TheOpacity of the Intergalactic Medium Measured along Quasar Sightlines at $z \sim 6$

Anna-Christina Eilers,1,2, Frederick B. Davies,3, and Joseph F. Hennawi3

1 Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany; eilers@mpia.de
2 International Max Planck Research School for Astronomy & Cosmic Physics at the University of Heidelberg, Germany
3 Physics Department, University of California, Santa Barbara, CA 93106-9530, USA

Received 2018 March 8; revised 2018 July 9; accepted 2018 July 18; published 2018 August 30

Abstract

We publicly release a new sample of 34 medium resolution quasar spectra at $5.77 \leq z_{	ext{em}} \leq 6.54$ observed with the Echellette Spectrograph and Imager on the Keck telescope. This quasar sample represents an ideal laboratory to study the intergalactic medium (IGM) during the end stages of the epoch of reionization, and constrain the timing and morphology of the phase transition. For a subset of 23 of our highest signal-to-noise ratio spectra ($S/N > 7$, per 10 km s$^{-1}$ pixel), we present a new measurement of the Ly$\alpha$ forest opacity spanning the redshift range $4.8 \leq z \leq 6.3$. We carefully eliminate spectral regions that could be causing biases in our measurements due to additional transmitted flux in the proximity zone of the quasars, or extra absorption caused by strong intervening absorption systems along the line of sight. We compare the observed evolution of the IGM opacity with redshift to predictions from a hydrodynamical simulation with uniform ultraviolet background (UVB) radiation, as well as two semi-numerical patchy reionization models, one with a fluctuating UVB and another with a fluctuating temperature field. Our measurements show a steep rise in opacity at $z \gtrsim 5.0$ and an increased scatter and thus support the picture of a spatially inhomogeneous reionization process, consistent with previous work. However, we measure significantly higher optical depths at $5.3 \lesssim z \lesssim 5.7$ than previous studies, which reduces the contrast between the highest opacity Gunn–Peterson troughs and the average opacity trend of the IGM, which may relieve some of the previously noted tension between these measurements and reionization models.

Key words: dark ages, reionization, first stars – intergalactic medium – methods: data analysis – quasars: absorption lines

1. Introduction

Determining when and how the epoch of reionization proceeded is one of the major goals of observational cosmology today. During this early evolutionary phase of our universe, the cosmic “dark ages” following recombination ended, and the intergalactic medium (IGM) transitioned from a neutral state into the ionized medium that we observe today due to the ultraviolet radiation of the first stars, galaxies, and quasars. The details of the reionization process not only reflect the nature of these primordial objects, but also the formation of large-scale structure and are therefore a subject of major interest. Despite much progress in the last decade, there are still crucial yet unanswered questions regarding the exact timing and morphology of reionization.

The most compelling constraints to date on the end of the reionization epoch come from the evolution of the Lyman-alpha (Ly$\alpha$) forest opacity, observed in the spectra of $z \gtrsim 6$ quasars (Fan et al. 2006; Becker et al. 2015; Bosman et al. 2018). The detection of transmitted flux spikes in the high redshift quasar spectra and the absence of large Gunn–Peterson (GP) troughs (Gunn & Peterson 1965) below $z \lesssim 5.0$ indicate that the epoch of reionization must be completed by that time. Both the steep rise in the observed opacity around $z \gtrsim 5.5$ as well as the increased scatter of the measurements suggest a qualitative change in the ionization state of the IGM (Becker et al. 2015), provoked by a decrease in the ionizing ultraviolet background (UVB) radiation (Calverley et al. 2011; Wyithe & Bolton 2011).

The inferred rapid decline of the UVB radiation has been interpreted as an indication for the end stages of reionization (e.g., Fan et al. 2006; Bolton & Haehnelt 2007; Calverley et al. 2011). Fan et al. (2006) argue further that the increased scatter in the opacity measurements around $z \gtrsim 5.5$ could be explained by strong variations in the UVB radiation field as expected in patchy and inhomogeneous reionization scenarios.

However, Lidz et al. (2006) have argued that large-scale density fluctuations alone could explain the significant variations between sightlines. They calculated the scatter arising solely from density fluctuations while assuming a uniform UVB, which gives results comparable to the observations by Fan et al. (2006). If this were correct, the evidence for patchy reionization based on the observations of the mean opacity would be significantly weakened.

Measurements of the Ly$\alpha$ forest opacity along additional quasar sightlines by Becker et al. (2015) finally showed that the observed scatter in the optical depth measurements significantly exceeds not only fluctuations expected from the density field alone, but also fluctuating UVB models with a spatially uniform mean free path of ionizing photons. They posited that, if a fluctuating UVB was in fact the source of the large scatter in optical depth, the mean free path must be spatially variable, supporting the interpretation of probing the end stages of an inhomogeneous hydrogen reionization period. Subsequently, Davies & Furlanetto (2016) modeled the UVB with a spatially varying mean free path and found that the additional fluctuations were sufficient to explain the extra scatter in the optical depth measurements (see also Chardin et al. 2017).

An alternative explanation for the observed scatter in the mean opacity was presented by D’Aloisio et al. (2015), who showed that residual spatial fluctuations in the temperature field could result in extended opacity variations. The spatially varying temperature field is a natural consequence of an extended an inhomogeneous reionization process, wherein regions that reionized early have had time to cool, while...
regions that reionized late are still hot. The amplitude of the resulting opacity variations then depends directly on the timing and duration of the reionization process.

The largest sample of quasar sightlines to date used for IGM opacity measurements was recently presented by Bosman et al. (2018). They compared their findings to IGM models that included either a fluctuating UVB or temperature fluctuations and concluded that neither fully captures the observed scatter in IGM opacity.

In this paper we present a new data set of 34 high redshift quasar spectra at $5.77 \leq z_{em} \leq 6.54$, which we make publicly available, and which presents an ideal laboratory for studying the epoch of reionization and setting constraints on the timing and the morphology of the reionization process. For a subset of 23 quasar sightlines we present new measurements of the evolution of the mean opacity of the IGM within the Ly$\alpha$ forest between $4.8 \leq z \leq 6.3$. We carefully mask all spectral regions that could be biasing our measurements—the region in the immediate vicinity of the quasar, its so-called proximity zone that where the transmitted flux is enhanced due to the radiation from the quasar itself, as well as all patches exhibiting additional absorption due to intervening absorption systems such as damped Ly$\alpha$ absorbers (DLAs), along the line of sight to the quasars. Additionally we correct for possible offsets in the zero level of the quasar spectra due to potential systematics in the data reduction procedure.

We then compare our opacity measurements to predictions from a hydrodynamical simulation for three different cases—assuming a uniform UVB radiation field, a fluctuating UVB field, and a fluctuating temperature field—in order to describe the observed evolution of the IGM opacity, and to assess the excess of inhomogeneities in the density, radiation or temperature field that would be required to explain our measurements.

This paper is organized as follows: in Section 2 we describe our data set and its properties. In this section we also present the data release of this quasar sample within the igmspec database. The methods we use to continuum normalize the quasar spectra and measure the IGM opacity are described in Section 3. The results of the opacity measurements and their evolution with redshift are presented in Section 4. We compare our measurements to different outputs from a hydrodynamical simulation, which are described in Section 6. The implications for the epoch of reionization are discussed in Section 7, and we conclude in Section 8 with a summary of the main results. Throughout the paper we assume a cosmology of $h = 0.685$, $\Omega_m = 0.3$ and $\Omega_{\Lambda} = 0.7$, which is consistent within the 1$\sigma$ errorbars in Planck Collaboration et al. (2016).

## 2. High Redshift Quasar Sample

In this section we describe the properties of the data sample of quasar spectra that we use for our analysis of the IGM opacity. This data set has been previously introduced in Eilers et al. (2017a), in which we conducted a detailed analysis of the proximity zones of the quasars. We briefly summarize the details of the observations in Section 2.1, and the data reduction procedure in Section 2.2. The properties of this quasar sample are described in Section 2.3, before presenting the data release at the end of Section 2.4.

### 2.1. Properties of the Data Set

Our complete data set consists of 34 quasar spectra at redshifts $5.77 \leq z_{em} \leq 6.54$, observed at optical wavelengths (4000–10000 Å) with the Echellette Spectrograph and Imager (ESI; Shethis et al. 2002) at the Keck II Telescope from 2001–2016. The data was collected from the Keck Observatory Archive and complemented with our own observations of four objects (PSO J036+03, PSO J060+25, SDSS J0100+2802, and SDSS J1137+3549) that we observed in 2016 January. All observations were obtained using slit widths ranging from 0.75–1.0, resulting in a resolution of $R \approx 4000$–5400. The total exposure times vary from 0.3 hr $\leq t_{exp} \approx 25$ h, resulting in median signal-to-noise ratios (S/N) in the quasar continuum, at rest-frame wavelengths of 1250–1280 Å between $S/N \lesssim 112$ per pixel. The details of the individual observations are shown in Table 1 of Eilers et al. (2017a).

