affects the size of reionized regions (visible in Figures 5.3 and 5.4), having the proper clustering affects the coherent scale of reionization. The size of reionized regions also affects the heating of the IGM, as larger reionization regions encompass regions of moderate- to low-density earlier than smaller regions. The reason for this is that the relatively fast timing of recombination means that moderate- to high-density gas quickly recombines, and requires additional radiation in order to re-reionize. For comparatively large regions, more of the gas being ionized is low-density, so there is less recombination. Also worth nothing is that the relatively high clustering leads to early overlap of reionized regions, which again reflects on the timing of ionization reaching regions of low-density.

Figure 5.4 shows visualizations of Simulations H1, H2, H3, and H4, all at redshift $z \sim 3.5$. Note that, though the underlying gas and large-scale structure are largely similar (as can be

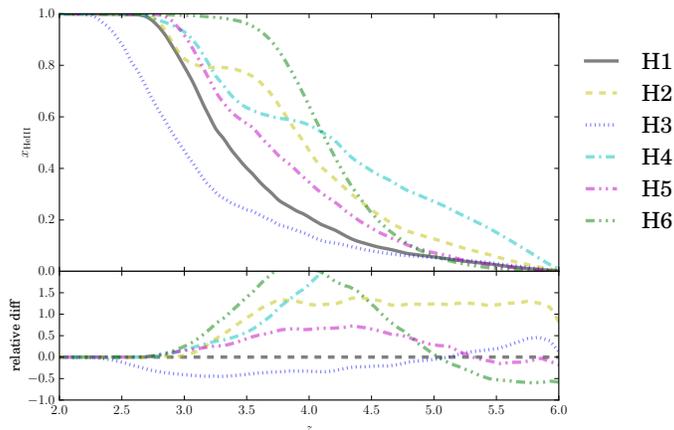


Figure 5.2: The helium ionization fraction x_{HeIII} as a function of redshift for the different quasar models explored in this work. The top panel shows the ionization fraction, and the bottom panel shows the relative difference compared to the fiducial simulation (H1). The different models presented here are described in Sec. 5.2.3, and summarized in Table 5.1. Among the models shown for comparison is a simulation with a uniform UV background from Haardt & Madau (2012) (Simulation H6). Note that most of the simulations reach a 99% ionization fraction in the range of $2.7 \lesssim z \lesssim 3$, which is consistent with observational findings (Dixon & Furlanetto, 2009; Worseck et al., 2011). In all simulations, the duration of helium reionization is typically $0.8 \lesssim \Delta_{50} \lesssim 1$, with the notable exceptions of Simulations H4 and H6. The durations for these simulations are significantly longer and shorter, respectively, of the other ones. The former has a long duration due to the early onset of relatively massive quasars, while the latter assumes a quasar emissivity that rises sharply starting at redshift $z \sim 5$. See the text in Sec. 5.3 for additional discussion.

seen by comparing Column 2 between the different rows), the ionization and temperature distribution are very different for the different simulations. The differences are driven by the different quasar models used in the simulations. Of particular interest is the difference between Simulations H2 and H4 (Rows 2 and 4). When performing simple photon-counting calculations, as seen in Figure 5.1, both of these simulations should produce a similar number: Simulation H2 has a factor of 2 increase in the total number of quasars at a given epoch, whereas Simulation H4 has a factor of 2 increase in the number of photons produced per quasar.

Despite this similarity, there are significant differences between the simulations, most notably the ionization fraction (Column 1). Additionally, from Columns 3 and 4, it can be seen that Simulation H2 has greater quasar activity at a given redshift. Part of the differences between the simulations can be attributed to the method by which quasars are populated in the volume: as explained in Sec. 5.2.1 (and more in-depth in Paper I), quasars are placed in halos using abundance matching. Thus, when the amplitude of the luminosity function is increased, sources of the same luminosity are placed in lower-mass halos. In addition to making rare objects more common, there are more sources in general. This feature leads to a greater number of photons intersecting gas cells that have not been previously exposed to quasar radiation.

Conversely, in Simulation H4, the number of photons produced per source is increased, but the total number of sources is the same as in Simulation H1. (Indeed, the same quasar catalog is used in the two simulations, with only the normalization of quasar radiation changed between the two. Note how the general morphology of ionized regions in Column 1, and the instantaneous quasar activity in Columns 3 and 4, are very similar between Rows 1 and 4.) Although there are two times as many photons being produced per source, the long mean free path of helium-reionizing photons means that not all photons are absorbed. Further, due to spectral filtering of the radiation from quasars, the photons with energy $h\nu \sim 54.4$ eV will be readily absorbed before more energetic ones, changing the effective SED of the quasar sources (Meiksin et al., 2010). The higher energy ones typically are not absorbed, leading to the large discrepancy in neutral fraction observed between these simulations. Thus, although a simple semi-analytic calculation would yield the same reionization time for these two simulations, we can see that a full treatment leads to important differences between the two cases.

The ionization fraction observed in our simulations is worth comparing with the results of McQuinn et al. (2009) and Compostella et al. (2013), hereafter M09 and C13. The duration of reionization in our simulations is comparable to the models explored in M09 (as seen in their Fig. 3). However, the reionization histories in C13 are much briefer than the ones seen here. This is largely due to the fact that the quasar population in their fiducial reionization model does not include sources for $z > 4$. The authors include an additional “extended” model which includes sources beginning at $z = 5$, which shows a duration of reionization more comparable to the ones in M09 and this work. Observations from McGreer et al. (2013) show a

non-negligible population of high-redshift quasars, which in Simulation H1 drives the ionization fraction of helium to have a value of a few percent at $z \sim 5$, with the volume being nearly a quarter ionized by $z \sim 4$. Thus, future studies should include high-redshift quasars as an important part of helium reionization.

5.4 The Temperature History of the IGM

One important impact of helium reionization on the IGM is the temperature feedback. Since quasars emit a hard spectrum with many energetic photons, and the IGM is in a highly ionized state, the excess energy remaining after photoionization is converted into heat in the gas. Though secondary ionizations are possible (*e.g.*, Shull 1979; Furlanetto & Stoever 2010), due to the ionization level of the IGM, their impact is negligible for helium reionization (McQuinn et al., 2009). Photoheating from radiation from quasars increases the average temperature of the IGM by $\sim 10,000$ K, and as will be seen in Secs. 5.4.1 and 5.4.2, contains important information about the history of helium reionization.

5.4.1 Temperature-density relation

The relationship between the temperature of the IGM T and the baryon overdensity $\Delta \equiv \Delta_b$ is an important measure of the state of the IGM, and is intimately related to the reionization process. One can write the relationship between temperature and density as a power law, and fit for the two parameters that define it (Hui & Gnedin, 1997):

$$T(\Delta) = T_0 \Delta^{\gamma-1}, \quad (5.5)$$

where T is the gas temperature, and T_0 and γ define the power law relation between the gas density and temperature. This is the so-called temperature-density relation, also sometimes called the equation of state of the IGM (though note that it is not a true equation of state). Hui & Gnedin (1997) showed that at late times following hydrogen reionization, the slope of the relation approaches $\gamma = 1.62$. In general, this relationship should hold for the low-density gas in the IGM where adiabatic cooling/heating and a uniform radiation field following reionization are the dominant sources of temperature change. However, the addition of heat from helium reionization changes the slope of this relation, as well as the overall amplitude.

Figure 5.5 shows the temperature-density relation for the gas in the different simulations. The relationship is shown at several different redshifts, in order to demonstrate several different effects that reionization has on the IGM temperature. In particular, the general trend is indicative of an “inside-out” reionization scenario. In such a scenario, the radiation from sources (quasars, in this case) propagate outward, and are absorbed in high-density regions near sources before low-density ones, depositing heat as the radiation is absorbed. Because the gas is reionized at different

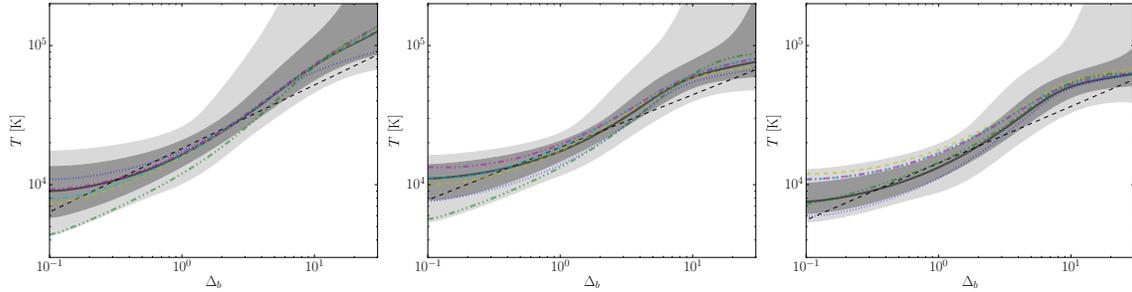


Figure 5.5: A comparison of the gas temperature T as a function of baryon density Δ . The panels, from left to right, are the temperature-density relation as measured from the simulations at $z \sim 2.5, 3,$ and 3.5 . The gray shaded regions correspond to the 68th and 95th percentiles of Simulation H1. Note that the overall amplitude of the relation rises as redshift decreases, showing that the overall temperature of the IGM increases as helium reionization heats up the volume. In addition, the slope of the relation becomes steeper as the simulations evolve. The temperature of relatively dense regions ($\Delta \gtrsim 10$) continues to rise even after helium reionization is largely finished. This is due to recombination of the gas, followed by additional reionization, adding more heat to the gas. Conversely, once ionization is finished, the low-density regions ($\Delta < 1$) cool adiabatically, with little heat input. The dashed black line in each figure is the best-fit power law relation given by Equation (5.5) for Simulation H1. This should be compared with the gray solid line, which actually follows the relationship for each density value Δ . Note that, in general, the simple power law does not accurately capture the relationship between density and temperature. See the text for additional discussion.

times, and is dominated by adiabatic cooling following reionization, the relative temperature between different gas densities reflects the reionization history. In particular, the temperature of underdense regions can in fact be higher than mean-density ones, due to the radiation from quasars tending to reach those regions at a later redshift. In the meantime, the gas from high-density regions has additional time to adiabatically cool. Because the amount of heat deposited in the gas from photoionization does not depend on the density, the gas from higher density regions may be a lower temperature than the low-density gas when the low-density gas is first reionized. Thus, the temperature-density relation can be relatively flat for medium- to low-density gas, and even turn over such that low-density regions have a *higher* temperature than mean-density ones (*e.g.*, as in Trac et al. 2008 for hydrogen reionization). The simulations presented here do not exhibit this inversion, due to both the longer mean free path of helium ionizing photons and the relatively smaller amount of adiabatic cooling experienced by gas at this redshift.² Nevertheless, several of our simulations, and

²The adiabatic cooling of gas causes the temperature to decrease as $T \propto (1+z)^2$; thus, a duration of reionization in redshift space of $\Delta z \sim 1$ at the higher redshift of hydrogen reionization leads to a

Simulation H5 at $z \sim 3$ in particular, show a relatively flat relation for underdense regions.

Another feature in Figure 5.5 is the evolution of regions of high density ($\Delta \gtrsim 10$). In these regions, the density of gas is high enough that an appreciable fraction of the doubly ionized helium can recombine with electrons to form singly ionized helium. Once the gas has recombined, it can undergo an additional reionization event, which will deposit additional heat into the gas. As can be seen in the Figure, the higher density regions show higher temperatures as redshift decreases, even after helium reionization is nominally finished. Thus, the temperature of these different regions at the same redshift can somewhat break the degeneracy between the different reionization scenarios. Since these differences are visible in higher density gas, it may be possible to observe these differences in the Lyman- β forest, since these observations saturate at higher densities than Lyman- α (Dijkstra et al., 2004; Iršič & Viel, 2014).

Figure 5.6 shows the evolution of the parameters of the temperature-density relation given in Equation (5.5) as a function of redshift for the different simulations. As can be seen by the general structure of Figure 5.5 and noted in C13, fitting the entire temperature-density relation to a single power law may not be the most optimal parameterization, due to the wide dispersion of temperatures at a given density value. We should note that part of the difficulty in fitting the result to a power law comes from the approximate nature of the relation: for large values of Δ , the approximation breaks down. Further, the resolution of the simulations does not capture all of the structure of the IGM, which leads to smoothing at certain scales. Nevertheless, we present these results for the sake of comparison.

In general, we see a similar trend to Figure 5.5, where the temperature value at mean density T_0 increases as reionization proceeds, reaches a peak value, and then decreases again. This is a general trend seen in the thermal evolution of the IGM, and is explored more below in Sec. 5.4.2. Another general trend is the evolution of the power law index γ which is roughly consistent between simulations. We reproduce the observation of M09 that $\gamma \sim 1.3$ during the bulk of helium reionization for our different scenarios. In the lower panel of Figure 5.6 we show the value of $\gamma = 1.62$, which is the asymptotic value of the IGM from Hui & Gnedin (1997) following hydrogen reionization without additional sources of photoheating.

As can be seen from Figure 5.6, simulations that include a patchy hydrogen reionization are not consistent with this value, though Simulation H6, which features a significantly earlier hydrogen reionization epoch, approaches this value. However, once helium reionization begins, there is a notable flattening of the temperature-density relation (where $\gamma = 1$ represents the limit of an isothermal gas). As a larger portion of the volume becomes ionized, denser regions will recombine and undergo additional reionization events, leading to additional heat being deposited at those densities. Conversely, low-density regions are dominated by adiabatic cooling. This leads to an overall steepening of the slope γ , a trend seen at low redshifts following

larger relative change in temperature than the lower redshift of helium reionization.

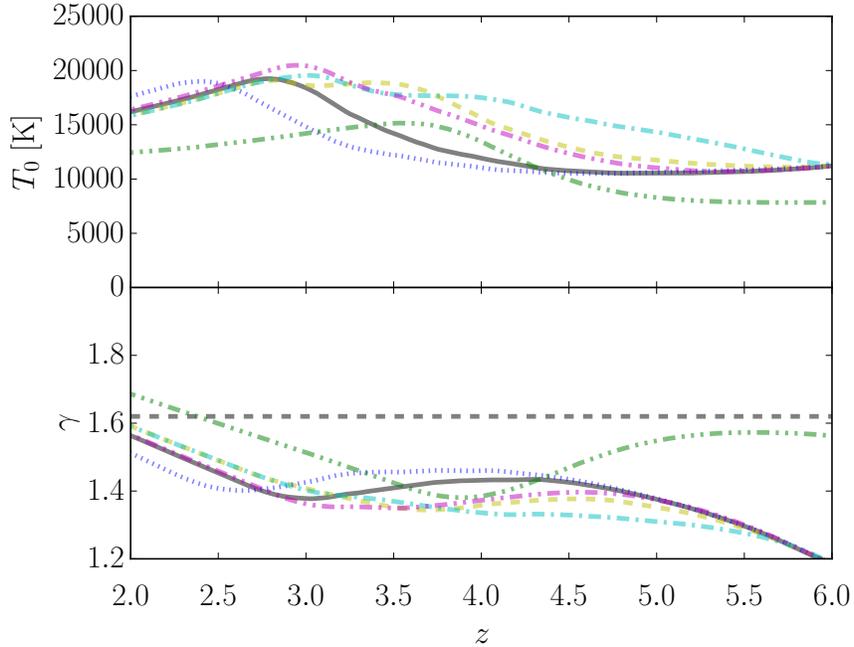


Figure 5.6: The parameters of the IGM power law temperature-density relation in Equation (5.5) as a function of redshift for the different simulations. The top panel shows the temperature parameter T_0 , and the bottom panel shows γ . In the panel for γ , we have shown the line of $\gamma = 1.62$, the predicted slope from Hui & Gnedin (1997) for the relation following hydrogen reionization. At early times, Simulation H6 approaches this value, but then deviates from it following helium reionization. The other simulations do not approach this value, likely due to the fact that helium reionization begins before a steady state can be established. As discussed in the text, the initial flattening of the slope ($\gamma < 1.62$) is due to the inside-out nature of reionization, and the later steepening ($\gamma \sim 1.62$) due to establishing equilibrium with the radiation field. For the T_0 parameter, the rise and fall of the value is consistent with the rise and fall seen in Figure 5.5. See the text for further discussion.

the completion of helium reionization. These trends are also visible in Figure 5.5. Note in particular that at $z \sim 3$, most of the simulations have a comparable value of γ . Indeed, the shape of these temperature-density relations in the central panel of Figure 5.5 is similar, albeit with different vertical offsets.

The results of M09 and C13 are largely consistent with the findings presented here. Before helium reionization begins, the temperature-density relation tightly follows a power law expression. Once helium reionization begins, the distribution of temperature as a function of density becomes highly variable, with a large dispersion forming for a given density value. This dispersion signifies the inhomogeneous reionization process, and is a general feature of helium reionization. Additionally, as in

C13, we find that the overall relation between temperature and density is ill-fit by a single power law. C13 finds that the temperature-density relation flattens out and begins to turn over at $\Delta \sim 10$ (as can be seen in their Fig. 8). Although we do not see a turn-over in our measurements, it is still clear that using a single power law to characterize the relationship between temperature and density is insufficient for the IGM following reionization.

5.4.2 Temperature at mean density

An important marker of the progress of helium reionization is the temperature at mean density of the simulation ($\Delta \sim 1$), since the temperature in these regions is dominated by adiabatic cooling of the Universe and heating from radiative transfer (Hui & Gnedin, 1997). It is the interplay of these two factors that determines the temperature of these regions of average density. The average temperature of these regions show two characteristic “bumps” as a function of redshift: one initial increase from $T \sim 200$ K to $T \sim 10^4$ K as a result of hydrogen reionization at $8 \lesssim z \lesssim 10$, and a subsequent increase in temperature from $T \sim 10^4$ K to $T \sim 2 \times 10^4$ K at $2 \lesssim z \lesssim 3.5$ as a result of helium reionization (Furlanetto & Oh, 2008b; Puchwein et al., 2015). In between the two epochs of reionization, and following helium reionization, adiabatic cooling dominates, and so the average temperature decreases. The locations and widths of these features can provide valuable insight into the timing and duration of reionization.

Previous studies of the mean temperature of the IGM, both semi-analytic (Furlanetto & Oh, 2008b) and using simulations with a uniform UVB (Puchwein et al., 2015; Bolton et al., 2016) have shown the general picture of the IGM temperature should hold, and can therefore be used to extract information about reionization. For our purposes here, we concern ourselves

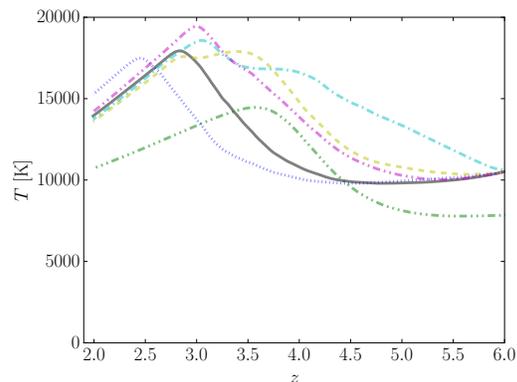


Figure 5.7: The median temperature at mean density ($0.95 \leq \Delta \leq 1.05$) of the IGM as a function of redshift. The temperature at mean density is significantly higher for the patchy hydrogen reionization scenarios than for the one with a uniform UVB (Simulation H6), due to the significantly earlier hydrogen reionization. Once quasar activity begins, the temperature of the IGM rises due to photoheating. The peak corresponds to an ionization fraction of $x_{\text{HeIII}} \sim 0.90-0.95$, marking the tail-end of reionization. Following helium reionization, the mean-density gas begins to adiabatically cool again, leading to the peak structure seen in the Figure. See the text for additional discussion.

primarily with this second epoch of heating in the IGM, corresponding to helium reionization.

