

# Appendix: A consistent and robust measurement of the thermal state of the IGM at $2 \leq z \leq 4$ from a large sample of Ly $\alpha$ forest spectra: Evidence for late and rapid HeII reionization

Prakash Gaikwad<sup>1,2\*</sup>, Raghunathan Srianand<sup>3</sup>, Martin G. Haehnelt<sup>1,2</sup> and Tirthankar Roy Choudhury<sup>4</sup>

<sup>1</sup>*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*

<sup>2</sup>*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge, CB3 0HA, UK*

<sup>3</sup>*Inter-University Centre for Astronomy and Astrophysics (IUCAA), Post Bag 4, Pune 411007, India*

<sup>4</sup>*National Centre for Radio Astrophysics, Tata Institute of Fundamental Research, Pune 411007, India*

## APPENDIX A: COADDITION OF OBSERVED SPECTRA

Most of the QSOs in the KODIAQ DR2 sample were observed more than once with different exposure times. In this section, we discuss our procedure to coadd the spectra. Let a QSO be observed  $N$  times with exposure times  $t_1, t_2, \dots, t_N$ . Let  $F_i(\lambda)$ ,  $\sigma_i(\lambda)$  be the value of the normalized flux and associated error at wavelength  $\lambda$  in the  $i^{\text{th}}$  observation. The coadded flux  $F_{\text{coadd}}(\lambda)$  and coadded variance  $\sigma_{\text{coadd}}^2$  of the flux is given by,

$$F_{\text{coadd}}(\lambda) = \frac{\sum_{i=1}^N t_i w_i F_i(\lambda)}{\sum_{i=1}^N t_i w_i} \quad (\text{A1})$$

$$\sigma_{\text{coadd}}^2(\lambda) = \frac{\sum_{i=1}^N t_i w_i \sigma_i^2(\lambda)}{\sum_{i=1}^N t_i w_i}$$

where  $w_i$  is weight for the  $i^{\text{th}}$  pixel.  $w_i = 0$  for bad pixels and  $w_i = 1$  for good pixels. For the KODIAQ DR2 sample, we define a pixel as a bad pixel if it satisfies any of the following criteria, (i) pixels with negative error i.e.,  $\sigma_i(\lambda) \leq 0$ , (ii) pixels with  $\sigma_i(\lambda) \gg 1$  (i.e., SNR per pixel  $\ll 1$ ) (iii)  $F_i(\lambda) = 0$  (due to sky background  $F_i(\lambda) \neq 0$  for good pixel), (iv) negative outliers i.e., pixels with  $F_i(\lambda) \leq -3\sigma_i(\lambda)$  and (v) positive outliers i.e., pixels with  $F_i(\lambda) \geq 1 + 3\sigma_i(\lambda)$ ,

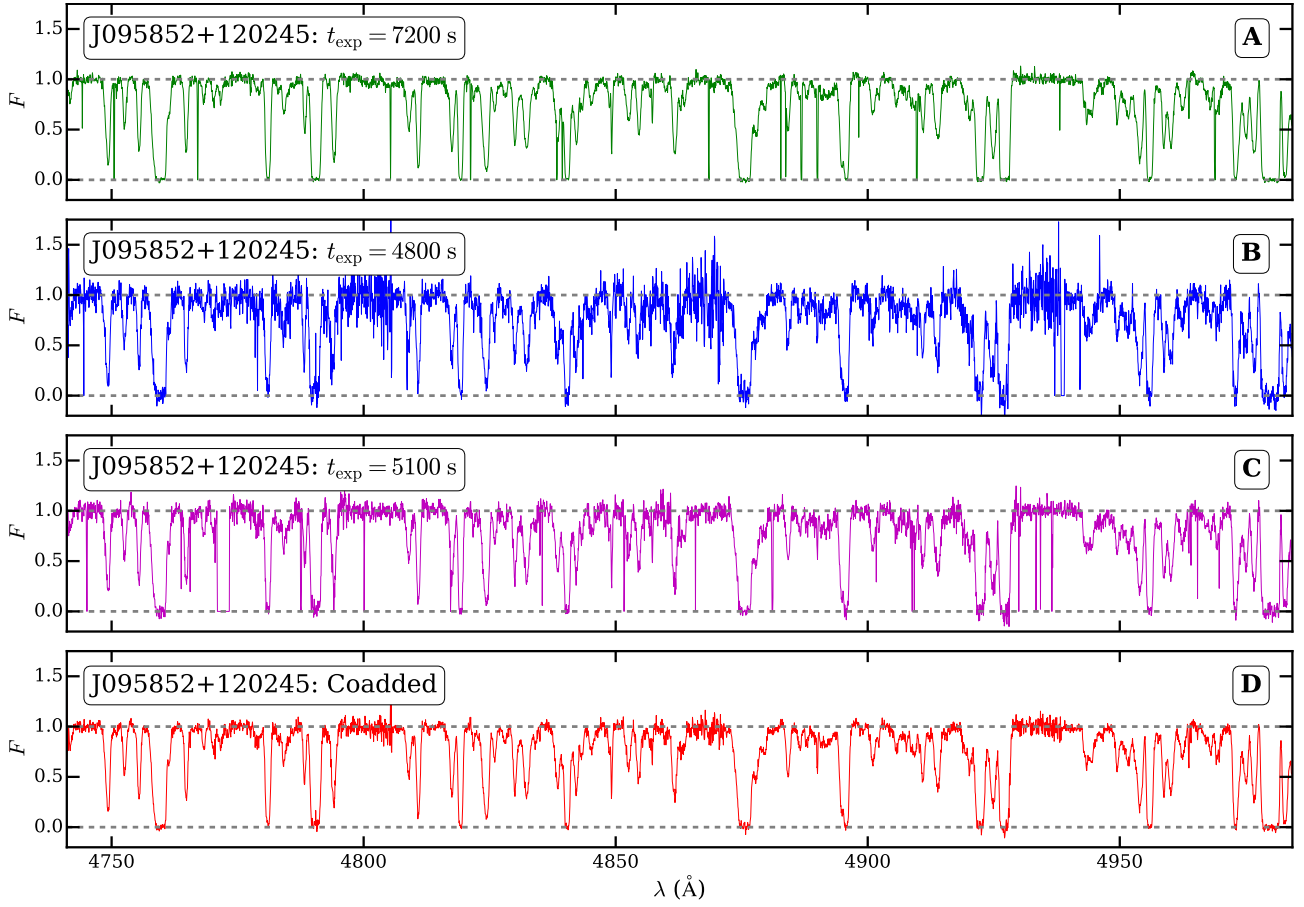
Fig. A1 shows an example of coadded spectrum obtained using our method. Spectrum obtained in individual exposures show many bad pixels due to observational systematics. Our coaddition method accounts for these effects as shown in panel D of Fig. A1.

In our analysis, we also exclude sightlines that contain large spectral gaps. We decide whether a sightline contains a large spectral gap or not in the following way. For a given emission redshift of QSO, we identify redshift bins that are spanned by QSO absorption spectra excluding proximity regions and Ly $\beta$ , O VI emission lines. The choice of QSO

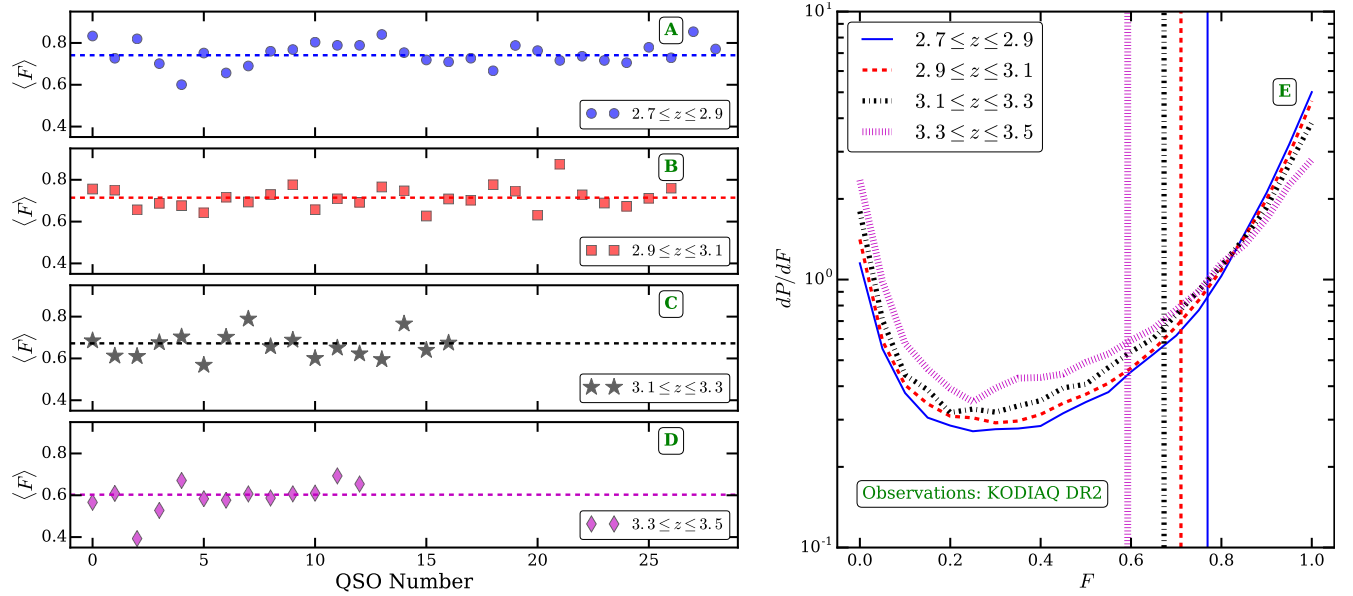
proximity region size corresponds to  $\sim 1166 \text{ \AA}$  at  $z = 3$ . To avoid possible contamination from intrinsic O IV absorption we consider only regions with rest wavelength greater than  $1050 \text{ \AA}$  at the quasar emission redshift. If the spectral gap is more than 25 percent of the redshift bin we mark it as a large gap and exclude the sightline from the analysis for that redshift bin. We note that the number of such sightlines are small ( $\leq 5$ ).

## APPENDIX B: EXCESS MEAN FLUX $\langle F \rangle$ AT $Z = 3.2$

In this section we test if the observed excess in mean flux at  $3.1 \leq z \leq 3.3$  in the KODIAQ DR2 sample is due to the presence of any unusual QSO sightlines. To check if there are any outliers in this particular redshift bin, we plot the mean flux along individual sightlines in Fig. B1 for four redshift bins. Fig. B1 illustrates that the distribution of mean flux along sightlines is random around the mean flux (dashed line in panel A to D) of the entire sample for all redshift bins. We also do not see any outlier at  $3.1 \leq z \leq 3.3$ . Since the mean flux along a sightline is an average quantity, we also check if there is any significant deviation in the observed FPDF evolution in panel E of Fig. B1. The observed FPDF at  $3.1 \leq z \leq 3.3$  also does not show any unusual deviation from other redshift bins. The excess in  $\langle F \rangle$  at  $3.1 \leq z \leq 3.3$  may thus be real in the KODIAQ DR2 sample. Another reason for the excess  $\langle F \rangle$  could be that the number of sightlines at  $3.1 \leq z \leq 3.3$  are limited in our sample. Our preliminary analysis using UVES SQUAD DR1 (Murphy et al. 2019) while showing distinct change in the slope at  $z \sim 3.2$  does not show any prominent excess. We plan to investigate this further in our future work. However, we emphasize that the effect of excess in  $\langle F \rangle$  at  $3.1 \leq z \leq 3.3$  has a marginal effect on our measurements of thermal parameters. This is because for any thermal parameter variation we rescale the optical depth in our simulations to match the mean flux. Thus, even if there are any systematics in the  $\langle F \rangle$  evolu-



**Figure A1.** Panel A, B and C show the observed spectrum of QSO J095852+120245 at three different epochs having exposure times of 7200 s, 4800 s and 5100 s, respectively. Many pixels in the individual exposures are bad pixels due to observational systematics. Panel D shows the coadded spectrum obtained by our method as discussed in online supplementary appendix A. Our coaddition procedure takes care of bad pixels due to observational systematics.



**Figure B1.** Panels A to D show the observed mean flux along individual QSO sightlines in different redshift bins. The dashed line in each panel shows the mean flux of the sample at the given redshift. We exclude the possibility of larger mean flux at  $3.1 \leq z \leq 3.2$  due to any outliers. Panel E shows the observed FPDF in four redshift bins. The vertical lines in panel E show the observed mean flux in the corresponding redshift bins. Similar to panel A-D, the FPDF comparison in panel E does not show any unusual evolution.

tion in the KODIAQ DR2 sample, this does not significantly change the main results of this work.

## APPENDIX C: COMPARISON OF OBSERVED FLUX STATISTICS ERRORS

In §2.1 we discuss the consistency of observed flux statistics from this work with that in the literature. In this section, we compare the errors of the observed FPDF, FPS and CDDF from this work with measurements in the literature. The errors for observed statistics are normally determined using the bootstrap method (except in Rollinde et al. 2013). For a fair comparison, we also compute the errors using the bootstrap method. Fig. C1 shows a comparison of the errors for different statistics. In general, the errors in different statistics from this work are either consistent or slightly smaller than those in the literature.

Panel A1 in Fig. C1 shows that the error in our observed FPDF is smaller than that in Kim et al. (2007); Calura et al. (2012); Rollinde et al. (2013). The smaller errors are due to the large number of QSOs in our sample. The large error in the FPDF of Rollinde et al. (2013) is due to the difference in their method of calculating the errors. Rollinde et al. (2013) argue that the errors computed using the bootstrap method are underestimated and may not account for the cosmic variance accurately. Hence they derive the errors from large number of simulated mocks. Furthermore, their sample size was limited to  $< 5$  at the redshift of interest. We have calculated the errors on our other observed statistics using both methods and accounted for cosmic variance (see §4.2 for details).

Panel B1 in Fig. C1 shows a comparison of the error of the FPS from this work with that from Walther et al. (2018). The errors (for the no metal case) are similar in the two works. This is expected because the number of QSOs per redshift bin are similar in both cases. Similarly, the data quality is the same because the subsets of QSOs are both drawn from the KODIAQ DR2 survey. The consistency in FPS and its error of this work with that from Walther et al. (2018) reflects the consistency of both the data and our method of calculating the FPS. This consistency is important because the FPS is one of the statistics used in this work to measure thermal parameters.

In Fig. C2, we compare the observed FPS measurements from this work with those from Croft et al. (2002); Kim et al. (2004); Walther et al. (2018); Day et al. (2019). We find that our FPS measurements are generally in good agreement (maximum deviation is  $\sim 1.8\sigma$ ) with those from the literature at  $k \leq 0.1 \text{ s km}^{-1}$ . The large differences at  $k > 0.1 \text{ s km}^{-1}$  are due to differences in the method of computing the FPS. In particular Kim et al. (2004); Walther et al. (2018) subtract power due to white noise in the spectra, whereas we do not subtract the noise power. Furthermore the number of observed spectra, quality of spectra (S/N) and total observed redshift path length is quite different in Croft et al. (2002); Kim et al. (2004) from that in our analysis. Even with such differences the FPS in these work seem to be in good agreement with our measurements. We also like to emphasize that any systematic differences in the method of computing the FPS does not significantly affect the measurements of thermal parameters. This is because

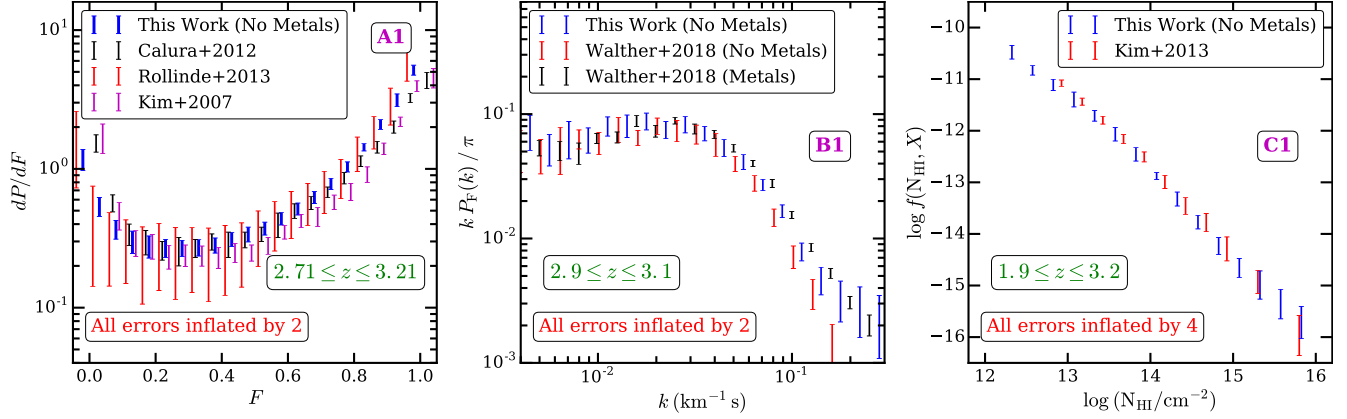
we use the same method to compute the FPS statistics in simulations and observations.

In panel C1 of Fig. C1, we compare the errors of our CDDF with that from Kim et al. (2013). Kim et al. (2013) measured the CDDF from 18 high resolution high S/N (usually  $> 50$ ) Ly $\alpha$  forest spectra. On the other hand we measure the CDDF from a sample of  $\sim 35$  high resolution but moderate S/N spectra. Thus the data size and data quality is different in both the works. These differences of the observed dataset are reflected in the CDDF errors. For example the errors in our CDDF at  $\log N_{\text{HI}} < 13.6$  is larger than that from Kim et al. (2013). This is because low  $N_{\text{HI}}$  systems are affected more by the S/N of the spectra. Since S/N of our sample is smaller than in Kim et al. (2013), our CDDF error at small  $\log N_{\text{HI}}$  is correspondingly large. At high  $\log N_{\text{HI}} > 13.6$ , S/N is no longer an issue since lines are detected with relatively high significance level. The errors in our CDDF at  $\log N_{\text{HI}} > 13.6$  are smaller than that from Kim et al. (2013) because the number of spectra in our sample are nearly twice that in Kim et al. (2013). Even though there are differences in data quality and sample size, the CDDF and the associated errors from the two works are in good agreement with each other. The agreement between the CDDF demonstrates the consistency of VIPER in fitting the Ly $\alpha$  forest with Voigt profiles. This consistency is important because we use the  $b$  parameter of Voigt profile fits to constrain thermal parameters. In summary, the observed statistics with associated errors from this work are in good agreement with that from the literature.