### 2.2. Data Reduction

A detailed description of the data reduction can be found in Eilers et al. (2017a) and will be summarized here only briefly. All spectra were reduced uniformly using the ESIRedux pipeline developed as part of the XIDL suite of astronomical routines in the Interactive Data Language. This pipeline employs standard data reduction techniques comprising the following: images are overscan subtracted, flat fielded using a normalized flat field image, and then wavelength calibrated by means of a wavelength image constructed from afternoon arc lamp calibration images. After identifying the objects in the science frames, we subtracted the background using B-spline fits (Kelson 2003; Bernstein et al. 2015) to the object-free regions of the slit. The profiles of the science objects were also fit with B-splines, and an optimal extraction was performed on the sky-subtracted frames. We combined 1D spectra of overlapping echelle orders to produce a spectrum for each exposure, and co-added individual exposures into a final 1D spectrum. For a more detailed description of the applied algorithms, see Bochanski et al. (2009). We further optimized the XIDL ESI pipeline to improve the data reduction for quasars at high redshift by differentiating two images (ideally taken during the same observing run) with similar exposure times, analogous to the standard difference imaging techniques performed for near-infrared observations, in order to improve the sky subtraction especially in the reddest echelle orders, which are affected by fringing. This procedure requires dithered exposures for which the trace of the science object lands at different spatial locations on the slit. Dithered exposures have the additional advantage that different parts of the fringing pattern are being sampled, and hence a combination of different exposures further reduces fringing issues in the data. However, since not every observer dithered their object along the slit, it was not possible for us to apply this procedure to $\approx$10% of the exposures that we took from the archive.

We co-added exposures from different observing runs taken by various PIs to maximize the S/N. To combine the data from different observing runs, we weighted each 1D spectrum by its squared signal-to-noise ratio (S/N$^2$), determined in the quasar

---

5 https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOALogin
6 http://www2.keck.hawaii.edu/inst/esi/ESIRedux/
7 http://www.ucolick.org/~xavierIDL/
continuum region of each spectrum, i.e., at wavelengths longer than the Ly\(\alpha\) emission line. This ensures that spectral regions with low or no transmitted flux, which are common in high redshift quasar spectra, obtain the same weight as regions with more transmitted flux.

All final reduced quasar spectra are shown in Figure 1, sorted by their emission redshift.

### 2.3. Quasar Properties

The properties of all quasars in our data sample, such as the emission redshift \(z_{\text{em}}\), the absolute magnitude \(M_{1450}\) at \(\lambda_{\text{rest}} = 1450 \text{ Å}\) in the rest frame, the size of the proximity zone \(R_p\) of the quasars, and the S/N of their spectra, are presented in Table 1.

The proximity zone measurements are taken from Eilers et al. (2017a). The uncertainties of these measurements arise solely from uncertainties in the redshift estimate, since these errors provide the largest source of uncertainty for the proximity zone measurements. For one object, SDSS J0840 +5624 we do not report a measurement of its proximity zone, because associated absorption systems located in the immediate vicinity of the quasar prematurely truncate its proximity zone (see Appendix A in Eilers et al. 2017a). Thus, we assume conservatively a region of \(\Delta \nu = 5000 \text{ km s}^{-1}\) to be within the influence of the quasar’s radiation, resulting in an effective \(R_p = 7.39 \pm 0.30\).

For three objects (ULAS J0203–0012, SDSS J0353+0104, and SDSS J1048+4637) no estimates of their proximity zones have been determined. These quasars show broad absorption line (BAL) features, which make a precise and unbiased estimate of their proximity zones unfeasible. Due to the additional absorption features in their spectra we exclude these objects from the analysis of the IGM opacity.

### 2.4. Data Release

Our new data set, including the final co-added spectra and their noise vectors together with the estimated quasar continua (see Section 3.1) as well as additional meta data on the quasars, will be available via the igmspec database\(^8\) (Prochaska 2017). A catalog of the data release comprising the main properties of the data set is shown in Table 4 in Appendix A.

### 3. Methods

In order to measure the IGM opacity the quasar spectra must be normalized to their unabsorbed continua. In this section we explain our method for continuum normalizing the quasar spectra (Section 3.1) and present the details of the IGM opacity measurements (Section 3.2). In the last part of this section (Section 3.3) we describe a procedure for correcting small offsets in the zero level of the quasar spectra.

#### 3.1. Quasar Continuum Normalization

We fit the quasar continua and normalize the spectra in a similar manner as that previously conducted by Eilers et al. (2017a). We summarize the main steps here briefly, but refer the reader to the previous paper for more details.

All spectra were normalized to unity at \(\lambda_{\text{rest}} = 1280 \text{ Å}\) in a spectral region that is free of emission lines. The quasar continuum was then estimated with principal component spectra (PCS) from a principal component analysis (PCA) of lower redshift quasar spectra (Pâris et al. 2011). The concept of PCA is to represent each continuum spectrum \(|q_\lambda\rangle\) at wavelength \(\lambda\) by a reconstructed spectrum comprising a mean spectrum \(|\mu_\lambda\rangle\) and a sum of \(m\) weighted PCS \(|\xi_i,\lambda\rangle\), i.e.,

\[
|q_\lambda\rangle \approx |\mu_\lambda\rangle + \sum_{i=1}^{m} \alpha_i |\xi_i,\lambda\rangle,
\]

where the index \(i\) refers to the \(i\)th PCS and \(\alpha_i\) denotes its weight.

Since the quasars in our data sample are all at \(z_{\text{em}} \sim 6\), they experience substantial absorption due to intervening neutral hydrogen within the IGM blueward of the Ly\(\alpha\) emission line. Thus, the continuum estimate was performed solely on the unabsorbed continuum spectrum redward of the Ly\(\alpha\) emission line with a set of PCS from Pâris et al. (2011) covering the wavelengths 1215.67 Å \(\leq\lambda_{\text{rest}} \leq 2000\) Å. We take the model that minimizes \(\chi^2\) using the noise vector from the spectra as the best estimate.

In order to obtain coefficients \(\alpha\) for a set of PCS that cover the entire spectral region between 1020 Å \(\leq\lambda_{\text{rest}} \leq 2000\) Å, we use a projection matrix \(P\) to project the estimated coefficients for the PCS redward of Ly\(\alpha\) onto this new set of coefficients for the entire spectrum. The projection matrix \(P\) has been computed by Pâris et al. (2011) using the set of PCS for both the red wavelength side only and the whole spectral region covering wavelengths blueward and redward of Ly\(\alpha\). Hence,

\[
\alpha = P \cdot \alpha_{\text{red}}.
\]

This new set of coefficients together with Equation (1) provides a continuum model for each quasar covering all wavelengths 1020 Å \(\leq\lambda_{\text{rest}} \leq 2000\) Å. Pâris et al. (2011) estimate that the median uncertainty on the transmitted flux in the Ly\(\alpha\) forest to be \(|\Delta F_{\text{forest}}| \approx 5\%\). However, since we do not take all PCA components into account and do not have the full spectral coverage up to \(\lambda_{\text{rest}} = 2000\) Å to estimate the continua, the uncertainty on the continua in our quasar spectra is most likely higher, i.e., \(|\Delta F_{\text{forest}}| \approx 10\%–20\%\).

For an estimate of the continua at lower wavelength we take the composite quasar spectrum provided by Shull et al. (2012), constructed from 22 low redshift quasars observed with the Cosmic Origins Spectrograph on the Hubble Space Telescope (HST) that extends from 550–1750 Å in the rest frame, and rescale the composite spectrum to match the PCA-constructed continuum model at \(\lambda_{\text{rest}} = 1020\) Å. We augment the continuum model by simply appending the composite spectrum at wavelengths \(\lambda_{\text{rest}} < 1020\) Å. Note, however, that we do not use the spectrum at \(\lambda_{\text{rest}} < 1020\) Å for the analysis of the Ly\(\alpha\) forest opacity. A few example quasar spectra from our data set and its continuum model are shown in Figure 2, and all remaining quasar spectra that we analyzed and their estimated continua are shown in Figure 12 in Appendix C.

Note that in some cases the N\(\acute{\text{v}}\) line at \(\lambda_{\text{rest}} = 1240\) Å is not very well represented by the continuum fit. This behavior occurs when the N\(\acute{\text{v}}\) line is slightly blue- or redshifted compared to the systemic redshift of the quasar and those shifts are not accounted for in the PCA basis. Additionally, the

---

8 http://specdb.readthedocs.io/en/latest/igmspec.html
continuum for quasar spectra with very weak emission lines, such as SDSS J0100+2802 or SDSS J1148+5251, is not very well captured by the PCA.