Figure 5.7 shows the median temperature at mean density of the different simulations. In order to compute the temperature at mean density, at each time step in the simulation we find the median temperature (as well as the ± 68 th and 95th percentiles) of all gas cells that have $0.95 \leq \Delta \leq 1.05$. At high redshift ($z \gtrsim 6$), the simulations have largely the same temperature, because the IGM temperature is dominated by hydrogen reionization. As explained in Sec. 5.2.3, all of the simulations with explicit quasar sources use a semi-analytic method for calculating patchy hydrogen reionization. The exception to this is Simulation H6, which uses the uniform UV background of HM12 for both hydrogen and helium reionization. Notably, the timing of hydrogen reionization is significantly earlier than for the patch hydrogen method used ($z_{\text{re}} \sim 13$ for HM12 compared to $z_{\text{re}} \sim 8$ for the patchy hydrogen), so the IGM has had additional time to adiabatically cool. This leads to the lower initial temperature at $z \sim 6$ seen in Figure 5.7.

Also note that in Figure 5.7, the temperature of the IGM peaks at a redshift that corresponds to 90-95% helium ionization level. This is consistent with the idea that the gas at mean density composes a large fraction of the volume of the simulation volume, and so will preferentially reionize later than regions of high density. This “inside-out” picture of reionization is similar to that of hydrogen reionization (Loeb & Furlanetto, 2012). Following this peak in the IGM temperature, the adiabatic cooling of the Universe becomes the dominant mechanism, because this comparatively low-density gas generally does not recombine (due to recombination being $\propto \rho_g^2$, as shown in Equation (5.2)).

5.5 Lyman- α Forest Measurements

An important observational tool used to understand helium reionization is the Lyman- α forest. Observationally, there have been many rich data sets using the Lyman- α forest, especially for cosmological measurements. The BOSS sample (Lee et al., 2013) has been used to observe the baryon acoustic oscillation (BAO) feature (Busca et al., 2013; Slosar et al., 2013), as well as generate one-dimensional power spectra (Palanque-Delabrouille et al., 2013), which have been used to constrain neutrino masses and other cosmological parameters (Palanque-Delabrouille et al., 2015). High-resolution measurements from Keck-HIRES and Magellan-MIKE (Lu et al., 1996; Becker et al., 2007, 2011b; Calverley et al., 2011) have given us information about the temperature history of the IGM.

Synthetic Lyman- α spectra can be created for the H I and He II densities. (See Chapter 6 for further discussion of the He II Lyman- α forest). In the following analysis we have drawn the spectra along the x -axis of the simulation, though we find nearly identical results when projecting along different axes. Once these spectra have been calculated, they can be used to measure the effective optical depth τ_{eff} of the volume,

compute the flux PDF, and calculate one-dimensional power spectra. To generate a synthetic sightline, we define a set of pixels along a line of sight in the simulation volume, such that the number of pixels is equal to the number of grid cells. For the resolution level discussed in these simulations, this means $N_{\text{pix}} = 2048$.

For each pixel i , the optical depth of the pixel τ_i is calculated due to the contributions of every other pixel according to the formula (Bolton et al., 2009b):

$$\tau_i = \frac{c\sigma_\alpha dR}{\pi^{1/2}} \sum_{j=1}^{N_{\text{pix}}} \frac{n_{\text{HI}}(j)}{b_{\text{HI}}(j)} H(a, x), \quad (5.6)$$

where $\sigma_\alpha = 4.479 \times 10^{-18} \text{ cm}^{-2}$ is the cross-section of the Lyman- α transition, $b_{\text{HI}} = \sqrt{2k_B T/m_{\text{H}}}$ is the Doppler parameter, dR is the (physical) width of the pixel, and $H(a, x)$ is the Voigt-Hjerting function (Hjerting, 1938):

$$H(a, x) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2}}{a^2 + (x - y)^2} dy, \quad (5.7)$$

where $x = [v_{\text{H}}(i) - u(j)]/b_{\text{HI}}(j)$ is the difference in redshift space between pixels i and j relative to the Doppler broadening, $u(j) = v_{\text{H}}(j) + v_{\text{pec}}(j)$ is the total velocity difference of Hubble flow plus peculiar velocity, $a = \Lambda_\alpha \lambda_\alpha / 4\pi b_{\text{HI}}(j)$ represents the gas damping, where $\Lambda_\alpha = 6.265 \times 10^8 \text{ s}^{-1}$ is the damping constant and $\lambda_\alpha = 1215.67 \text{ \AA}$ is the wavelength corresponding to the Lyman- α transition. In order to efficiently compute the Voigt-Hjerting function, we use the analytic approximation provided by Tepper-García (2006).

As can be seen from Equations (5.6-5.7), the thermal properties of the gas enter in the form of the Doppler parameter b . This term increases as the temperature of the gas increases, and serves to broaden the apparent width in velocity space of a particular gas parcel. The tendency of absorption features to widen in velocity space as the temperature increases can be used to learn about the thermal state of the IGM. More approximately, the local optical depth of the IGM will depend on the average temperature of the volume. We will further discuss some of the implications of this process below in Sec. 5.5.2.

5.5.1 Effective Optical Depth

Once the optical depth for each pixel has been calculated, the corresponding flux is given simply by $F_i = \exp(-\tau_i)$. We can then define the effective optical depth of the volume by averaging over all values of the flux:

$$\langle F \rangle = \exp(-\tau_{\text{eff}}). \quad (5.8)$$

Note that in general $\tau_{\text{eff}} \neq \langle \tau \rangle$. The effective optical depth as a function of redshift has been measured to high precision as a volume-averaged quantity for the H I forest

(Lee et al., 2015) and for individual objects of the He II forest (Worseck et al., 2014). In Lee et al. (2015), the BOSS survey measures more than 50,000 quasar spectra at intermediate-to-high redshift, and has a formula for the evolution of the effective optical depth as a function of redshift $\tau_{\text{eff}}(z)$.

In general, cosmological simulations of the Lyman- α forest must renormalize the flux level measured in order to match the observed optical depth measurements (see, *e.g.*, Bolton et al. 2009b). This is due to the fact that the resolution of these simulations is typically not high enough to capture the small-scale, high-absorption Lyman-limit systems (LLSs) and damped Lyman- α systems (DLAs) that can lead to cosmological simulations predicting too high of a value of τ_{eff} (though see McQuinn et al. 2009 for attempts to account for these systems in simulations). Typically, this renormalization of Lyman- α spectra is done in post-processing when the sightlines are generated.

Figure 5.8 shows τ_{eff} for all of the simulations presented in this work. As noted in Sec. 5.2.3, this quantity is matched by construction for all of the simulations. Note that, in general, the agreement is excellent. For redshifts $z \lesssim 6$ (the nominal end of hydrogen reionization, after which $\tau \lesssim 1$), all of the simulations match the observed value from Lee et al. (2015) to within a few percent. This matching allows for more straightforward comparison between the simulations, as well as with observations.

As explained in Secs. 5.2.3 and 5.2.4, our simulations change the value of Γ_{gal} on-the-fly in order to match the value of τ_{eff} as specified by Lee et al. (2015). By ensuring that all of our simulations match the same value of τ_{eff} , we are better able to compare them with each other and with the observations. Previous studies of the Lyman- α forest (Theuns et al., 2002; Ciardi et al., 2003; Dall’Aglio et al., 2008; Faucher-Giguère et al., 2008) have reported a dip in τ_{eff} at $z \sim 3.2$. In some of these works, the authors cited this dip as evidence of helium reionization, due to an

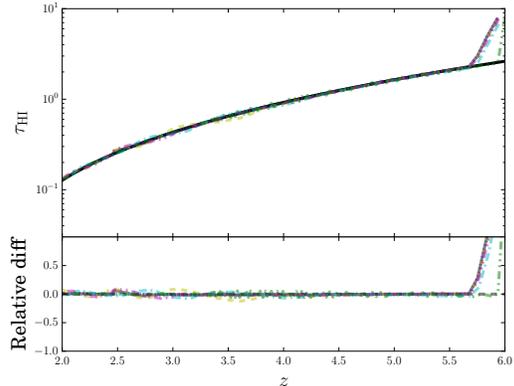


Figure 5.8: The effective optical depth of hydrogen $\tau_{\text{eff,HI}}$ as a function of redshift for the different simulations. The solid black line is the observational data from Lee et al. (2015). The linestyles for the simulations are the same as in Fig. 5.2. As discussed in Sec. 5.5.1, the amount of hydrogen-ionizing radiation from galaxies Γ_{gal} is modified while the simulation is running, so that this quantity is matched by construction. This avoids the requirement of renormalizing the simulations in post-processing. Note that for the HM12 simulation, Γ_{gal} was specified by the model of Haardt & Madau (2012) rather than being modified to match $\tau_{\text{eff,HI}}$, leading to the apparent difference between the simulations.

increased IGM temperature decreasing the optical depth. By matching the τ_{eff} of Lee et al. (2015), which does not contain this dip, it is possible that we would miss this feature. We explore this possibility more in Appendix C.1.

When comparing against the simulations of M09 and C13, we notice that in all cases, τ_{eff} is comparable to the most recent determinations of the H I Lyman- α forest for the then state-of-the-art measurements. Our simulations are the only ones that renormalize Γ_{HI} in real time, so we are able to match the value of τ_{eff} by construction. Nevertheless, our values of Γ_{HI} are comparable to those in M09 and C12, as well as HM12. We again note that the relative uncertainty on Γ_{HI} is much larger than that of τ_{HI} , and so to generate more realistic comparisons with measurements of the H I forest, we advocate for matching the value of τ_{eff} by construction, as we have done here.

5.5.2 Flux PDF

Another statistic related to the Lyman- α forest is the flux PDF. This measurement is carried out by taking the flux value of each of the pixels in the sightlines of the Lyman- α forest and creating a normalized PDF of their values. The result gives additional information about the distribution of gas in the IGM. The flux PDF is also dependent on the resolution of the measurement. For instance, compare the results from a relatively high-resolution measurement (Calura et al., 2012) with that of a relatively low-resolution measurement (Lee et al., 2015). In the lower resolution case, the pixels of extreme absorption or emission become averaged, and the flux PDF tends toward the mean. Thus, the measured PDF is resolution dependent.

From a simulation point of view, the resolution of the gas grid (and to a lesser extent, the radiation grid) affects the resolution of the Lyman- α forest. For the default-resolution grid at $z \sim 3$, a single gas cell has an equivalent velocity width of $\Delta v = 7.3$ km/s. This resolution level is significantly greater than that of BOSS ($\Delta v \sim 69$ km/s, Lee et al. 2015), though not as good as Keck-HIRES ($\Delta v \sim 6.6$ km/s, Lu et al. 1996).

Figure 5.9 shows the flux PDF of the H I Lyman- α forest as a function of redshift including the various simulations. The figure also includes the measurements of Calura et al. (2012). Note that the spectra from Calura et al. (2012) were taken at UVES, with a FWHM of 6.7 km/s, slightly higher than the resolution of our simulations. As a result, the different resolution may have a non-trivial impact on the shape of the resulting flux PDF. Note that, largely, the flux PDF is identical for different simulations at the same redshift, despite having different He III ionization fractions and thermal histories. This result implies that to given the same underlying gas structure, the flux PDF depends on having the same value of τ_{eff} . Given the same large-scale structure and τ_{eff} , our result shows that helium reionization is largely undetectable in the hydrogen flux PDF.

Nevertheless, there are still several trends that are visible upon closer inspection.

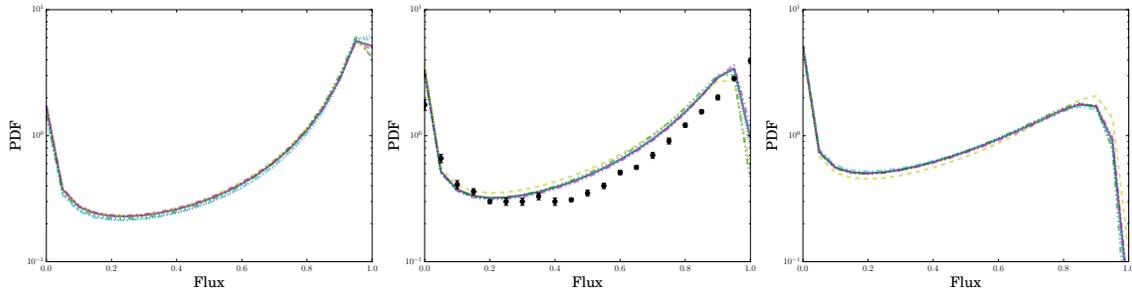


Figure 5.9: A comparison of the flux PDF of the H I Lyman- α forest at $z \sim 2.5$, 3, and 3.5, from left to right. All of the simulations have the same color scheme as in Figure 5.2. In the middle panel, the data points are taken from the results of Calura et al. (2012), at $z \sim 2.9$. Note that largely, all of the simulations show a nearly identical distribution of fluxes. This result implies that the flux PDF is only weakly sensitive to the temperature information of the IGM, since largely the only difference between the simulations (besides the helium ionization fraction) is the temperature. The flux PDF is instead more sensitive to τ_{eff} and, observationally, the continuum-level uncertainty of the Lyman- α forest. See the text in Sec. 5.5.2 and Appendix C.2 for additional details.

Note that after helium reionization is largely completed at $z \sim 2.5$, the value of the flux PDF in the highest transmission bin of $F \sim 1$ are ordered by the helium ionization fraction: Simulation H3 has the highest value in this bin, and Simulation H6 has the lowest. Note further that helium reionization is still ongoing for Simulation H3, whereas for the other simulations reionization is largely over (Fig. 5.2).

We can understand this trend by employing the so-called fluctuating Gunn-Peterson approximation (FGPA; Croft et al. 1998). The FGPA assumes that the gas of the IGM accurately follows a temperature-density relation of the form found in Equation (5.5), and is in photoionization equilibrium with a uniform ionization background. Under these assumptions, the local optical depth of the IGM τ_{HI} can be expressed in terms of the gas density, mean temperature of the IGM, and the H I photoionization rate, along with other cosmological parameters. In particular, it can be shown that the optical depth is related to the temperature as $\tau_{\text{HI}} \propto T^{0.7}$. Thus, for reionization histories with a higher average temperature, there is an increased local value of τ , leading to an overall lower flux value in high-density regions. Therefore, the comparatively high value for the flux PDF in the bin where $F \sim 1$ for Simulation H3 can be interpreted as conveying information about the thermal state of the IGM. Indeed, Lee et al. (2015) have proposed using the flux PDF to learn information about the thermal state of the IGM at different redshifts.

One point to note is that there is a visible difference between the observations of Calura et al. (2012) at $z \sim 2.9$ and the results from the simulations at $z \sim 3$, shown in the middle panel in the plot. The flux PDF at intermediate flux values in

the simulations is higher than that of the observations, until the highest bin (where there is almost total flux transmission). Part of the difference can be attributed to the fact that the simulations and the observations are normalized to different values of τ_{eff} : the simulations use the value from Lee et al. (2015), whereas the observational results determine the parameters for $\tau_{\text{eff}}(z)$ based on their measurements. At $z \sim 3$, the results for τ_{eff} from Lee et al. (2015) are higher than those from Calura et al. (2012) by about 30%. This result accounts for some of the difference in the flux PDF, but not all of it. (See Appendix C.2 for further discussion on the effect of renormalization.) Alternatively, as noted in Calura et al. (2012), the continuum-level estimation of the observational Lyman- α forest can significantly affect the shape of the flux PDF. As shown in Figure 8 of Calura et al. (2012), increasing the continuum level by 5% modifies the shape of the flux PDF to be comparable to the ones seen in the simulations. There are also some discrepancies with previous theoretical explorations of the predicted flux PDF that show a greater distribution of pixels with $F \sim 1$ (Bolton & Haehnelt, 2007; Becker et al., 2007). However, these simulations employ semi-analytic models similar to the FGPA to determine the optical depth as a function of gas temperature, and so do not capture the inhomogeneous ionization field or instantaneous ionization state of the gas. Thus, a combination of changing τ_{eff} of the simulations and the continuum-level of the observations can bring the simulations and observations into agreement.

5.5.3 One-dimensional flux power spectra

In addition to the statistics already discussed, the one-dimensional flux power spectrum can provide valuable information about underlying dark matter density distributions. To calculate the one-dimensional flux pdf, we first define a “flux overdensity” δ_F for each pixel:

$$\delta_F \equiv \frac{F}{\langle F \rangle} - 1, \quad (5.9)$$

where $\langle F \rangle$ is the average flux for all pixels in the volume (which is typically close to the average flux within a given sightline, due to the length of the sightlines). Having defined this quantity, a Fourier transform is applied to each sightline, so that we have $\delta_F(k)$. The one-dimensional power spectrum $P_{1D}(k)$ is the average power per k -mode: $P_{1D}(k) = \langle |\delta(k)|^2 \rangle$. In the following analysis, we look primarily at the dimensionless power spectrum:

$$\Delta_{1D}^2(k) = \frac{k}{\pi} P_{1D}(k). \quad (5.10)$$

Previous studies have shown that the one-dimensional power spectrum can be used to measure the three-dimensional power spectrum (Croft et al., 1998; McDonald et al., 2005; McDonald & Eisenstein, 2007), though here we explore only the one-dimensional power spectrum. As with the flux PDF, the amplitude of the one-dimensional power spectrum on large scales is largely similar between the different reionization scenarios

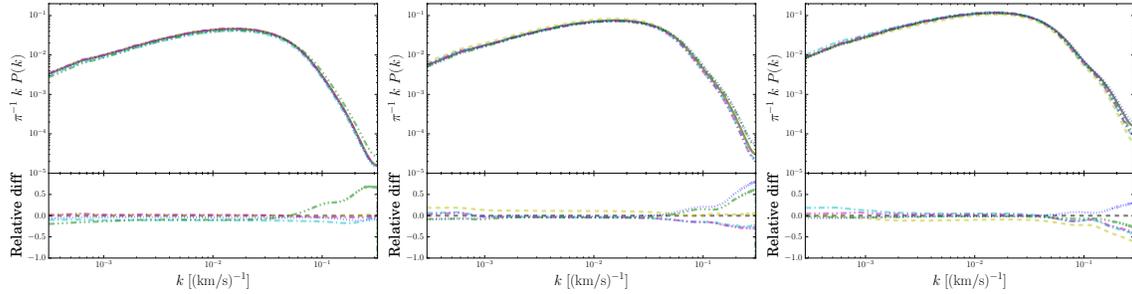


Figure 5.10: A plot of the one-dimensional H I Lyman- α forest at $z \sim 2.5, 3,$ and 3.5 , from left to right. As with the flux PDF in Figure 5.9, the effects of helium reionization are largely invisible in the hydrogen forest. Most of the differences between simulations are visible at small scales. Specifically, simulations in which the average temperature of the IGM is higher show less power on small scales. This is due to the thermal motion of the gas washing out some of the small-scale structure. The overall amplitude of the power spectrum tends to decrease with redshift, since the total number density of hydrogen is decreasing. See the text for additional discussion.

at the same redshift. However, there are significant differences on small scales ($k \gtrsim 0.1 \text{ (km/s)}^{-1}$). This is likely due to the differences in the thermal histories of the IGM. Notice that, in particular at $z \sim 2.5$, Simulation H6 shows a greater amplitude than many of the other simulations, and also has a cooler temperature (see Fig. 5.7). The cooler temperature is correlated with additional power at small scales, which is consistent with additional structure as a result of cooler gas.