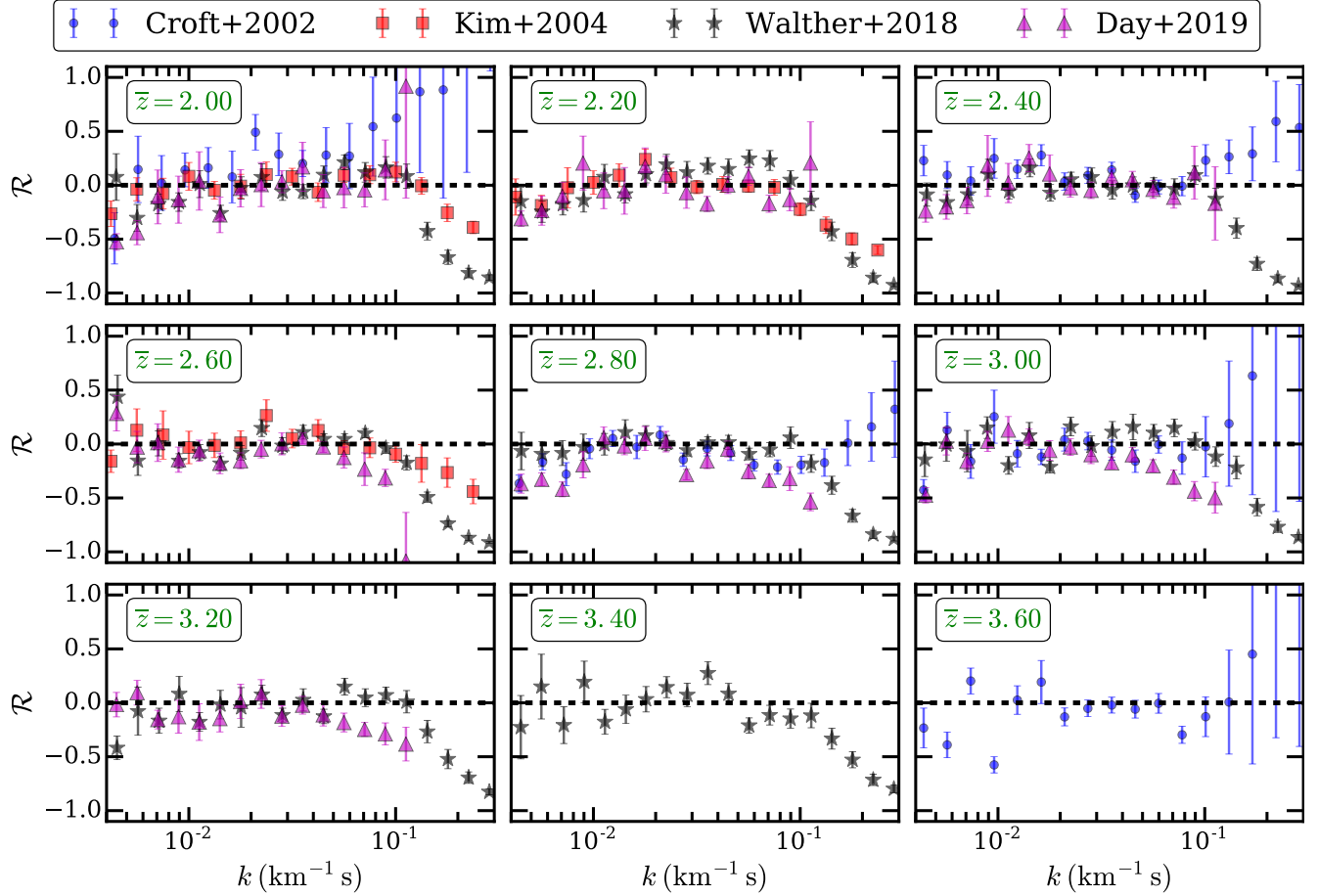
## APPENDIX D: VARIATION IN THERMAL PARAMETERS

In this section, we explain our method of generating different thermal histories for a given UVB model. We vary the thermal parameter evolution by modifying the H I, He I and He II photo-heating rates of the KS19 UVB. Following the approach of Becker et al. (2011), we scale the photo-heating rates ( $\epsilon_i$ ) by two free parameters  $a, b$  such that  $\epsilon_i = a \Delta^b \epsilon_i^{\text{KS19}}$  where  $i \equiv [\text{H I}, \text{He I}, \text{He II}]$ . A change in factor  $a$  ( $b$ ) leads to variation in  $T_0$  ( $\gamma$ ) while keeping the evolution of  $\gamma$  ( $T_0$ ) relatively unchanged. It is important to note that the variation in thermal parameters obtained by scaling the photo-heating rate is not physical. Ideally, one would like to vary the properties of the ionizing sources such as, luminosity function, QSO SED index, redshift of reionization etc., and generate the UVB model from a cosmological 1D radiative transfer code. However, such an approach cannot produce a large variation in  $T_0$  and  $\gamma$  (but see Oñorbe et al. 2019). Our approach, even though not physical, is useful in practice to probe a large range of the parameters  $T_0$  and  $\gamma$  on a finely sampled grid. Fig. D1 shows examples of the evolution in  $T_0$  and  $\gamma$  obtained by varying  $a$  and  $b$  for several thermal histories. The case  $a = 1$  and  $b = 0$  shows the  $T_0$  and  $\gamma$  evolution obtained using the default photo-heating rates from the KS19 UVB. Note that the shock heated gas ( $T > 10^5 \text{ K}$ ) is mostly unaffected by the photo-heating rate scaling, while the low density gas responds to this scaling differently and changes normalization ( $T_0$ ) and slope ( $\gamma$ ) of the TDR.

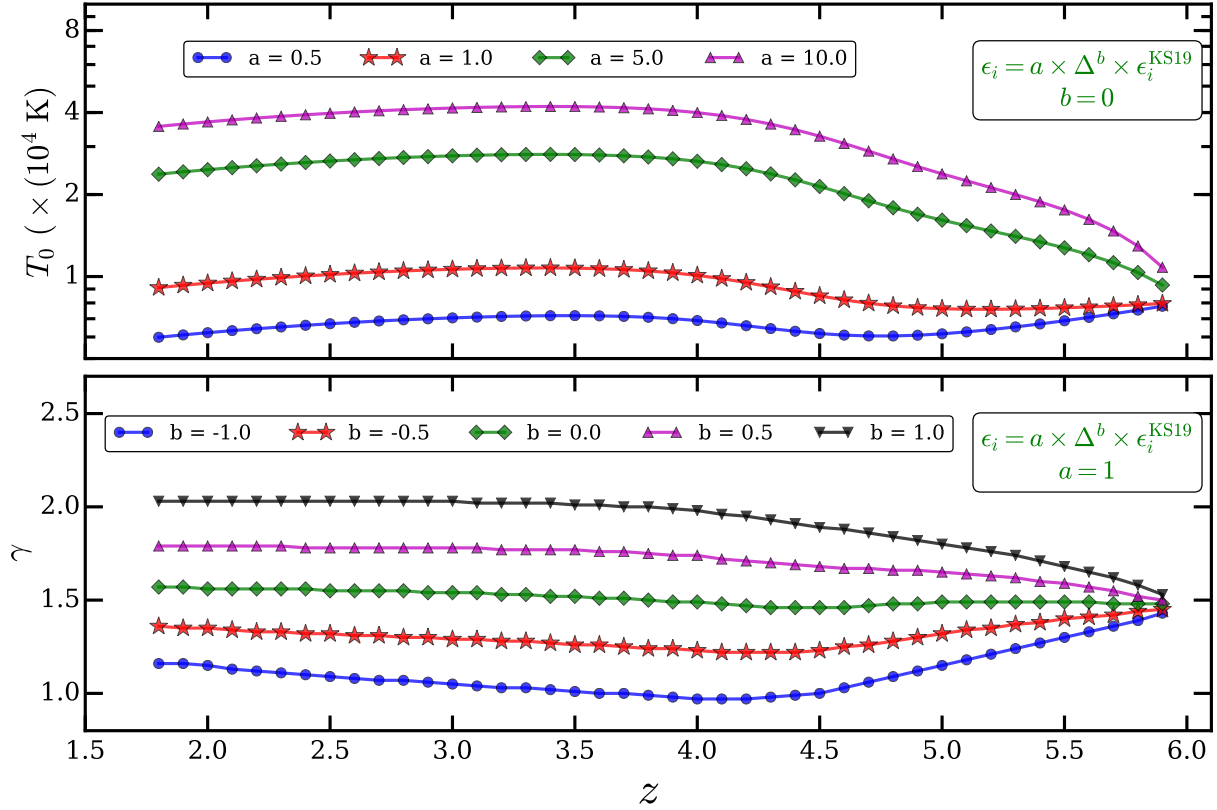
When measuring the thermal parameters, it is important to sample the  $T_0 - \gamma$  plane densely enough. A coarse



**Figure C1.** Each panel is same as the corresponding panel in Fig. 3 except that the errors of the observed FPDF, FPS and CDDF statistics are shown. For visual purposes, the errors of the FPDF, FPS and CDDF statistics are inflated by a factor of 2, 2 and 4 respectively. For a fair comparison, the errors from our sample are calculated using the bootstrap method. In general the errors of the FPDF, FPS and CDDF statistics from our sample are smaller because of the large sample size. The errors of the FPDF in Rollinde et al. (2013) are computed from simulated mock spectra and hence are larger than that of the other measurements. The errors of the FPS statistics from this work are similar to those in Walther et al. (2018, for the no metal case) due to similar sample size and data quality. The errors of our CDDF at  $\log N_{\text{HI}} < 13.6$  are larger due to the somewhat smaller S/N of the spectra in our sample. At  $\log N_{\text{HI}} > 13.6$ , the errors of our CDDF are smaller than that in Kim et al. (2013).



**Figure C2.** Each panel shows residuals between observed FPS calculated in this work and corresponding observed FPS in the literature (Croft et al. 2002; Kim et al. 2004; Walther et al. 2018; Day et al. 2019). The observed number of spectra, redshift path length, quality of spectra and methods of computing flux power spectrum are different in each work. Even with such differences the measured FPS from this work is generally in good agreement with the measurements in the literature.



**Figure D1.** The top and bottom panels show the variation in thermal parameters  $T_0$  and  $\gamma$  obtained by rescaling the H I, He I and He II photo-heating rates of the KS19 UVB with our post-processing module CITE. A simple constant scaling of photo-heating rates (by factor  $a$ ) leads to variation in  $T_0$  while  $\gamma$  changes relatively little. A density dependent scaling (by a factor  $b$ ) leads to a variation in  $\gamma$  for similar values of  $T_0$ . CITE evolves the thermal and ionization state of the IGM by using the outputs of a GADGET-3 simulation. A self-consistent GADGET-3 simulation has been performed with the KS19 UVB (the curves with  $a = 1$  and  $b = 0$ ). For generating the variation in thermal parameters, we assume ionization equilibrium (but no thermal equilibrium) in CITE. We generate 999 different thermal parameter evolution histories.

sampling of  $T_0 - \gamma$  plane can lead to incorrect best fit values and excessive smoothing of the  $\chi^2$  field which, in turn, can result in underestimation of the uncertainties of the thermal parameters. On the other hand a very dense sampling of  $T_0 - \gamma$  would be computationally very expensive. We found it to be the best compromise to choose the  $T_0 - \gamma$  grids such that  $T_0$  is varied from  $\sim 6000$  K to  $\sim 24000$  K in steps of 500 K while  $\gamma$  is varied from  $\sim 0.7$  to  $\sim 2.0$  in steps of 0.05 at  $z = 3$ <sup>1</sup>. We have thus simulated the Ly $\alpha$  forest for  $37 \times 27 = 999$  different thermal histories. The computational time required to run the thermal histories and extract mock spectra for 999 UVB models is  $\sim 2.25$  million cpu hours.

## APPENDIX E: THE SENSITIVITY OF Ly $\alpha$ FOREST STATISTICS TO $T_0, \gamma$

### E1 Kernel Density Estimation

Kernel density estimation is a non-parametric method to estimate the probability distribution function. KDEs are a

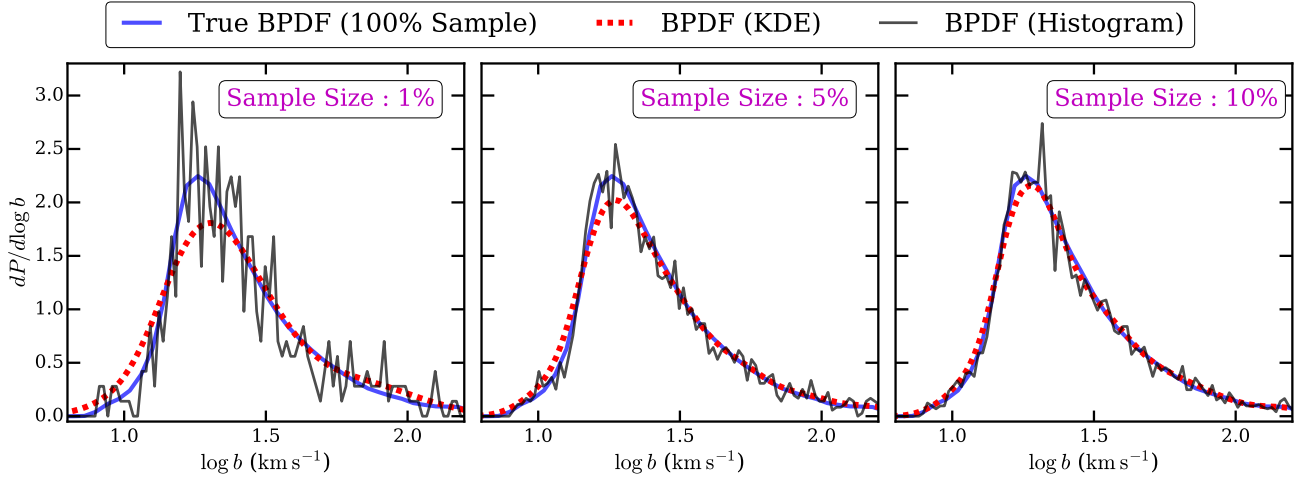
generalization of histograms with the advantage that the loss of information due to binning of the data is minimal with the KDE method. Unlike histograms, KDE is not sensitive to the choice of bins. The PDF calculated using the KDE method converges faster than the simple histogram method.

We use KDE to estimate the PDF of wavelet amplitudes, curvature and  $b$  parameters. There are 3 main reasons for using KDE instead of histograms (i) the number of observed sightlines per redshift bins are usually limited (ii) the observed spectra have finite S/N and (iii) characterizing the shape of the PDF is important as it is sensitive to thermal and ionization parameters.

When calculating PDFs with KDE, we chose a Gaussian function as a kernel. The bandwidth ( $h$ ) of the kernel is set by Scott's rule,  $h = 3.5 \sigma n^{-1/3}$ , where  $\sigma$  is the variance of the quantity whose PDF is to be estimated and  $n$  is number of samples. Fig. E1 illustrates the advantage of using KDE over histograms to estimate the  $b$  distribution (BPDF) from one of our models. First we generate a true BPDF by fitting Voigt profiles to 500 sightlines. We then randomly chose a fraction of sightlines and estimate the BPDF using histograms and KDE. The left, middle and right panel in Fig. E1 show the comparison of the BPDF using the histogram and KDE methods for 1%, 5% and

<sup>1</sup> The corresponding  $T_0$  and  $\gamma$  at other redshifts differ by  $\sim 7$  percent





**Figure E1.** The figure illustrates the convergence of the BPDF using the histogram method (black solid curve) and kernel density estimation (KDE, red dashed line). The true BPDF (blue solid line) is calculated from Voigt profile fits to 500 sight-lines at  $2.9 \leq z \leq 3.1$ . The left, middle and right panel show the BPDF calculated using the histogram method and KDE for 1%, 5% and 10% of sightlines. The PDF calculated using KDE converges faster with a smaller sample than the histogram method. For the KDE method, we use a Gaussian kernel with bandwidth given by Scott's rule. The choice of bins and bin width is the same for all cases.

10% of the total sample. Fig. E1 clearly demonstrates that with KDE the estimated BPDF reaches a good approximation of the true BPDF with a smaller sample than with the histogram method. PDFs estimated using KDE are smoother and converge faster than PDFs estimated with the histogram method<sup>2</sup>.

## APPENDIX F: RESOLUTION AND CONVERGENCE TESTS

The size of the simulation box used in this work (L10N512) is 10 Mpc/h and is likely to be smaller than the typical size of He III bubbles. As discussed in §3 the choice of our simulation box is motivated by the trade-off between dynamic range needed to resolve the Ly $\alpha$  forest at  $2 \leq z \leq 4$  and the large thermal parameter space to probe. However, it is important to demonstrate the sufficient convergence of our simulation by comparing the properties of the Ly $\alpha$  forest from this work with that from a range of simulations in which box sizes and particle numbers are varied. We illustrate the convergence of our simulation in Fig. F1 to F6. We compare the FPDF, FPS, curvature PDF, wavelet PDF, BPDF and CDDF statistics from our model with that from the Sherwood simulation suite Bolton et al. (2017) at  $2.9 \leq z \leq 3.1$ . For a fair comparison we chose thermal parameter  $T_0, \gamma$  consistent with that from the Sherwood simulations. For the models shown in Fig. F1 and F2, we post-process the simulated spectra to match the observed properties of our sample at  $2.9 \leq z \leq 3.1$ . Note that the initial conditions of the density field used for the L10N512 simulation in this work is different from those used in the Sherwood simulation suite. When we rescale the optical depth to match the

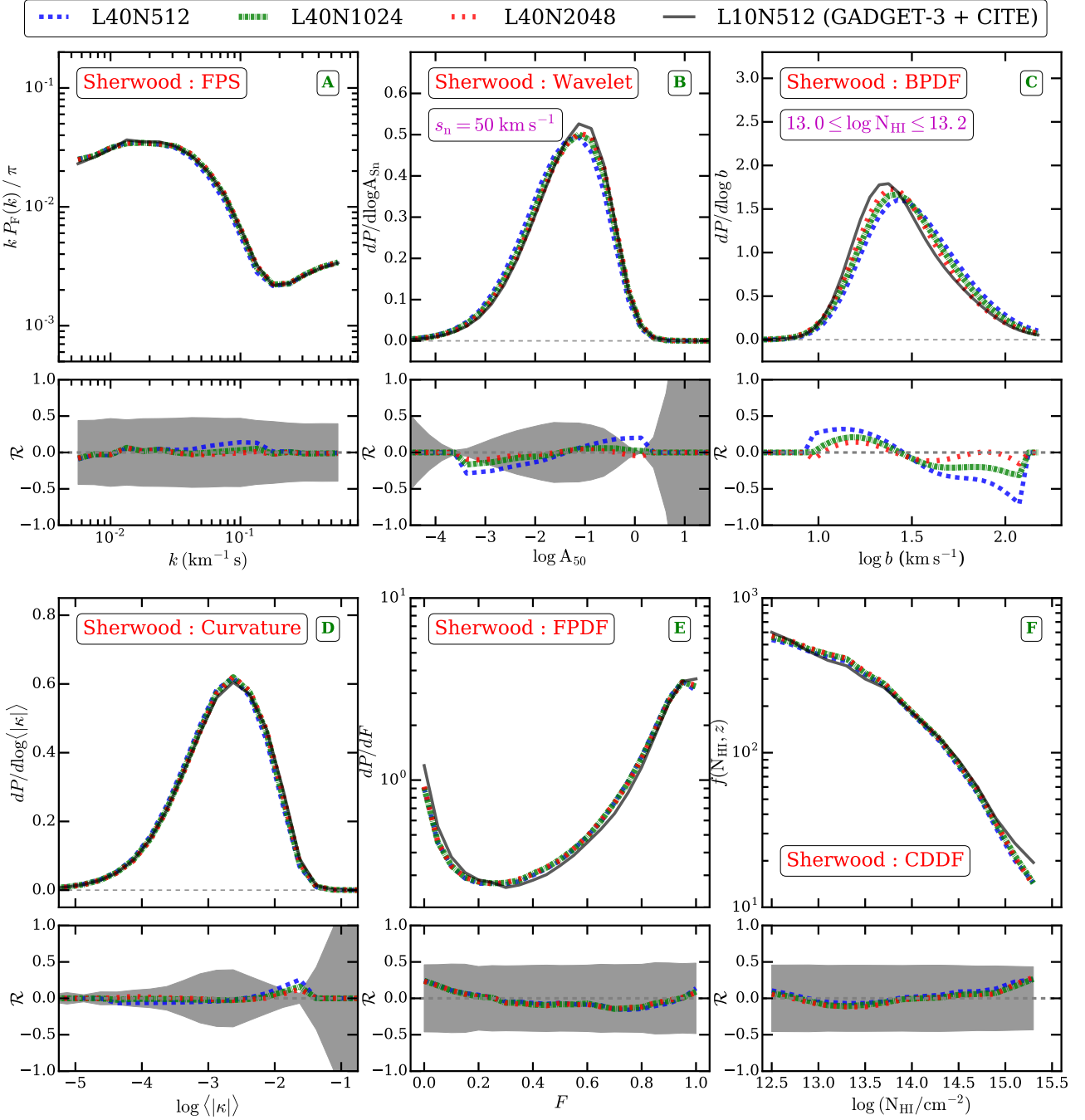
mean flux, we find that the amount of rescaling required is very similar in all the models. Fig. F1 to F6 shows a good agreement between our model statistics with that from the L10N512 Sherwood simulation suggesting that our method of varying the thermal history is consistent with the self-consistent GADGET-3 simulation. The slight mismatch between our Ly $\alpha$  forest statistics and that from the Sherwood simulation suite are primarily due to CITE and not due to differences in initial conditions or the UVB models used. In Gaikwad et al. (2018), we have shown this explicitly by comparing the Ly $\alpha$  forest statistics from spectra post-processed with CITE and those from self-consistent simulations for the same initial condition and UVB model. However, these differences are smaller than the observational uncertainties and are compensated by the ability of our method to probe the large thermal parameter space which would be computationally expensive for simulations with larger dynamic range/particle numbers. Fig. F2, F4 and F6 also illustrate that the statistics are converged with regards to box sizes. We compare the statistics from the L10N512, L20N512, L40N1024 and L80N2048 Sherwood simulations. The good agreement among the Sherwood models L40N512, L40N1024 and L40N2048 (Sherwood) indicates that the statistics are converged with regard to resolution/ particle mass.