3.2. Measuring the Optical Depth of the IGM

We exclude BAL quasars (ULAS J0203+0012, SDSS J0353+0104, and SDSS J1048+4637) from our IGM optical
depth sample to avoid potential contamination by broad non-IGM absorption in the Ly$\alpha$ forest, and quasars with only very low S/N data, i.e., $S/N < 7$, whose spectra are more subject to systematic errors. Our final IGM optical depth sample then consists of 23 quasar spectra out of the original 34.
We estimate the mean opacity of the IGM by means of the effective optical depth, which is defined as

$$\tau_{\text{eff}} = -\ln \left\langle F \right\rangle,$$

where \(F\) is the continuum normalized flux. The effective optical depth is measured in discrete spectral bins along the line of sight of each quasar. We chose fixed comoving bins of size 50 comoving Mpc h\(^{-1}\) (cMpc h\(^{-1}\)) (Becker et al. 2015), which contains a similar path length as the bins of size \(\Delta z = 0.15\) at \(z \sim 5-6\) previously applied by Fan et al. (2006).

In order to avoid biases in the measurement of the opacity of the IGM, we mask the spectral region around each quasar that is strongly influenced and ionized by the quasar’s own radiation. We use the measurements for the proximity zones \(R_p\) as an estimate the influenced region. However, the proximity zone is defined such that it does not completely reach out to the ionization front expanding from the quasar. Thus, the influence of the quasar’s radiation is expected to be still present outside of its measured proximity zone (Eilers et al. 2017a), since the radiation of the quasar still dominates the UVB radiation, i.e., \(\Gamma_{\text{QSO}} \gg \Gamma_{\text{UVB}}\) at \(R_p\). Hence, we mask an additional 2.5 proper Mpc (pMpc) around each quasar, i.e., the masked region measures \(R_p + 2.5\) pMpc, in order to eliminate all enhanced transmission due to the quasar’s radiation.

Thus, we choose the maximum wavelength that we consider for opacity measurements to lie just blueward of this masked region. The minimum wavelength we consider for measurement to lie just blueward of this masked region. The last column states whether we used the quasar sightline for the IGM opacity analysis in this paper.

Notes. The columns show the object name, the coordinates of the quasar given in RAhms and DECdms, the emission redshift and the quasar’s magnitude \(M_{\text{1450}}\), the measurements for their proximity zones, and the S/N of the quasar spectrum. The last column states whether we used the quasar sightline for the IGM opacity analysis in this paper.

\(^{a}\) Median S/N per 10 km s\(^{-1}\) pixel; estimated between 1250 Å \(\leq \lambda_{\text{rest}} \leq 1280\) Å.

\(^{b}\) Estimate of the proximity zone conservatively assuming \(\Delta v = 5000\) km s\(^{-1}\), because the proximity zone is prematurely truncated due to associated absorption systems (see Appendix A in Eilers et al. 2017a).
same redshift as a spectral region showing complete absorption in the forest of the spectrum. Additionally, we searched the literature for DLAs and low-ion metal absorbers along the quasar sightlines in our sample. For each absorber, we conservatively masked the spectral region around the absorption system within $\pm 30 \, \text{Å}$ in the observed wavelength frame in the Ly$\alpha$ forest at the corresponding wavelength. Spectral bins containing absorbers were then excluded from the IGM opacity measurements. Table 2 shows a compilation of all identified absorbers along the line of sight to the quasars in our sample.

### Table 2: Intervening Low-ion Absorption Systems along the Line of Sight

| Object         | $z_{\text{em}}$ | $z_{\text{abs}}$ | References$^a$ |
|----------------|-----------------|------------------|----------------|
| SDSS J1148+5251 | 6.4189          | 6.0115           | 1              |
| SDSS J2054--0005 | 6.0391          | 5.9776           | 1              |
| SDSS J1630+0412 | 6.065           | 5.8865           | 1              |
| SDSS J1137+3549 | 6.03            | 5.0124           | 1              |
| SDSS J1623+3112 | 6.2572          | 5.8415           | 1              |
| SDSS J0840+5624 | 5.8441          | 5.5940           | 2              |
| SDSS J0002+2550 | 5.82            | 4.914            | 3              |
| SDSS J0100+2802 | 6.3258          | 6.1437           | 3              |
| SDSS J0818+1722 | 6.02            | 5.7911           | 1              |

Note. The columns show the name of the object and its emission redshift, the absorption redshift of the intervening low-ion absorber, and the reference therefor.

$^a$ Reference for low-ion absorbers. 1: Becker et al. (2011), 2: Ryan-Weber et al. (2009), 3: this work.

Table 2 Intervening Low-ion Absorption Systems along the Line of Sight

Note that most of the identified absorbers have already been found by other authors, since most quasars in our data set were previously known and observed.

Additionally we mask all spuriously high pixels within the Ly$\alpha$ forest of the quasar spectra by checking for single pixels showing $F > 1$ in the continuum normalized spectra. Because sky-subtraction systematics occasionally result in large negative sky-subtraction residuals, we also mask all negative flux pixels with the $2.5\%$ lowest S/N to avoid biases due to large uncertainties in pixels that fall onto sky lines and have large negative residuals.

After masking all low-ion metal absorption systems, proximity zones, and spuriously high and negative pixels, the combined usable path length for the opacity measurements is 6350 cMpc h$^{-1}$ for the 23 quasar sightlines in our data sample. The spectral regions in which we measure the mean flux and calculate its effective optical depth are shown for each quasar as the dark and light blue colored bars in Figure 3. Masked regions are shown in white. The gray regions show path lengths that are in principle usable but are not used, because the remaining unmasked region would be smaller than our chosen bin size of 50 cMpc h$^{-1}$.

### 3.3. Correcting for Offsets in the Zero Level of the Spectra

The noise for pixels with no intrinsic flux should be symmetrically distributed around zero, since pixels with zero flux have equal probability to be scattered into positive or negative values. This idea was applied by McGreer et al. (2011), for example, to estimate the number of so-called “dark pixels” that have a flux value consistent with zero. However, a detailed inspection of the quasar spectra in our data set reveals that the zero level in some spectral regions can be slightly biased, i.e., we do not observe a symmetric distribution around...
zero in the noise, possibly caused by sky-subtraction systematics present in a small fraction of exposures. But even tiny offsets in the zero level can cause large differences in the opacity estimates, especially in regions with very little transmitted flux. Because these are the regions we are particularly interested in, we correct for small offsets in the zero level of the spectra. These offsets are calculated and applied to each 50 cMpc $h^{-1}$ spectral bin individually in order to avoid correlations between the optical depth measurements.

A detailed description of this correction procedure can be found in Appendix B. We estimate that the systematic error in the mean flux $\langle F_{\alpha} \rangle$ due to corrections of the zero-level offset is

$$\sigma_{F_{\alpha}} \approx 0.0067,$$

i.e., less than $<1\%$, and thus constitutes only a minor correction to the optical depth measurements.

### 4. Results

We compute the effective optical depth $\tau_{\text{eff}}$ from the measurements of the observed mean flux $\langle F_{\text{obs}} \rangle$ in bins of 50 cMpc $h^{-1}$ using Equation (3). We list all measurements of the mean flux within the Ly$\alpha$ forest for each bin along all 23 quasar sightlines in our data sample in Table 5 in Appendix D. All spectral bins indicating the respective measurements of $\langle F_{\text{obs}} \rangle$ and $\tau_{\text{eff}}$ are shown in Figure 13 in Appendix E. If the mean flux is detected with less than $2\sigma$ significance, or if we measure a negative mean flux, we adopt a lower limit on the optical depth at $\tau_{\text{eff}} = -\ln(2\sigma_{F_{\text{obs}}})$ consistent with previous works. Note that we do not include the systematic uncertainty of the mean flux ($\sim 10\%$–$20\%$, see Section 3.1) arising from the PCA continuum estimate. This uncertainty of the mean flux would lead to an additional uncertainty of $\tau_{\text{eff}}^{10\%}$ of $\sim 2\%$–$5\%$, when most of the flux is absorbed in the quasar spectra (see also Figure 7 in Becker et al. 2015). The results of the optical depth measurements within the Ly$\alpha$ forest, plotted as a function of redshift, are shown in Figure 4.