Figure 5.10 shows the one-dimensional power spectrum of the Lyman- α forest for redshifts $z \sim 2.5, 3,$ and 3.5 . As can be seen in Figure 5.9, most of the differences between simulations are largely invisible in the H I Lyman- α forest. Most of the differences between simulations are visible at small scales. In general, the simulations that have a hotter average temperature of the IGM show less structure at small scales. This is due to the decrease in clumping that results from the increased thermal motion of the gas. On large scales, the differences between the simulations are typically less than 10%.

One important point to note is that the small-scale structure of the one-dimensional power spectrum is its dependence on the thermal history of the gas. The power spectrum is sensitive not only to the current temperature of the IGM, but also its past temperature, a phenomenon first pointed out in Gnedin & Hui (1998). The power on small scales is set by Jeans smoothing in the gas, which is caused by the propagation of pressure waves in the gas and hence depends on the sound speed in the gas. Because the sound speed depends on the temperature of the gas (for an ideal gas, $c \propto T^{1/2}$), the thermal history sets the maximum scale over which a pressure wave can travel in the IGM. Note that in middle panel of Figure 5.10 at $z \sim 3$, on small scales the simulations with the most power are Simulation H6 and Simulation H3. According

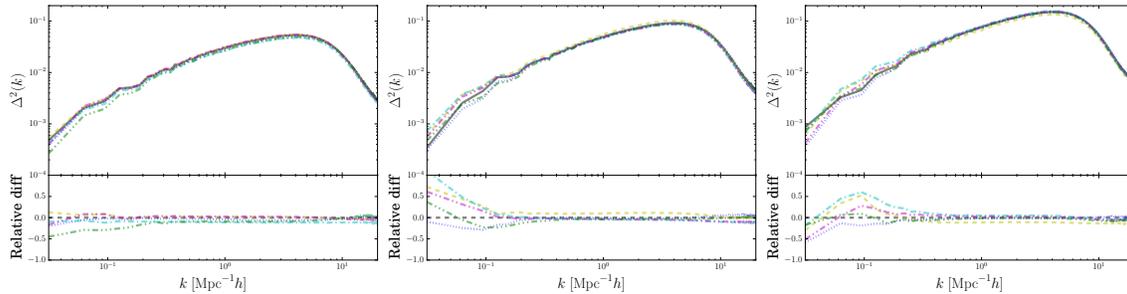


Figure 5.11: The three-dimensional H I Lyman- α forest flux power spectrum at $z \sim 2.5, 3,$ and 3.5 , from left to right. The only significant differences between the simulations emerge on large-scales. This result is consistent with previous studies that different thermal histories have on the three-dimensional power spectrum (McDonald, 2003). The trend on large scales may be related to the large-scale bias of sources, leading to different biases in the radiation field. See the discussion in Sec. 5.5.4 for further details.

to Figure 5.7, the temperature of the mean-density gas is similar between the two simulations. However, in the case of Simulation H6, the temperature is decreasing after having reached an earlier peak, whereas in Simulation H3, the temperature is increasing from a relatively cool phase after hydrogen reionization. Accordingly, there is additional power in the smallest scales for Simulation H3, which is consistent with the findings of Gnedin & Hui (1998).

5.5.4 Three-dimensional flux power spectra

We have also made predictions for the full three-dimensional flux power spectrum of the H I Lyman- α forest. To compute this quantity, we have generated the full number of sightlines in the volume of N_{grid}^2 , which provides the full three-dimensional information about the volume. Several previous studies (Croft et al., 1998) instead differentiated the one-dimensional power spectrum to extract the three-dimensional information. Our approach of using the full set of correlations present in the underlying density field, as well as yielding the power spectrum at finer resolution in k -space. The information contained in the three-dimensional flux power spectrum can contain information about the state of the gas of the IGM (Pichon et al., 2001; McDonald, 2003; Caucci et al., 2008; Cisewski et al., 2014; Ozbek et al., 2016), which would provide an exciting window into the IGM at high redshift. Additionally, several previous studies have started to measure the full three-dimensional power spectrum using quasar sightlines from SDSS (Slosar et al., 2011; Lee et al., 2014), which have provided important insight. In principle, like the one-dimensional flux power spectrum, the three-dimensional flux power spectrum can reveal important information about the thermal history of the IGM (Gnedin & Hui, 1998), as well as the large-scale

distribution of matter.

Figure 5.11 shows the three-dimensional H I Lyman- α forest flux power spectrum. The general shape of the power spectrum is similar to that of the one-dimensional version seen in Figure 5.10, though the drop in power at high- k is not as pronounced. More importantly, there are observable differences on large scales between the different reionization histories, which can differ by up to a factor of 2. Importantly, the gas power spectrum of all of the simulations is essentially identical on large scales, so all differences are due to the different ionization histories of the IGM rather than the underlying matter or gas distribution.

The differences in power at large scales is likely due to correlations present in the radiation field present in the IGM. As noted in McDonald (2003), differences in the thermal state of the IGM (either the temperature T_0 or the slope γ) only lead to differences at the $\sim 10\%$ level, which is consistent with the results seen in Figure 5.11. The differences on large scales are significantly larger than this, and further do not seem to be correlated with particular values of T_0 and γ . Indeed, when comparing with the values in Figure 5.6, the power on large scales does not seem to be correlated with either value, further demonstrating that the thermal history alone is not responsible for the differences on large scales.

Proper characterization of the full three-dimensional power spectrum is important for measurements of the BAO from the Lyman- α forest (Busca et al., 2013; Slosar et al., 2013). As can be seen in Figure 5.11, there are differences on large scales, in some cases as large as a factor of two between the different reionization scenarios. Thus, properly understanding the impact that the reionization of helium has on the three-dimensional power spectrum is important for systematic errors for the BAO measurement.

5.6 Conclusion

In this chapter, we have presented a new suite of simulations that couple N -body methods, hydrodynamics, and radiative transfer simultaneously in order to study helium reionization. Some of the most important observational implications that helium reionization leaves on the low-density gas of the IGM come from the dramatic increase in temperature from the photoheating of the gas. Using the results of the simulations, we summarize here several conclusions that we can make:

1. In addition to changing the ionization fraction of helium as a function of redshift $x_{\text{HeIII}}(z)$, helium reionization also leaves an important signature on the thermal history of the IGM. This finding is consistent with previous studies of helium reionization, which suggest using the temperature of the IGM to learn about helium reionization. We show that the peak in the temperature at mean density as a function of redshift $T(z)$ is a relatively robust signifier of helium reionization, occurring when the volume is 90-95% ionized by volume. The redshift

interval over which the temperature of the IGM increases can be used to determine the duration of reionization, though this measurement is observationally less straightforward.

2. Observations of synthetic H I Lyman- α sightlines show that many statistics concerning the forest are generally similar for different ionization states of helium. In particular, the one-dimensional power spectrum and the flux PDF are more sensitive to the effective optical depth τ_{eff} of the volume, rather than the IGM temperature or helium ionization. However, there are still noticeable differences for regions of high-transmission, which becomes clear when examining the flux PDF.
3. One exception to the previous point is the three-dimensional flux power spectrum, which shows differences on large scales by as much as a factor of 2. Previous studies have attempted to measure this quantity (Slosar et al., 2011; Lee et al., 2014), though the error bars are still significant.

In future studies, we plan to investigate the effect that anisotropic sources has on helium reionization. The effect was discussed briefly in McQuinn et al. (2009), though we plan to explore this aspect more thoroughly. In addition, we plan to look for observational signatures in the Lyman- β and Lyman- γ forests. Because these transitions saturate at much higher neutral hydrogen densities, they can give additional information about the thermal state of the IGM at higher densities. This type of comparison can provide an additional observational tool for understanding helium reionization, and provide another point of comparison with observations.

Chapter 6

The Helium Lyman- α Forest

6.1 Introduction

In the past several decades, there has been an interest in understanding the reionization of helium, using semi-analytic methods (Gleser et al., 2005; Furlanetto & Oh, 2008b,a, 2009; Dixon et al., 2014), numerical simulations (McQuinn et al., 2009, 2011; Compostella et al., 2013, 2014; Puchwein et al., 2015; Bolton et al., 2016), and observations (Jakobsen et al., 1994; Reimers et al., 1997; Zheng et al., 2008; Dixon & Furlanetto, 2009; Syphers & Shull, 2014; Worseck et al., 2014). Helium reionization is thought to be driven by highly energetic photons emitted by quasars. Due to photoheating of gas in the intergalactic medium (IGM) from these high-energy photons, helium reionization leaves an important signature on the thermal state of the IGM. However, such temperature measurements are difficult to make, and have large systematic or statistical uncertainties (Schaye et al., 1999; McDonald et al., 2001; Becker et al., 2011a; Boera et al., 2014). Further, these methods for determining the temperature rely on correctly calibrating the state of the hydrogen Lyman- α forest with the gas temperature, which is fraught with difficulty.

A more appealing approach is to measure the ionization state of helium more directly. Just as the Lyman- α transition from neutral hydrogen (H I) appears as absorption spectra of radiation from distant quasars, so too does it appear for singly ionized helium (He II). This feature appears at 304 Å in the rest-frame of the absorbing gas, a shift of a factor of 4 in frequency space compared to the hydrogen transition due to the additional proton in the helium nucleus. As with the H I Lyman- α forest, the very high transition strength means a very small amount of singly ionized helium can lead to total absorption of the incoming radiation. Typically, neutral fractions of $f_{\text{HeII}} \gtrsim 10^{-3}$ can produce a Gunn-Peterson trough (Gunn & Peterson, 1965), making detection of the early stages of helium reionization difficult. Despite this difficulty, measuring the ionization status of helium from the He II Lyman- α forest provides a more direct probe than using temperature measurements or the H I Lyman- α forest.

Part of the difficulty in observing the He II Lyman- α forest lies in contamination

of high-density systems at lower redshift. Lyman-limit systems (LLSs) and damped Lyman- α systems (DLAs) which are at intermediate redshift (say at z_{LLS}) between the comparatively high-redshift IGM gas we are interested in observing (say at z_{IGM}) and observers on Earth can absorb much of the radiation above the ionization potential of hydrogen at 912 Å. Thus, if $912(1 + z_{\text{LLS}}) \gtrsim 304(1 + z_{\text{IGM}})$, then the lower-redshift LLS or DLA will obfuscate the He II Lyman- α forest of interest. Due to the inferred abundance of LLSs and DLAs at low redshift, only a small number of quasar sightlines are suitable for measuring the He II Lyman- α forest (Møller & Jakobsen, 1990; Zheng et al., 2005). Indeed, despite having more than 150,000 quasar sightlines from BOSS alone (Dawson et al., 2013), to date there have been only about 50 sightlines for which the He II Lyman- α forest has been measured (Syphers et al., 2009b,a, 2012). These measurements have provided much insight to the general picture of helium reionization: at redshifts $z > 3$, a Gunn-Peterson trough has been detected (Jakobsen et al., 1994; Zheng et al., 2008; Syphers & Shull, 2014); below this redshift, helium reionization becomes patchy, showing extended regions of absorption and transmission corresponding to the ionization level of the gas (Reimers et al., 1997); finally, by redshift $z \sim 2.7$, helium appears to be totally reionized (Dixon & Furlanetto, 2009; Worseck et al., 2011). However, information beyond this general picture is difficult to glean from the current limited set of He II spectra. To this end, measurements providing additional information about helium reionization is an important application of current and ongoing research.

In Chapter 4, we provided a method by which simulation volumes can be populated with quasars in order to reproduce the quasar luminosity function (QLF) at various redshift epochs (Masters et al., 2012; Ross et al., 2013; McGreer et al., 2013) as well as quasar clustering (White et al., 2012). In Chapter 5, we presented a new suite of large-scale simulations with the purpose of exploring helium reionization. These simulations include N -body, hydrodynamics, and radiative transfer solved simultaneously, which allows us to capture the evolution of the IGM with newfound accuracy. Based on the snapshots of these simulations, we are able to generate synthetic Lyman- α sightlines for H I and He II. In this paper, we present results specifically about the He II spectra, and discuss ways to learn about the timing of helium reionization.

We organize the rest of this chapter as follows. In Sec. 6.2, we briefly discuss our suite of simulations. In Sec. 6.3, we discuss the He II Lyman- α forest, and various measurements that can be made using the spectra. In Sec. 6.4, we discuss prospects for detecting helium reionization properties given the current measurements. In Sec. 6.5, we summarize our findings. Throughout this work, we assume a Λ CDM cosmology with $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.045$, $h = 0.7$, $\sigma_8 = 0.8$, and $Y_{\text{He}} = 0.24$. These values are consistent with the *WMAP*-9 year results (Hinshaw et al., 2013).

6.2 Radiation-Hydrodynamic Simulations

In Chapter 5, we present a new suite of hydrodynamic simulations with radiative transfer, conducted with the goal of studying helium reionization. Here, we summarize the properties of the simulations that are relevant to the results presented here. The simulations were run using the RadHydro code with volumes of 200 comoving $h^{-1}\text{Mpc}$ on a side. The large-scale structure is solved using a particle mesh (PM) scheme with 2048^3 dark matter particles. The hydrodynamics is solved using a fixed-grid Eulerian scheme with 2048^3 grid cells. Radiative transfer is solved using ray tracing to propagate photons from quasars and a uniform UV background for radiation from galaxies. The grid for radiative transfer contains 512^3 resolution elements, which is coarser than the resolution of the gas grid by a factor of 4. The simulation code has already been used to study hydrogen reionization (Trac & Cen, 2007; Trac et al., 2008; Battaglia et al., 2013b).

The simulations contain two features in particular that bear mentioning. First, the simulations include a patchy model for hydrogen reionization developed in Battaglia et al. (2013b). The midpoint of reionization has been set such that $\bar{z}_{\text{re}} = 8$, but in general, regions of high-density undergo reionization before regions of low density. By incorporating an “inside-out” reionization scenario, we ensure that the thermal state of the IGM before helium reionization accurately reflects the impact of hydrogen reionization. Second, the contribution of galaxies to the UV background Γ_{gal} is modified on-the-fly in order to reproduce the observed effective optical depth τ_{eff} of the H I Lyman- α forest, defined as:

$$\langle F \rangle_{\text{HI}} = e^{-\tau_{\text{eff,HI}}}, \quad (6.1)$$

where $\langle F \rangle_{\text{HI}}$ is the average flux of the H I Lyman- α forest of the volume, with an analogous definition for He II. Values of $F \sim 0$ represent total absorption, and values of $F \sim 1$ represent total transmission. Specifically, we match τ_{eff} as determined by BOSS (Lee et al., 2015). Modifying Γ_{gal} while the simulations are running means we do not need to renormalize the Lyman- α forest in post-processing, as previous studies of the Lyman- α forest have done (*e.g.*, Bolton et al. 2009a). This feature allows us to more easily compare the results between simulations and observations.

As explained in detail in Chapter 4, the simulation volumes are populated with quasars such that the observed QLF is matched between redshifts $2 \leq z \leq 6$ (Masters et al., 2012; Ross et al., 2013; McGreer et al., 2013), as well as the clustering measurements at $z \sim 2.4$ (White et al., 2012). For individual quasar objects, we use the SED from Lusso et al. (2015), which has a spectral index of $\alpha = 1.7$ ($f_{\nu} \propto \nu^{-\alpha}$) for $\lambda \leq 912 \text{ \AA}$, and a spectral index of $\alpha = 0.61$ for $\lambda > 912 \text{ \AA}$.

In Chapter 5, we present a series of 6 simulations, with different quasar properties. We will now briefly summarize each of these simulations. Simulation H1 is the fiducial reionization model, which uses the QLF combining the various measurements at different epochs and the SED of Lusso et al. (2015). Simulation H2 increases the

amplitude of the QLF by a factor of 2, leading to an earlier reionization scenario. Simulation H3 decreases the amplitude of the QLF by a factor of 2, leading to a late reionization time. Simulation H4 increases the normalization of the SED by a factor of 2, so that a given quasar with a given magnitude M will have a luminosity at 912 \AA L_{912} that is two times greater as that provided by Lusso et al. (2015). Simulation H5 uses a slightly different method for combining the QLF from different epochs than Simulation H1, but uses the same SED. Simulation H6 uses the same quasar population as simulation H5, but instead of renormalizing Γ_{gal} to match observations, uses Γ_{gal} as specified by Haardt & Madau (2012). By extension, the values of τ_{eff} for the H I Lyman- α forest do not match at all redshifts for this simulation. Finally, Simulation H7 does not have explicit quasar sources, but instead uses a uniform UV background with the photoionization and photoheating rates as specified by Haardt & Madau (2012). As with Simulations H1-H5, these rates are renormalized in order to match the observed H I τ_{eff} . See Paper II for further details about each of the simulations.

6.3 He II Lyman- α Forest Measurements

At each time step in the simulation, we generate synthetic Lyman- α sightlines on-the-fly for H I and He II. The measurements of $\tau_{\text{eff,HI}}$ for the H I sightlines allow for modifying Γ_{gal} to ensure that the value is matched at all times in the simulation. For the He II sightlines, we are able to follow the evolution of $\tau_{\text{eff,HeII}}$ in order to understand the evolution of the volume. We will now turn to specific observables related to the He II Lyman- α forest.

6.3.1 Effective Optical Depth

The effective optical depth τ_{eff} , as noted in Eqn. (6.1), is defined in terms of the average flux in the volume. As with the H I Lyman- α transition, the strength of the He II transition ensures that only a very small amount of singly-ionized helium is necessary to completely absorb incoming radiation. As a result, measurement of $\tau_{\text{eff,HeII}}$ is most sensitive to the end of reionization. Further, due to the very large comoving size of He II regions (typically tens of Mpc in diameter), there is a large variation between different sightlines in the simulation volume, or even along a given sightline. This variation is especially prevalent while helium reionization is proceeding. In other words, due to the large coherence of the doubly ionized regions, the observed optical depth can vary greatly from sightline to sightline, and so one expects there to be a large variance in the measurements. This variation is in addition to any inherent variance in $\tau_{\text{eff,HeII}}$, primarily due to density fluctuations.