## APPENDIX G: COMPARISON OF BEST FIT STATISTICS WITH OBSERVATIONS

Similarly to Fig. 12, we show the comparison of FPS, wavelet, BPDF and curvature statistics of our best fit models with those obtained from observations at different redshift bins in Fig. G1 to G9. In general, the predictions of the best fit models are in good agreement with the observed statistics. We summarize the reduced  $\chi^2$  between best fit model and observations in Table G1. The reduced  $\chi^2$  is close to 1 in most cases.

We also plot a comparison of the observed  $b - \log N_{\text{HI}}$

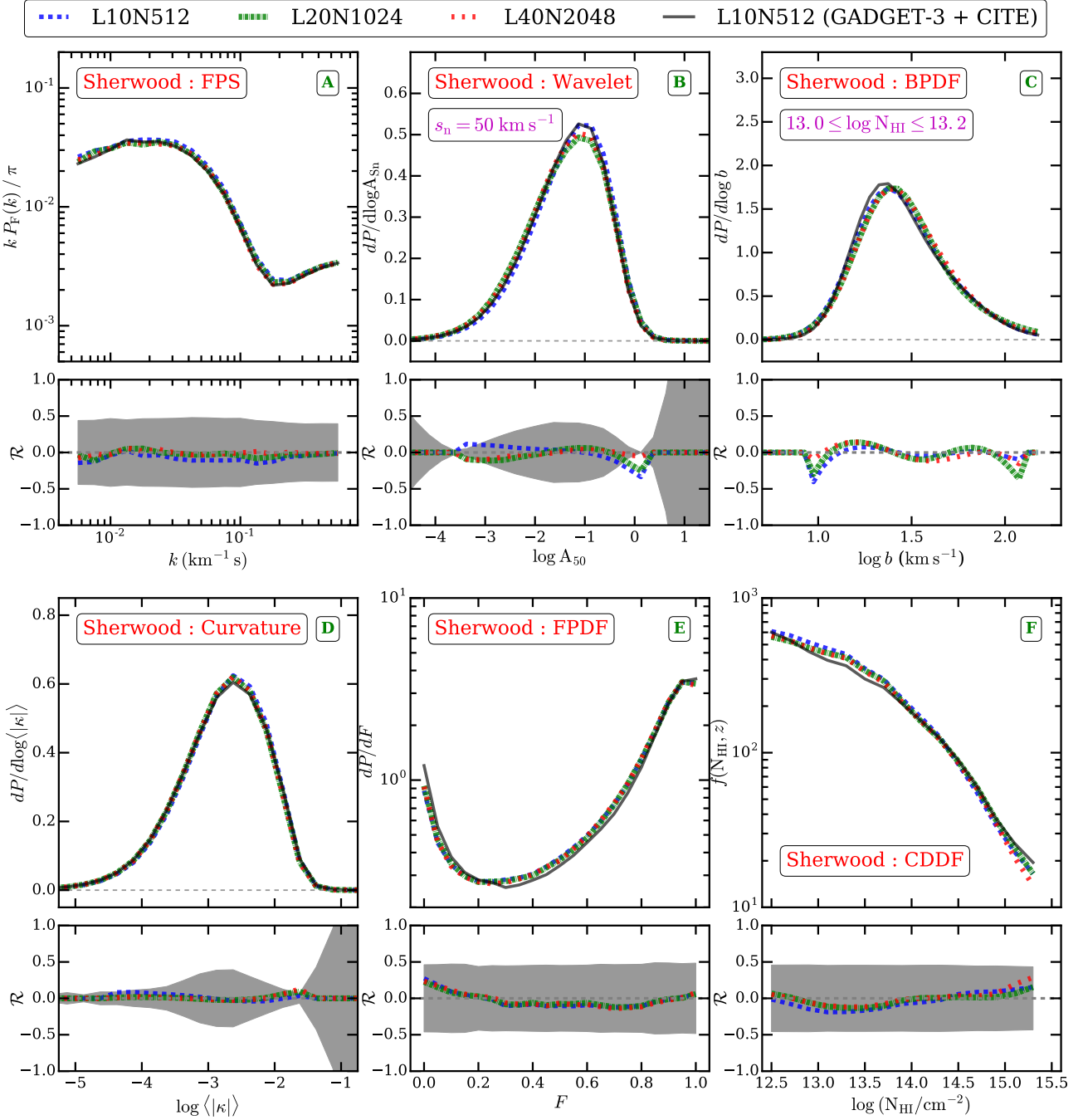
<sup>2</sup> The bin center and bin width is fixed in Fig. E1. One can use coarse binning with the histogram method to estimate PDFs. However, the coarse binning can lead to loss of information with regard to the detailed shape of the PDF.



**Figure F1.** Here we show the effect of resolution on Ly $\alpha$  statistics by comparing these statistics for a range of simulations from the Sherwood simulation suite. All other parameters such as  $T_0$ ,  $\gamma$ ,  $\Gamma_{\text{HI}}$  and box size ( $40 h^{-1}$  cMpc) are the same for all the models. The flux statistics of our default simulation are sufficiently converged with regard to resolution. For comparison, we also show the flux statistics using our GADGET-3 + CITE method. The flux statistics calculated for our post-processed simulation is consistent with that for the corresponding self-consistent simulation from the Sherwood simulation suite. The difference between the GADGET-3 + CITE model and the self-consistent Sherwood simulation suite are within the observational uncertainty for each statistics (see Fig. 12). All statistics are shown at  $2.9 \leq z \leq 3.1$ . The shaded region shows the typical observational uncertainty for each statistics.

contours (that includes 90 percent of points) with that from 5 mock samples (out of 100) in Fig. G10. The mock  $b - \log N_{\text{HI}}$  distribution corresponds to the best fit thermal parameters model. Traditionally, the lower  $b$  envelope in the  $b - \log N_{\text{HI}}$  distribution is used to measure the thermal parameters. The limitations of this method are described in

§4.1.4. Fig. G10 shows that the 2D observed  $b - \log N_{\text{HI}}$  distribution is similar to that from the mock samples at all redshifts. However, in general we see a mismatch between observed  $b - \log N_{\text{HI}}$  and best fit  $b - \log N_{\text{HI}}$  distribution at low and high  $\log N_{\text{HI}}$ . The low  $\log N_{\text{HI}}$  systems are usually contaminated due to the finite S/N of the spectra. Note



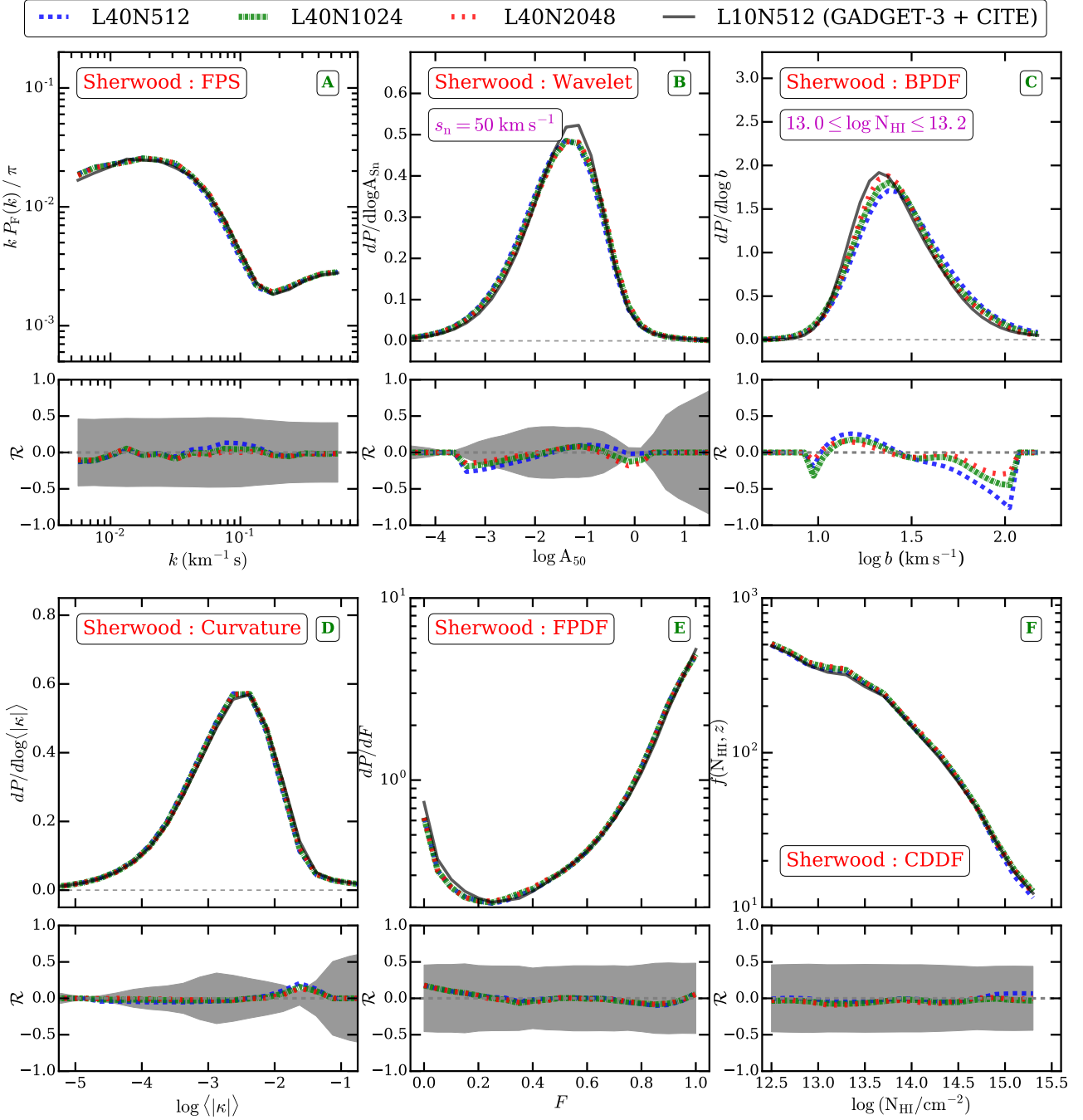
**Figure F2.** Same as Fig. F1 except that the effect of box size (resolution is same) on Ly $\alpha$  statistics is shown by comparing these statistics for a range of simulations from the Sherwood simulation suite. The flux statistics of our default simulation are sufficiently converged with regard to box size. For comparison, we also show the flux statistics using our GADGET-3 + CITE method. In this figure the resolution of the Sherwood model is same as that of the GADGET-3 + CITE model.

here that the Voigt decomposition of high  $\log N_{\text{HI}}$  systems may not be unique. This effect is more prominent at higher ( $z > 3.3$ ) redshifts. Due to combination of these effects the mock  $b - \log N_{\text{HI}}$  distributions further show some differences compared to observations. It is noteworthy that the shape of the contours for lower  $b$  values from the simulations are in good agreement with that from observations. The variations in the  $b - \log N_{\text{HI}}$  distribution in different mock samples is also small indicating that the distribution is converged with

respect to cosmic variance. In summary, the  $b - \log N_{\text{HI}}$  distribution from our best fit models is in general good agreement with that from observations.

In Fig. 16 (panel B1), we show a comparison of  $T_0$  from our method using the characteristic density with the results from Becker et al. (2011); Boera et al. (2014). To get the  $T_0$  evolution in our model we assume the best fit  $\gamma$  evolution (with errors). To perform a fair comparison with the Becker et al. (2011); Boera et al. (2014) method, we show



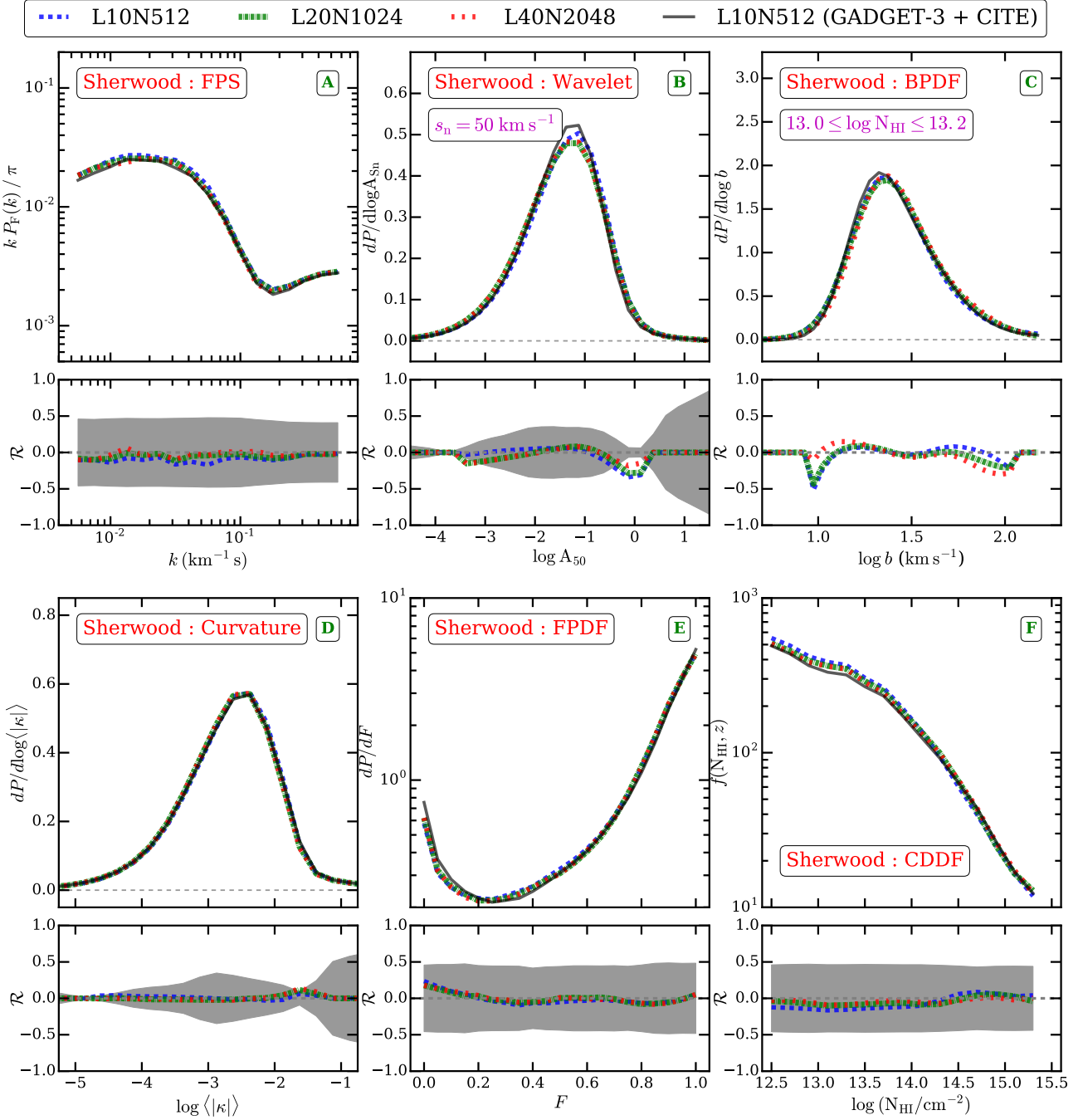


**Figure F3.** Same as Fig. F1 except for  $1.9 \leq z \leq 2.1$ .

the temperature at the characteristic density from our work and Becker et al. (2011); Boera et al. (2014) in Fig. G11. Our temperature measurements from the curvature and wavelet statistics are in good agreement with that from the literature at  $z > 2.4$ . However, our measurements at  $z < 2.4$  are systematically lower than that from the literature, but are still consistent within  $1.5\sigma$ . It is interesting to note that the characteristic density predicted from our model ( $\bar{\Delta} \sim 4.85$ ) is also smaller at  $z \sim 2$  compared to that from Becker et al. (2011,  $\bar{\Delta} \sim 6$ ).

## APPENDIX H: SOURCES OF UNCERTAINTY IN $T_0$ , $\gamma$ MEASUREMENTS

In addition to statistical uncertainty, the  $T_0$ ,  $\gamma$  measurements are also affected by the modeling and observational uncertainties. We have considered four main sources of uncertainties contributing to the total uncertainty on the measured parameters. These are uncertainties due to (i) continuum placement, (ii) contamination by unidentified metal lines, (iii) (numerical) modeling and (iv) cosmological parameters. The first two uncertainties are observational uncertainties while the later two uncertainties are modeling



**Figure F4.** Same as Fig. F2 except for  $1.9 \leq z \leq 2.1$ .

uncertainties. We discuss below the contribution of each to the measurement uncertainty in detail.

### H1 Continuum placement uncertainty

When comparing simulated Ly $\alpha$  forest spectra with observations, it is customary to use normalized observed spectra ( $F_{\text{norm}} = F_{\text{unnorm}}/F_{\text{cont}}$ ). The continuum fitting of spectra is usually based on high order polynomial fitting connecting points in the spectrum believed to have no absorption. The uncertainty in continuum placement ( $\delta F_{\text{cont}}/F_{\text{cont}}$ ) affects

the observed flux statistics of Ly $\alpha$  forest data and influences the inferred physical parameters. In the KODIAQ DR2 sample, all the spectra are normalized by the procedure described in O’Meara et al. (2015, 2017). The continuum placement uncertainty in the KODIAQ DR2 sample is of the order of a few percent (see O’Meara et al. 2015). We assume a conservative value for the continuum placement uncertainty of  $\pm 5$  percent. If  $F_{\text{default}}$  is the flux corresponding to the default continuum and  $\delta F_{\text{cont}}/F_{\text{cont}}$  is the uncertainty in continuum placement then we define a low /high continuum as  $F_{\text{default}}/[1 \pm (\delta F_{\text{cont}}/F_{\text{cont}})]$ .

We quantify the effect of continuum placement uncer-

**Table G1.** Reduced  $\chi^2$  ( $\chi^2$  per degree of freedom) between data and best fit  $T_0$  and  $\gamma$  measurements for various statistics.

Redshift	$b$ distribution <sup>[a]</sup>	Flux power spectrum	Wavelet PDF <sup>[b]</sup>	Curvature PDF	Joint Analysis
$2.0 \pm 0.1$	1.45	1.15	1.70	1.51	1.50
$2.2 \pm 0.1$	1.51	1.33	1.80	1.43	1.57
$2.4 \pm 0.1$	1.85	1.75	1.17	1.64	1.34
$2.6 \pm 0.1$	1.02	0.85	1.94	1.78	1.40
$2.8 \pm 0.1$	1.77	1.61	1.01	0.45	1.21
$3.0 \pm 0.1$	1.35	0.71	1.29	0.92	1.10
$3.2 \pm 0.1$	1.53	0.87	1.23	1.29	1.24
$3.4 \pm 0.1$	1.45	0.71	1.45	0.69	1.12
$3.6 \pm 0.1$	1.61	0.60	0.54	0.82	0.94
$3.8 \pm 0.1$	1.72	0.73	0.72	1.23	1.14

<sup>a</sup> The reduced  $\chi^2$  is calculated by adding the  $\chi^2$  between model and data  $b$  distribution in different  $\log N_{\text{HI}}$  bins.

<sup>a</sup> The reduced  $\chi^2$  is calculated by adding the  $\chi^2$  between model and data wavelet PDF for different wavelet scales ( $s_n$ ).