In Figure 4 our new measurements are shown in dark blue in both panels. For the majority of quasars the spectra that we use in our analysis the IGM opacity has been analyzed before. However, we have co-added the data from multiple observation runs (see Section 2.2 for details) in order to achieve higher S/N. However, for one object, ULAS J0148+0600, the data obtained by Becker et al. (2015) with VLT/X-Shooter in a 10 hr observation, has a higher S/N than our spectrum. This sightline exhibits a particularly deep GP trough (Becker et al. 2015), and hence the enhanced S/N results in more stringent opacity limits, representing the strongest fluctuations in the IGM opacity at this redshift. In order to model the IGM fluctuations correctly, it is important to include these outliers (Chardin et al. 2015; D’Aloisio et al. 2015; Davies & Furlanetto 2016). Thus, we construct a master compilation of opacity measurements, presented in the left panel of Figure 4, and replace our optical depth measurements within the Ly$\alpha$ forest along just the sightline of ULAS J0148+0600 with the more precise measurements obtained by Becker et al. (2015) (orange data points). This mainly adds the two most stringent limits at $z = 5.634$ and $z = 5.796$ to our analysis, since the better data quality results in higher sensitivity in the GP troughs. The lower redshift $\tau_{\text{Ly} \alpha}$ measurements for this object are consistent with our measurements. This master compilation is shown in the left panel of Figure 4. Additionally, we also show the lower redshift $\tau_{\text{Ly} \alpha}$ measurements from Becker et al. (2015) (green data points).

We present the average opacity evolution by calculating the mean flux and the bootstrapped error on the mean in bins of $\Delta z = 0.25$ within the Ly$\alpha$ forest from the master compilation set, shown in the left panel of Figure 4. We then compute the binned opacity values via $\tau_{\text{eff}} = -\ln(F)$, where $F$ is the mean flux computed in these bins. The uncertainties on the opacity values with uncertainties also determined via bootstrapping are shown as the red data points and tabulated in Table 3. Similar to the individual $\tau_{\text{eff}}$ measurements, we adopt a limit if the mean flux in the bin is measured with less than $2\sigma$ significance (where $\sigma$ is here the bootstrap errors on the mean flux).

Note that we do not take any systematic errors on the mean flux measurements into account that could, for instance, result from uncertainties in the continuum estimation. The dark gray regions give an estimate of the additional scatter expected due to continuum uncertainties of $\sim 20\%$, which are negligible at high redshift, where the transmitted flux is low and the scatter...
is dominated by fluctuations along different sightlines (Becker et al. 2015; Eilers et al. 2017b).

The right panel of Figure 4 compares our data set to opacity measurements from additional sightlines from the literature that are not in our data sample. The additional data points come from the sightlines of SDSS J0144–0125 and SDSS J1436+5007 (Fan et al. 2006), CFHQS J1509–1749 (Willott et al. 2007), ULAS J1120+0641 (Barnett et al. 2017), PSO J006.1240+39.221 (Tang et al. 2017), and J0332–4701, J0330–4025, J0410–4414, J0454–4448, J0810+5105, J1257+6349, J1609+3041, J1621+5155, and J2310+1855 (Bosman et al. 2018). In most of these analyses the bins were chosen to be \( \Delta z = 0.15 \), following Fan et al. (2006). This bin size covers roughly the same spectral region as the chosen bin size of 50 cMpc \( h^{-1} \) in our analysis and the one by Becker et al. (2015) at \( z \sim 6 \), but in the redshift interval of \( 5 \leq z \lesssim 7 \), the bin size changes quite significantly. Overall the agreement between the various measurements with our new analysis is good, but we chose not to add these measurements to the master compilation, because of the different path lengths used to construct the measurements, very low S/N data, or the variety of different instruments and data reduction pipelines used to obtain the spectra, which enlarges the systematic uncertainties on these measurements (see Section 5).

5. Comparison to Other Studies

For several quasar sightlines in our data sample the optical depth has been measured previously by Fan et al. (2006) and Becker et al. (2015), and more recently by Bosman et al. (2018). However, the quality of the data and the methods to analyze the data differ. Here, we carry out a detailed comparison of our methods and measurements to previous work, and discuss potential systematic uncertainties (Section 5.1) and resulting discrepancies in the cumulative distribution functions (CDFs) of the optical depth (Section 5.2).

5.1. Estimating Systematic Uncertainties

We compare the measurements of the IGM opacity for the 16 quasar sightlines that are both part of our analysis and the data set of Fan et al. (2006) and are not BAL quasars. The spectra from Fan et al. (2006) partially overlap with our data set, but six quasars were observed with a different telescope and instrument (MMT Red Channel, Hobby-Eberly Telescope, and KP 4 m MARS) and eight quasars have additional Keck/ESI data. We have reduced and stacked all of the Keck/ESI observations from

### Table 3

Measurements of the Average Flux and Optical Depth within the Ly\( \alpha \) Forest of Our Master Compilation Sample

| \( z_{\text{obs}} \) | \( \langle F \rangle \) | \( \sigma_F \) | \( \langle \tau_{\text{eff}} \rangle \) | \( \sigma_{\tau_{\text{eff}}} \) |
|---|---|---|---|---|
| 4.0 | 0.4046 | 0.0151 | 0.9049 | 0.0372 |
| 4.25 | 0.3595 | 0.0112 | 1.0230 | 0.0311 |
| 4.5 | 0.2927 | 0.0190 | 1.2286 | 0.0651 |
| 4.75 | 0.1944 | 0.0150 | 1.6381 | 0.0770 |
| 5.0 | 0.1247 | 0.0132 | 2.0818 | 0.1060 |
| 5.25 | 0.0795 | 0.0078 | 2.5321 | 0.0984 |
| 5.5 | 0.0531 | 0.0058 | 2.9357 | 0.1090 |
| 5.75 | 0.0182 | 0.0045 | 4.0057 | 0.2469 |
| 6.0 | 0.0052 | 0.0043 | \( >4.7595 \) | \( \ldots \) |
| 6.25* | −0.0025 | 0.0007 | \( >6.5843 \) | \( \ldots \) |

Note. The columns show the mean redshift \( z_{\text{obs}} \) of the redshift bins of size \( \Delta z = 0.25 \), the averaged flux \( \langle F \rangle \) and its uncertainty \( \sigma_F \) determined via bootstrapping, and the mean optical depth \( \langle \tau_{\text{eff}} \rangle \) in that redshift bin and its error \( \sigma_{\tau_{\text{eff}}} \), also determined via bootstrapping.

*a Note that this redshift bin only contains two measurements.
the archive and thus the spectra analyzed in this paper have an improved quality due to their longer exposure time.

Additionally, our methods to analyze the data differ. For instance, Fan et al. (2006) applied a power law to the red side of the quasar spectra to estimate the quasar continua, whereas we estimated the quasar continua by a PCA (see Section 3.1). In our analysis we mask all spectral regions containing low-ion metal absorption systems, while in previous work it has been argued that those have a negligible influence and can thus be ignored (Fan et al. 2006; Becker et al. 2015).

All these differences contribute to the systematic error of the opacity measurements. We attempted to assess these systematic uncertainties by comparing the results from our analysis to those of Fan et al. (2006) along the sightlines that are part of both data sets. To this end we measure the mean flux in the same redshift bins as Fan et al. (2006) with a fixed bin size of $\Delta z = 0.15$, and compare our measurements to those of Fan et al. (2006) in Figure 5. We observe a large scatter in the distribution and a systematic offset toward lower mean flux values in our measurements, much larger than the formal measurement uncertainties. The offset is strongest at lower redshifts with higher mean flux values, i.e., lower optical depths.

We estimate the systematic error arising due to different observations, different data reduction pipelines, and different analyses by the median of the distribution of measured flux differences $\Delta (F) = \langle F_{\text{Eilers et al. 2018}} \rangle - \langle F_{\text{Fan et al. 2006}} \rangle$. We find a median systematic error of

$$\sigma_{\Delta (F)} \approx -0.023,$$

with a large scatter of $\approx 0.026$ determined from the mean of the 16th and 84th percentiles of the distribution. A detailed investigation of a few of the largest outliers in this distribution suggests that differences in the spectra itself, due to the different instruments with which they were observed and potentially due to differences in the data reduction, cause the largest discrepancies.

Recently, a similar analysis of the Ly$\alpha$ optical depths measured from a comparable quasar sample at $z_{\text{med}} > 5.7$ was presented by Bosman et al. (2018). Of the 62 sightlines they analyzed, 22 satisfy our quality criteria, namely, that they are non-BAL quasars with $S/N > 7$.9 Out of these, 17 objects overlap with our sample. Although the Bosman et al. (2018) sample is comparable to ours in size and partially overlapping, their methods differ in a variety of important aspects from ours. As in Fan et al. (2006) different data reduction pipelines have been used to reduce the data, the quasar continuum estimation methods differ, and while both our study and that of Bosman et al. (2018) mask the proximity zone regions, we adapt the excluded region dependent on the actual measured proximity zone size $R_p$ (see Section 3.2), whereas their analysis excludes a fixed spectral range until $\lambda_{\text{med}} = 1178$ Å, which corresponds to $\Delta R_p = 13.3$ pMpc at $z = 6$. Finally, we have masked strong absorbers and account for small zero-level offsets, whereas they do not.