Figure 6.1 shows $\tau_{\text{eff,HeII}}$ as a function of redshift averaged over the whole simulation volume. The Figure also includes observational data from Worseck et al. (2014). These quasar spectra were taken using the cosmic origins spectrograph (COS) on

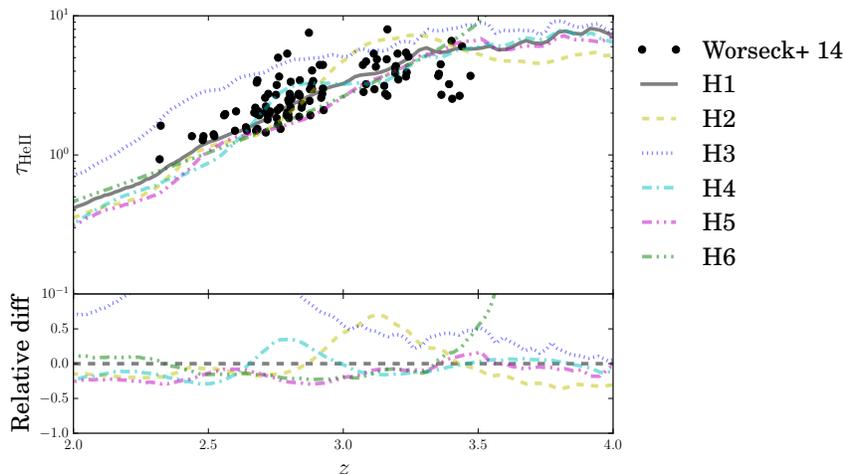


Figure 6.1: The effective optical depth of singly ionized helium $\tau_{\text{eff,HeII}}$ as a function of redshift for the different simulations. The black dots represent the observational data from Worseck et al. (2014), which are taken from HST/COS data. The top panel shows the results for the different simulations presented in Chapter 5, and the bottom panel shows the relative difference to the fiducial simulation. As can be seen, there is a large degree of scatter in the measurements. By extension, none of the simulations is clearly disfavored. See Sec. 6.3.1 for additional discussion.

the Hubble Space Telescope (HST). These spectra measure $\tau_{\text{eff,HeII}}$ for segments of about 10 proper Mpc. The large sightline-to-sightline variation is evident in the observational data, which show very different values of $\tau_{\text{eff,HeII}}$ at the same redshift. The results from most of the simulations are largely consistent with the data at the redshifts for which data is available ($2.5 \lesssim z \lesssim 3.5$). The main exception to this is Simulation H3, which completes reionization at a significantly later time than the other simulations. Quantitatively, the redshift when the volume reaches an ionization fraction $x_{\text{HeIII}} \equiv n_{\text{HeIII}}/n_{\text{He}}$ of 99% is $z_{99} \sim 2.23$, compared to the timing of reionization in the fiducial scenario of $z_{99} \sim 2.69$. Given that this simulation completes reionization significantly later than the other ones, it is not that surprising that the value of $\tau_{\text{eff,HeII}}$ in Simulation H3 is significantly higher than that of the other ones.

6.3.2 Flux PDF

As discussed in Chapter 5, another tool for analyzing the ionization state of the medium is the flux PDF. This measurement captures the relative number of pixels with high transmission. As with $\tau_{\text{eff,HeII}}$, this statistic is most sensitive to the tail-end of reionization. Due to the low number of He II pixels with high transmission ($F \gtrsim 0.5$) before the end of reionization ($x_{\text{HeIII}} \gtrsim 0.99$), the flux PDF cannot provide detailed information while reionization is underway. Nevertheless, it can still provide

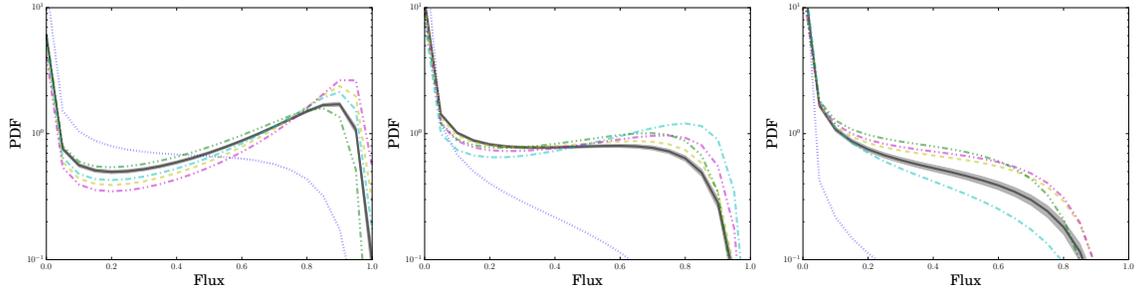


Figure 6.2: The flux PDF of the He II Lyman- α forest at redshifts $z \sim 2.3$, 2.5, and 2.7, from left to right. Note that the flux PDF is very sensitive to the tail end of reionization: most of the simulations have an ionization fraction $x_{\text{HeIII}} \gtrsim 0.99$ at $z \sim 2.7$, and yet have a comparatively low number of pixels with high transmission ($F \gtrsim 0.5$). Nevertheless, the ionization fraction can be determined from the overall shape of the PDF: note that the flux PDF of Simulation H3, which has a very late reionization time, at $z \sim 2.3$ looks comparable to the other simulations at earlier times $z \sim 2.7$. The shaded regions show the error in the measurement of Simulation H1 computing using bootstrap resampling. For more details, see the discussion in Sec. 6.3.2.

valuable information about the timing of reionization.

Figure 6.2 shows the He II flux PDF as measured by the different simulations. The panels, from left to right, show the volume at redshifts $z \sim 2.3$, 2.5, and 2.7. Note that at redshift $z \sim 2.7$, most of the simulations are 99% reionized. Despite this fact, there are comparatively few pixels with high transmission: for the fiducial case of Simulation H1, more than 90% of the pixels have flux of $F \leq 0.5$. This relatively strong absorption is related to the strength of the Lyman- α transition, where only a small amount of He II is necessary to absorb most of the incoming radiation. Note, though, that measurement of the flux PDF can still be an important marker of the timing of reionization. As noted above, Simulation H3 completes reionization at $z_{99} \sim 2.23$, which is evident in the very different shape of the flux PDF. The flux PDF of H3 at $z \sim 2.3$ is comparable to that of, *e.g.*, Simulation H1 at $z \sim 2.7$. Thus, by measuring the redshift when the central portion of the flux PDF is relatively flat (*e.g.*, when $\text{PDF}(F = 0.25) \approx \text{PDF}(F = 0.75)$), one can determine the timing of 99% ionization.

The shaded error regions in Figure 6.2 show 1σ uncertainties, and are calculated using bootstrap resampling of 50 sightlines and computing the variance within each flux bin. This number was chosen to coincide roughly with the current number of He II sightlines (as noted in Sec. 6.1). Note that these error regions are generally small, showing that only a few sightlines are necessary to determine the shape of the flux PDF. As with the H I flux PDF (and further discussed in Appendix C.2, the continuum level of the measured spectra can have a dramatic effect on the shape of

the flux PDF. Accordingly, uncertainty in this level can lead to systematic shifts in the calculated flux PDF. Nevertheless, Figure 6.2 shows that given low systematic uncertainties, the flux PDF of the He II Lyman- α forest can be a powerful tool for determining the ionization state of helium in the IGM.

6.3.3 One-d flux power spectra

In addition to the flux PDF, the one-dimensional power spectrum of the He II Lyman- α forest can be used to learn important information about the ionization state of the IGM. The overall amplitude of the power spectrum as well as the shape as a function of k will change as the ionization state and the size of ionized regions change. As with the one-dimensional power spectrum for H I, the amplitude on large scales is related to the overall ionization state, with higher amplitudes corresponding to increased ionization.

Figure 6.3 shows the one-dimensional power spectrum of the He II Lyman- α forest. The primary differences between the different simulations are in the amplitude of the power spectra. At a given redshift, the amplitude of the power spectrum is directly related to the value of τ_{eff} . Note that Simulation H3 has the largest value of τ_{eff} at a given redshift (*c.f.*, Figure 6.1), and also has the largest amplitude in Figure 6.3. This can be understood in terms of the amplitude of fluctuations in the flux field: when the IGM has a relatively low value of τ_{eff} , then all points in the volume have a similarly (high) value of flux. Conversely, when τ_{eff} is higher, there are more variations between regions of high flux and low flux. Thus, the overall amplitude of the power spectrum is higher.

At $z \sim 2.7$, the difference in the power spectrum amplitude at all scales is an order of magnitude larger than that of the other simulations. Such a dramatic difference should be detectable, and would allow for a straightforward determination of the ionization state of the IGM. Most importantly, the differences in the amplitudes of the IGM as a function of redshift are clear and pronounced, even for reionization histories that are not fully reionized. Hence, the one-dimensional power spectrum can be a window into helium reionization at times prior to 99% ionization.

As with Figure 6.2, the shaded regions show 1σ uncertainty in the measurements using 50 sightlines and bootstrap resampling. At the earliest redshift ($z \sim 3.2$), there is a relatively large uncertainty. The uncertainty is large enough that at most scales, several of the reionization histories are expected to lie within 1σ of each other. However, simulations with vastly different values of τ_{eff} (as in Simulations H2 and H3 in Figure 6.1) still show a strong difference in amplitude of the power spectrum. Thus, the one-dimensional power spectrum can serve as another measurement of the overall opacity of the volume.

However, at late times, the uncertainty of the power spectra decreases noticeably. As a result, in principle it becomes easier to distinguish the histories. On the other hand, there is significant overlap in several of the histories, which is due to having

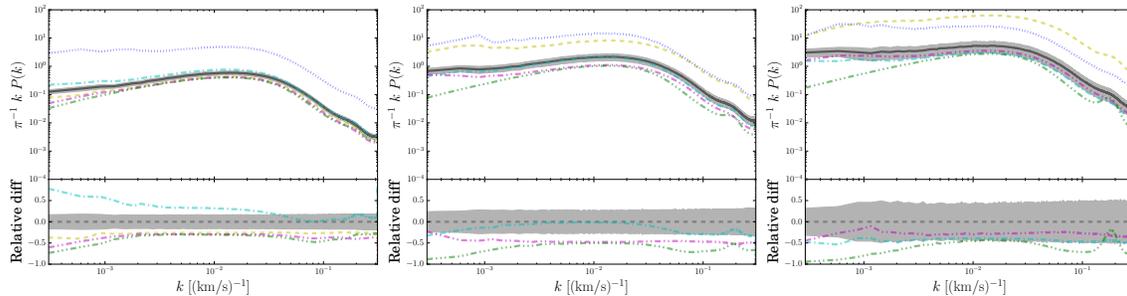


Figure 6.3: The one-dimensional flux power spectrum of the He II Lyman- α forest. From left to right, the plots are at redshift $z \sim 2.7$, $z \sim 3.0$ and $z \sim 3.2$. There is a marked difference in the overall amplitudes of the power spectrum. These differences are largely due to the difference in τ_{HeII} at a given redshift. Compare to Fig. 6.1, and note that the amplitude of the power spectrum largely tracks the values of τ_{eff} . See the text for additional discussion.

comparable values of τ_{eff} . Some of the largest differences that remain are at large scales. As with the H I Lyman- α forest (and discussed in Sec. 5.5.4), these differences might be attributable to the large-scale differences in the radiation field. Because the radiation field is highly non-uniform for He II ionization level (as opposed to the H I forest that had a uniform background component from galactic radiation), the large-scale power may reflect the degree of bias in the sources. Note in particular that Simulation H6, which has only a uniform UV background and no explicit sources, has consistently the lowest large-scale power, despite having one of the earliest reionization times. At all redshifts considered, this simulation shows a lack of power compared to the simulations with explicit sources. Thus, the large-scale power may be a way to learn about the bias of sources of helium reionization.

6.4 Discussion

One very pertinent question with these measurements is the degree to which the reionization history can be determined with a limited number of observations. As discussed in Sec. 6.1, to date there have been only about 50 observations of the He II Lyman- α forest (Syphers et al., 2012). We have shown in Secs. 6.3.2 and 6.3.3 that the flux PDF and one-dimensional power spectrum provide significant information about the ionization state of helium. However, it is reasonable to wonder to what extent current observations are able to determine the ionization state of the IGM.

To this end, we have used bootstrap resampling using 50 sightlines to estimate the standard deviation for our different scenarios. Figures 6.2 and 6.3 show the 1σ dispersion as measured for 50 sightlines. As noted in the earlier discussion, the ionization state of helium may be readily detectable in the flux PDF measurement. Even accounting for uncertainty in the continuum level of the forest, the shape of

the flux PDF varies strongly as a function of ionization fraction. This variation in shape is significantly larger than the inherent variation of the flux PDF, and so even with comparatively few sightlines, a meaningful determination of the ionization level of helium may be possible given the current data.

6.5 Conclusion

To date, the He II Lyman- α forest has largely only been used to determine the value of $\tau_{\text{eff,HeII}}$. As can be seen from Figure 6.1, there is a very large dispersion in this measurement, owing to the large sightline-to-sightline variations. Thus, determining the reionization history from this quantity alone is very difficult, and leads to large uncertainties in the determination of the redshift of reionization. Additionally, as discussed earlier, this measurement is largely sensitive to the tail-end of reionization, and does not yield much information about the intermediate stages of the reionization process. Accordingly, new applications of the He II Lyman- α forest would be beneficial for learning more about the timing and duration of reionization.

To this end, we have presented the flux PDF and the one-dimensional power spectrum as ways to break the degeneracy present in τ_{eff} . These differences are generally quite large between different simulations, in some cases being larger than an order of magnitude. Future surveys will hopefully be able to take advantage of these pronounced differences, and begin to measure the timing and duration of helium reionization.

Chapter 7

Conclusions

Over the past two decades, our understanding of reionization has grown by leaps and bounds. Pure analytic theory provided the general framework of the reionization picture as it is understood today. Cosmological simulations have grown in their scale and complexity to provide predictions to unprecedented accuracy. And observational evidence that promises to manifest just over the horizon should dovetail with the predictions made thus far, and provide concrete evidence for many of the reionization features described by theory and simulations. Meanwhile, there are several key extensions of the work presented in this thesis that can be extended to make further discoveries.

7.1 Future Observations of Reionization

7.1.1 Hydrogen Reionization

Observations of hydrogen reionization have been ongoing for many years, and are reaching a point where the a signal is likely to be detected in the next ten years. Most of the radio telescopes in the current generation of reionization-era experiments are interferometers, which consist of a series of antennas whose signals are cross-correlated. Some of the currently operating telescopes are the Murchison Widefield Array (MWA, Bowman et al. 2005), the Low Frequency Array (LOFAR, Harker et al. 2010), and the Precision Array for Probing the Epoch of Reionization (PAPER, Parsons et al. 2010). To date, the results from these experiments have provided only upper-limits to the key observables related to hydrogen reionization. For instance, the 64-element configuration of the PAPER array has placed upper-limits on the 21 cm 3D power spectrum (Ali et al., 2015) and the thermal state of the IGM (Pober et al., 2015) at specific redshifts relevant to reionization. By assuming some underlying model for the evolution of the 21 cm signal at different epochs, some constraints can be placed on the redshift dependence of the power spectrum (Jacobs et al., 2015). Despite these recent advances, there has yet to be a conclusive detection of the 21 cm

power spectrum.

Next-generation experiments promise to provide a positive detection of the power spectrum. The Hydrogen Epoch of Reionization Array (HERA, Pober et al. 2014) will likely provide such a detection within the next several years. The sensitivity of the instrument will provide sufficient signal-to-noise so as to overcome the strong foreground signal. Additionally, the proposed array configurations allow for potential detection across a variety of redshifts, so the evolution of the 21 cm power spectrum should be detected as well. This evolution will be invaluable for providing the best picture of reionization observed to date, and will allow for detailed testing of various reionization models discussed in the literature. In particular, the ionization fraction and thermal state of the IGM inferred from these measurements should settle the question on early reionization caused by X-ray binaries (Mirabel et al., 2011) or primordial quasars (Madau & Haardt, 2015).

The measurements from HERA also promise to provide information about astrophysics, as well as improve the constraints on cosmological parameters. One broader application of the 21 cm power spectrum measurement is to learn about the sources of ionizing radiation. Specifically, several parameters of semi-analytic models, such as 21CMFAST (Mesinger et al., 2011), can be relatively well constrained given current forecasts for measurements from HERA (Liu & Parsons, 2016). As to cosmological parameters, measurements of the CMB made by the *WMAP* or Planck satellites typically employ a seven-parameter Λ CDM model. One of these parameters is τ_{reion} , which is the optical depth of photons to the CMB. This parameter is degenerate with several other ones, most prominently A_s , the primordial amplitude of the power spectrum. Note, though, that this degeneracy exists only when analyzing the TT spectrum, and this degeneracy is broken somewhat by the low- ℓ polarization (EE and TE) data. Accurately determining τ_{reion} from HERA measurements promises to further reduce the uncertainty of the remaining cosmological parameters by supplying a known value (or very narrow range of “prior” values) to the MCMC method (Liu et al., 2016).

Even further in the future, the Square Kilometer Array (SKA, Koopmans et al. 2015) will provide exquisite measurements of hydrogen reionization, for a large range of redshifts and k -modes. The increased sensitivity of the array will allow for extracting the full tomographic information of the 21 cm signal. These series of slices will allow for three-dimensional reconstruction of the density field as traced by 21 cm as a function of frequency (and hence redshift). Additionally, the SKA should also allow for the direct imaging of the 21 cm sky, which represents the ultimate aim of radio astronomy efforts. The SKA is also likely the practical terrestrial limit of sensitivity of radio interferometers. The detection power of future interferometers could be helped by constructing an interferometer on the (radio) dark side of the moon, but is far outside the scope of this work.

As a complementary measurement to the power spectrum of the 21 cm field, the global signal provides important information about the evolution history during the

Dark Ages. Prevailing theories of reionization predict a dip in the global signal as initial the gas couples to Lyman- α radiation from the first stars, and then becomes heated by photoionization. This transition is expected to appear at redshifts significantly higher than hydrogen reionization, for $30 \gtrsim z \gtrsim 15$. Several experiments have been proposed to observe this portion of the reionization history, to determine the temperature of the IGM before reionization. Some of the experiments include the Experiment to Detect the Global Epoch-of-Reionization Signature (EDGES, Bowman et al. 2008; Bowman & Rogers 2010), Large-aperture Experiment to Detect the Dark Ages (LEDA, Greenhill & Bernardi 2012), SCI-HI (Voytek et al., 2014), and the satellite-based Dark Ages Radio Explorer (DARE, Mirocha et al. 2015). By detecting the 21 cm signal during the pre-reionization era, these experiments will provide important information about the initial conditions that led to reionization, such as when star formation first happened. These experimental efforts are exciting probes of the very young hydrogen sky, and will help bridge the gap between the CMB era and reionization.

7.1.2 Helium Reionization

Observational efforts to detect helium reionization have different constraints from hydrogen reionization, but in some sense can be just as difficult to detect. As I discussed in Chapters 5 and 6, the clearest and most direct probe of helium reionization is detecting the He II Lyman- α forest. By measuring its associated statistics, such as the flux PDF or one-dimensional power spectrum, the ionization state of the IGM can be determined. By tracking how these quantities change with redshift, it will be possible to learn about the timing and duration of reionization. These details will also indirectly probe the features of quasars, in new and interesting ways.

One of the largest obstacles to measuring the He II Lyman- α forest is dense systems of hydrogen gas between the background quasar and observers on Earth. Due to the additional proton in the helium nucleus, the relevant wavelength of the Lyman- α transition is increased in energy by a factor of four compared to the H I Lyman- α transition (since to first-order the energy difference $\Delta E \sim Z^2$, where Z is the number of protons). Accordingly, the wavelength of the He II Lyman- α transition is 304 Å, which is much higher in frequency space than the H I Lyman- α transition at 1216 Å. Accordingly, a dense Lyman-limit system (LLS) or damped Lyman- α system (DLA) at an intervening lower redshift has the possibility to absorb these higher-energy photons through photoionization. Any photons with a wavelength of $\lambda \leq 912$ Å in the rest frame of the LLS can be absorbed by the neutral gas, thus removing it from the quasar spectrum. For instance, suppose that a background quasar emits radiation at a rest wavelength of $z_q = 3.2$, and observers on Earth are interested in detecting the He II Lyman- α forest in the vicinity of the quasar. The absorption in the quasar spectrum from He II appears at $\lambda = 304(1 + z_q) \approx 1277$ Å in the rest frame on Earth. Let us assume that as the quasar radiation propagates toward Earth, it encounters a

comparatively low-redshift LLS at $z_{\text{LLS}} = 0.4$. This dense cloud of neutral hydrogen will absorb any radiation where $\lambda \leq 912(1 + z_{\text{LLS}}) \approx 1277 \text{ \AA}$! Thus, the He II forest will be spoiled by the intervening LLS. Thus, to observe the He II spectra, the distant quasar spectra cannot encounter a low-redshift LLS or DLA, because the spectra will suffer irreparable contamination.