**Table H1.**  $T_0$  and  $\gamma$  measurements and the associated error budget

Redshift $z \pm dz$	Statistical uncertainty <sup>[a]</sup> ( $\sigma_{\text{stat}}$ )	Model uncertainty <sup>[b]</sup> ( $\sigma_{\text{model}}$ )	Cosmological parameters <sup>[c]</sup> ( $\sigma_{\text{cosmo}}$ )	Metal line contamination <sup>[d]</sup> ( $\sigma_{\text{metal}}$ )	Continuum uncertainty <sup>[e]</sup> ( $\sigma_{\text{cont}}$ )	Best fit $\pm$ Total uncertainty <sup>[f]</sup> ( $\sigma_{\text{tot}}$ )
Contribution of various uncertainties to the total uncertainty on $T_0$ (in K) measurements						
$2.0 \pm 0.1$	1254	285	31	125	100	$9500 \pm 1393$
$2.2 \pm 0.1$	831	330	36	125	125	$11000 \pm 1028$
$2.4 \pm 0.1$	952	382	41	100	100	$12750 \pm 1132$
$2.6 \pm 0.1$	894	405	44	125	400	$13500 \pm 1390$
$2.8 \pm 0.1$	848	442	48	125	375	$14750 \pm 1341$
$3.0 \pm 0.1$	827	442	48	125	375	$14750 \pm 1322$
$3.2 \pm 0.1$	1045	382	41	100	375	$12750 \pm 1493$
$3.4 \pm 0.1$	796	338	37	125	250	$11250 \pm 1125$
$3.6 \pm 0.1$	885	308	33	125	125	$10250 \pm 1070$
$3.8 \pm 0.1$	690	278	30	100	125	$9250 \pm 876$
Contribution of various uncertainties to the total uncertainty on $\gamma$ measurements						
$2.0 \pm 0.1$	0.081	0.018	0.003	0.010	0.012	$1.500 \pm 0.096$
$2.2 \pm 0.1$	0.080	0.017	0.003	0.013	0.050	$1.425 \pm 0.133$
$2.4 \pm 0.1$	0.095	0.016	0.003	0.010	0.025	$1.325 \pm 0.122$
$2.6 \pm 0.1$	0.092	0.015	0.003	0.025	0.025	$1.275 \pm 0.122$
$2.8 \pm 0.1$	0.094	0.015	0.003	0.010	0.012	$1.250 \pm 0.109$
$3.0 \pm 0.1$	0.106	0.015	0.003	0.013	0.012	$1.225 \pm 0.120$
$3.2 \pm 0.1$	0.115	0.015	0.003	0.013	0.013	$1.275 \pm 0.129$
$3.4 \pm 0.1$	0.096	0.016	0.003	0.012	0.010	$1.350 \pm 0.108$
$3.6 \pm 0.1$	0.088	0.017	0.003	0.012	0.010	$1.400 \pm 0.101$
$3.8 \pm 0.1$	0.066	0.018	0.003	0.038	0.062	$1.525 \pm 0.140$

<sup>a</sup> Statistical uncertainty corresponds to uncertainty from joint constraints of respective parameters.

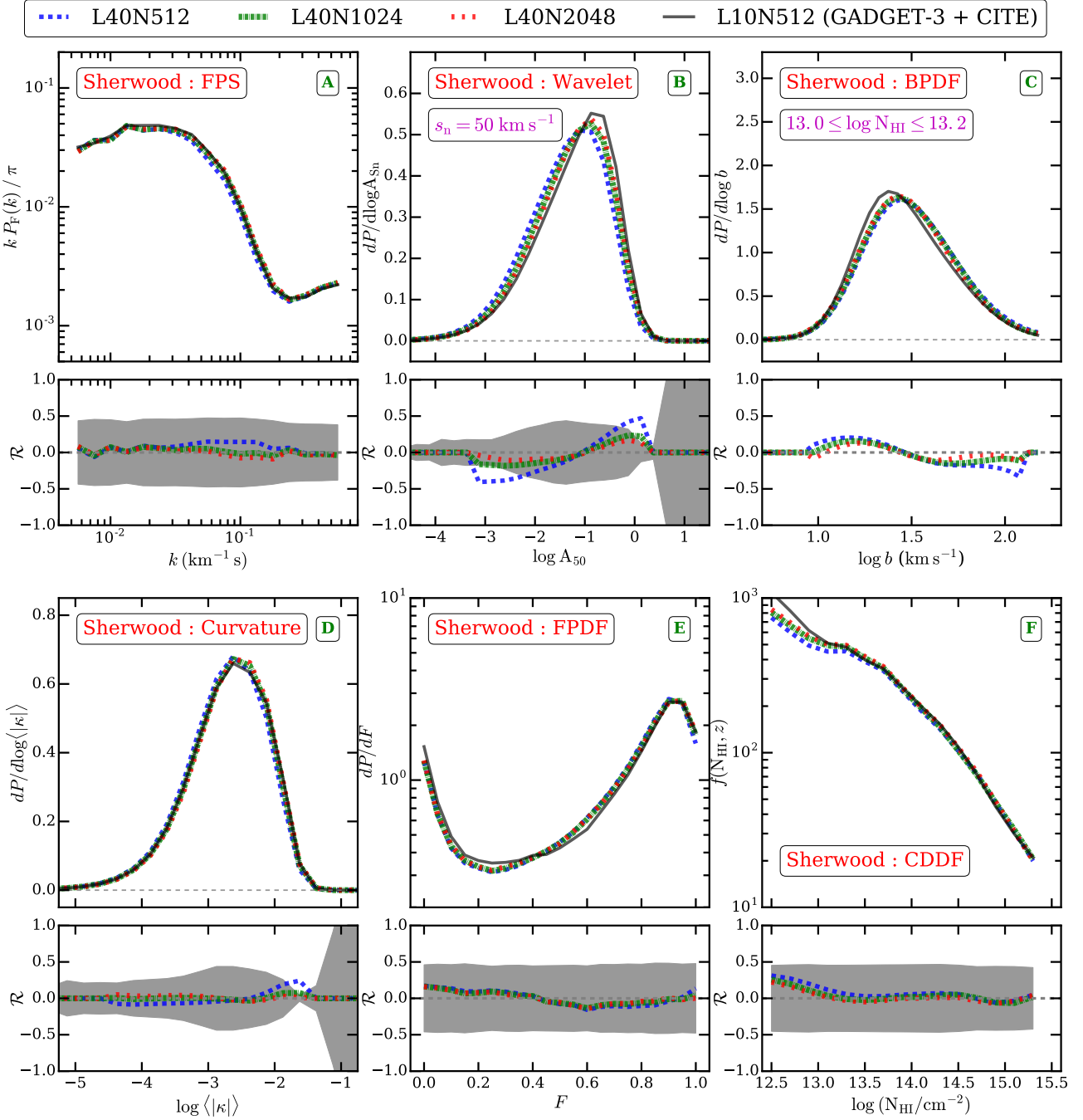
<sup>b</sup> Model uncertainty corresponds to uncertainty in modeling of the Ly $\alpha$  forest in our simulation.

<sup>c</sup> Cosmological parameter uncertainty is calculated using the Gunn-Peterson approximation. Note that the GP approximation does not account for the uncertainty in  $\sigma_8$ .

<sup>d</sup> Contribution of unidentified metal lines to the uncertainty in derived quantity.

<sup>e</sup> Continuum placement uncertainty ( $\sigma_{\text{cont}}$ ) is assumed to be  $\pm 5$  percent. For  $T_0$  and  $\gamma$  measurements,  $\sigma_{\text{cont}}$  includes the uncertainty due to the mean transmitted flux for a given continuum. For  $\log \Gamma_{12}$  measurements,  $\sigma_{\text{cont}}$  include the uncertainty due to variation of  $T_0$  and  $\gamma$  parameters.

<sup>f</sup> Total uncertainty is obtained by  $\sigma_{\text{tot}} = \sqrt{\sigma_{\text{stat}}^2 + \sigma_{\text{model}}^2 + \sigma_{\text{cosmo}}^2 + \sigma_{\text{metal}}^2} + \sigma_{\text{cont}}$



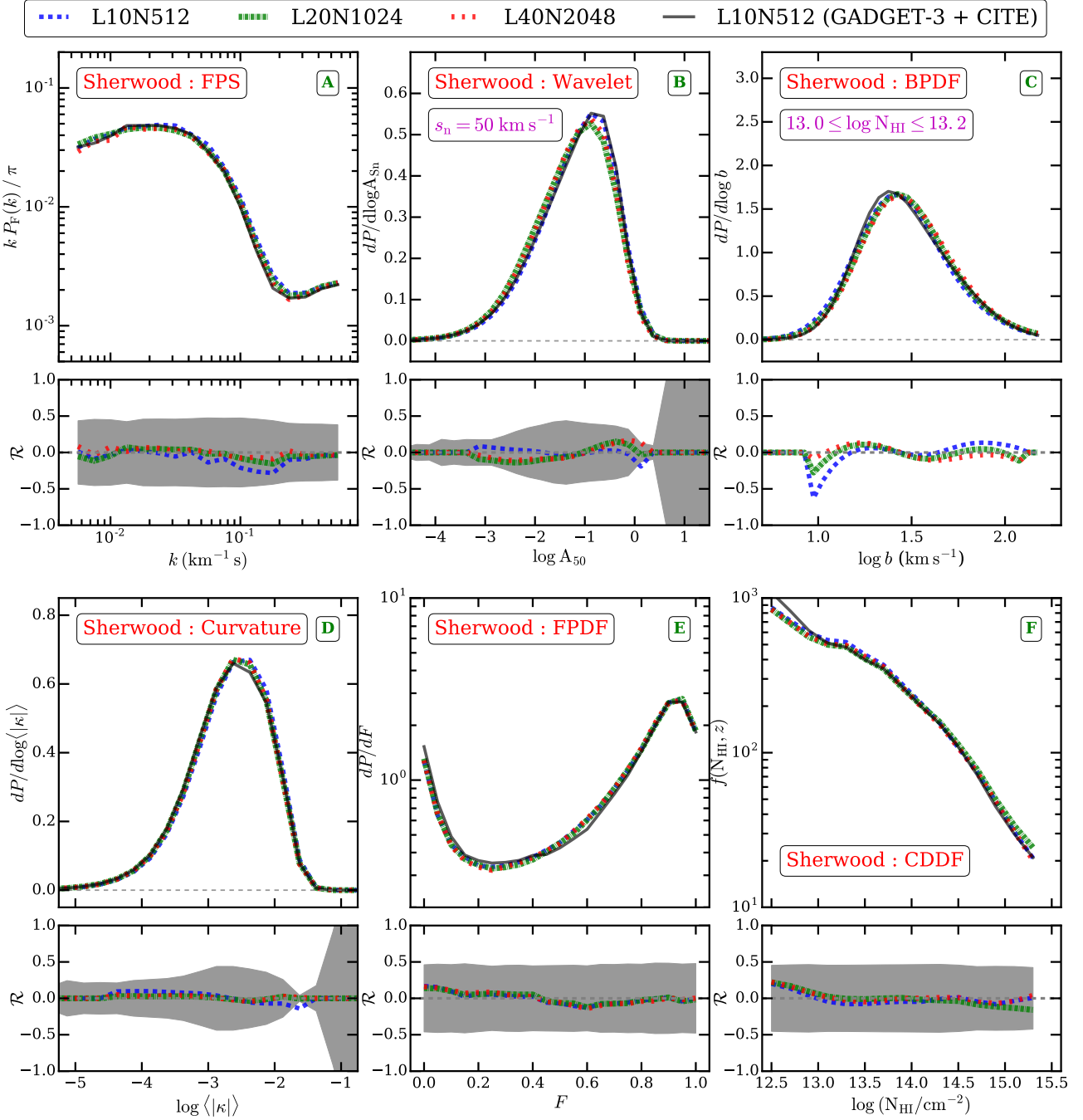
**Figure F5.** Same as Fig. F1 except for  $3.7 \leq z \leq 3.9$ .

tainty by rescaling the observed flux, deriving all the statistics from the rescaled spectra and constraining the  $T_0$ ,  $\gamma$  parameters using our simulations. Fig. 15 shows the effect of continuum placement on the (joint)  $T_0$  and  $\gamma$  constraints. For low (high) continuum placement,  $T_0$  and  $\gamma$  are systematically underpredicted (overpredicted) relative to measurements with the default continuum. One can qualitatively understand this systematic effect of continuum placement on flux. Since dividing a flux by a low (high) continuum stretches (compresses) the flux in the vertical direction, this artificially reduces (increases) the line widths and hence one gets systematically smaller (higher) temperature. Fig. H2

shows a similar effect of continuum placement uncertainty in the other redshift bins. The uncertainty in  $T_0$  and  $\gamma$  measurement due to continuum placement is systematic in nature hence we add this uncertainty in the total error budget (see Table H1).

## H2 Metal line contamination uncertainty

The observed Ly $\alpha$  forest is usually contaminated by absorption lines arising from metals. For a given temperature, the metal lines are typically narrower than H I lines because for thermally broadened lines, the line width  $b \propto m^{-1/2}$  and  $m$



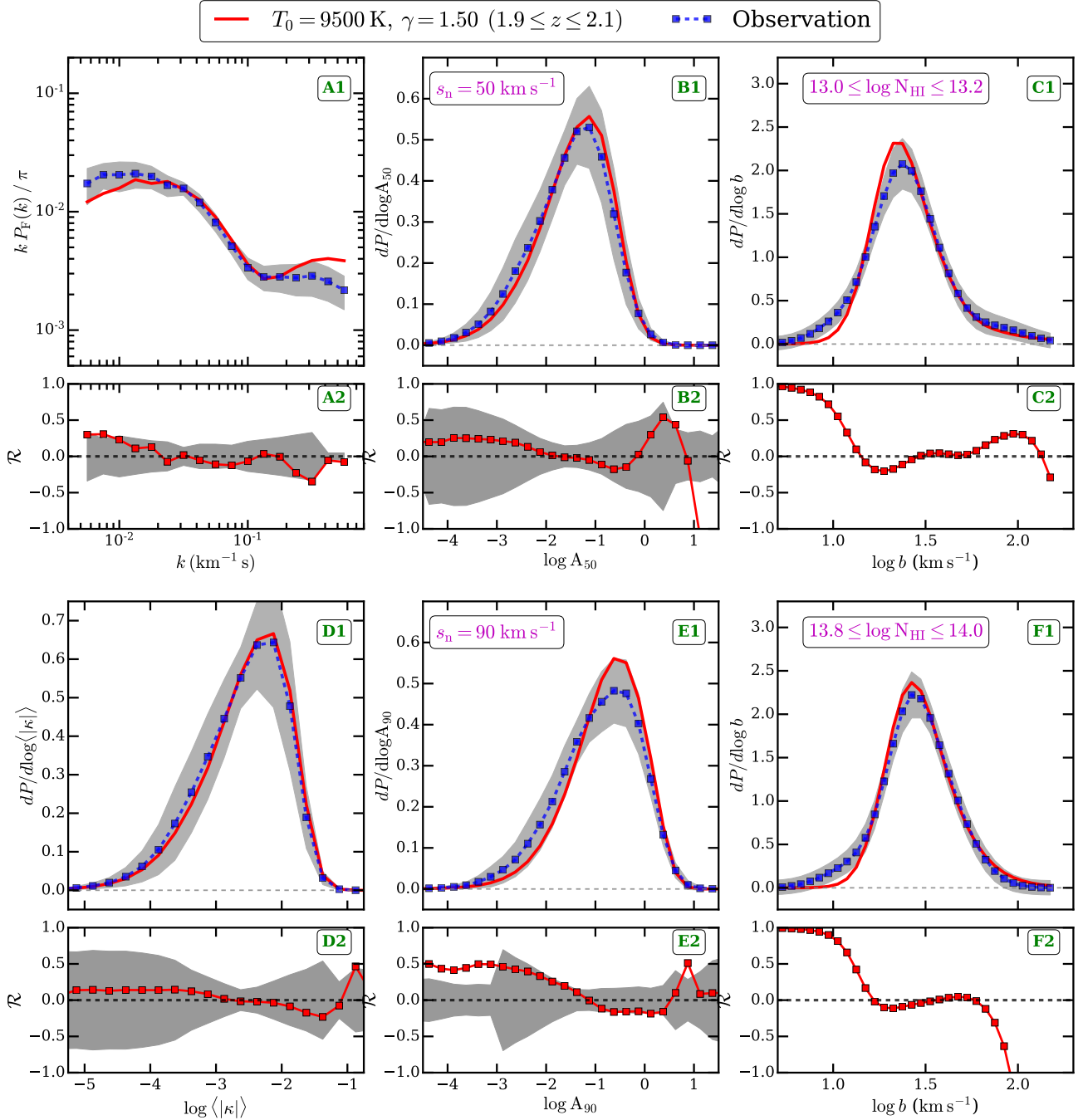
**Figure F6.** Same as Fig. F2 except for  $3.7 \leq z \leq 3.9$ .

is larger for metals than hydrogen. The presence of such narrower metal lines can potentially bias the temperature measurement to lower values. In this work, we do not model the intergalactic metal absorption systems as our simulations neither follow chemical evolution of metals nor include feedback processes. To mitigate this, we chose QSO sightlines that do not contain DLAs or sub-DLAs. Thus our sample selection criteria minimizes the number of metal lines contaminating the Ly $\alpha$  forest. Furthermore, we identify metal lines and replace them with continuum and add noise. We use manual and automated approaches to identify metal lines. We find that both approaches give consistent results.

In the manual approach, we first find the redshift of absorption systems using part of the spectra redward of the H I emission line. Corresponding to this redshift, we look for other transitions contaminating Ly $\alpha$  forest. Due to insufficient wavelength coverage, it is possible that we may not identify all the metal lines. Hence we manually look at each spectrum and look for obvious metal line systems. We then replace metal lines with continuum and add noise<sup>3</sup>.

<sup>3</sup> We account for the effect of metal lines blended with H I absorption lines.





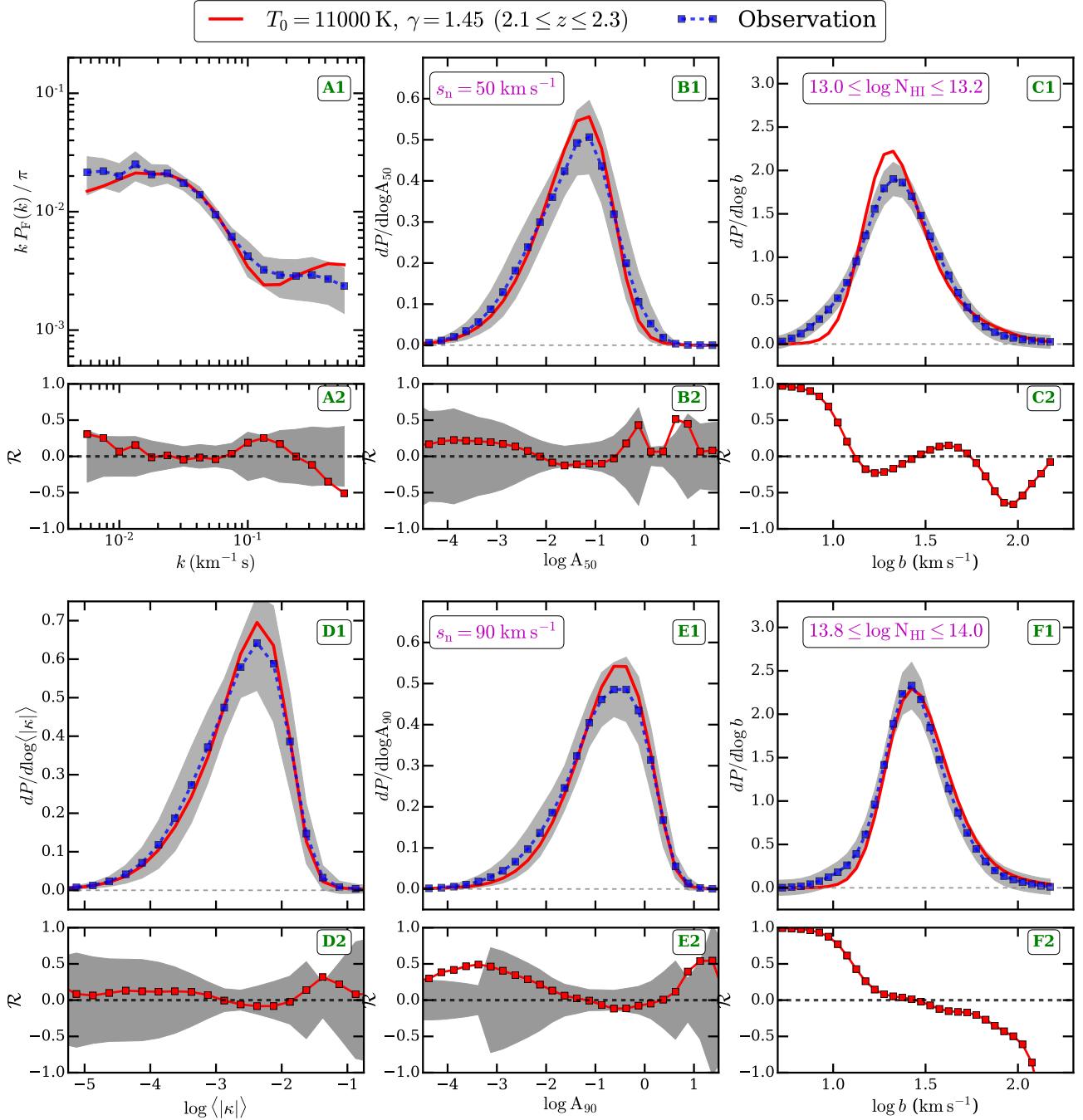
**Figure G1.** Same as Fig. 12 except for  $1.9 \leq z \leq 2.1$ .