5.2. Comparison of the CDFs

In Figure 6 we compare the CDF from our measurements shown in blue to the CDF from previous studies by Fan et al. (2006) and Becker et al. (2015), which are shown as the gray curves, and by Bosman et al. (2018) shown in yellow, in different redshift bins centered around $5.0 < z < 6.0$. We show the so-called GOLD sample from Bosman et al. (2018), including 33 quasar spectra for which they applied a data quality cut of $S/N > 11.2$ per 60 km s$^{-1}$ pixel to their sample, which would imply a quality cut on our sample of $S/N > 4.6$ per 10 km s$^{-1}$ pixel.

While previous work noticed an increased scatter in the opacity measurements only at $z \gtrsim 5.5$, we also see evidence for increased scatter at $5.0 < z < 5.5$. We see systematic differences toward higher optical depths in our work compared to others, most strikingly in the $5.3 < z < 5.7$ bins whose excess fluctuations have been the focus of several works. However, in most redshift bins the measurements agree within $1 \sigma$ uncertainties (shown as the shaded regions in Figure 6), which we determined via bootstrap resampling, the only exception being the redshift bin at $z = 5.4$ and $z = 5.6$, where our results are slightly more discrepant with previous studies.

A discrepancy between the Fan et al. (2006) and Becker et al. (2015) measurements in this bin was previously noted by Chardin et al. (2017), and in particular it seems that our (higher) $\tau_{\text{eff}}$ measurements are more consistent with the data from Becker et al. (2015) than those from Fan et al. (2006).

6. Simulations of the IGM

We would like to compare our measurements of the IGM opacity to expectations from simulations. For this purpose we use a hydrodynamical simulation, which we briefly describe in section 6.1. In this simulation we use a uniform UVB radiation field and thus this simulations provides a good approximation for opacity fluctuations in the IGM long after the epoch of reionization, when we expect to have a uniform UVB. We use two more sophisticated models to compare our measurements with more realistic conditions in the post-reionization IGM. To this end, we use two semi-numerical models with fluctuating UVB and temperature field, which we describe in Section 6.2. In Section 6.3 we explain how we compute the Ly$\alpha$ optical depth from the skewers through the simulation box.
6.1. Nyx Hydrodynamical Simulation

In this work we employ a Nyx hydrodynamical simulation (Almgren et al. 2013) 100 cMpc \( h^{-1} \) on a side with 4096\(^3\) dark matter particles and gas elements on a uniform Eulerian grid, designed for precision studies of the Ly\(\alpha\) forest (Lukić et al. 2015). We extracted 1000 random skewers of density, temperature, and velocity along the directions of the grid axes from simulation outputs at \( z = 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, \) and 6.5. We then computed the resulting Ly\(\alpha\) forest spectra in 50 cMpc \( h^{-1} \) bins consistent with the scale of the IGM opacity measurements presented here. For redshift bins in between the simulation outputs, we take the closest output and rescale the density field by \((1+z)^3\), accordingly.

The simulation adopted the uniform UVB model of Faucher-Giguère et al. (2009), resulting in a “vanilla” IGM model, which (uniformly) reionized at early times (\( z_{\text{reion}} > 10 \)). Thus, any deviations of the distribution of IGM opacity observations from those predicted in the simulation likely represent spatial inhomogeneities in the UVB (Davies & Furlanetto 2016) or thermal state related to a more recent epoch of reionization (D’Aloisio et al. 2015).

6.2. Semi-numerical Models with Fluctuating UVB and Temperature Fields

We also compare our observations at \( z = 5.7 \) to the semi-numerical fluctuating UVB and fluctuating IGM temperature models from Davies et al. (2018a), which we describe briefly below.

The Davies et al. (2018a) semi-numerical simulation consists of a cosmological volume, 780 cMpc on a side, with a 2048\(^3\) density field computed via the Zel’dovich approximation (Zel’dovich 1970) and dark matter halos (\( M_{\text{halo}} \geq 2 \times 10^9 \ M_{\odot} \)) populated using the excursion set formalism as in Mesinger & Furlanetto (2007). Ionizing...
luminosities were assigned to halos by first abundance matching to the (non-ionizing) UV luminosity function (Bouwens et al. 2015) and then allowing the ratio of ionizing to non-ionizing luminosities to vary as a free parameter.

UVB fluctuations in this volume were computed on a $156^3$ grid following the method of Davies & Furlanetto (2016), with a spatially varying mean free path of ionizing photons. To construct a fluctuating IGM temperature field, the reionization redshifts of each density cell in the semi-numerical simulation were computed with 21cmFAST (Mesinger et al. 2011) and the subsequent cooling from an initial post-reionization temperature of 30,000 K was computed via numerical integration of the IGM thermal evolution (as in Upton Sanderbeck et al. 2016). Finally, Lyα forest sightlines were then computed using the fluctuating GP approximation (Weinberg et al. 1997) applied to the Zel’dovich approximation density field, with a normalization factor applied to the optical depth of each pixel to account for the approximate nature of the method.

6.3. Calculating the Lyα Optical Depth from Simulated Skewers

We then extract 1000 skewers through the various simulation boxes to compute the optical depths and compare the results to our measurements. Because the exact strength of the UVB radiation $\Gamma_{UVB}$ is unknown, we have to rescale the optical depth in each skewer at each pixel $i$, i.e., $\tau_{Ly, unscaled}^i$, to match the mean optical depth corresponding to the observed mean flux value ($F_{obs}$) of our measurements, which in turn depends on the exact value of $\Gamma_{UVB}$. Hence, at each redshift we determine a scaling factor $A_0$ that solves the following equation:

$$
\langle \exp[-\tau_{Ly}^i]\rangle = \langle \exp[-A_0 \times \tau_{Ly, unscaled}^i]\rangle = \langle F_{obs} \rangle.
$$

We then average the rescaled flux at each pixel $\langle \exp[-\tau_{Ly}^i]\rangle$ within each skewer of size 50 cMpc $h^{-1}$, and determine the 68% and 95% of the distribution of $\tau_{Ly}^i$. This gives an estimate of the expected scatter within the Lyα optical depth measurements given a uniform UVB and IGM thermal state, that is, arising from density fluctuations alone.

7. Implications for the Epoch of Reionization

In order to assess the implications of our opacity measurements for the epoch of reionization, we compare our measurements to the opacity distributions from the Nyx hydrodynamical simulation (Section 6.1). The Nyx simulation was computed with a uniform UVB, which we would expect long after the epoch of reionization or when assuming no signatures of an inhomogeneous reionization process. The evolution of the optical depth distributions from these simulations, with the mean fluxes matched to our measurements, are shown as the gray region in Figure 4. The width of the light (68th percentile) and medium (95th percentile) gray regions corresponds to the scatter in $\tau_{eff}$ expected due to fluctuations in the underlying density field alone. The dark gray regions indicate the additional scatter expected from approximately 20% uncertainties in the quasar continuum estimation. These have been calculated by dividing the mean flux in each spectral bin by $(1 + \Delta C)$, where the continuum uncertainty $\Delta C$ was drawn randomly from a normal distribution with $\sigma = 0.2$ and $\mu = 0$.

corresponding to approximately 20% uncertainties in the continuum estimate. As expected continuum uncertainties matter only very little at higher redshifts, when the mean transmitted flux is low and fluctuations between different quasar sightlines dominate, and the scatter at high redshift can thus not be explained by continuum uncertainties.

The measurements show a steep rise in $\tau_{eff}$ for $z \gtrsim 5$ and an increased scatter in the distribution of measurements. At lower redshifts ($z \lesssim 5$) the scatter in the observed $\tau_{eff}$ decreases rapidly and becomes consistent with the expectations from density fluctuations alone. It is evident, however, that at high redshifts the scatter in the optical depths measurements significantly exceeds the scatter expected from density fluctuations alone, i.e., the scatter represented by our hydrodynamical simulation with uniform UVB. The tiny rare flux spikes observed in the Lyα forests of SDSS J0100+2802 and SDSS J1148+5251 at redshifts $5.8 \lesssim z \lesssim 5.9$ are in strong contrast to the abundant transmitted flux along the sightlines toward SDSS J1306+0356 or SDSS J2054+0005 at similar redshifts, for example. We show the respective spectral regions exhibiting very high (upper panels), average (middle panels), and very low (lower panels) optical depths at similar redshifts of the aforementioned sightlines in Figure 7.

The discrepancy between our measured opacity distribution and the expectation from simulations with a uniform UVB becomes even more evident in Figure 6, where we show the cumulative distributions of $\tau_{Ly}^{unscaled}$ from our master compilation in different redshift bins. The CDFs of $\tau_{eff}$ from our hydrodynamical simulation + uniform UVB are shown as the dashed red curves, while we have rescaled the pixel optical depths (Section 6.3) to match the mean $\tau_{eff}$ in each redshift bin. This model with a uniform UVB is clearly not a good match to the observations. While they are more consistent with the measurements at lower redshift ($z \sim 5$), there are large discrepancies at higher redshifts ($z \gtrsim 5.6$) between the simulated and the observed CDF, a point previously noted by Fan et al. (2006) and Becker et al. (2015). While it may seem that the model provides a better fit in the highest redshift bin at $z \sim 6.0$, this apparent agreement is misleading, and arises due to the fact that we show limits on the optical depth in the same way as measurements, and the bin at $z \sim 6$ contains several limits.