This issue of the He II spectra being clouded by low-redshift hydrogen was explored quantitatively by Møller & Jakobsen (1990). The authors concluded that one can expect the majority of quasar spectra to intersect at least one LLS, given the assumed distribution as a function of redshift. Optimistically, one would expect $\sim 1\%$ of quasar sightlines showing less than 50% absorption. In practice, the actual number of sightlines that are clean enough to be used for measuring the He II Lyman- α forest is less than 0.1%: BOSS has measured spectra for 150,000 objects (Dawson et al., 2013), but to date only about 50 He II spectra have been measured (Syphers et al., 2009b,a). Future experiments may be able to find additional candidates for measuring the helium forest; in the meantime, additional efforts must be devised.

In lieu of making measurements of the entire IGM, it can be fruitful to instead focus on individual quasars. By learning more about the engines that drive helium reionization, better models can be used when performing calculations or simulations. One of the most uncertain properties of quasars (as it relates to helium reionization) is the lifetime. There is a fundamental challenge associated with determining quasar lifetimes, in that changes in quasar luminosity are not directly observable. (Perhaps measurements by future astronomers made in 10 million years will be compared to those of the SDSS observations...) Therefore, indirect methods of probing quasar lifetimes must be devised.

One potential method for investigating the lifetime of quasars is to examine the spectra of quasars that have little angular separation on the sky. Two quasars that have similar positions on the sky (though perhaps wildly different redshifts) will have tell-tale signs in their spectra. Specifically, the spectra of the more distant quasar at redshift z_1 will have less absorption in the vicinity of the closer quasar at redshift z_2 , relative to the value inferred by $\tau_{\text{eff,HI}}(z_2)$. The decrease absorption is due to the increased local ionization rate induced by the quasar at z_2 . By performing a statistical measurement of these quasar proximity zones (and their typical size), the average lifetime of quasars can be determined (Lidz et al., 2007; Khrykin et al., 2016). These types of calculations can lead to better models for helium reionization calculations.

7.2 Future Research Directions

There are several interesting extensions that can be made to the simulations presented here. The current understanding of hydrogen and helium reionization would stand to benefit significantly by further refinement of the numerical simulations performed here. I will now discuss several possible avenues for further investigation, which could provide further understanding of reionization. I will be focusing primarily to the

extensions of the work on helium reionization simulations, though see Sec. 3.7 for possible extensions to the work on hydrogen reionization presented here.

One extension to the helium reionization simulations presented here is diversifying the quasar population. For the work in Chapters 5 and 6, all of the quasars had the same spectral index, as well as the same peak luminosity as a function of halo mass. Additionally, all of the sources were populated in halos using the methods outlined in Chapter 4 using abundance matching. Further, all of the sources emitted radiation isotropically, with no beaming or directional emission.

Adding variety to the quasar population would generate an interesting comparison point with the models discussed in this work. This additional study could help show to what extent variation within a quasar population can affect the major properties of reionization. Specifically, it would reveal to what extent the hydrogen reionization timing and duration are fixed by the QLF and average spectral slope alone, or whether the population variance can play a significant role in the observables. More concretely, the temperature of the IGM following reionization depends on the hardness of the emitted quasar spectra; because many of the observables in the H I Lyman- α forest are strongly dependent on the temperature of the IGM, the spectral index variation can have a dramatic effect. The quasar population could be made even more realistic with the inclusion of different spectral indices for radio-loud and radio-quiet quasars, with the number of each type represented accordingly. A potential limitation to this diversification is the uncertainty of the high-redshift quasar population, so further observations with regard to these quantities would be tremendously useful.

As mentioned in Chapter 4, abundance matching is not the only method by which dark matter halos can be populated by quasars. Other methods, such as using HOD modeling (Richardson et al., 2012) or using the dynamical quantities of the halos to inform quasar properties (Cen & Safarzadeh, 2015). An alternative mass-to-light ratio as determined by these other methods could potentially change the conclusions about helium reionization. However, the impact of these differences would likely be minimal, due to the fact that the timing and duration of helium reionization are largely fixed by the quasar luminosity function. Given that these alternative methods would presumably match the QLF observations, the overall conclusions presented here should remain intact. Nevertheless it would serve as an interesting check on the robustness of the conclusions to change the quasar modeling.

Another potential change to the quasar population could be the introduction of anisotropic emission, or “beaming” to quasar radiation. If the quasars have some finite opening angle, then there could be interesting effects on the topology of helium reionization. Recent measurements have hinted at this possibility (Borisova et al., 2015), so implementing simulations that include this feature could provide very interesting results. It is quite likely that the one-point statistics (such as the average neutral fraction as a function of redshift, and $\tau_{\text{eff,HeII}}$) would be the same as in the case of isotropic emission, given the same total photon output over the lifetime of a given quasar. However, some two-point statistics, such as the one-dimensional and

three-dimensional power spectra of the He II Lyman- α forest, could change. If the differences prove to be significant, then measuring these quantities could provide an alternative method by which one can measure helium reionization properties.

The thermal history of the IGM can be further constrained by looking at higher-order Lyman series transitions in the H I forest. Lyman- β and Lyman- γ (a transition of $n = 1 \rightarrow n = 3$ and $n = 1 \rightarrow n = 4$, respectively) will also produce a forest in quasar spectra. As with the Lyman- α forest, the amount of radiation absorbed depends on the density of neutral hydrogen present locally in the IGM. However, because these transitions are not as strong as the Lyman- α one, the absorption signal saturates at much higher densities of hydrogen. Thus, instead of saturating for overdensities near mean density, the absorption would saturate for greater overdensities (Iršič & Viel, 2014). Accordingly, it would allow for probing the temperature at these higher overdensities, providing additional information about the average temperature-density relation (Boera et al., 2014). Using simulations to observe these potential differences would allow for better understanding of some of the systematic effects that may be present from such measurements, as well as providing guidance for calibrating future measurements.

An additional extension of this work as applied to the thermal history of the IGM is to provide better calibration of temperature measurements of the IGM from the H I Lyman- α forest. In particular, the curvature statistic κ (Becker et al., 2011a) has been used as a proxy for the temperature of the IGM near mean density. One of the systematic uncertainties of this measurement is the calibration: the relation between the curvature κ and the gas temperature is established by the use of hydrodynamic simulations using radiative transfer. However, the radiation field employed in the calibration assumes is uniform, and does not include any spatial inhomogeneities. Consequently, the spatial variations in the temperature field are not accurately included in the simulations. As shown in Chapter 5, helium reionization is very anisotropic, and the heating of the IGM is not well-modeled by using a uniform field. Further, the calibration process requires applying a power-law relation of temperature-density to extrapolate from moderate overdensities (where the curvature statistic is most sensitive) to mean density. As discussed in Sec. 5.4.1, the IGM is not well-parameterized as a single power law. Thus, the simulations presented here can be used to better apply calibrate measurements of the thermal state of the IGM, and in particular the curvature statistic.

The future of hydrogen and helium reionization represent very exciting fronts in astrophysics. The combination of theory, computation, and experiment have led to exciting new discoveries about these formative periods of the Universe's history. In the coming years, further investigation promises to provide the first concrete evidence of several of the most important features of hydrogen and helium reionization. The future of reionization is incredibly bright, not only at the prospect of learning about the first and brightest radiation sources in the Universe, but also because of learning about the fundamental understanding about the Universe.

Appendices

Appendix A

Additional Material for the Light Cone Effect

A.1 Exclusion of $k_{\perp} = 0$ Modes

Modes where $k_{\perp} = 0$ correspond to the total flux at a particular frequency defined by k_{\parallel} . For radio interferometers, this mode is inaccessible, since interferometers only measure fluctuations relative to a background level. Alternatively, to probe modes where $k_{\perp} = 0$, the antennas would have to have no separation between them, which is not possible. These modes would not be detectable in most experiments proposing to measure the 21 cm brightness temperature (Datta et al., 2012).

In order to determine how the exclusion of the $k_{\perp} = 0$ modes changed our predictions, we performed the preceding analysis both including and excluding these modes. Removing these modes is roughly equivalent to subtracting the mean temperature from each 2D slice in the xy -plane. Accordingly, the variance measured by the power spectrum has three components: the change in the average neutral fraction, the change in this average temperature as a function of redshift, and the average HII region bubble size as a function of redshift. The removal of $k_{\perp} = 0$ essentially eliminates the variance due to the changing average temperature, but it does not eliminate the contributions from changing neutral fraction contribution or the bubble size.

Throughout the analysis, we computed different statistics both including and excluding modes where $k_{\perp} = 0$. In general, we find that removal of this mode causes the light cone case to appear similar to the coeval case. However, performing the analysis with $k_{\perp} = 0$ included has theoretical interest, since it explicitly demonstrates that the light cone effect shifts power from small scales to large scales. Plots similar to Figures 3.5 and 3.6, but with all of the Fourier modes included, are shown in Figures A.1, A.2, and A.3.

The inclusion of all Fourier modes in the analysis produces a signal that deviates by up to two orders of magnitude for large scales ($k \lesssim 0.05 h/\text{Mpc}$). This deviation is with respect to both the light cone effect without these modes, and the coeval case.

The dramatic increase in power at these scales is largely due to the combined change in neutral fraction during reionization. In other words, since there is a significant change in the mean temperature when examining large scales, there is much excess power on these scales. Note that the ringing in the case of the short reionization scenario is due to the sharp discontinuity between the front and back of the box.

This effect also introduces a strong anisotropy in the signal. When analyzing the signal using the power wedges analysis presented in Sec. 3.4.4, we found that the modes parallel to the line of sight contributed about an order of magnitude more power than modes perpendicular, with all of this excess being due to the $k_{\perp} = 0$ mode.

A.2 Additional Figures

In addition to the plot presented in Figure 3.5, we also computed the light cone effect for all box sizes and reionization histories with $f_{\text{HI}} = 0.75, 0.25$. These plots are shown in Figures A.1 and A.3. Also, as mentioned in Appendix A.1, these plots include all Fourier modes. As in the case of $f_{\text{HI}} = 0.5$ in Figure 3.5, the light cone effect is still pronounced, though not quite as prominently. As before, the light cone effect is larger for bigger scales, and is most evident in the 500 Mpc/ h sub-box size. We conclude that regardless of the precise details of reionization, the light cone effect is an essential consideration for the 3D power spectrum of large volumes.

Figure A.4 shows the anisotropic power spectrum for the medium and small box sizes. At small scales, there is more power in $k_{\perp} > k_{\parallel}$ modes, with the exception of $k_{\perp} = 0$. Another interesting feature of these plots is how the shape of the isopower contours changes when the light cone effect is included. As discussed in §3.4.3, the difference in the extent in redshift space and extent along the line of sky changes the amount of power for a given overall k . Also, as can be seen in Figure 3.10, the anisotropy not including the $k_{\perp} = 0$ mode is greater for shorter reionization scenarios. For the smallest sub-box size, there is almost no anisotropy in most of the plot, because the extent in redshift space is small compared to the duration of reionization.

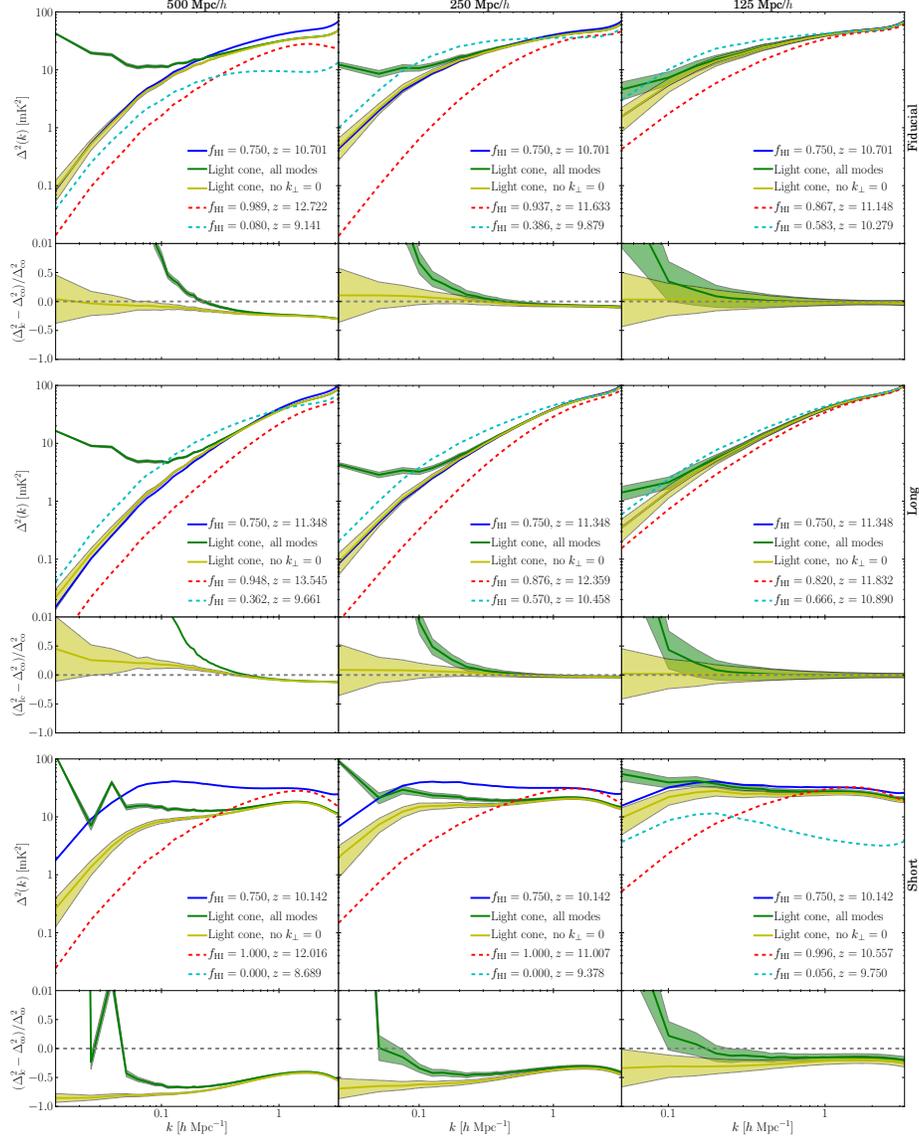


Figure A.1: The same plot as in Figure 3.5, but at a 75% ionization fraction and with all Fourier modes included. The inclusion of all Fourier modes produces a dramatic increase in the power spectrum, especially at small k -modes. (See the text in Appendix A.1 for more discussion.) We also find for the long and fiducial reionization scenarios that there is more power on large scales for the light cone than the coeval case. In general, the light cone effect at this neutral fraction is less pronounced, though still very significant. As in the main case of 50% ionization, the effect is most noticeable for large box sizes. By extension, in the small sub-box case, the effect is still not very significant, as is the same for 50% ionization fraction. One can also see that the shape of the power spectrum has changed dramatically in the case of short reionization.

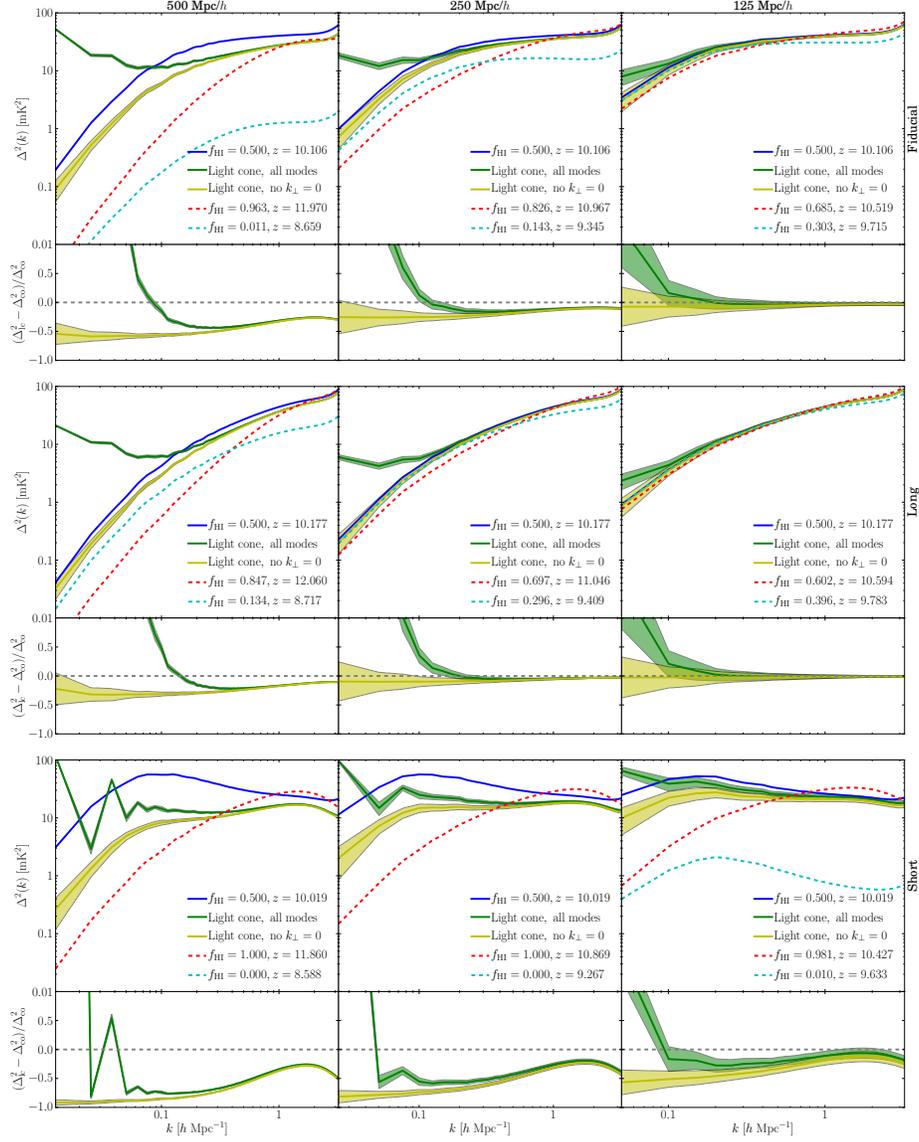


Figure A.2: The same plot as in Figure 3.5, but with all Fourier modes included. Note that the inclusion of the $k_{\perp} = 0$ mode still dramatically increases the power at small k -modes. On small scales, the inclusion of these modes do not change the signal significantly.