The manual method is tedious and somewhat subjective. Hence, we have also used an automatic method to identify the metal lines based on line widths. Since metal lines are narrower than corresponding H I lines, we fit all the spectra with multi-component Voigt profiles using VIPER. We treat all the lines with  $b \leq 8 \text{ km s}^{-1}$  as metal lines<sup>4</sup>.

<sup>4</sup> The narrowest line that can be detected in spectra is limited by the line spread function (FWHM  $\sim 6 \text{ km s}^{-1}$ ) of the spectrograph.

We replace these metal lines with continuum and add noise accounting for any blending effects.

Fig. 15 and Fig. H2 shows the effect of metal contamination on  $T_0$  and  $\gamma$  constraints in all the redshift bins. As expected, when we use flux statistics for the observed spectra without subtracting metal contamination, we get systematically smaller  $T_0$  and  $\gamma$ . Even though we identify and remove the metal lines, it is still possible that some of the metal lines are not identified by our method. Since the metal lines contamination introduces an uncertainty of  $\sim 2$  percent, we add this uncertainty in the total error budget of constrained parameters (see Table H1).

Figure G2. Same as Fig. 12 except for  $2.1 \leq z \leq 2.3$ .

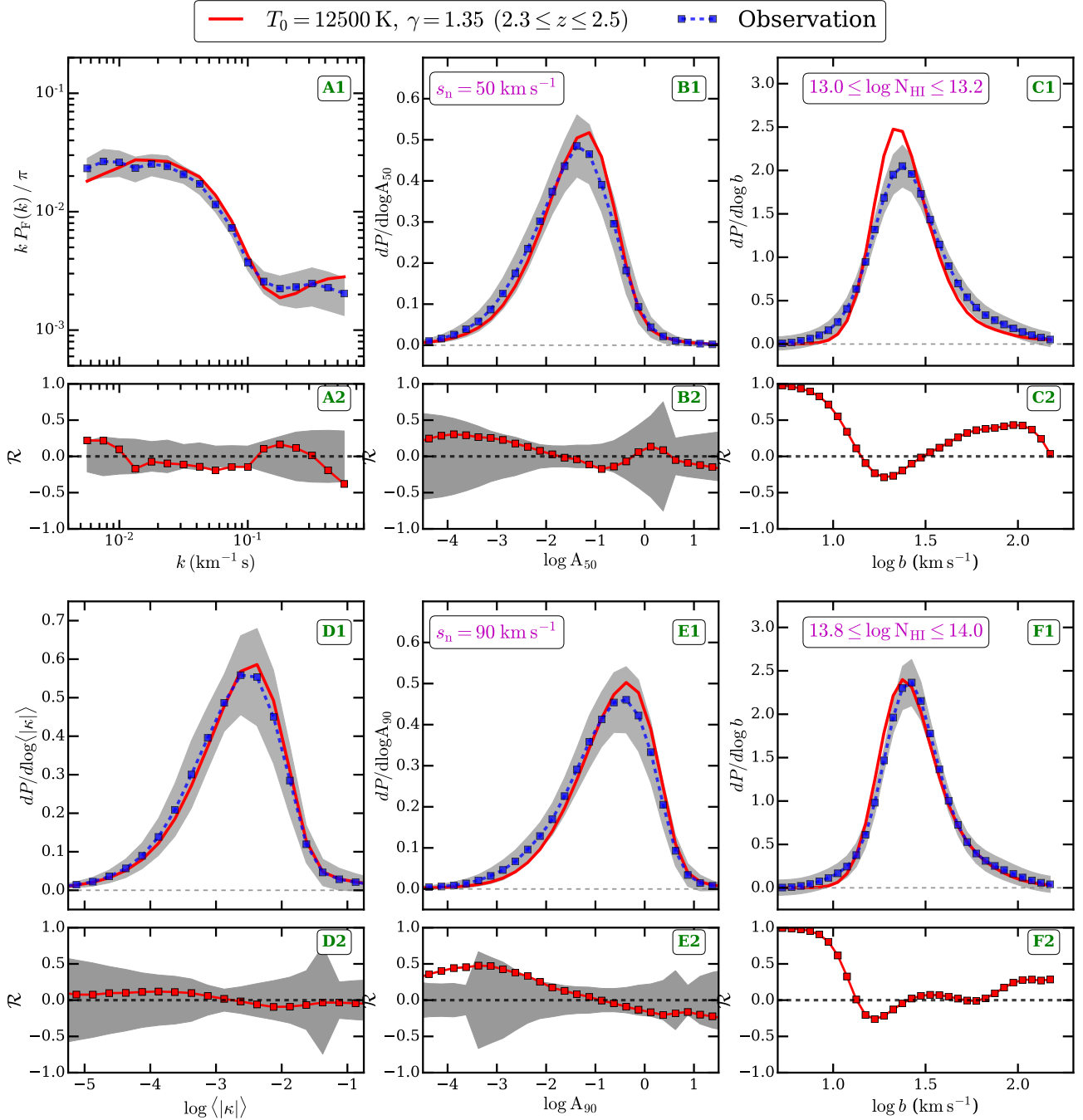
### H3 Cosmological parameter uncertainty

All the results presented in this work assume a flat  $\Lambda$ CDM cosmology consistent with [Planck Collaboration et al. \(2014\)](#). However the Ly $\alpha$  optical depth depends on cosmological parameters  $\Omega_m$ ,  $h$  and  $\Omega_b$  as

$$\tau \propto (\Omega_b h^2)^2 \Omega_m^{-0.5} \frac{T_0^{-0.7} \Delta^{2-0.7(\gamma-1)}}{\Gamma_{\text{HI}}} \quad (\text{H1})$$

where we assume that recombination rate scales with temperature as  $T^{-0.7}$ . It would be computationally expensive to perform all the analysis with varying cosmological parameters.

Hence we use a simplified approach of propagating the error using the above equation. For this we use the  $1\sigma$  uncertainty on  $\Omega_m$ ,  $h$  and  $\Omega_b$  from [Planck Collaboration et al. \(2014\)](#). The cosmological parameters derived from the CMB are correlated with each other. Some of these correlation may lead to a cancellation of errors. However, while performing the error analysis, we always add errors due to different cosmological parameters. Thus the assumed uncertainty due to cosmological parameters is conservative. We do not account for the uncertainty in  $\sigma_8$  and  $n_s$  as this would require to perform self-consistent GADGET-3 simulations. However it is important to note that, we find that the uncertainty



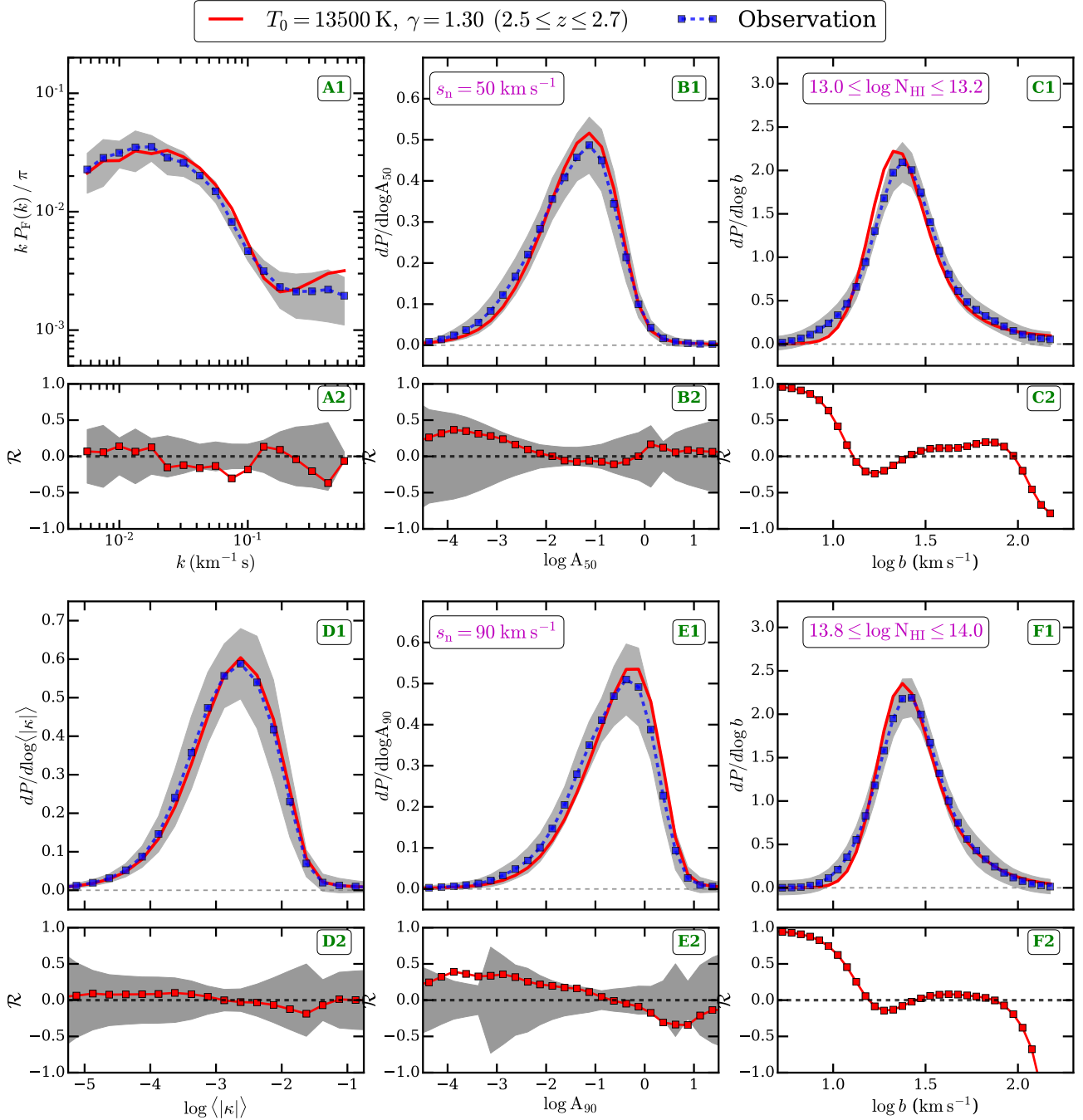
**Figure G3.** Same as Fig. 12 except for  $2.3 \leq z \leq 2.5$ .

in  $T_0$  and  $\gamma$  measurements due to cosmological parameters uncertainty is small,  $\leq 0.5$  percent (see Table H1).

#### H4 Modeling uncertainty due to Jeans smoothing

In the simulations, we vary the thermal state of the IGM by post-processing GADGET-3 simulations using CITE. Pressure smoothing effects due to the hydrodynamical response of the gas are important during reionization (Park et al. 2016; D’Aloisio et al. 2020). As shown in Gaikwad et al. (2018), we account for the pressure smoothing of the gas by

convolving the SPH kernel with a Gaussian kernel that depends on the pressure smoothing scale (Gnedin & Hui 1998; Kulkarni et al. 2015). This approach is based on the ansatz that the Ly $\alpha$  absorbers are in local hydrostatic equilibrium (Schaye 2001). However, this assumption may not be valid during the process of reionization as the photo-heating time scales are typically shorter than the dynamical response time of the gas. In addition to this, the Ly $\alpha$  forest traces the neutral gas in large scale filaments which are undergoing gravitational collapse onto nearby over-dense objects. The hydrostatic equilibrium approximation may thus not be correct. To quantify the effect of this ansatz, we perform a



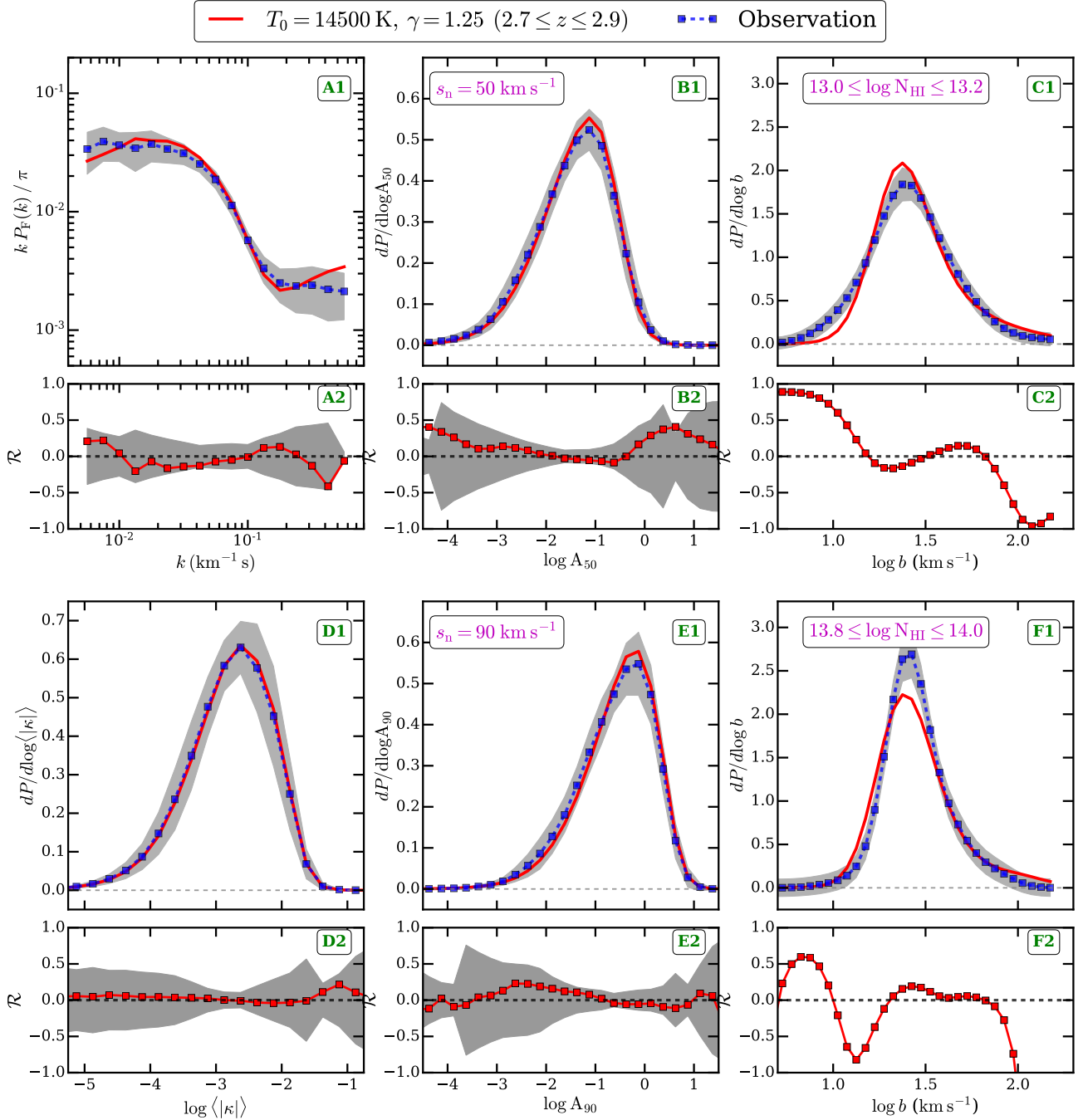
**Figure G4.** Same as Fig. 12 except for  $2.5 \leq z \leq 2.7$ .

self-consistent GADGET-3 simulation with enhanced photoheating rates where  $T_0$  is doubled. The dynamical effect of pressure smoothing is thus captured accurately in this self-consistent simulation. We compute the Ly $\alpha$  forest statistics for this self-consistent model and assume it as the fiducial model. We constrain  $T_0$ ,  $\gamma$  and  $\Gamma_{\text{HI}}$  using our model, i.e. GADGET-3 (at lower temperature) + CITE (with convolution of SPH kernel with Gaussian). We find that the fiducial  $T_0$ ,  $\gamma$  and  $\Gamma_{\text{HI}}$  are recovered within 2.5, 1.3 and 0.6 percent. Such a good level of agreement is partly expected because the smoothing of transmitted flux at these redshifts is dominated by temperature, while the pressure smoothing effects

are sub-dominant. We account for this modeling uncertainty in the total error budget of derived parameters (see Table H1). We also refer the reader to §3 for a related discussion.

## APPENDIX I: THERMAL PARAMETER EVOLUTION IN UVB MODELS

In this section we discuss the effect of uncertainty in observed and modeling parameters in the UVB models on the thermal parameter evolution during He II reionization.



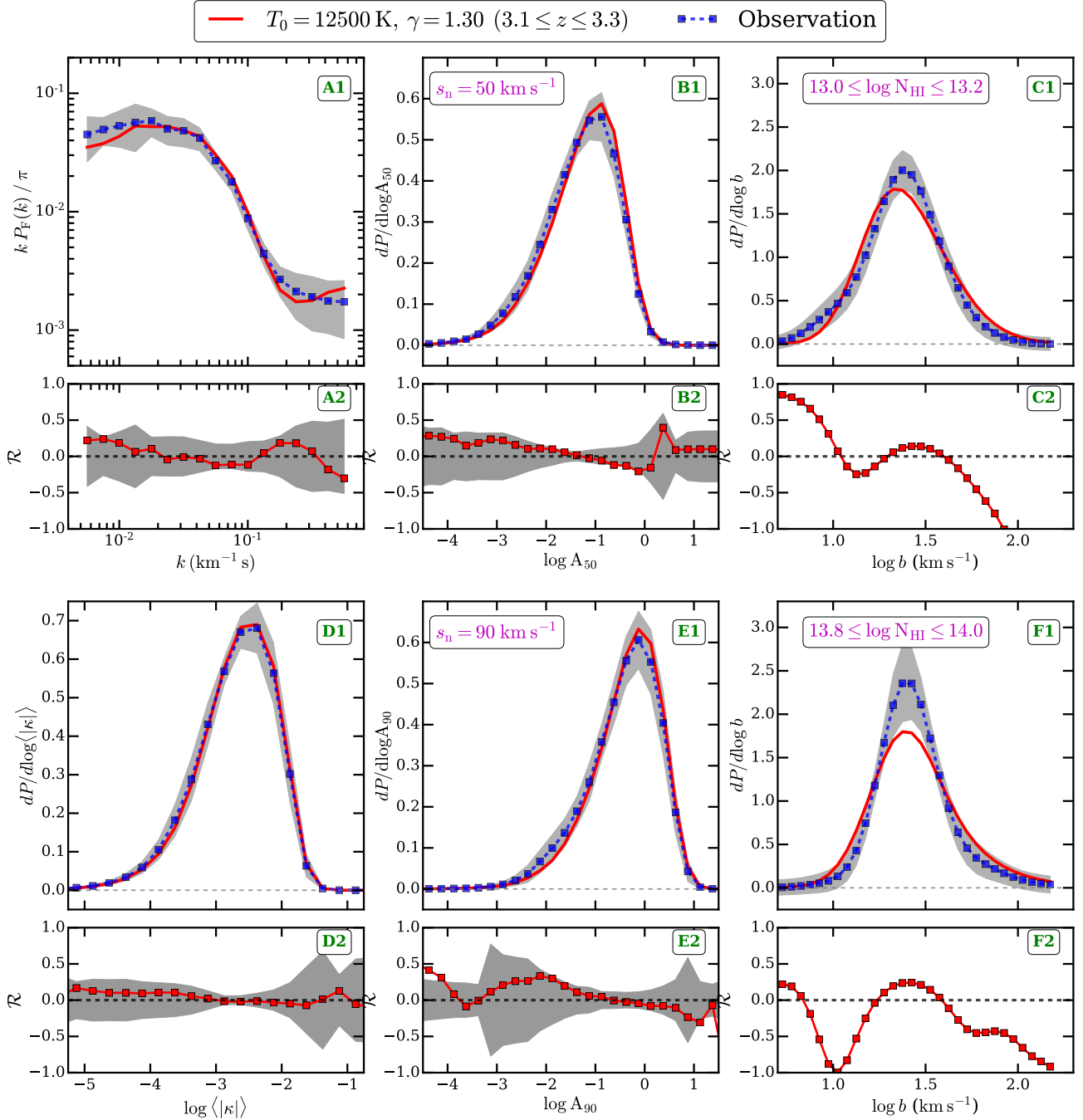
**Figure G5.** Same as Fig. 12 except for  $2.7 \leq z \leq 2.9$ .