7.1. Comparison to Patchy Reionization Models

Multiple scenarios have been proposed to explain the increased scatter in the optical depth relative to the fluctuations one would expect from the density field of the IGM alone. One possible explanation is that the UVB is strongly fluctuating, either due to coupled variations in the mean free path of ionizing photons (Davies & Furlanetto 2016; D’Aloisio et al. 2018) or a rare source population, such as quasars (Chardin et al. 2015, 2017). Another possibility is that the thermal state of the IGM is highly inhomogeneous (Lidz & Malloy 2014; D’Aloisio et al. 2015). Such fluctuations can arise as a result of an extended and patchy reionization process, where different regions in the universe were reionized (and simultaneously photoheated) at different redshifts $z_{reion}$. The regions that reionized earlier would have had time to cool down, while regions that were reionized at a later time would still be relatively hot.

In Figure 8 we compare our measurements at $5.6 < z < 5.8$ to the semi-numerical models with a fluctuating UVB and fluctuating temperature field (see Section 6.2), where we have rescaled the opacities in the Lyα forest skewers from that work to...
match the mean $\tau_{\text{eff}}$ we have measured in this bin. Note that while previously the CDF of optical depths at this redshift bin containing the strong outliers in the opacity measurements in the GP trough along the sightline of ULAS J0148+0600 (Becker et al. 2015) was very challenging to reproduce in simulations (D’Aloisio et al. 2015; Davies & Furlanetto 2016; Chardin et al. 2017; Davies et al. 2018a; Keating et al. 2018), these outliers are now easier to explain because the mean $\tau_{\text{eff}}$ in our measurements is substantially higher than the Fan et al. (2006)+Becker et al. (2015) compilation, and hence these data points do not represent such strong deviations from the mean of the distribution anymore. This first comparison of our measurements to the two semi-numerical models with a fluctuating UVB and a fluctuating thermal state of the IGM shows that both models can reproduce the observations. A more detailed comparison to these models will be part of future work.

8. Summary

We present a new data set of 34 quasar spectra at $5.77 \leq z_{\text{em}} \leq 6.54$ that we make publicly available via the igmspec database. The spectra have all been observed with ESI on the Keck telescopes, and exposures from different observing runs have been co-added, resulting in a very rich and homogeneous data set, with a total of $\sim$180 hr of telescope time.

For a subsample of 23 quasar spectra, that do not show BAL features and have good quality data (i.e., $S/N > 7$), we measure the IGM opacity by means of the effective optical depth of the Ly$\alpha$ forest in bins of 50 cMpc $h^{-1}$, covering a redshift range of $4.0 \leq z \leq 5.5$. Our results are in qualitative agreement with previous studies (Fan et al. 2006; Becker et al. 2015; Bosman et al. 2018), showing a steep rise in opacity and increased scatter within the measurements at high redshift. However, while previous work noticed an increased scatter at $z \geq 5.5$, we also see evidence for increased scatter at $5.0 < z < 5.5$. A detailed comparison in the optical depth in several redshift bins, shows systematic differences toward higher optical depths in our work compared to others, most strikingly at $5.3 < z < 5.7$. The discrepancies, however, between our measurements and previous work are mostly within the $\sim 1\sigma$ uncertainties, which we determined via bootstrap resampling.

Our work improves upon previous studies in several aspects. We considered possible contamination due to intervening low-ion metal absorption systems such as DLA’s that have previously been ignored and carefully masked all regions in the Ly$\alpha$ forest that are affected by these high H$\text{I}$ column density absorption systems. We also corrected for small offsets in the zero level of the quasar spectra, introduced presumably by improper sky subtraction of a few individual exposures. Finally and most importantly, we considered a very homogeneously reduced data sample that minimizes systematic effects due to the use of different telescopes and detectors,
data reduction pipelines. We present a master compilation sample, including mainly our newly analyzed sample with the exception of the sightline of ULAS J0148+0600 taken from Becker et al. (2015), which has a larger sensitivity in the prominent GP trough due to the higher S/N spectrum observed with VLT/X-Shooter.

We compare our measurements to a large-volume hydrodynamical simulation with a uniform UVB. As noted previously by Fan et al. (2006) and Becker et al. (2015), we find that the spread in observed $\tau_{\text{eff}}$ cannot be explained by fluctuations of the underlying density field alone, and thus our results support an inhomogeneous reionization scenario. Whether temperature fluctuations in the IGM, a fluctuating UVB, or a combination of both can best explain this increased scatter in opacity, remains an open question. A preliminary comparison of our measurements to semi-numerical simulations of UVB and IGM temperature fluctuations shows good agreement for both scenarios.

This work presents a crucial ingredient in constraining the end stages of the epoch of reionization at $5.0 \lesssim z \lesssim 6.0$, when the physical conditions of the post-reionization IGM can be directly measured via absorption spectroscopy of high redshift quasars. The past several years have seen an impressive fivefold increase in the number of $z > 6$ quasars from deep wide-field optical and infrared surveys, which are enabling precise measurements of the Ly$\alpha$ forest absorption at $5 < z < 6.5$ (Becker et al. 2015; Gnedin et al. 2017; Davies et al. 2018a). The requirement that reionization models reproduce these high-precision measurements provides an important low redshift anchor point that all models must reproduce, and can dramatically narrow the family of viable reionization models. Statistical analyses of the Ly$\alpha$ forest, such as measurements of the power spectrum (Oñorbe et al. 2017; D’Aloisio et al. 2018), or the probability distribution function (PDF) of the IGM opacity (Davies et al. 2018b), set further constraints on the reionization process, allowing us to develop accurate models about the early evolutionary stages of our universe.

We would like to thank the anonymous referee for insightful and constructive feedback. We thank José Oñorbe for helpful discussions regarding the hydrodynamical simulations, Zarija Lukic for providing the Nyx simulation, and Vikram Khaire for valuable feedback. Furthermore, we appreciate the help of Xiaohui Fan for making his quasar spectra available to us for a data comparison.

The data presented in this paper were obtained at the W.M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W.M. Keck Foundation.

This research has made use of the Keck Observatory Archive, which is operated by the W.M. Keck Observatory and the NASA Exoplanet Science Institute, under contract with the National Aeronautics and Space Administration.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We are most fortunate to have had the opportunity to conduct observations from this mountain.

Software: ESIRedux (http://www2.keck.hawaii.edu/inst/esi/ESIRedux/), XIDL (http://www.ucolick.org/xavier/IDL/), igmspec (http://specdb.readthedocs.io/en/latest/igmspec.html), Nyx (Almgren et al. 2013), emcee (Foreman-Mackey et al. 2013), numpy (van der Walt et al. 2011), scipy (Jones et al. 2001), matplotlib (Hunter 2007), astropy (The Astropy Collaboration et al. 2018), and seaborn (Waskom et al. 2017).

Appendix A
Catalog of the Data Release

In Table 4 we present the catalog with the properties of our data set that will be available together with the final co-added spectra, their noise vectors, and their continuum estimates via the igmspec database. Igmspec is a database of publicly

---

[^10]: http://specdb.readthedocs.io/en/latest/igmspec.html
available ultraviolet, optical, and near-infrared spectra that probe the IGM. It provides ~500,000 spectra from various data sets, including the Sloan Digital Sky Survey (SDSS), 2dF QSO Redshift Survey (2QZ), and data from the HST, the Keck Telescopes, the Very Large Telescope, and more. The database is part of the specdb repository\(^\text{11}\) within the specdb Github organization\(^\text{12}\), which provides software developed in Python for accessing and interacting with the quasar spectra.

### Appendix B

**Details of the Correction of Zero-level Offsets in the Quasar Spectra**

In order to correct for possible offsets in the zero level of the quasar spectra, we examine the negative pixels in each spectral bin, which should appear to be a truncated Gaussian distribution, providing an estimate of the noise level in the spectra. To this end, we take all flux pixels \(F\) below zero in each spectral bin of 50 cMpc \(h^{-1}\), and calculate the CDF, ignoring correlations between neighboring pixels.