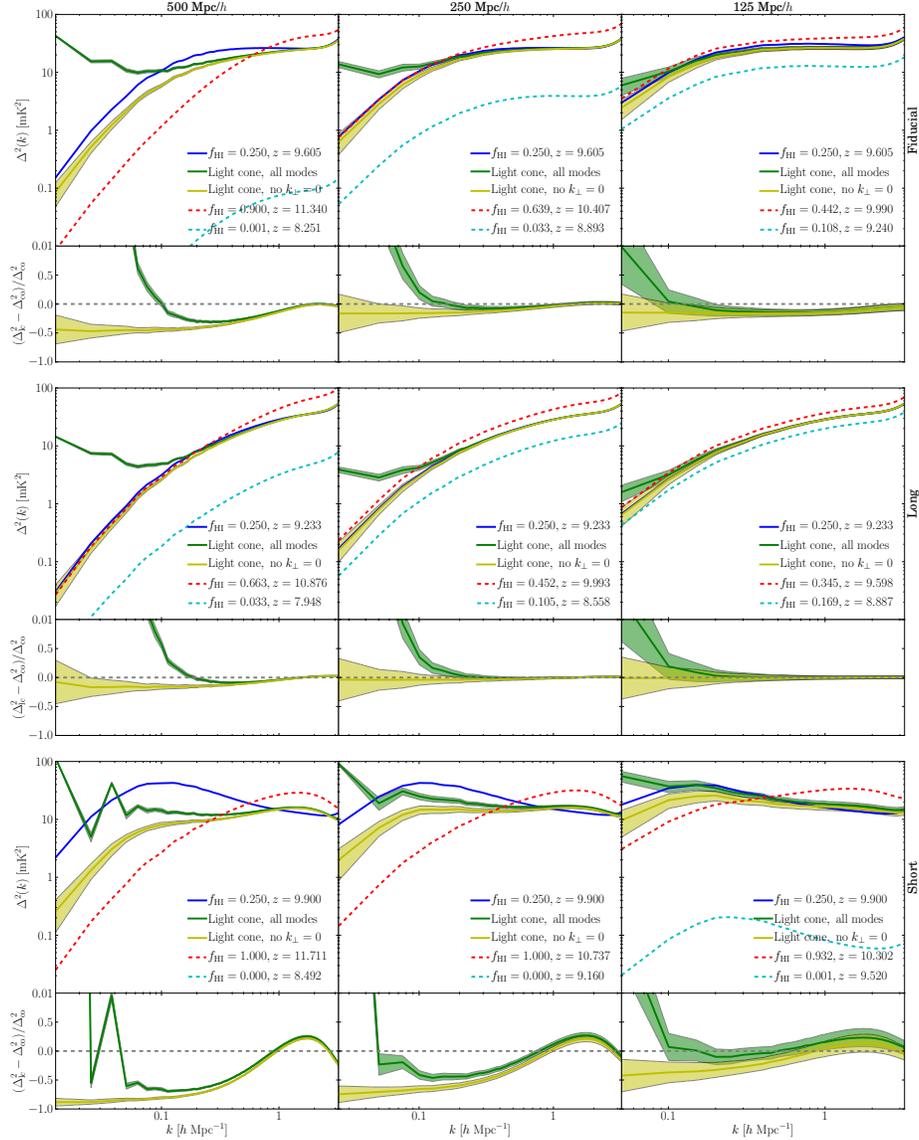


Figure A.3: The same plot as in Figures 3.5 and A.1, but at a 25% ionization fraction. As with the case of a 75% ionization fraction presented in Figure A.1, in general the difference between the light cone and coeval cases is not as great as 50% ionization. Nevertheless, it is still an important feature, and especially on the largest scales. One of the major implications is that the light cone effect is very important at large scales across a large ionization fraction range. Additionally, as in the coeval case, the light cone signal peaks at roughly a 50% ionization fraction.

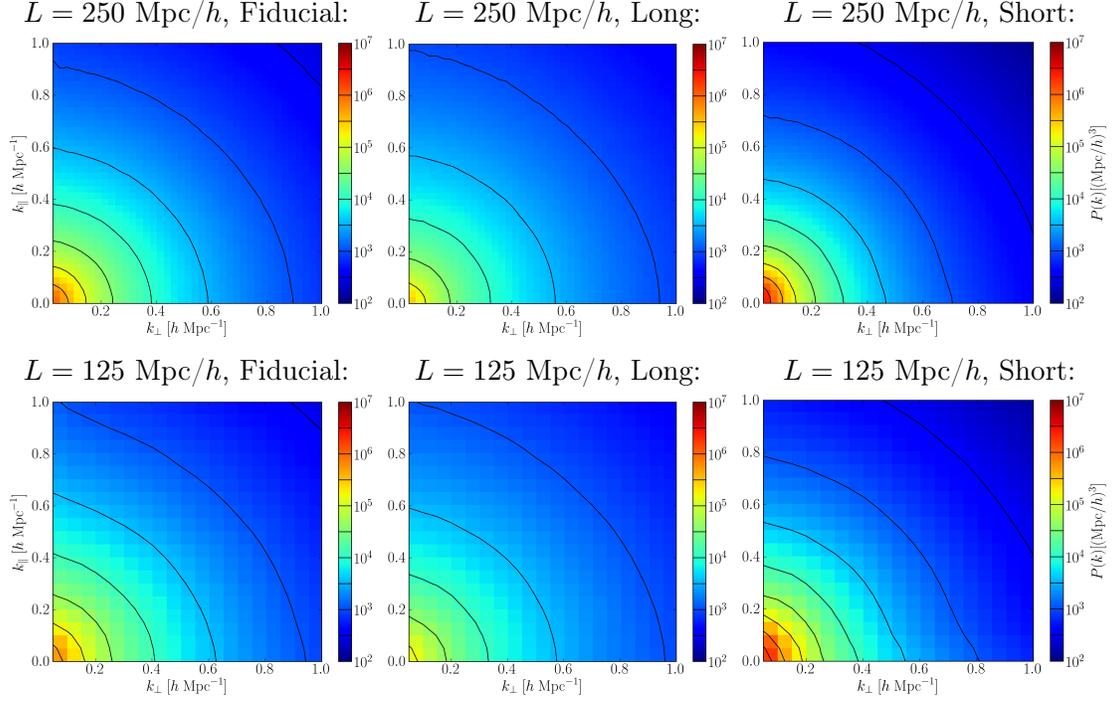


Figure A.4: A comparison of the anisotropic power spectrum across different reionization histories and sub-box sizes. One interesting aspect in these plots is how the isopower lines are shaped: when the light cone effect is included, there is a change in the semi-circular contours. The central portion, near values where $k_{\perp} \sim k_{\parallel}$, has more power than regions where one component is much larger than the other. This is an interesting and subtle change in the contribution to the power introduced in the light cone.

Appendix B

Additional Material for Quasars as Radiation Sources

B.1 Fitting the parameters of the quasar luminosity function

In order to construct a QLF informed by the observations at all redshifts relevant to helium reionization, we have combined the measurements of R13, M12, and M13. We will now briefly summarize the relevant findings of each paper. In all three results, the QLF is parameterized as a double-power law, according to Eqn. (4.12). R13 uses quasars identified from SDSS-III Data Release 9 (DR9), and provides a luminosity-evolution density-evolution (LEDE) model in which the base-10 logarithm of the QLF normalization, $\log_{10} \phi^*$, and the break magnitude M^* , evolve linearly with redshift, as parameterized in Eqns. (4.13-4.14). The parameters α and β are fixed as a function of redshift. Nominally, the LEDE fit is valid over the redshift range $2.2 \leq z \leq 3.5$. M12 uses data from the COSMOS survey, and measures the four QLF parameters at $z \sim 3.2$ and $z \sim 4$. M13 uses quasars identified in SDSS data in Stripe 82 (S82), and reports the four QLF parameters at $z \sim 5$. For all three results, the parameters themselves and their associated 1σ uncertainties are reported. In the M13 results, the authors actually provide three different fits to the observed results. In their fiducial result, they fix the value of β , and fit for the three parameters $\log_{10} \phi^*$, M^* , and α . In a second set of parameters, the authors fix the value of α and find the best-fit values for the other three quantities. Finally, the authors fix M^* , α , and β , and only fit for $\log_{10} \phi^*$. The best-fit values for the parameters change significantly in some cases between the different fits. More importantly, none of these fits seem to be ruled out conclusively by the data presented in M13, and so we incorporate all of the fits in our results.

As explained in Sec. 4.2.5, our goal is to combine the observational data from different epochs. For redshifts $z \leq 3.5$, the parameters from R13 are used. At higher

redshift, the parameters are assumed to vary linearly in redshift. The equations for the parameters are given in Eqns. (4.16a-4.16d). The constant values are taken to be those of R13 at $z = 3.5$, and the slope of the redshift evolution is allowed to take on a range of values. We will now discuss each of the four parameters in turn.

For the parameter $\log_{10} \phi^*$, the fiducial value for the slope c_1 is chosen to reproduce the average of the three reported values of M13 at $z \sim 5$. As discussed in M13, the fits from R13 extrapolated to $z \sim 5$ do not reproduce the overall normalization well, and predict too high of a number density. Thus, a steeper value than the one of R13 is necessary. The range of values for c_1 are chosen to bracket the range of best-fit values reported by M13.

For the parameter M^* , the fiducial value of the slope c_2 is chosen to reproduce the average of the three reported values of M13 at $z \sim 5$. The slope is allowed to take on a range of values that bracket the three reported values of M13. Also note that we have converted between magnitude systems using $M_i(z = 2) = M_{1450} - 1.486$, which assumes that the spectral index $\alpha = 0.5$. If instead the value of $\alpha = 0.61$ is used instead, as suggested by Lusso et al. (2015) and used in the calculations of Sec. 4.5, then conversion is $M_i(z = 2) = M_{1450} - 1.681$. Further, if the SED from Shang et al. (2011) is used, the conversion is $M_i(z = 2) = M_{1450} - 2.139$. The reason for the differences is that the K -corrections depend on the spectral index of the SED (see Eqn. 3 of Richards et al. 2006). By extension, the QLF can be affected when combining different data sets. However, to be consistent with previous works that have combined disparate data sets in this manner (*e.g.*, R13 and M13), we use the conversion given by assuming $\alpha = 0.5$.

For the parameter α , the fiducial value of the slope c_3 is chosen to reproduce the

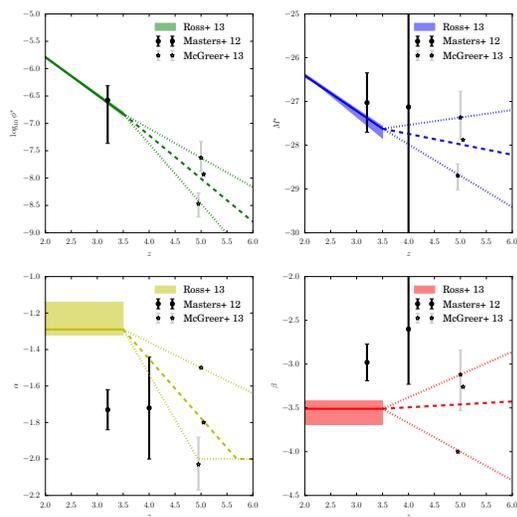


Figure B.1: A plot of the evolution of the QLF parameters as a function of redshift: the base-ten logarithm of ϕ^* (top left), the break magnitude M^* (top right), the faint-end slope α (bottom left), and the steep-end slope β (bottom right). Best-fit values and associated 1σ errors from R13, M12, and M13 are represented as the solid lines with shaded error regions, dark-gray triangles, and light-gray stars, respectively. For the M13 data, all three sets of parameters provided by the authors are plotted at $z \sim 5$, slightly offset for visual clarity. The dashed lines for $z > 3.5$ show the fiducial evolution of the QLF, and the dotted lines show the bracketing ranges of values explored. See the text in Appendix B.1 for further details.

average of the three reported values of M13 at $z \sim 5$. As with the other parameters discussed, a range of values is also explored which brackets all of the reported values of M13. Further, the value of α is bounded to lie where $\alpha > -2$. For $\alpha \leq -2$, the QLF does not converge for low-luminosity objects, and a cutoff luminosity must be specified below which quasars do not contribute significantly to helium reionization. To avoid defining such a cutoff luminosity, the value of α is bounded. As a practical matter, the ultimate goal of this project is to study helium reionization using full numerical simulations, where the minimum resolved halo mass will set the lower-limit of quasar luminosities.

Finally, for the parameter β , the fiducial value for the slope c_4 is chosen to reproduce the average of the values from M13 at $z \sim 5$. The range of slopes is chosen to bracket the values reported by M13. For the fiducial choice of slope, the value of β does not vary significantly with redshift. This range of values incorporates much of the parameter space constrained by M13, without the values of β becoming arbitrarily steep. However, the choice of β ultimately does not significantly affect the ionization level predicted by Eqn. (4.23).

As a final note, the values of α and β at $z \sim 3.2$ from M12 are nominally inconsistent with the combined results from R13. However, when looking at the results for individual redshift bins at $z \sim 3.2$ (*e.g.*, Fig. 15 from R13), the uncertainties for the R13 values are significantly larger, and the results are largely consistent at 1σ . The values of $\log_{10} \phi^*$ and M^* from M12 at $z \sim 3.2$ are consistent with the results from R13, and those at $z \sim 4$ are consistent with the linear redshift evolution given by the requirement of matching the M13 data. Note that in Fig. B.1, we do not plot the value of $\log_{10} \phi^*$ at $z \sim 4$ from M12, because the reported lower-bound of the error bars is larger than the best-fit value, which must be positive. Despite this fact, the best-fit value is very close to the fiducial linear evolution given here.

Figure B.1 shows the measured parameters as a function of redshift, as well as the assumed high- z evolution for each parameter. The solid lines and shaded regions show the best-fit parameters from R13, and the individual points with error bars show the results from M12 and M13. The dashed lines show the fiducial choices for the parameters, which are chosen as outlined above. The dotted lines show the full range of parameters explored. The range of parameter combinations is applied to helium reionization in Figure 4.9 in the top-left panel. Note that, as discussed in Sec. 4.5, this uncertainty primarily affects the early stages of reionization. Due to the recombination term in the calculation of the volume filling fraction and the fact that all reionization histories use the parameters of R13 at $z \leq 3.5$, the high- z values for the QLF do not ultimately affect the timing of reionization significantly; nevertheless, the different reionization scenarios can leave unique observable signatures on the IGM.

Redshift selection	z_{eff}	Light curve	L_{eff}^* ^{1,2}	t_{eff}^* ³
High-z	2.51	Lightbulb	12.92	7.62
		Exp	12.40	7.14
Fiducial	2.39	Lightbulb	13.29	7.77
		Exp	13.05	7.18
Low-z	2.28	Lightbulb	13.17	7.84
		Exp	13.15	7.29

Table B.1: A list of the best-fit parameters for our quasar model as a function of redshift.

B.2 Bias as a function of redshift

In addition to reproducing the “fiducial” sample from the BOSS results, the quasars from the constructed catalogs were also partitioned by redshift into a “high-redshift” and “low-redshift” sample in an analogous manner to the auxiliary BOSS samples. In the case of the BOSS results, the “fiducial” sample is actually the combination of the “high-redshift” and “low-redshift” samples, so these two datasets are statistically independent of each other, but not the fiducial sample. For the purposes of comparing with the quasar catalogs, though, it is possible to compute $\xi(s)$ at distinct points in redshift, and compare with the BOSS results. The central redshifts for the high-redshift and low-redshift samples are $z = 2.51$ and 2.28 , respectively. Then an analysis similar to the above was performed, but at these additional redshifts. This procedure yields further constraints on the bias as a function of redshift in terms of the model parameters t_0 and γ . Figure B.2 is similar to Figure 4.4, and shows how the selection of models varies as a function of redshift. In general, we find that the choice of parameters for our model t_0 and γ evolve slightly with redshift. In general, the BOSS measurements show an increase in bias with decreasing redshift. In order to accommodate this increased bias, the model parameters must vary slightly. In general, the model favors quasars with increased lifetimes as redshift decreases. Despite this evolution with redshift, the relationship between $\log_{10}(t_0)$ and γ remains fairly linear, and it is still possible to parameterize these models in terms of the characteristic lifetime and luminosity factors t_{eff} and L_{eff} as defined in Eqn. (4.21).

Table B.1 summarizes the changes in best-fit parameters as a function of redshift. Interestingly, these values change somewhat: as structure continues to build, models with increasingly higher bias values are preferred. The fact that the best-fit values change demonstrates that the passive evolution of an increased clustering signal within a given model is not sufficient: rather, this redshift evolution introduces additional

¹ L_{eff} and t_{eff} as defined in Eqns. (4.21-4.22)

² $L_{\text{eff}}^* = \log_{10}(L_{\text{eff}}/L_{\odot})$

³ $t_{\text{eff}}^* = \log_{10}(t_{\text{eff}}/\text{yr})$

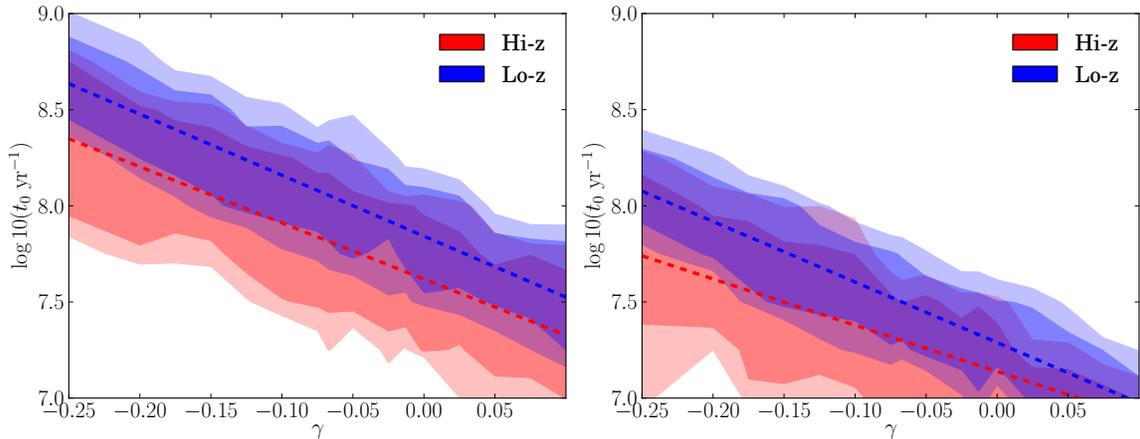


Figure B.2: Parameter space evolution of t_0 and γ from Eqn. (4.3) as a function of redshift for the lightbulb model (left) and the exponential model (right). As redshift decreases, the space of preferred models shifts slightly toward ones with higher intrinsic clustering. This is in addition to the passive evolution in clustering signal that each individual model experiences, which constrains the space of applied models somewhat. Nevertheless, the results are consistent with there being no redshift evolution.

constraints that we can use to select the most appropriate model. Nevertheless, the results are consistent with no redshift evolution. The results of White et al. (2012) also suggest that redshift evolution is minimal. Extending the clustering measurements to a larger redshift range could provide important constraints on the properties of quasar hosts.

B.3 Bias as a function of luminosity

We can also examine the dependence of bias as a function of quasar luminosity. In the preceding analysis, we looked at the fiducial luminosity selection of the BOSS measurements for clustering, $-25 \geq M_i \geq -27$. In order to break the degeneracy in Fig. 4.4, we explored the implications of measuring the clustering of quasars with different luminosity cuts. We examined a high-luminosity cut $M_i \leq -27$, and a low-luminosity cut $-23 \geq M_i \geq -25$. Unfortunately, since the simulation volumes are only $1 (h^{-1}\text{Gpc})^3$, there are an insufficient number (~ 400) of high-luminosity objects to constrain the two-point correlation function.