## I1 Effect of uncertainty in $T_0 - \gamma$ at $z = 6$ on thermal parameter evolution at $2 \leq z \leq 4$

In this section we discuss the effect of varying the thermal parameters  $T_0$  and  $\gamma$  at  $z = 6$  (the initial redshift chosen for CITE in this work) on the thermal evolution at  $2 \leq z \leq 4$ . Recently Gaikwad et al. (2020) measured  $T_0$  ( $\gamma$ ) at  $z = 6$  and found it to be larger than predicted by the HM12; OH17; KS19; FG20 UVB models. Since we run CITE from  $z = 6$  to  $z = 2$ , it is possible that an uncertainty in  $T_0$  and  $\gamma$  at  $z = 6$  affects the thermal parameter evolution at  $2 \leq z \leq 4$ . Fig. 11 shows that the effect of even a large uncertainty in thermal

parameters at  $z = 6$  is subdominant at lower redshifts. This is because the lower redshift thermal evolution is mainly affected by the He II reionization prescriptions used in these models. Thus the effect of uncertainty in  $T_0, \gamma$  at  $z = 6$  is less than 4 percent (maximum) at  $2 \leq z \leq 4$ . Note that even though we show this result only for the FG20 UVB models, the results are similar for the other UVB models.

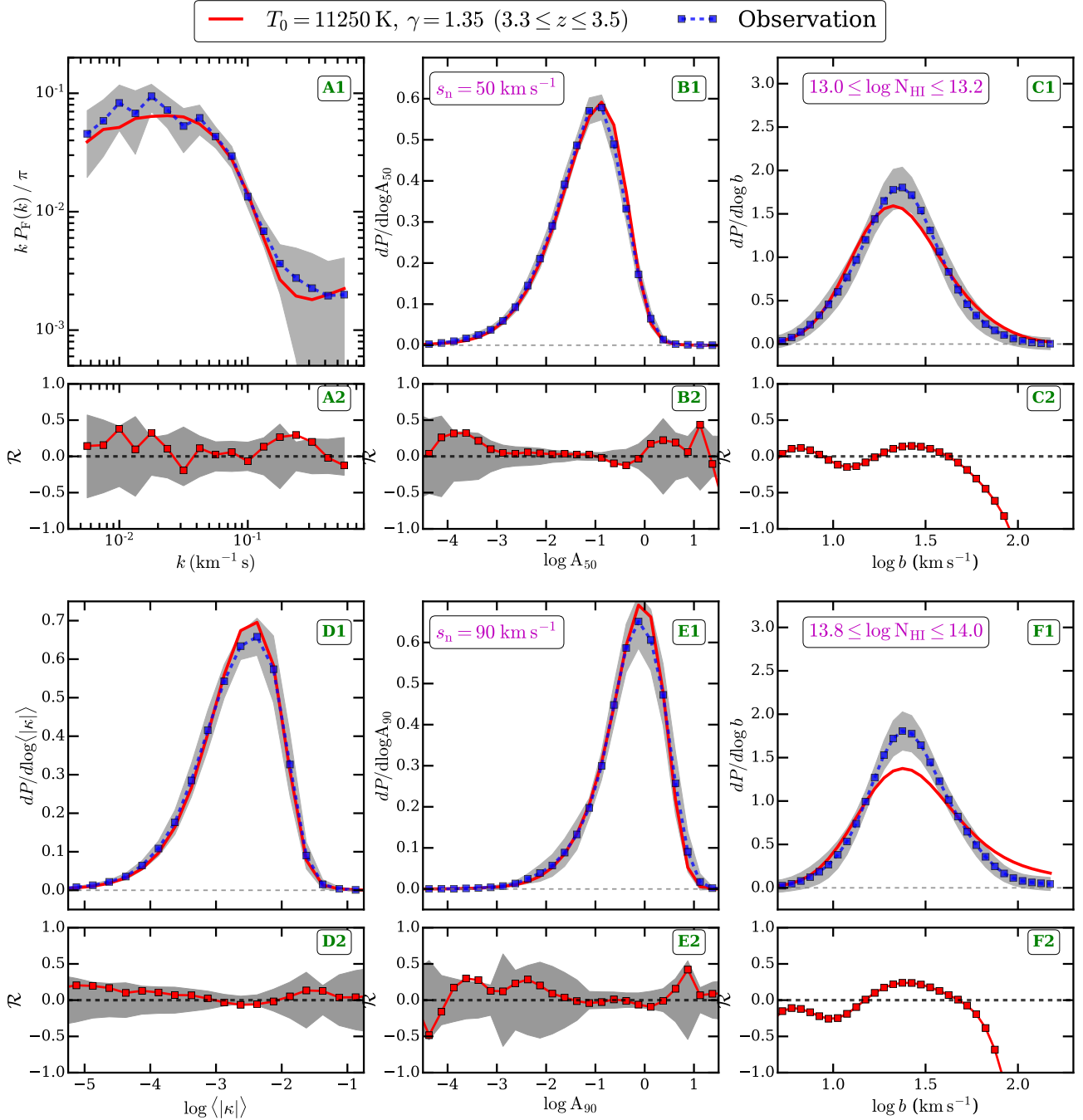


Figure G6. Same as Fig. 12 except for  $3.1 \leq z \leq 3.3$ .

## I2 Effect of QSO spectral energy distribution on thermal parameter evolution at $2 \leq z \leq 4$

The observed QSO spectral index ( $\alpha$ ,  $f_\nu \propto \nu^{-\alpha}$ ) is another uncertain quantity in UVB calculations. The observations of stacked QSO spectra in the far ultra-violet suggests that QSO spectral indices vary from  $\alpha = 1.4$  to  $2.0$  (Lusso et al. 2014; Stevans et al. 2014; Shull & Danforth 2020). Recently Gaikwad et al. (2019) showed that the effect of the QSO spectral index on thermal parameter evolution is degenerate with the redshift of He II reionization in KS19 such that a flatter QSO SED corresponds to an earlier He II reioniza-

tion. They further showed that  $\alpha = 1.8$  is consistent with the evolution of  $\tau_{\text{eff,HeII}}$  from Worseck et al. (2019). In Fig. 13, we show the comparison of our thermal parameter measurements with that predicted by the KS19 UVB for different QSO spectral indices. Irrespective of the QSO spectral index, the total emissivity is kept constant in all the models. The shape of the thermal parameter evolution is hence similar for all the models. The redshift of maximum  $T_0$  occurs at slightly higher redshift and is slightly higher for flatter (small  $\alpha$ ) QSO SED. The earlier temperature rise is mainly because of the  $f_{\text{HeII}}$  evolution (see panel B1 Fig. 13) in these models which is also different such that He II reionization

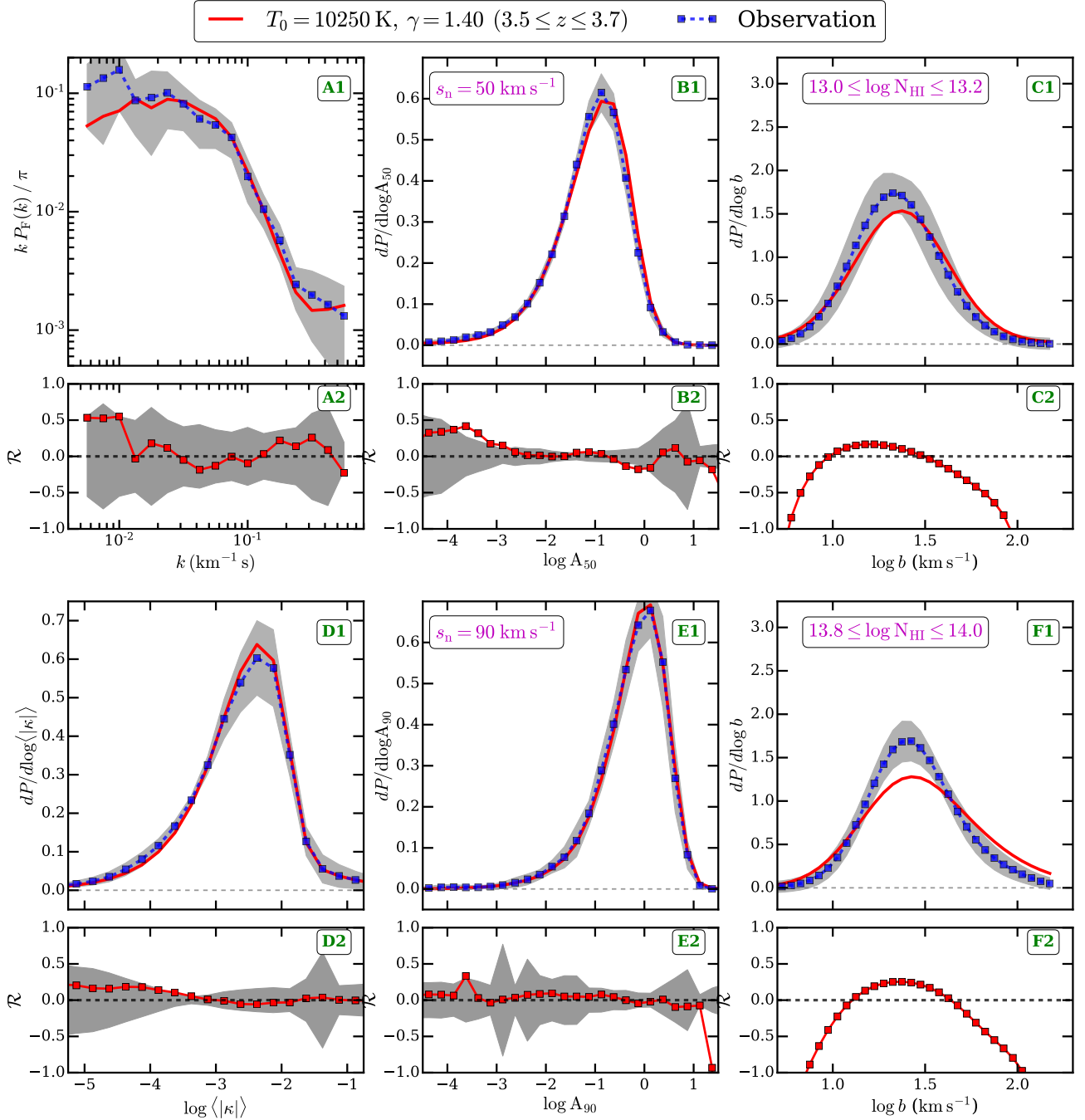


**Figure G7.** Same as Fig. 12 except for  $3.3 \leq z \leq 3.5$ .

occurs slightly earlier for flatter QSO SED (Gaikwad et al. 2019). However, even after changing the QSO SED spectral index in the observationally allowed range, the thermal parameter evolution in the KS19 UVB model is significantly different from our measurements. This suggests that additional modifications (such as  $\eta$  evolution) may be needed for the KS19 UVB models to become consistent with our temperature measurements.

### I3 Modifications to UVB models

Recently FG20 updated their Faucher-Giguère et al. (2008) UVB model with updated constraints on galaxy and AGN luminosity functions (Kulkarni et al. 2019), intergalactic H I and He II opacity models, and calibrations to measurements of  $\Gamma_{\text{HI}}$  and  $\Gamma_{\text{HeII}}$  etc. In their UVB model, AGN (galaxy) dominate the H I ionizing background at  $z \leq 3$  ( $z > 3$ ). Their UVB model crucially depends on parameters related to the reionization history mainly the total temperature increments due to H I ( $\Delta T_{\text{HI}}$ ) and He II ( $\Delta T_{\text{HeII}}$ ) reionization heating. The H I and He II photo-heating rates are directly

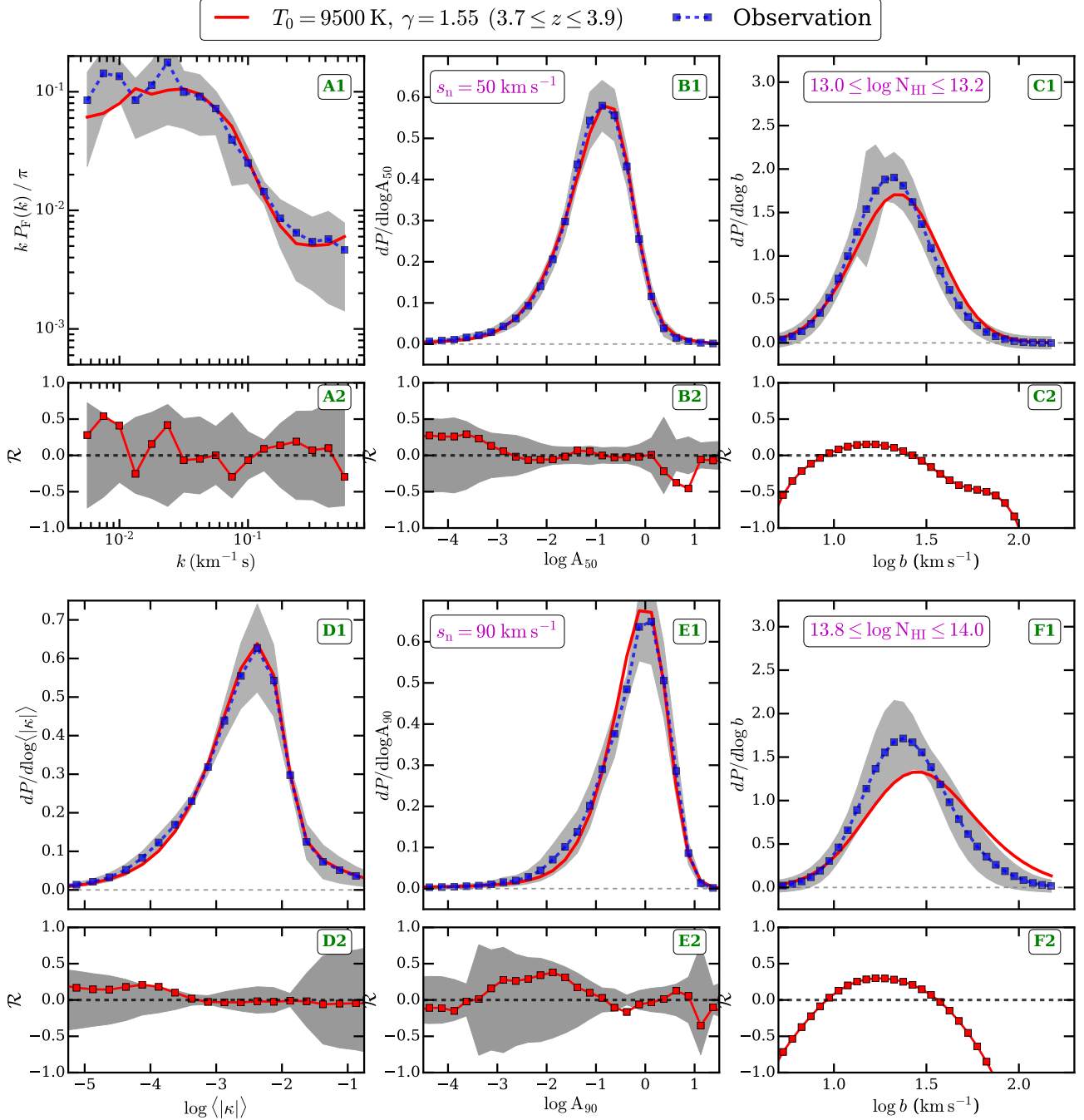


**Figure G8.** Same as Fig. 12 except for  $3.5 \leq z \leq 3.7$ .

proportional to  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$  respectively. Since the timing and amount of photo-heating during HeII reionization is poorly constrained at present, they parameterize this ignorance in  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$ . OH17; FG20 assumed  $\Delta T_{\text{HI}} = 20000 \text{ K}$  and  $\Delta T_{\text{HeII}} = 15000 \text{ K}$  and obtain a  $T_0$  and  $\gamma$  evolution similar to previous measurements. However, both groups assumed that photo-ionization equilibrium is valid at all times. But as shown by Puchwein et al. (2015); Gaikwad et al. (2018); Puchwein et al. (2019), this is not a good assumption during HeII reionization as the photo-ionization time scales are shorter than the recombination time scales. Thus non-equilibrium ionization evolution is more physically

motivated. Puchwein et al. (2015); Gaikwad et al. (2018) shows that the  $T_0$  of the IGM in non-equilibrium is significantly larger than in equilibrium calculations. Thus when we use non-equilibrium ionization evolution for OH17; FG20, which are calibrated to match observation assuming photo-ionization equilibrium, we get significantly larger temperature predictions for their UVB models. In order to match the observations with non-equilibrium ionization evolution, we need to reduce  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$ . The reduction of  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$  by a factor of  $f_{\text{scale}}$  is equivalent to reducing HI, HeI and HeII photo-heating rates by a factor  $f_{\text{scale}}$ .