In the case of no offset in the zero level within a spectral bin, the estimated mean of the CDF \(\mu_{CDF}\) should be equal to zero, as shown by the toy model example in the lower middle panel of Figure 9. The upper panels show the PDF of the same

\[\text{Table 4} \quad \text{Catalog for the igmspec Database}\]

| Object | RA (degree) | Decl. (degree) | \(z_{em}\) | \(M_{1450}\) | S/N | \(R(\Delta\lambda/\lambda)\) | Telescope | Instrument |
|--------|-------------|----------------|------------|-------------|-----|-------------------------|-----------|------------|
| SDSS J0002+2550 | 0.6641 | 25.843 | 5.82 | −27.31 | 57 | 5400 | Keck II | ESI |
| SDSS J0005−0006 | 1.4681 | −0.1155 | 5.844 | −25.73 | 13 | 4000 | Keck II | ESI |
| CFHQS J0050+3445 | 13.7621 | 34.756 | 6.253 | −26.7 | 15 | 4000 | Keck II | ESI |
| SDSS J0100+2802 | 15.0542 | 28.0405 | 6.3258 | −29.14 | 35 | 4000 | Keck II | ESI |
| ULAS J0418+0600 | 27.1568 | 0.6056 | 5.98 | −27.39 | 20 | 4000 | Keck II | ESI |
| ULAS J0203+0012 | 30.8849 | 0.2081 | 5.72 | −26.26 | 5 | 4000 | Keck II | ESI |
| CFHQS J0210−0456 | 32.555 | −4.9391 | 6.4323 | −24.53 | 1 | 5400 | Keck II | ESI |
| PSO J0226−0302 | 36.5078 | 3.0498 | 6.5412 | −27.33 | 9 | 4000 | Keck II | ESI |
| CFHQS J0227−0605 | 36.9304 | −6.0917 | 6.2 | −25.28 | 3 | 4000 | Keck II | ESI |
| SDSS J0303−0019 | 45.8808 | −0.3202 | 6.078 | −25.56 | 2 | 4000 | Keck II | ESI |
| SDSS J0353+0104 | 58.4572 | 1.068 | 6.072 | −26.43 | 10 | 4000 | Keck II | ESI |
| PSO J0402+2451 | 60.5529 | 24.8658 | 6.18 | −26.95 | 11 | 4000 | Keck II | ESI |
| SDSS J0818+1722 | 122.6142 | 17.3811 | 6.02 | −27.52 | 5 | 4000 | Keck II | ESI |
| SDSS J0836+0054 | 120.1827 | 0.9148 | 5.81 | −27.75 | 108 | 4000 | Keck II | ESI |
| SDSS J0840+5624 | 130.1471 | 56.4056 | 5.8441 | −27.24 | 23 | 4000 | Keck II | ESI |
| SDSS J0842+1218 | 130.6226 | 12.314 | 6.069 | −26.91 | 8 | 4000 | Keck II | ESI |
| SDSS J0927+2001 | 141.8409 | 20.0232 | 5.7722 | −26.76 | 6 | 4000 | Keck II | ESI |
| SDSS J1030+0524 | 157.613 | 5.4153 | 6.309 | −26.99 | 20 | 4000 | Keck II | ESI |
| SDSS J1048+4637 | 162.1878 | 46.6218 | 6.2284 | −27.24 | 42 | 4000 | Keck II | ESI |
| SDSS J1137+3549 | 174.3239 | 35.8325 | 6.03 | −27.36 | 18 | 4000 | Keck II | ESI |
| SDSS J1148+5251 | 177.0694 | 52.864 | 6.4189 | −27.62 | 28 | 4000 | Keck II | ESI |
| SDSS J1250+3310 | 192.7164 | 31.5061 | 6.15 | −26.53 | 7 | 4000 | Keck II | ESI |
| SDSS J1306+0356 | 196.5345 | 3.9907 | 6.016 | −26.81 | 28 | 4000 | Keck II | ESI |
| ULAS J1319+0950 | 199.7971 | 9.8476 | 6.133 | −27.05 | 8 | 4000 | Keck II | ESI |
| SDSS J1335+3533 | 203.9617 | 35.5544 | 5.9012 | −26.67 | 5 | 4000 | Keck II | ESI |
| SDSS J1411+1217 | 212.7937 | 12.2807 | 5.904 | −26.69 | 25 | 4000 | Keck II | ESI |
| SDSS J1602+4228 | 240.7249 | 42.4736 | 6.09 | −26.94 | 16 | 4000 | Keck II | ESI |
| SDSS J1623+3112 | 245.8825 | 31.2001 | 6.2572 | −26.55 | 9 | 4000 | Keck II | ESI |
| SDSS J1630+4012 | 247.6412 | 40.2027 | 6.065 | −26.19 | 11 | 4000 | Keck II | ESI |
| CFHQS J1641+3755 | 250.3405 | 37.9223 | 6.047 | −25.67 | 2 | 4000 | Keck II | ESI |
| SDSS J2054−0005 | 313.527 | −0.0874 | 6.0391 | −26.21 | 12 | 4000 | Keck II | ESI |
| CFHQS J2229−1457 | 337.2569 | 14.9525 | 6.1517 | −24.78 | 2 | 4000 | Keck II | ESI |
| SDSS J2315−0023 | 348.944 | −0.3995 | 6.117 | −25.66 | 12 | 4000 | Keck II | ESI |
| CFHQS J2329−0301 | 352.2845 | −3.033 | 6.417 | −25.25 | 2 | 4000 | Keck II | ESI |

Note. The columns show the object name, the coordinates RA, and decl. of the quasar given in degrees, the emission redshift and the quasar’s magnitude \(M_{1450}\), the S/N of the data, the telescope and instrument with which the spectra are observed, and the spectral resolution of the data. Note that in some cases we co-added data with different spectral resolution.

---

\(^{11}\) http://specdb.readthedocs.io/en/latest/

\(^{12}\) https://github.com/specdb/specdb
will likewise be below or above zero. It is clear that we can obtain a handle on these systematic offsets by examining the purely negative pixels and fitting a truncated CDF model.

We use a Markov chain Monte Carlo (MCMC) approach making use of the implementation of the affine-invariant ensemble sampler emcee\(^{13}\) (Foreman-Mackey et al. 2013) to estimate the mean \(\mu_{\text{CDF}}\) of the best fitting model CDF

\[
  \text{CDF}_{\text{model}}(F) = \frac{A_{\text{CDF}}}{2} \left[ 1 + \text{erf} \left( \frac{F - \mu_{\text{CDF}}}{\sqrt{2} \sigma_{\text{CDF}}} \right) \right],
\]

while marginalizing over the width of the distribution \(\sigma_{\text{CDF}}\) and the scaling factor \(A_{\text{CDF}}\). The likelihood function just maximizes the least squares between the CDF model and the measured CDF, i.e.,

\[
  \ln \mathcal{L} = -0.5 (\text{CDF}_{\text{model}} - \text{CDF}_{\text{data}})^2.
\]

We then take the median of the resulting posterior PDF as the best estimate for \(\mu_{\text{CDF}}\).

The three free parameters of the CDF\(_{\text{model}}\), \(\mu_{\text{CDF}}, \sigma_{\text{CDF}},\) and \(A_{\text{CDF}}\) are highly degenerate with each other because we are fitting only a small part of the CDF when taking solely flux pixels below zero into account, i.e., \(F < 0\). Thus, we have to apply strict priors to break this degeneracy. The priors we chose are flat priors within the intervals \(\mu_{\text{CDF}} \in [-0.05, 0.05]\), since we expect the total offset of the zero level to be less than \(\pm 5\%\), \(\sigma_{\text{CDF}} \in [0.75 \sigma_{\text{eff}}, 1.25 \sigma_{\text{eff}}]\), which takes into account the noise vector at each pixel \(i\) of each quasar spectrum to estimate \(\sigma_{\text{eff}} = \left( \sum \sigma_i^2 / N \right)^{1/2}\), and \(A_{\text{CDF}} \in [0.45 N_{F < 0}, 0.55 N_{F < 0}]\). The upper and lower boundaries for \(A_{\text{CDF}}\) result from the fact that the number of pixels with flux below zero, i.e., \(N_{F < 0}\), in an unbiased case should be exactly half of the pixels, i.e.,

0.5 \(N_{F < 0}\). In the presence of the possible offsets in the zero level of the spectra we allow \(A_{\text{CDF}}\) to deviate from the unbiased case.

Figure 10 shows two examples of the procedure. Both panels show the CDF for two spectral bins along the sightline of SDSS J0840+5624. We show the actually measured CDF of all negative flux pixels (blue curves) and overplot the best fitted CDF (red dashed curves, with mean \(\mu_{\text{CDF}}\) indicated by the red dashed-dotted lines) and the ideal CDF with no zero-level offset, i.e., \(\mu_{\text{CDF}} = 0\), and the variance given by the noise of the data, i.e., \(\sigma_{\text{CDF}} = \sigma_{\text{eff}}\). We can see that we obtain small zero-level offsets of about \(\Delta (F^{16}) \sim \mu_{\text{CDF}} \approx 0.3\%–0.4\%\).