When fitting the functional form of the two-point correlation function, a power law is used:

$$\xi(s) = \left(\frac{s}{s_0} \right)^\beta. \quad (\text{B.1})$$

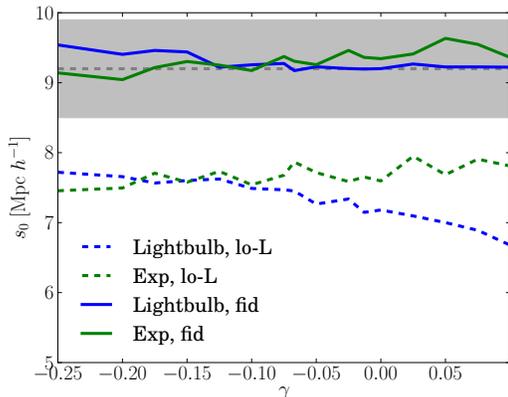


Figure B.3: The best fit parameter s_0 for the two-point correlation function in the form $\xi(s) = (s/s_0)^{-2}$ as a function of power-law index γ from Eqn. (4.3) for the lightbulb and exponential cases using a fiducial (solid) and low-luminosity (dashed) luminosity selection. The gray shaded region shows the BOSS measurement for the fiducial luminosity cut. For the low-luminosity quasars, we see opposite trends for the two models. For the lightbulb, more negative values of γ mean that dimmer quasars have longer lifetimes, which combined with abundance matching implies they have more massive hosts. They therefore have larger values of s_0 compared to more positive values of γ . In the exponential case, larger values of γ show more clustering because the bright quasars are longer lived, and are more likely to be included in the low-luminosity cuts while they are below their peak luminosity. Since they are abundance matched to more massive, highly clustered hosts, this leads to the behavior seen. See the text for further discussion.

In the exponential case, the opposite trend is observed due to the presence of high- L_{peak} interlopers. For positive values of γ , brighter quasars have longer lifetimes, and are more likely to be included in the low-luminosity selection. Since these hosts are abundance matched to occupy more

Fits to the function are made for cases where the exponent β is allowed to vary, and others with a fixed value of $\beta = -2$ as in White et al. (2012). In both cases, the clustering length s_0 increases for larger values of the bias. To fit the best parameters, the parameters s_0 and β that minimized the $\chi^2 = \delta^T C^{-1} \delta$ value were found, where δ is defined as the difference between the average $\xi(s)$ and the functional form and C is the covariance matrix, calculated in the same way as in Sec. 4.3.2. These fits were made for the best-fit models defined in Eqn. (4.21) using the values in Table B.1.

Figure B.3 shows the value of the correlation length fits s_0 for the fiducial luminosity cut $-25 \geq M_i \geq -27$ (solid lines) and the low-luminosity cut $-23 \geq M_i \geq -25$ (dashed lines) for the lightbulb and exponential models. The data are somewhat noisy, owing to the comparatively large shot-noise error in the correlation function measurement. However, there does seem to be a trend emerging: in the lightbulb case, for more negative values of γ , the bias is larger, with the opposite trend for the exponential case. In the lightbulb case, this can be explained by noting, as in Sec. 4.3.2, that in abundance matching longer lifetimes leads to a larger bias in the host halos. For negative values of γ , less luminous quasars have longer lifetimes. Subsequently these quasars are being hosted in more massive halos. This means the clustering is stronger for large negative values of γ , implying a larger value of s_0 .

massive, more clustered halo hosts, this leads to a stronger clustering signal, and a larger value of s_0 . The evolution is not as strong as in the lightbulb case, however. In principle, the clustering measurement in different luminosity ranges could help break the degeneracy of best-fit models.

Unfortunately, in practice this type of measurement might be difficult to actually make. The change in bias between the extreme values of γ is not very significant, and the measurement is very noisy. The shaded gray region in Figure B.3 shows the current 1σ bounds from the BOSS measurement, which has a larger spread than the variation in s_0 as a function of γ . Nevertheless, this ratio is a possible way to break the degeneracy between the different models.

Appendix C

Additional Material for Signatures of Quasar Activity on the IGM

C.1 Renormalizing τ_{eff}

In several previous observational studies of the HI Lyman- α forest (Theuns et al., 2002; Bernardi et al., 2003; Dall’Aglio et al., 2008; Faucher-Giguère et al., 2008), there was a reported dip in the effective optical depth τ_{eff} at $z \sim 3.2$. It was proposed that this dip could be related to helium reionization. Several subsequent studies (Bolton et al., 2009a,b; McQuinn et al., 2009; Compostella et al., 2013) did not reproduce this feature. In particular, the functional form of $\tau_{\text{eff}}(z)$ from Lee et al. (2015) does not include this feature. As explained in Sec. 5.2.3, the usual approach taken in the simulations is to renormalize the photoionization rate of galaxies Γ_{gal} in order to reproduce $\tau_{\text{eff}}(z)$ by construction. As a result, the potential dip at $z \sim 3.2$ would not appear. To study whether this feature emerges from the simulations without renormalization, we have run Simulation H7, which uses the same simulation parameters as H5, but with Γ_{gal} provided by the model of HM12. To isolate the contribution of the galaxies, the authors of HM12 have furnished a series of photoionization rates and photoheating rates which only include the contribution from galaxies, and does not include quasars (P. Madau, private communication). Thus, we are able to determine if the dip in τ_{eff} can be reproduced in our simulations.

C.2 Renormalizing the Lyman- α flux PDF

In Sec. 5.5.2, we discussed the results of measuring the Lyman- α forest flux PDF for the different simulations. In order to compare against observation, Figure 5.9 shows the measurement of the flux PDF from Calura et al. (2012) at $z \sim 2.9$. There is a noticeable difference in the shape between the observational and simulated results. As explained in Sec. 5.5.2, there is a difference in the measured τ_{eff} of the measurements

compared to our simulations, which used the more recent measurements of Lee et al. (2015) to define the value of $\tau_{\text{eff}}(z)$ that the simulations matched. In order to investigate whether the difference in the flux PDF shape could be attributed entirely to the different value of τ_{eff} , we adjusted the average flux absorption $\langle F \rangle$ of the volume to match the lower value of τ_{eff} from Calura et al. (2012).

Figure C.1 shows the flux PDF of Simulation H6 renormalized to have the same value of τ_{eff} as Calura et al. (2012), and the measurements. The figure also shows the original flux PDF for the default normalization. Note that the value of τ_{eff} reported by Lee et al. (2015) is greater than the value reported by Calura et al. (2012). Accordingly, when the spectra have been renormalized to have the same value of τ_{eff} , some of the discrepancy between the simulations and the measurements has been removed. Nevertheless, there is still some tension between the measurements, especially for the bins of high flux ($F \sim 1$). Thus, this difference cannot be attributed entirely to the difference in τ_{eff} . As noted in Calura et al. (2012), the placement of the continuum-level can have a significant effect on the shape of the flux PDF. Additionally, the difference in effective resolution between the simulations and the observations may also play some role.

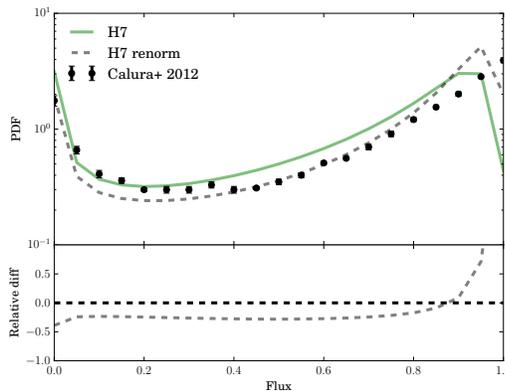


Figure C.1: A comparison of the Lyman- α forest flux PDF of Simulation H6 using the default value of τ_{eff} from Lee et al. (2015) (labeled “fiducial” in the Figure) and a renormalized value of τ_{eff} from Calura et al. (2012), as well as the data from Calura et al. (2012). Note that the discrepancy at high flux values is less in the case of the renormalized value of τ_{eff} , which is consistent with the fact that the value of τ_{eff} is smaller for Calura et al. (2012). Nevertheless, the difference in the normalization cannot account for all of the discrepancy. As shown in Figure 8 of Calura et al. (2012), the continuum uncertainty can have a significant effect on the shape of the flux PDF. Thus, a proper estimation of the continuum level for observations is essential for understanding the flux PDF.

C.3 The Quasar Luminosity Function

In Chapter 4, we provide a method for parameterizing the quasar luminosity function (QLF) as a function of redshift that combines measurements from Ross et al. (2013), Masters et al. (2012), and McGreer et al. (2013) (hereafter referred to as R13, M12,

and M13). These observations provide fits for the QLF at redshifts $2.2 \lesssim z \lesssim 3.5$, $z \sim 3.2$ and $z \sim 4$, and $z \sim 5$, respectively. All three works parameterize the QLF as a double-power law, defined by the same four parameters: ϕ^* , the overall amplitude of the QLF with units of $\text{Mpc}^{-1} \text{mag}^{-1}$; α , the slope of the faint-end of the QLF; β , the slope of the bright-end; and M^* , the so-called break magnitude where the QLF transitions between the slopes α and β . Mathematically, the QLF can be written as:

$$\phi(M) = \frac{\phi^*}{10^{0.4(\alpha+1)(M-M^*)} + 10^{0.4(\beta+1)(M-M^*)}}. \quad (\text{C.1})$$

To combine the R13, M12, and M13 data sets into a single set of quantities, we cast the 4 parameters of the QLF (ϕ^* , M^* , α , and β) as quantities that have evolution in redshift. We define these parameters as:

$$\log_{10} \phi^*(z) = \log_{10} \phi_0^* + c_1(z - 3) + c_2(z - 3)^2, \quad (\text{C.2a})$$

$$M^*(z) = M_0^* + c_3(z - 3), \quad (\text{C.2b})$$

$$\alpha(z) = \alpha_0 + c_4(z - 3), \quad (\text{C.2c})$$

$$\beta(z) = \beta_0 + c_5(z - 3). \quad (\text{C.2d})$$

For the case of the overall normalization $\log_{10} \phi^*$, we include quadratic evolution with redshift. One should note that the comoving number density of quasars is not monotonic, and peaks at $z \sim 2$ (*e.g.*, Fig. 20 of Richards et al. 2006). Accordingly, at high redshifts, there is a significant decrease in the overall amplitude in the QLF, and the redshift evolution is not well fit by a single linear term. Thus, observations suggest that the redshift evolution of this parameter is not purely linear over such a large span in redshift. (See Fig. C.2 for a comparison between a linear and quadratic fit.) The other parameters have redshift evolution that are fit adequately with simple linear evolution in redshift, and so we only include linear terms to avoid over-fitting.

We will now briefly summarize the relevant findings of R13, M12, and M13. In all three results, the QLF is parameterized as a double-power law, according to Eqn. (C.1). R13 uses quasars identified from SDSS-III Data Release 9 (DR9), and provides a luminosity-evolution density-evolution (LEDE) model in which the base-10 logarithm of the QLF normalization, $\log_{10} \phi^*$, and the break magnitude M^* , evolve linearly with redshift. The parameters α and β are fixed as a function of redshift. Nominally, the LEDE fit is valid over the redshift range $2.2 \leq z \leq 3.5$. M12 uses data from the COSMOS survey, and measures the 4 QLF parameters at $z \sim 3.2$ and $z \sim 4$. M13 uses quasars identified in SDSS data in Stripe 82 (S82), and reports the 4 QLF parameters at $z \sim 5$. For all three results, the parameters themselves and their associated 1σ uncertainties are reported. The one exception to this is the value of β from the M13 measurements, which was fixed to a value of $\beta = -4$. The authors report that the value was fixed during the fits since allowing the bright-end slope to take on any value would result in arbitrarily steep value of β . The authors of M13 state this is due to the low number count of objects at very bright magnitudes. In

order to prevent the value from being fixed in our composite QLF, we parameterize β as being an upper limit, with 1σ scatter above the value of $\beta = -4$ of $\sigma = 0.4$. This value is inferred from Figure 18 of M13, which shows the joint likelihood of β and M^* , the break magnitude. At 68% confidence, the authors report $\beta < -3.6$.

It should also be noted that M12 and M13 use different magnitude conventions from the data in R13. Rather than reporting $M_i(z = 2)$, the absolute i -band magnitude at $z = 2$, M12 and M13 report magnitudes as M_{1450} , the absolute magnitude at 1450 \AA . In order to convert between these two systems, we follow the convention of R13, and use $M_i(z = 2) = M_{1450} - 1.486$ (Ross et al., 2013, Appendix B). We should note, though, that this conversion assumes a power-law slope of $\alpha = 0.5$ ($f_\nu \propto \nu^{-\alpha}$), and changes slightly for different spectral indices. Ultimately, the conversion between different magnitude systems is not important for our overall conclusions, since for most of our simulation models, the observables we are most interested in (especially the peak in the IGM temperature, Figure 5.7) are dominated by the QLF at redshifts $z \leq 3.5$. At these redshifts, the QLF is determined with very small statistical uncertainty by the measurements of R13, and thus no conversion between magnitude systems is necessary.

C.3.1 Model Q1

To combine the data from the different data sets, we fit for the four QLF parameters independently as a function of redshift. The parameters are assumed to vary linearly in redshift, except for the base-10 logarithm of the normalization, which includes quadratic evolution. As explained above, we would expect that a purely linear fit of this quantity should

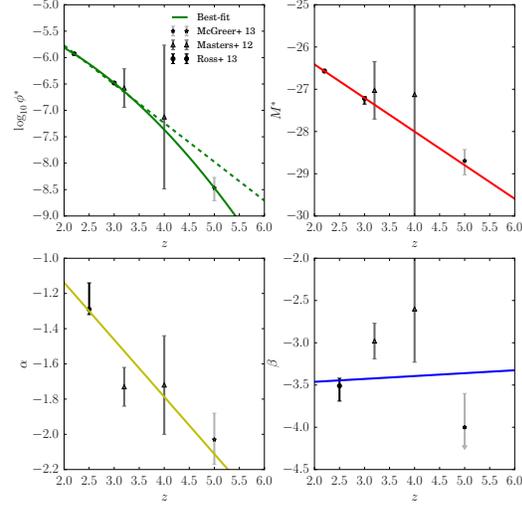


Figure C.2: A plot of the evolution of the QLF parameters as a function of redshift: the base-ten logarithm of ϕ^* (top left), the break magnitude M^* (top right), the faint-end slope α (bottom left), and the steep-end slope β (bottom right). Best-fit values and associated 1σ errors from R13, M12, and M13 are represented as the black circles, dark-gray triangles, and light-gray stars, respectively. The solid lines show the parameterization of the parameters given by Equations (C.2a–C.2d) based on these data, reproduced in Table C.1. For the evolution of $\log_{10} \phi^*$, the dashed line shows the best-fit assuming only linear evolution in z instead of quadratic, motivating an empirical need for quadratic evolution. See the text in Appendix C.3 for further details.

not be adequate over such a large range in redshift, since the total quasar number density peaks around $z \sim 2$ and turns over. The equations for the parameters are given in Equations (C.2a–C.2d), and the resulting best-fit values for the parameters and uncertainties given in Table C.1. Instead of fitting for the evolution of the four parameters independently, it would be better to find a simultaneous fit to all of the data spanning the entire redshift range. However, many degeneracies exist between these parameters, and finding a simultaneous fit to adequately describe all of the data over a very large redshift range is difficult to achieve.

Parameter	Best-fit value
$\log_{10} \phi_0^*$	−6.48
c_1	−0.776
c_2	−0.109
M_0^*	−27.2
c_3	−0.795
α_0	−1.46
c_4	−0.324
β_0	−3.43
c_5	0.0342

Table C.1: A list of the best-fit parameters in Equations (C.2a– C.2d) given the data listed in R13, M12, and M13. These provide a fit to the luminosity function through redshift, and ensure that the abundance of quasars matches observations as nearly as possible. For additional details on the parameters and the fitting procedure, see the text in Appendix C.3.

evolution particularly for α . Therefore, our model includes redshift evolution in these parameters. At $z \sim 2.5$, there is very good agreement between the binned QLF and the fit model of R13. However, we note that the fit values are ultimately not very sensitive to the choice of redshift. Combined with the results from M12 and M13, this creates 4 data points to fit. The fit for all of the parameters is reasonably good, with the notable exception of the steep-end slope β . As noted earlier, constraining β is observationally difficult due to the low number count of objects. Also worth noting is that the fits of M12 do not constrain β with their data directly. Their measurements from the COSMOS field are primarily for faint objects, and are fainter than the break magnitude M^* . In order to determine β in their fits, M12 use measurements from Richards et al. (2006) to provide observations of bright objects. The overall result is little evolution in β over the redshift interval $2.5 \lesssim z \lesssim 5$, with perhaps a slight

Figure C.2 shows the measured parameters as a function of redshift, as well as the best-fit line for each parameter. As explained above, for the QLF normalization $\log_{10} \phi^*$ and break magnitude M^* , we include the parameters from R13 at $z = 2.2$, where the parameters are determined best, and at $z = 3$, in order to provide good constraints on the overall normalization at a slightly higher redshift. Combined with the two points from M12 ($z = 3.2$ and $z = 4$) and the single point from M13 ($z = 5$), there are 5 total data points being fit. For the cases of α and β from R13, we include them at $z = 2.5$, since there is no explicit redshift dependence included in the R13 fits. Nevertheless, when looking at the reported parameters from higher redshift data (and even when comparing with the binned data from $z \sim 3$ in the R13 data), there does seem to be redshift

steepening when moving to lower redshifts. This is the opposite trend of α , which shows a very clear trend of becoming shallower at lower redshifts. Nevertheless, due to the low overall amplitude of the luminosity function at high magnitudes, the precise value of β does not significantly affect the predictions for reionization.

C.3.2 Model Q2

As an alternative to finding the best-fit parameterizations is to simply interpolate between the values reported in R13, M12, and M13. To this end, we take the values for the parameters ϕ^* , M^* , α , and β reported by the different studies to be accurate for their respective redshift ranges. Specifically, we use the values reported by R13 for redshifts $z \leq 3.5$, the values of M12 at $z \sim 4$, and the values of M13 for redshifts of $z \geq 5$. In order to determine values of the parameters at intermediate redshifts, we linearly interpolate in redshift. This method produces a QLF which is consistent with the different measurements by construction, but can introduce some features into the QLF's evolution due to the naïve linear interpolation method. We therefore regard Model Q1 as our fiducial one, and present this one merely as a point of comparison.