Fig. 11 shows the variation in  $T_0$  and  $\gamma$  evolution with

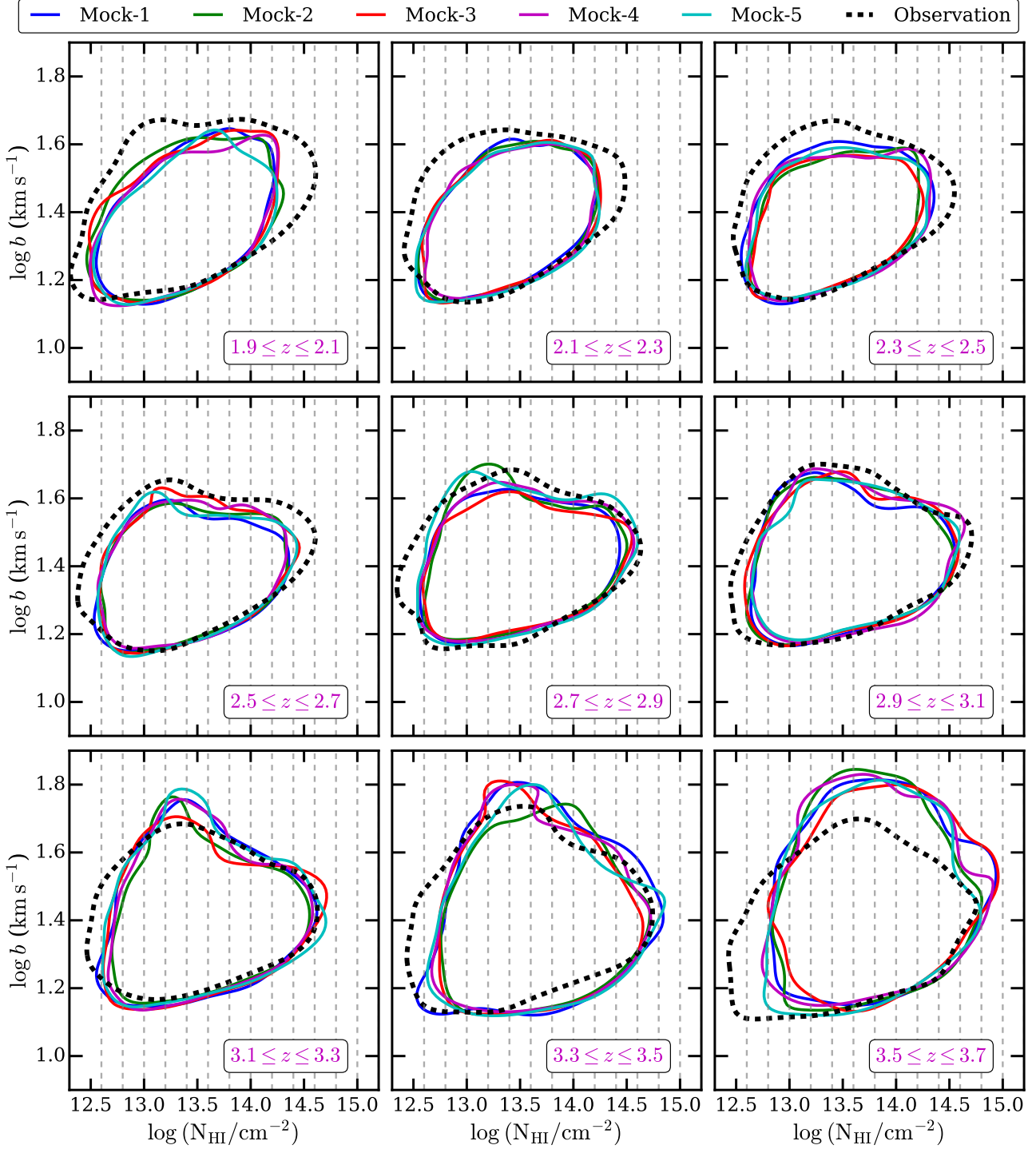


**Figure G9.** Same as Fig. 12 except for  $3.7 \leq z \leq 3.9$ .

the variation in  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$  predicted by the FG20 model using GADGET-3 + CITE. We show results in Fig. 11 for four  $f_{\text{scale}} = 1.0, 0.8, 0.7, 0.6$  for H I, He I and He II photo-heating rates. The main effect of changing this parameter is to reduce  $T_0$  systematically at all redshifts while  $\gamma$  changes very little. This is expected as scaling the parameters  $\Delta T_{\text{HI}}$ ,  $\Delta T_{\text{HeII}}$  is equivalent to scaling H I and He II photo-heating rates and such scaling changes  $T_0$  (keeping  $\gamma$  relatively same see Fig. D1). Fig. 11 shows that the  $T_0$  and  $\gamma$  evolution is consistent with our measurements if we use non-equilibrium ionization evolution with  $f_{\text{scale}} = 0.7$  parameters in FG20. For completeness, we summarize the best fit  $\Delta T_{\text{HI}}$  and  $\Delta T_{\text{HeII}}$

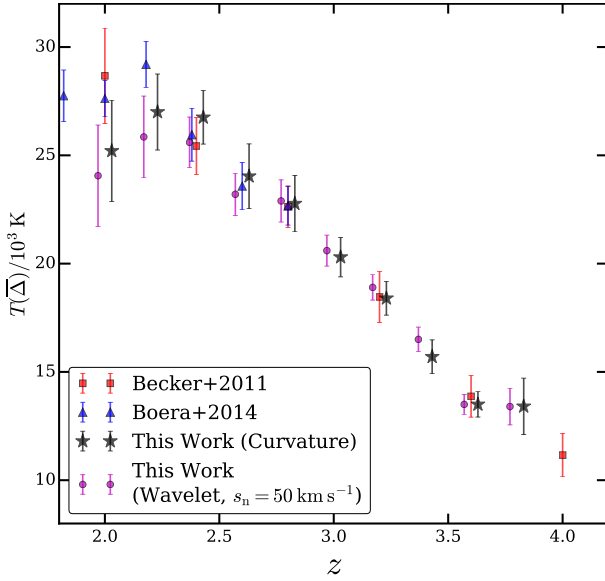
parameters used in OH17; FG20 in Table 11 that are consistent with our temperature measurements. In order to do a fair comparison with other UVB models, we show the  $T_0$  and  $\gamma$  evolution for varying the photo-heating rates in Fig. 12. We find that other UVB models match the  $T_0$  evolution in some redshift bins but are systematically above or below in other redshift bins. The evolution of  $\gamma$  is similar to that shown in Fig. 18.

Finally, we compare the evolution of  $T_0$  and  $\gamma$  over the full redshift range  $2 \leq z \leq 6$  and all the different flux statistics in Fig. 14. We emphasize that a considerable amount of work has also been done to measure the  $T_0$  and  $\gamma$  evolution

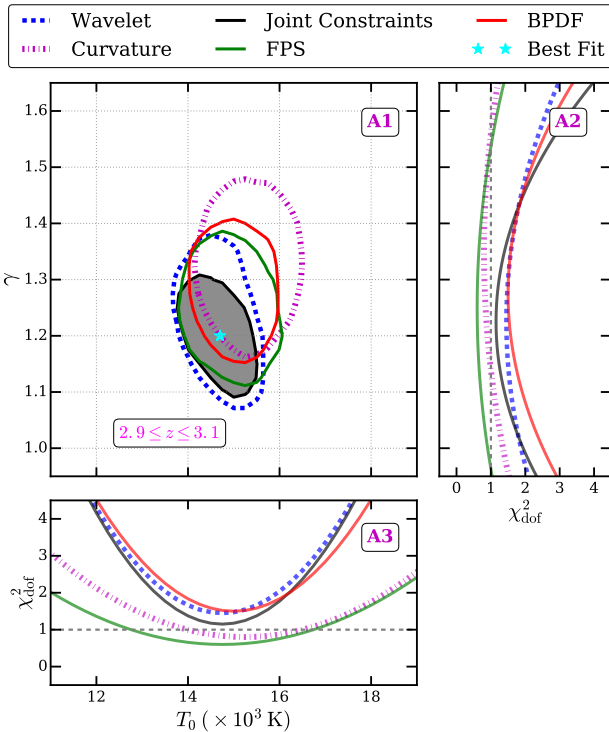


**Figure G10.** Each panel shows the contour enclosing 90 percent of the points in the  $b - \log N_{\text{HI}}$  plane. Black dashed contours show the distribution of  $b - \log N_{\text{HI}}$  points from observed samples in this work. The solid contours (blue, green, red, magenta and cyan) show the regions of the  $b - \log N_{\text{HI}}$  distribution for our mock samples. Each mock sample matches the total observed redshift path length in a given redshift bin. For visual purposes, we show contours for 5 mock samples. However, we have computed  $b$  distribution for 100 such mock samples. The thermal parameters of the mock samples in each redshift bin correspond to the best fit thermal parameters obtained in this work. The  $b - \log N_{\text{HI}}$  distributions from our mock samples is similar to the observed  $b - \log N_{\text{HI}}$  distribution. The shape of the contours for lower  $b$  values from the simulations are in good agreement with that from observations. Traditionally the lower envelope of  $b$  values is used to measure the thermal parameters. The variations in the  $b - \log N_{\text{HI}}$  distribution in different mock samples are small indicating that the distribution is converged with respect to cosmic variance.





**Figure G11.** The figure shows a comparison of temperature at the characteristic density,  $T(\bar{\Delta})$  from this work with that from Becker et al. (2011); Boera et al. (2014). Our temperature measurements from the curvature (black stars) and wavelet (magenta circles) statistics are in good agreement with those from the literature at  $z > 2.4$ . However, our measurements at  $z < 2.4$  are systematically lower than those in the literature, but are consistent within  $1.5\sigma$ .



**Figure H1.** Each panel is same as Fig. 11 except in panel A2 and A3, we show reduced  $\chi^2$  ( $\chi^2_{\text{dof}}$ ). The minimum  $\chi^2_{\text{dof}}$  is close to  $\sim 1$  suggesting that the errors estimated on statistics are not overestimated/underestimated.

**Table I1.** Heat injection parameters ( $\Delta T_{\text{HI}}$ ,  $\Delta T_{\text{HeII}}$ ) in UVB models consistent with the measured  $T_0$  and  $\gamma$  evolution from this work.

UVB model	$\Delta T_{\text{HI}}$ (K)	$\Delta T_{\text{HeII}}$ (K)
Oñorbe et al. (2017)	$15200 \pm 1800$	$11400 \pm 2100$
Faucher-Giguère (2020)	$13600 \pm 2400$	$10200 \pm 1800$

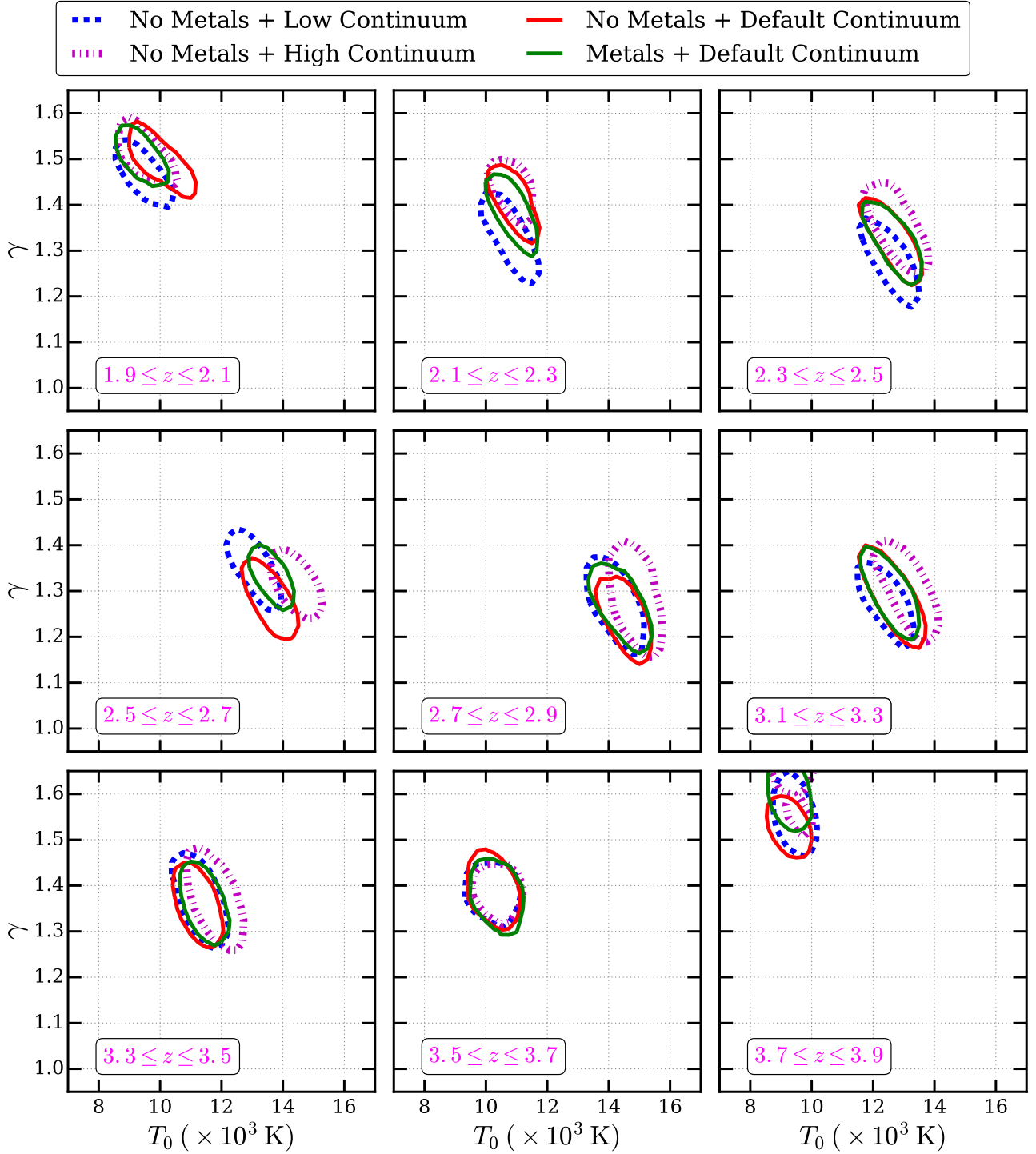
at  $z > 4$ . At these redshifts the thermal parameters have been measured using FPS Walthier et al. (2019); Boera et al. (2019), curvature statistics Becker et al. (2011), proximity effects Bolton et al. (2012) and the width distribution of transmission spikes Gaikwad et al. (2020). These measurements broadly agree with the theoretical  $T_0$  and  $\gamma$  evolution predicted by the different UVB models.

#### I4 The way forward for UVB models

We have rescaled the photo-heating rates in §6.2 to better match our temperature measurements. Even though this is a simple modification to the UVB models, this is not necessarily physical because photo-ionization and photo-heating rates are ultimately linked to the population of ionising sources (QSOs and /or galaxies) and IGM properties. In this section we discuss the possible modifications that may be needed in the various UVB models to match our measurements. Note, however, that implementing these in the UVB models is beyond the scope of present work hence we discuss qualitative features of a physically motivated UVB model that could match our measurements.

From Fig. 18, it is clear that the maximum in  $T_0$  occurs at systematically lower redshift than that from HM12; KS19 UVB models. However, the maximum value of  $T_0$  from these models seem to agree well with that from observations. This suggests that the normalization of the QSO emissivity in these models is roughly consistent with observations. However, the He II reionization in these models starts early and is more gradual than suggested by observations (see panel B1 in Fig. 19). Thus, in HM12; KS19 the redshift evolution of  $f_{\text{HeII}}$  needs to be modified. In KS19, this could be achieved by changing the ratio of He II and H I fractions ( $\eta$ ). Since there are only observational limits on  $f_{\text{HeII}}$  and the fluctuations in He II ionizing background are important at  $z > 3$ , there is scope for modification in the  $f_{\text{HeII}}$  evolution of the KS19 UVB. The  $f_{\text{HeII}}$  evolution can also be modified by varying QSO SED index ( $\alpha$ ,  $f_\nu \propto \nu^{-\alpha}$  Gaikwad et al. 2019). However, we show in online supplementary appendix I3 that the observed thermal parameter evolution can not be matched by simply varying the QSO SED index within the observationally allowed range of  $\alpha = 1.4 - 2.0$  (Lusso et al. 2014; Stevans et al. 2014; Shull & Danforth 2020). For the HM12 UVB model both the QSO luminosity functions and QSO SED index need to be updated to be compatible with the most recent compilations of the relevant observation (see e.g. Kulkarni et al. 2019).

In summary, the observed thermal parameter evolution from this work suggests a combination of changes to existing UVB models. For HM12; KS19 UVB models, modification in  $f_{\text{HeII}}$  is necessary. For the OH17; P19; FG20 UVB models modification of the ionizing emissivity from QSOs is needed.



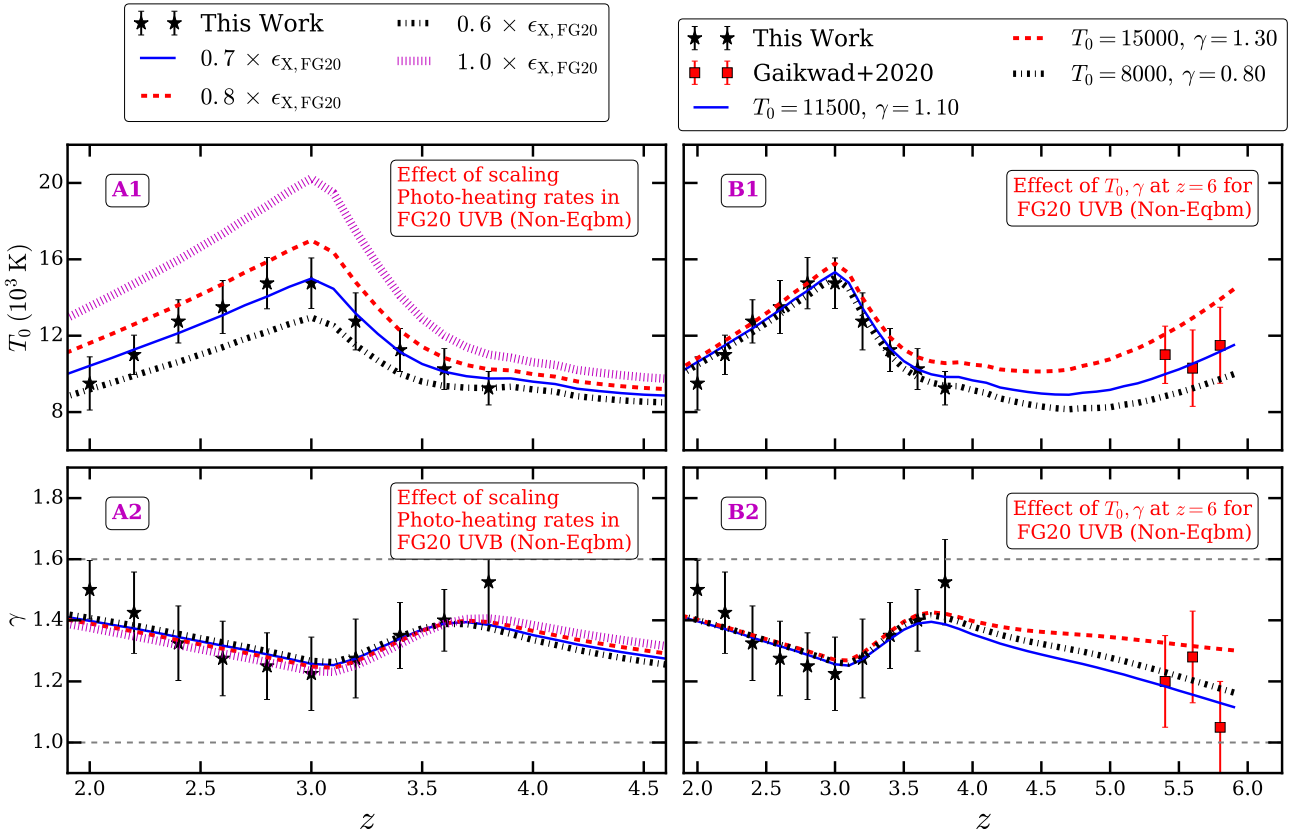
**Figure H2.** The figure shows the effect of continuum placement uncertainty and metal contamination on the  $T_0$  and  $\gamma$  constraints. Each panel is similar to panel A1 in Fig. 15 and shows the other 8 redshift bins spanning  $1.9 \leq z \leq 3.9$ .