In the end, we offset all pixels \(i\) within each spectral bin by the respective best estimate for \(\mu_{\text{CDF}}\), i.e., \(F_i^{\text{new}} = F_i - \mu_{\text{CDF}}\), and calculate the mean flux and the opacity from the offsetted pixels \(F_i^{\text{new}}\).

Figure 11 shows the difference in mean flux estimates due to corrections in the zero level. We estimate this systematic uncertainty in the mean flux measurements from calculating the 84th percentile and 16th percentile of this distribution and taking their average, which results in

\[
  \sigma_{\text{flux}} = 0.0067.
\]

**Appendix C**

**All Quasar Continuum Estimates**

In Figure 12 we show the estimates of the quasar continua for all objects, which we use for measurements of the IGM opacity that are not already shown in Figure 2.

---

\(^{13}\) http://dfm.io/emcee/current/
Figure 12. Same as Figure 3, but for the remaining 20 quasar spectra in our data set used for the IGM opacity measurements.
Appendix D
Mean Flux Measurements within the Lya Forest

Our measurements of the mean flux in the Lya forest are shown in Table 5 for spectral bins of size 50 cMpc h⁻¹ along all 23 quasar sightlines in our data sample.

Table 5
Mean Flux Measurements in the Lya Forest

| Object          | zstart | zabs | zend | \langle F\rangle | \sigma_f | \langle z\rangle |
|-----------------|--------|------|------|------------------|---------|-----------------|
| SDSS J002+2550  | 5.54   | 5.62 | 5.70 | 0.0472           | 0.0004  | 5.61            |
| SDSS J005−0006  | 5.39   | 5.47 | 5.54 | 0.0315           | 0.0005  | 5.45            |
| SDSS J1000+2802 | 5.60   | 5.68 | 5.76 | 0.0507           | 0.0016  | 5.68            |
| CFIQS J0550+3445| 5.45   | 5.53 | 5.60 | 0.0384           | 0.0028  | 5.56            |
| ULAS J0148+0600 | 5.80   | 5.88 | 5.96 | 0.0217           | 0.0031  | 5.94            |
| ULAS J1137+3549 | 5.64   | 5.72 | 5.80 | 0.0269           | 0.0022  | 5.83            |
| SDSS J1148+5251 | 5.49   | 5.56 | 5.64 | 0.0570           | 0.0020  | 5.65            |
| SDSS J1250+3130 | 5.34   | 5.41 | 5.49 | 0.1336           | 0.0200  | 5.51            |
| SDSS J1306+0356 | 5.19   | 5.26 | 5.34 | 0.1009           | 0.0014  | 5.33            |
| SDSS J1411+1217 | 5.81   | 5.89 | 5.98 | 0.0013           | 0.0003  | 5.86            |
| SDSS J1600+4228 | 5.65   | 5.73 | 5.81 | 0.0038           | 0.0003  | 5.70            |
| SDSS J1630+0412 | 5.50   | 5.57 | 5.65 | 0.0046           | 0.0003  | 5.55            |
| PSO J036−03     | 0.60   | 0.65 | 0.72 | 0.0018           | 0.0011  | 0.72            |
| PSO J060+25     | 0.57   | 0.62 | 0.68 | 0.0018           | 0.0009  | 0.65            |
| SDSS J0818+1722 | 0.57   | 0.63 | 0.73 | 0.0006           | 0.0006  | 0.70            |
| SDSS J0836+0054 | 0.52   | 0.59 | 0.65 | 0.0032           | 0.0028  | 0.63            |
| SDSS J0840+5624 | 0.49   | 0.56 | 0.63 | 0.0078           | 0.0041  | 0.67            |
| SDSS J0842+1218 | 0.54   | 0.62 | 0.70 | 0.1061           | 0.0003  | 0.71            |
| SDSS J1030+0524 | 0.53   | 0.62 | 0.69 | 0.0097           | 0.0004  | 0.70            |

Note. The different columns show the name of the object, the beginning of each redshift bin \( z_{\text{start}} \), the mean redshift of each bin \( \langle z \rangle \), and the end of each bin \( z_{\text{end}} \). The mean flux \( \langle F \rangle \) and the standard deviation \( \sigma_f \) are also given.

* Note that the measurements along this quasar sightline have been replaced by the ones from Becker et al. (2015) in our master compilation.
Figure 13. Spectral bins of 50 cMpc h⁻¹ along all quasar sightlines in our data sample, for which we measure $\tau_{\text{Ly}\alpha}^{\text{eff}}$ within the Ly$\alpha$ forest. The red data points show the measurements of $F_{\text{obs}}$ and the corresponding optical depth measurements are shown in the legend.
Appendix E
Spectral Bins along All Quasar Sightlines

In Figure 13 we show all spectral bins of 50 cMpc $h^{-1}$ along all 23 quasar sightlines in our data sample, indicating the respective optical depth measurements.

Figure 13. (Continued.)
References

Almgren, A. S., Bell, J. B., Lijewski, M. J., Lukić, Z., & Van Andel, E. 2013, ApJ, 765, 39
Barnett, R., Warren, S. J., Becker, G. D., et al. 2017, A&A, 601, A16
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015, MNRAS, 447, 3402
Becker, G. D., Sargent, W. L. W., Rauch, M., & Calverley, A. P. 2011, ApJ, 735, 93
Bernstein, R. M., Burles, S. M., & Prochaska, J. X. 2015, PASP, 127, 911
Boscaini, D., Menci, N., & Tozzi, P. 2013, ApJ, 767, 40
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, MNRAS, 412, 2543
Chardin, J., Haehnelt, M. G., Aubert, D., & Puchwein, E. 2015, MNRAS, 453, 2943
Chardin, J., Puchwein, E., & Haehnelt, M. G. 2017, MNRAS, 465, 3429
D’Aloisio, A., McQuinn, M., Davies, F. B., & Furlanetto, S. R. 2018, MNRAS, 473, 560
D’Aloisio, A., McQuinn, M., & Trac, H. 2015, ApJL, 813, L38
Davies, F. B., Becker, G. D., & Furlanetto, S. R. 2018a, ApJ, 860, 155
Davies, F. B., & Furlanetto, S. R. 2016, MNRAS, 460, 1328
Davies, F. B., Hennawi, J. F., Eilers, A.-C., & Lukić, Z. 2018b, ApJ, 855, 106
Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017a, ApJ, 840, 24
Eilers, A.-C., Hennawi, J. F., & Lee, K.-G. 2017b, ApJ, 844, 136
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gnedin, N. Y., Becker, G. D., & Fan, X. 2017, ApJ, 841, 26
Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633
Hunter, J. D. 2007, CSE, 9, 90
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python, http://www.scipy.org/
Keating, L. C., Puchwein, E., & Haehnelt, M. G. 2018, MNRAS, 477, 5501
Kelson, D. D. 2003, PASP, 115, 688
Lidz, A., & Malloy, M. 2014, ApJ, 788, 175
Lidz, A., Oh, S. P., & Furlanetto, S. R. 2006, ApJL, 639, L47
Lukić, Z., Stark, C. W., Nugent, P., et al. 2015, MNRAS, 446, 3697
McGreer, I. D., Mesinger, A., & Fan, X. 2011, MNRAS, 415, 3237
Mesinger, A., & Furlanetto, S. 2007, ApJ, 669, 663
Mesinger, A., Furlanetto, S., & Cen, R. 2011, MNRAS, 411, 955
Oñorbe, J., Hennawi, J. F., Lukić, Z., & Walther, M. 2017, ApJ, 847, 63
Páris, I., Petitjean, P., Rollinde, E., et al. 2017, A&A, 530, A50
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Prochaska, J. X. 2017, A&C, 19, 27
Ryan-Weber, E. V., Pettini, M., Madau, P., & Zych, B. J. 2009, MNRAS, 395, 1476
Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, PASP, 114, 851
Shull, J. M., Stevans, M., & Danforth, C. W. 2012, ApJ, 752, 162
Tang, J.-J., Goto, T., Ohyama, Y., et al. 2017, MNRAS, 466, 4568
The Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, arXiv:1801.02634
Upton Sanderbeck, P. R., D’Aloisio, A., & McQuinn, M. J. 2016, MNRAS, 460, 1885
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Waskom, M., Botvinnik, O., O’Kane, D., et al. 2017, mworksom/seaborn: v0.8.1 (September 2017), Zenodo, doi:10.5281/zenodo.883859
Weinberg, D. H., Hernquist, L., Katz, N., Croft, R., & Miralda-Escudé, J. 1997, in Structure and Evolution of the Intergalactic Medium from QSO Absorption Line System, ed. P. Petitjean & S. Charlot (Paris: Nouvelles Frontières), 133
Willott, C. J., Delorme, P., Omont, A., et al. 2007, AJ, 134, 2435
Wyithe, J. S. B., & Bolton, J. S. 2011, MNRAS, 412, 1926
Zel’dovich, Y. B. 1970, A&A, 5, 84