Bibliography

- Alcock, C., & Paczyński, B. 1979, *Nature*, 281, 358
- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, *ApJ*, 809, 61
- Barkana, R. 2006, *MNRAS*, 372, 259
- Barkana, R., & Loeb, A. 2001, *Phys. Rep.*, 349, 125
- . 2004, *ApJ*, 609, 474
- . 2005, *ApJL*, 624, L65
- . 2006, *MNRAS*, 372, L43
- Battaglia, N., Natarajan, A., Trac, H., Cen, R., & Loeb, A. 2013a, *ApJ*, 776, 83
- Battaglia, N., Trac, H., Cen, R., & Loeb, A. 2013b, *ApJ*, 776, 81
- Becker, G. D., & Bolton, J. S. 2013, *MNRAS*, 436, 1023
- Becker, G. D., Bolton, J. S., Haehnelt, M. G., & Sargent, W. L. W. 2011a, *MNRAS*, 410, 1096
- Becker, G. D., Rauch, M., & Sargent, W. L. W. 2007, *ApJ*, 662, 72
- Becker, G. D., Sargent, W. L. W., Rauch, M., & Calverley, A. P. 2011b, *ApJ*, 735, 93
- Bernardi, M., Sheth, R. K., SubbaRao, M., et al. 2003, *AJ*, 125, 32
- Bharadwaj, S., & Ali, S. S. 2004, *MNRAS*, 352, 142
- . 2005, *MNRAS*, 356, 1519
- Boera, E., Murphy, M. T., Becker, G. D., & Bolton, J. S. 2014, *MNRAS*, 441, 1916
- Bolton, J. S., & Haehnelt, M. G. 2007, *MNRAS*, 382, 325
- Bolton, J. S., Oh, S. P., & Furlanetto, S. R. 2009a, *MNRAS*, 395, 736

- . 2009b, *MNRAS*, 396, 2405
- Bolton, J. S., Puchwein, E., Sijacki, D., et al. 2016, ArXiv e-prints, arXiv:1605.03462
- Borisova, E., Lilly, S. J., Cantalupo, S., et al. 2015, ArXiv e-prints, arXiv:1510.00029
- Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2005, in *Bulletin of the American Astronomical Society*, Vol. 37, American Astronomical Society Meeting Abstracts, 1217
- Bowman, J. D., & Rogers, A. E. E. 2010, *Nature*, 468, 796
- Bowman, J. D., Rogers, A. E. E., & Hewitt, J. N. 2008, *ApJ*, 676, 1
- Boyle, B. J., Shanks, T., Croom, S. M., et al. 2000, *MNRAS*, 317, 1014
- Busca, N. G., Delubac, T., Rich, J., et al. 2013, *A&A*, 552, A96
- Calura, F., Tescari, E., D’Odorico, V., et al. 2012, *MNRAS*, 422, 3019
- Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, *MNRAS*, 412, 2543
- Caucci, S., Colombi, S., Pichon, C., et al. 2008, *MNRAS*, 386, 211
- Cen, R., & Safarzadeh, M. 2015, *ApJL*, 798, L38
- Ciardi, B., Ferrara, A., & White, S. D. M. 2003, *MNRAS*, 344, L7
- Cisewski, J., Croft, R. A. C., Freeman, P. E., et al. 2014, *MNRAS*, 440, 2599
- Cohen-Tannoudji, C., Diu, B., & Laloë, F. 2005, *Quantum Mechanics, Volume 2* (Wiley-VCH)
- Compostella, M., Cantalupo, S., & Porciani, C. 2013, *MNRAS*, 435, 3169
- . 2014, *MNRAS*, 445, 4186
- Conroy, C., & White, M. 2013, *ApJ*, 762, 70
- Cooray, A. 2004, *Phys. Rev. D*, 70, 063509
- Courbin, F., Faure, C., Djorgovski, S. G., et al. 2012, *A&A*, 540, A36
- Croft, R. A. C., Weinberg, D. H., Bolte, M., et al. 2002, *ApJ*, 581, 20
- Croft, R. A. C., Weinberg, D. H., Katz, N., & Hernquist, L. 1998, *ApJ*, 495, 44
- Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, *MNRAS*, 356, 415

- Croton, D. J. 2009, *MNRAS*, 394, 1109
- Dall’Aglio, A., Wisotzki, L., & Worseck, G. 2008, *A&A*, 491, 465
- Datta, K. K., Mellema, G., Mao, Y., et al. 2012, *Monthly Notices of the Royal Astronomical Society*, 424, 1877
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
- Di Matteo, T., Khandai, N., DeGraf, C., et al. 2012, *ApJL*, 745, L29
- Dijkstra, M., Lidz, A., & Hui, L. 2004, *ApJ*, 605, 7
- Dixon, K. L., & Furlanetto, S. R. 2009, *ApJ*, 706, 970
- Dixon, K. L., Furlanetto, S. R., & Mesinger, A. 2014, *MNRAS*, 440, 987
- Eddington, A. S. 1926, *The Internal Constitution of the Stars* (Cambridge: Cambridge University Press)
- Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, *AJ*, 132, 117
- Faucher-Giguère, C.-A., Prochaska, J. X., Lidz, A., Hernquist, L., & Zaldarriaga, M. 2008, *ApJ*, 681, 831
- Feng, Y., Di Matteo, T., Croft, R., & Khandai, N. 2014, *MNRAS*, 440, 1865
- Field, G. B. 1958, *Proceedings of the IRE*, 46, 240
- Flesch, E. W. 2015, *PASA*, 32, 10
- Fry, J. N., & Gaztañaga, E. 1993, *ApJ*, 413, 447
- Furlanetto, S. R., & Oh, S. P. 2008a, *ApJ*, 682, 14
- . 2008b, *ApJ*, 681, 1
- . 2009, *ApJ*, 701, 94
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *Phys. Rep.*, 433, 181
- Furlanetto, S. R., & Stoeber, S. J. 2010, *MNRAS*, 404, 1869
- Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, *Space Sci. Rev.*, 123, 485
- Gleser, L., Nusser, A., Benson, A. J., Ohno, H., & Sugiyama, N. 2005, *MNRAS*, 361, 1399
- Glikman, E., Djorgovski, S. G., Stern, D., et al. 2011, *ApJL*, 728, L26

- Gnedin, N. Y., & Hui, L. 1998, *MNRAS*, 296, 44
- Greenhill, L. J., & Bernardi, G. 2012, ArXiv e-prints, arXiv:1201.1700
- Gunn, J. E., & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Haardt, F., & Madau, P. 1996, *ApJ*, 461, 20
- . 2012, *ApJ*, 746, 125
- Harker, G., Zaroubi, S., Bernardi, G., et al. 2010, *MNRAS*, 405, 2492
- Hearin, A. P., Zentner, A. R., Berlind, A. A., & Newman, J. A. 2013, *MNRAS*, 433, 659
- Heavens, A. F., Matarrese, S., & Verde, L. 1998, *MNRAS*, 301, 797
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, 208, 19
- Hjerting, F. 1938, *ApJ*, 88, 508
- Hopkins, P. F., & Hernquist, L. 2009, *ApJ*, 698, 1550
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, 163, 1
- Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, *ApJS*, 175, 356
- Hopkins, P. F., Richards, G. T., & Hernquist, L. 2007, *ApJ*, 654, 731
- Hui, L., & Gnedin, N. Y. 1997, *MNRAS*, 292, 27
- Ikeda, H., Nagao, T., Matsuoka, K., et al. 2011, *ApJL*, 728, L25
- . 2012, *ApJ*, 756, 160
- Iliev, I. T., Mellema, G., Pen, U.-L., Bond, J. R., & Shapiro, P. R. 2008, *MNRAS*, 384, 863
- Iršič, V., & Viel, M. 2014, *J. Cosmology Astropart. Phys.*, 12, 024
- Jacobs, D. C., Pober, J. C., Parsons, A. R., et al. 2015, *ApJ*, 801, 51
- Jakobsen, P., Boksenberg, A., Deharveng, J. M., et al. 1994, *Nature*, 370, 35
- Jeeson-Daniel, A., Ciardi, B., & Graziani, L. 2014, *MNRAS*, 443, 2722
- Jena, T., Norman, M. L., Tytler, D., et al. 2005, *MNRAS*, 361, 70
- Jensen, H., Datta, K. K., Mellema, G., et al. 2013, ArXiv e-prints, arXiv:1303.5627

- Kashikawa, N., Ishizaki, Y., Willott, C. J., et al. 2015, *ApJ*, 798, 28
- Kaurov, A. A., & Gnedin, N. Y. 2014, *ApJ*, 787, 146
- Kazin, E. A., Sánchez, A. G., Cuesta, A. J., et al. 2013, *ArXiv e-prints*, arXiv:1303.4391
- Khrykin, I. S., Hennawi, J. F., McQuinn, M., & Worseck, G. 2016, *The Astrophysical Journal*, 824, 133
- Koopmans, L., Pritchard, J., Mellema, G., et al. 2015, *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, 1
- La Plante, P., & Trac, H. 2015, *ArXiv e-prints*, arXiv:1507.03021
- Lee, K.-G., Bailey, S., Bartsch, L. E., et al. 2013, *AJ*, 145, 69
- Lee, K.-G., Hennawi, J. F., Stark, C., et al. 2014, *ApJL*, 795, L12
- Lee, K.-G., Hennawi, J. F., Spergel, D. N., et al. 2015, *ApJ*, 799, 196
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
- Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006, *ApJ*, 641, 41
- Lidz, A., McQuinn, M., Zaldarriaga, M., Hernquist, L., & Dutta, S. 2007, *The Astrophysical Journal*, 670, 39
- Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Hernquist, L. 2008, *ApJ*, 680, 962
- Liu, A., & Parsons, A. R. 2016, *MNRAS*, 457, 1864
- Liu, A., Pritchard, J. R., Allison, R., et al. 2016, *Phys. Rev. D*, 93, 043013
- Liu, A., Tegmark, M., Bowman, J., Hewitt, J., & Zaldarriaga, M. 2009, *Mon. Not. R. Astron. Soc.*, 398, 401
- Loeb, A., & Furlanetto, S. R. 2012, *The First Galaxies in the Universe* (Princeton, NJ: Princeton University Press)
- Loeb, A., & Zaldarriaga, M. 2004, *Physical Review Letters*, 92, 211301
- Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M. 1996, *ApJ*, 472, 509
- Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, *MNRAS*, 449, 4204
- Madau, P., & Haardt, F. 2015, *ApJL*, 813, L8

- Madau, P., Haardt, F., & Rees, M. J. 1999, *ApJ*, 514, 648
- Madau, P., Meiksin, A., & Rees, M. J. 1997, *ApJ*, 475, 429
- Majumdar, S., Bharadwaj, S., & Choudhury, T. R. 2013, *MNRAS*, 434, 1978
- Mao, Y., Shapiro, P. R., Mellema, G., et al. 2012, *MNRAS*, 422, 926
- Martini, P. 2004, *Coevolution of Black Holes and Galaxies*, 169
- Martini, P., & Weinberg, D. H. 2001, *ApJ*, 547, 12
- Masters, D., Capak, P., Salvato, M., et al. 2012, *ApJ*, 755, 169
- McDonald, P. 2003, *ApJ*, 585, 34
- McDonald, P., & Eisenstein, D. J. 2007, *Phys. Rev. D*, 76, 063009
- McDonald, P., Miralda-Escudé, J., Rauch, M., et al. 2001, *ApJ*, 562, 52
- McDonald, P., Seljak, U., Cen, R., Bode, P., & Ostriker, J. P. 2005, *MNRAS*, 360, 1471
- McDonald, P., Seljak, U., Burles, S., et al. 2006, *ApJS*, 163, 80
- McGreer, I. D., Jiang, L., Fan, X., et al. 2013, *ApJ*, 768, 105
- McQuinn, M., Hernquist, L., Lidz, A., & Zaldarriaga, M. 2011, *MNRAS*, 415, 977
- McQuinn, M., Lidz, A., Zaldarriaga, M., et al. 2009, *ApJ*, 694, 842
- McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., & Furlanetto, S. R. 2006, *ApJ*, 653, 815
- Meiksin, A., Tittley, E. R., & Brown, C. K. 2010, *MNRAS*, 401, 77
- Mellema, G., Koopmans, L. V. E., Abdalla, F. A., et al. 2013, *Experimental Astronomy*, arXiv:1210.0197
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, 411, 955
- Mirabel, I. F., Dijkstra, M., Laurent, P., Loeb, A., & Pritchard, J. R. 2011, *A&A*, 528, A149
- Mirocha, J., Harker, G. J. A., & Burns, J. O. 2015, *ApJ*, 813, 11
- Møller, P., & Jakobsen, P. 1990, *A&A*, 228, 299
- Morales, M. F., & Wyithe, J. S. B. 2010, *ARA&A*, 48, 127

- Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, *ApJS*, 221, 27
- Natarajan, A., Battaglia, N., Trac, H., Pen, U.-L., & Loeb, A. 2013, *ApJ*, 776, 82
- Nusser, A. 2005, *MNRAS*, 364, 743
- Oesch, P. A., Brammer, G., van Dokkum, P. G., et al. 2016, *ApJ*, 819, 129
- O’shea, B. W., Bryan, G., Bordner, J., et al. 2005, in *Adaptive Mesh Refinement - Theory and Applications: Proceedings of the Chicago Workshop on Adaptive Mesh Refinement Methods*, Sept. 3–5, 2003, ed. T. Plewa, T. Linde, & V. Gregory Weirs (Berlin, Heidelberg: Springer Berlin Heidelberg), 341–349
- Osmer, P. S. 2004, *Coevolution of Black Holes and Galaxies*, 324
- Outram, P. J., Hoyle, F., Shanks, T., et al. 2003, *MNRAS*, 342, 483
- Ozbek, M., Croft, R. A. C., & Khandai, N. 2016, *MNRAS*, 456, 3610
- Paciga, G., Chang, T.-C., Gupta, Y., et al. 2011, *MNRAS*, 413, 1174
- Paciga, G., Albert, J. G., Bandura, K., et al. 2013, *MNRAS*, arXiv:1301.5906
- Palanque-Delabrouille, N., Yèche, C., Borde, A., et al. 2013, *A&A*, 559, A85
- Palanque-Delabrouille, N., Yèche, C., Baur, J., et al. 2015, *ArXiv e-prints*, arXiv:1506.05976
- Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, *The Astronomical Journal*, 139, 1468
- Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, *ApJ*, 788, 106
- Pen, U.-L., Chang, T.-C., Hirata, C. M., et al. 2009, *MNRAS*, 399, 181
- Peterson, J. B., Aleksan, R., Ansari, R., et al. 2009, in *Astronomy*, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey
- Pichon, C., Vergely, J. L., Rollinde, E., Colombi, S., & Petitjean, P. 2001, *MNRAS*, 326, 597
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, *ArXiv e-prints*, arXiv:1303.5076
- Pober, J. C., Parsons, A. R., Aguirre, J. E., et al. 2013, *ApJL*, 768, L36
- Pober, J. C., Liu, A., Dillon, J. S., et al. 2014, *ApJ*, 782, 66
- Pober, J. C., Ali, Z. S., Parsons, A. R., et al. 2015, *ApJ*, 809, 62

Porciani, C., Magliocchetti, M., & Norberg, P. 2004, MNRAS, 355, 1010

Porciani, C., & Norberg, P. 2006, MNRAS, 371, 1824

Pritchard, J., & Loeb, A. 2012, Rep. Prog. Phys., 75

Puchwein, E., Bolton, J. S., Haehnelt, M. G., et al. 2015, Monthly Notices of the Royal Astronomical Society, 450, 4081

Raičević, M., & Theuns, T. 2011, MNRAS, 412, L16

Reimers, D., Kohler, S., Wisotzki, L., et al. 1997, A&A, 327, 890

Richards, G. T., Strauss, M. A., Fan, X., et al. 2006, AJ, 131, 2766

Richardson, J., Zheng, Z., Chatterjee, S., Nagai, D., & Shen, Y. 2012, ApJ, 755, 30

Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71

Ross, N. P., McGreer, I. D., White, M., et al. 2013, ApJ, 773, 14

Rybicki, G. B., & Lightman, A. P. 1986, Radiative Processes in Astrophysics (Wiley-VCH), 400

Salpeter, E. E. 1964, ApJ, 140, 796

Sánchez, A. G., Kazin, E. A., Beutler, F., et al. 2013, ArXiv e-prints, arXiv:1303.4396

Santós, M. G., Amblard, A., Pritchard, J., et al. 2008, ApJ, 689, 1

Schaye, J., Theuns, T., Leonard, A., & Efstathiou, G. 1999, MNRAS, 310, 57

Schlegel, D., Abdalla, F., Abraham, T., et al. 2011, ArXiv e-prints, arXiv:1106.1706

Schmidt, M., & Green, R. F. 1983, ApJ, 269, 352

Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, AJ, 110, 68

Shang, Z., Brotherton, M. S., Wills, B. J., et al. 2011, ApJS, 196, 2

Shapiro, P. R., Mao, Y., Iliev, I. T., et al. 2013, Physical Review Letters, 110, 151301

Shen, Y., & Kelly, B. C. 2012, ApJ, 746, 169

Shen, Y., Strauss, M. A., Oguri, M., et al. 2007, AJ, 133, 2222

Shen, Y., Strauss, M. A., Ross, N. P., et al. 2009, ApJ, 697, 1656

Shull, J. M. 1979, ApJ, 234, 761

- Sijacki, D., Vogelsberger, M., Genel, S., et al. 2015, MNRAS, 452, 575
- Simha, V., Weinberg, D. H., Davé, R., et al. 2012, MNRAS, 423, 3458
- Slosar, A., Font-Ribera, A., Pieri, M. M., et al. 2011, J. Cosmology Astropart. Phys., 9, 001
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, J. Cosmology Astropart. Phys., 4, 26
- Springel, V. 2005, MNRAS, 364, 1105
- . 2010, MNRAS, 401, 791
- Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776
- Stevans, M. L., Shull, J. M., Danforth, C. W., & Tilton, E. M. 2014, ApJ, 794, 75
- Syphers, D., Anderson, S. F., Zheng, W., et al. 2009a, ApJS, 185, 20
- . 2012, AJ, 143, 100
- Syphers, D., & Shull, J. M. 2014, ApJ, 784, 42
- Syphers, D., Anderson, S. F., Zheng, W., et al. 2009b, ApJ, 690, 1181
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, ApJ, 565, 773
- Tepper-García, T. 2006, MNRAS, 369, 2025
- Theuns, T., Bernardi, M., Frieman, J., et al. 2002, ApJL, 574, L111
- Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, MNRAS, 301, 478
- Thorne, K. S. 1974, ApJ, 191, 507
- Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709
- Trac, H., & Cen, R. 2007, ApJ, 671, 1
- Trac, H., Cen, R., & Loeb, A. 2008, ApJL, 689, L81
- Trac, H., Cen, R., & Mansfield, P. 2015, ApJ, 813, 54
- Trac, H., & Pen, U.-L. 2004, New A, 9, 443
- Valageas, P., Clerc, N., Pacaud, F., & Pierre, M. 2011, A&A, 536, A95
- Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013, Phys. Rev. D, 88, 043502

- Voytek, T. C., Natarajan, A., Jáuregui García, J. M., Peterson, J. B., & López-Cruz, O. 2014, *ApJL*, 782, L9
- Warren, S. J., Hewett, P. C., & Osmer, P. S. 1994, *ApJ*, 421, 412
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2012, ArXiv e-prints, arXiv:1201.2434
- White, M., Pope, A., Carlson, J., et al. 2010, *ApJ*, 713, 383
- White, M., Myers, A. D., Ross, N. P., et al. 2012, *MNRAS*, 424, 933
- Windhorst, R. A., Cohen, S. H., Jansen, R. A., Conselice, C., & Yan, H. 2006, *New A Rev.*, 50, 113
- Worseck, G., Prochaska, J. X., Hennawi, J. F., & McQuinn, M. 2014, ArXiv e-prints, arXiv:1405.7405
- Worseck, G., Prochaska, J. X., McQuinn, M., et al. 2011, *ApJL*, 733, L24
- Wouthuysen, S. A. 1952, *AJ*, 57, 31
- Wyithe, J. S. B., & Loeb, A. 2003, *ApJ*, 595, 614
- Yu, Q., & Lu, Y. 2004, *ApJ*, 602, 603
- Yu, Q., & Tremaine, S. 2002, *MNRAS*, 335, 965
- Zahn, O., Lidz, A., McQuinn, M., et al. 2007, *ApJ*, 654, 12
- Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, *ApJ*, 608, 622
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, *ApJ*, 630, 1
- Zheng, W., Anderson, S. F., Kriss, G. A., et al. 2005, in *IAU Colloq. 199: Probing Galaxies through Quasar Absorption Lines*, ed. P. Williams, C.-G. Shu, & B. Menard, 484–486
- Zheng, W., Meiksin, A., Pifko, K., et al. 2008, *ApJ*, 686, 195