## APPENDIX J: CUMULATIVE ENERGY PARAMETER

The default photo-heating rates of all the UVB models predict systematically higher  $T_0$  than we measure at  $z > 3$  (see Fig. 18). Since we have changed the photo-heating rates to better match the observed  $T_0$  evolution, it is also interesting to have a look at the evolution of the cumulative energy ( $u_0$ )

deposited into IGM by He II reionization. We calculate the cumulative energy per unit mass ( $u_0$ ) deposited into IGM as

$$u_0 = \int_{z=6}^{z_{re}} \frac{\mathcal{H}}{\bar{\rho}} \cdot \frac{dz}{H(z)(1+z)} \quad (\text{J1})$$



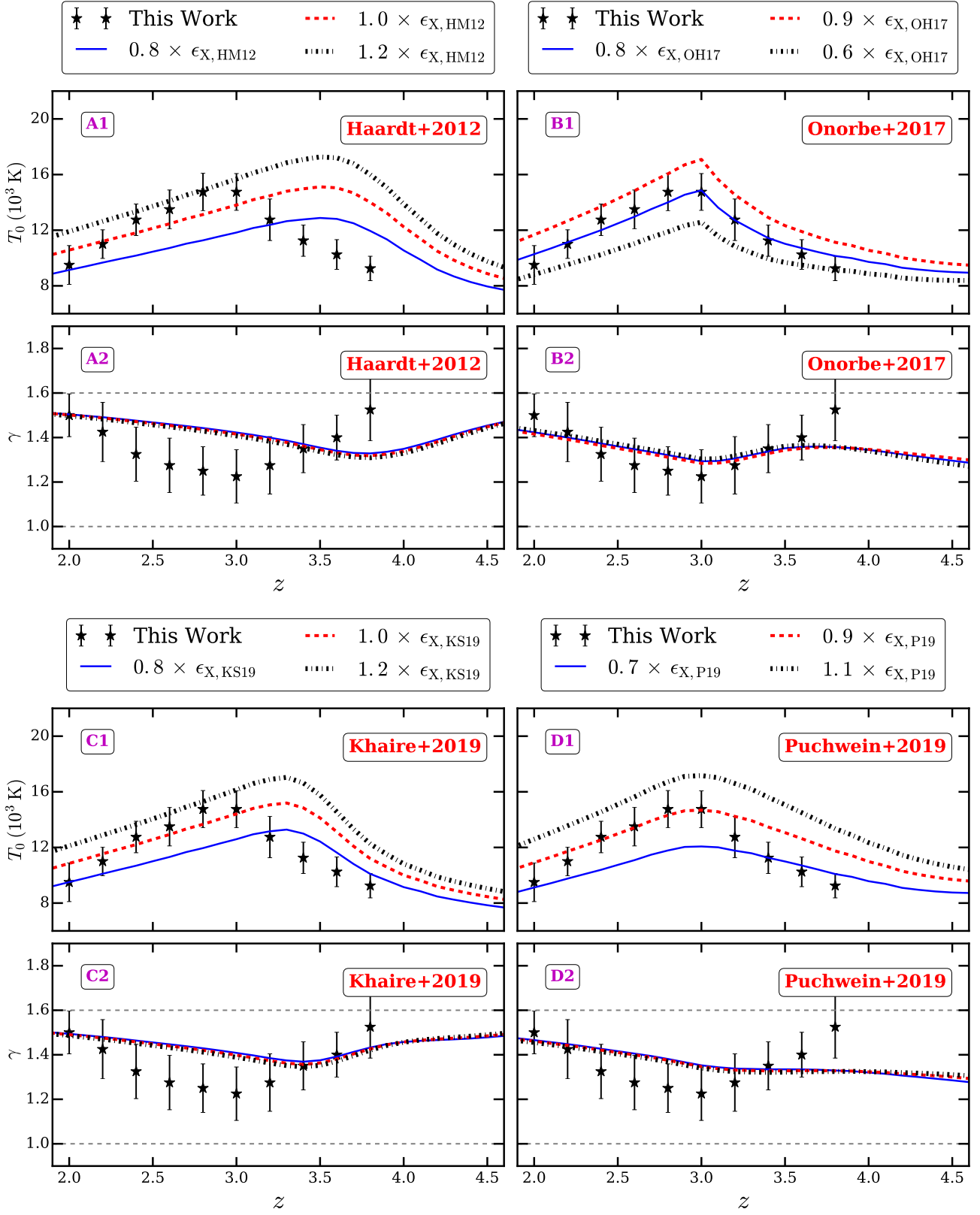
**Figure I1.** Each panel here is the same as Fig. 18 except the sensitivity of thermal parameters to the photo-heating rates and initial  $T_0$  and  $\gamma$  at  $z = 6$  is illustrated for the FG20 non-equilibrium model. Since the  $T_0$  evolution for the default FG20 model is systematically higher than our measurements, we rescale the H I, He I and He II photo-heating rates in panel A1 and A2. The observed thermal parameter evolution (along with  $1\sigma$  uncertainty) from this work is in good agreement with the non-equilibrium prediction of the FG20 UVB model if photo-heating rates are reduced by a factor of  $0.7 \pm 0.1$ . A simple scaling of photo-heating rates does not affect the  $\gamma$  evolution which is consistent with Fig. D1. Given that the uncertainty in thermal parameters after H I reionization is large, panel B1, B2 show the effect of initial  $T_0$  and  $\gamma$  at  $z \sim 6$  on the  $T_0, \gamma$  evolution at late times  $2 \leq z \leq 4$  (Gaikwad et al. 2020). The uncertainty in  $T_0, \gamma$  at  $z \sim 6$  has only a small (less than 3 percent at  $z \sim 3$ ) effect on the  $T_0$  and  $\gamma$  evolution at late times. This suggests that the thermal effect of H I reionization at  $2 \leq z \leq 4$  is sub-dominant relative to that of He II reionization.

where  $\mathcal{H} = \sum n_{\text{XI}} \epsilon_{\text{XI}}$  (with  $\text{XI} = \text{H I}, \text{He I}, \text{He II}$ ),  $\bar{\rho}, H(z), n_{\text{XI}}, \epsilon_{\text{XI}}$  are mean baryon density, Hubble constant, number density of species XI and photo-heating rate of specie XI (Nasir et al. 2016). To calculate  $u_0$  for a simulation box with given UVB and at a given redshift, we first calculate the volume average neutral fraction  $f_{\text{XI}}$ . We then use above expression to compute  $u_0$ . Given that there could be uncertainty in  $u_0$  after H I reionization, we chose  $u_0(z = 6) = 0$  as a reference point. The  $u_0$  defined in this work is mainly sensitive to heat deposited during He II reionization.

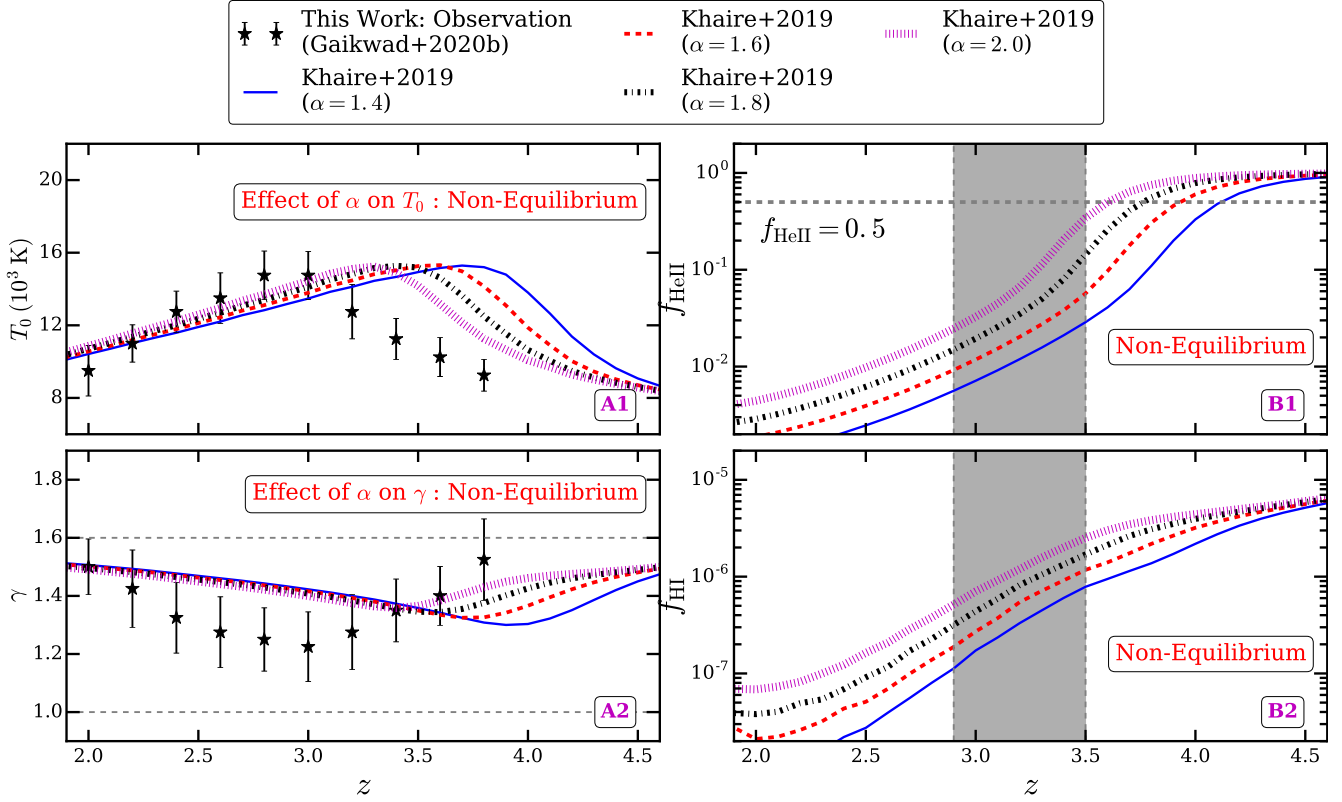
In Fig. J1 we show the evolution of  $u_0$  for all the UVB models with non-equilibrium ionization evolution. The observed  $u_0$  evolution is difficult to measure directly from observations. To nevertheless facilitate a comparison, we show the  $u_0$  evolution in Fig. J1 for the FG20 non-equilibrium UVB with scaled photo-heating rates that matches the observed  $T_0 - \gamma$  evolution from this work. We treat the  $u_0$  evolution corresponding to this model as that inferred from our temperature measurements (black dashed line with shaded region).

Fig. J1 also shows the  $u_0$  evolution in all the UVB

models assuming non-equilibrium ionization evolution. The photo-heating rates here are not scaled for the UVB models. At  $z = 3$  the  $u_0$  predicted by all the UVB models is systematically higher than the  $u_0$  evolution inferred from our temperature measurements (Table 4). Thus to match the observed  $T_0 - \gamma$  evolution smaller values of energy deposited per unit mass  $2.13 \pm 0.32 \text{ eV } m_p^{-1}$  at  $z \sim 3$  are required. It is interesting to note that  $u_0$  at  $z = 3$  from the KS19 model is close to that from a model which is consistent with our measurements. However, the  $T_0$  and  $\gamma$  evolution from the KS19 model is significantly different at  $z > 3$ . This is because  $u_0$  accounts for the heating of the IGM, but does not account for the cooling of the IGM due to recombination, Hubble expansion or inverse Compton cooling etc. Hence to match the  $T_0 - \gamma$  evolution, the whole of the  $u_0$  evolution is important and not only matching  $u_0$  at some redshift. The comparison of  $u_0$  evolution in Fig. J1 again suggests that rather rapid He II reionization is needed to match our measured  $T_0 - \gamma$  evolution. The  $u_0$  evolution in HM12; KS19; P19 is more gradual compared to the other two models.



**Figure I2.** Same as panel A1 and A2 in Fig. I1 except photo-heating rates are varied here for the [HM12](#); [OH17](#); [KS19](#); [P19](#) UVB models.

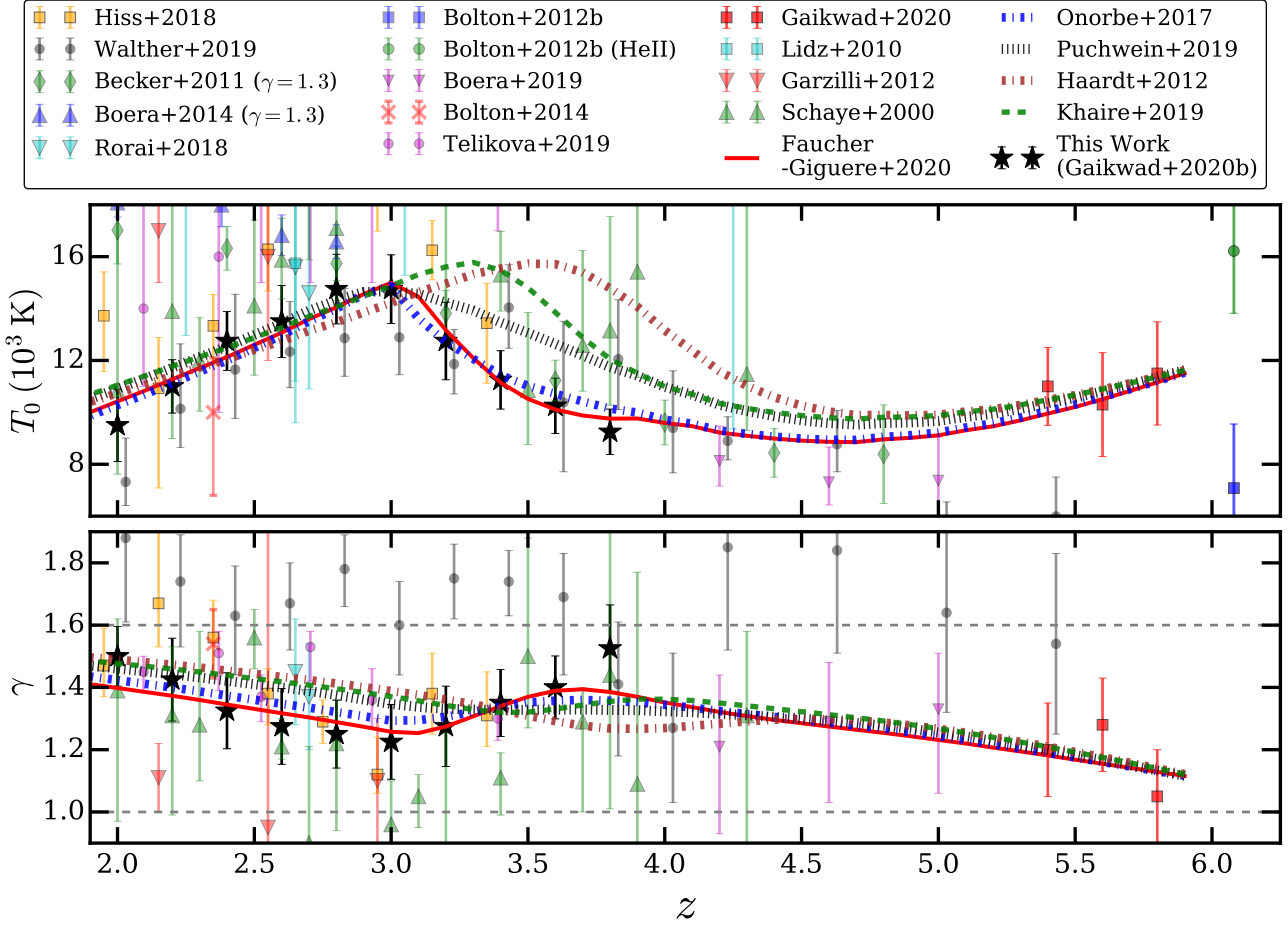


**Figure I3.** Each panel here is the same as Fig. 19 except the variation of thermal parameters and H I, He II fractions with QSO spectral index ( $f_\nu \propto \nu^{-\alpha}$ ) in KS19 is shown. Flatter slopes of the QSO spectra correspond to earlier He II reionization consistent with Gaikwad et al. (2019). All the models predict similar  $T_0$  and  $\gamma$  evolution at  $z < 3$ .

## REFERENCES

- Becker G. D., Bolton J. S., Haehnelt M. G., Sargent W. L. W., 2011, *MNRAS*, **410**, 1096
- Boera E., Murphy M. T., Becker G. D., Bolton J. S., 2014, *MNRAS*, **441**, 1916
- Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, *ApJ*, **872**, 101
- Bolton J. S., Becker G. D., Raskutti S., Wyithe J. S. B., Haehnelt M. G., Sargent W. L. W., 2012, *MNRAS*, **419**, 2880
- Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, *MNRAS*, **464**, 897
- Calura F., Tescari E., D’Odorico V., Viel M., Cristiani S., Kim T. S., Bolton J. S., 2012, *MNRAS*, **422**, 3019
- Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, *ApJ*, **581**, 20
- D’Aloisio A., McQuinn M., Trac H., Cain C., Mesinger A., 2020, arXiv e-prints, p. [arXiv:2002.02467](https://arxiv.org/abs/2002.02467)
- Day A., Tytler D., Kambalur B., 2019, *MNRAS*, **489**, 2536
- Faucher-Giguère C.-A., 2020, *MNRAS*, **493**, 1614
- Faucher-Giguère C.-A., Lidz A., Hernquist L., 2008, *Science*, **319**, 52
- Gaikwad P., Choudhury T. R., Srianand R., Khaire V., 2018, *MNRAS*, **474**, 2233
- Gaikwad P., Srianand R., Khaire V., Choudhury T. R., 2019, *MNRAS*, **490**, 1588
- Gaikwad P., et al., 2020, arXiv e-prints, p. [arXiv:2001.10018](https://arxiv.org/abs/2001.10018)
- Gnedin N. Y., Hui L., 1998, *MNRAS*, **296**, 44
- Haardt F., Madau P., 2012, *ApJ*, **746**, 125
- Khaire V., Srianand R., 2019, *MNRAS*, **484**, 4174
- Kim T. S., Viel M., Haehnelt M. G., Carswell R. F., Cristiani S., 2004, *MNRAS*, **347**, 355
- Kim T. S., Bolton J. S., Viel M., Haehnelt M. G., Carswell R. F., 2007, *MNRAS*, **382**, 1657
- Kim T. S., Partl A. M., Carswell R. F., Müller V., 2013, *A&A*, **552**, A77
- Kulkarni G., Hennawi J. F., Oñorbe J., Rorai A., Springel V., 2015, *ApJ*, **812**, 30
- Kulkarni G., Worseck G., Hennawi J. F., 2019, *MNRAS*, **488**, 1035
- Lusso E., et al., 2014, *ApJ*, **784**, 176
- Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, *MNRAS*, **482**, 3458
- Nasir F., Bolton J. S., Becker G. D., 2016, *MNRAS*, **463**, 2335
- Oñorbe J., Hennawi J. F., Lukić Z., 2017, *ApJ*, **837**, 106
- Oñorbe J., Davies F. B., Lukić Z., Hennawi J. F., Sorini D., 2019, *MNRAS*, **486**, 4075
- O’Meara J. M., et al., 2015, *AJ*, **150**, 111
- O’Meara J. M., Lehner N., Howk J. C., Prochaska J. X., Fox A. J., Peebles M. S., Tumlinson J., O’Shea B. W., 2017, *AJ*, **154**, 114
- Park H., Shapiro P. R., Choi J.-h., Yoshida N., Hirano S., Ahn K., 2016, *ApJ*, **831**, 86
- Planck Collaboration et al., 2014, *A&A*, **571**, A16
- Puchwein E., Bolton J. S., Haehnelt M. G., Madau P., Becker G. D., Haardt F., 2015, *MNRAS*, **450**, 4081
- Puchwein E., Haardt F., Haehnelt M. G., Madau P., 2019, *MNRAS*, **485**, 47
- Rollinde E., Theuns T., Schaye J., Pâris I., Petitjean P., 2013, *MNRAS*, **428**, 540
- Schaye J., 2001, *ApJ*, **559**, 507
- Shull M., Danforth C., 2020, arXiv e-prints, p. [arXiv:2007.02948](https://arxiv.org/abs/2007.02948)
- Stevens M. L., Shull J. M., Danforth C. W., Tilton E. M., 2014,





**Figure 14.** Comparison of  $T_0 - \gamma$  measurements from this work with measurements in the literature. We also show the theoretical evolution of thermal parameters predicted by the modified OH17; P19; FG20 UVB models. For the OH17; P19; FG20 UVB models, we have rescaled the photo-heating rates by a factor of 0.8, 0.9, 0.7, respectively.

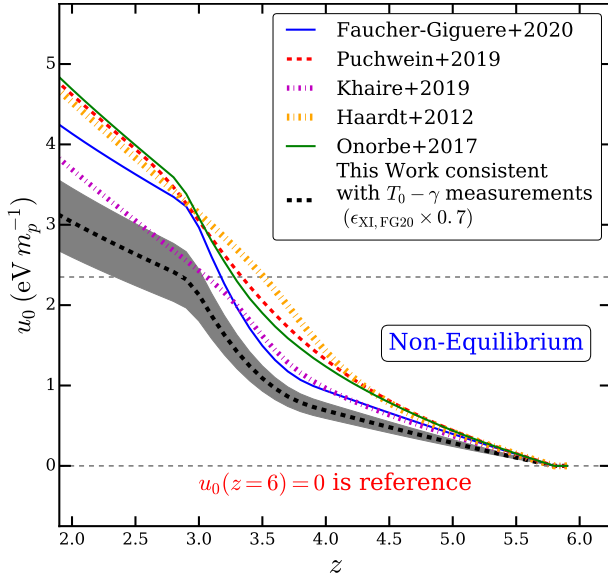
[ApJ, 794, 75](#)

Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O’Meara J., 2018, [ApJ, 852, 22](#)

Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, [ApJ, 872, 13](#)

Worseck G., Davies F. B., Hennawi J. F., Prochaska J. X., 2019, [ApJ, 875, 111](#)

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.



**Figure J1.** The evolution of the cumulative energy deposited into the IGM per unit mass ( $u_0$ ) for different non-equilibrium UVB models is shown. We chose  $u_0(z = 6) = 0$  as reference point to minimize the effect of uncertainty in  $u_0$  due to H I reionization. The  $u_0$  described in this work is primarily due to He II reionization. The black dashed line and the shaded region show the  $u_0$  evolution consistent with the measured  $T_0 - \gamma$  evolution from this work. Without re-scaling the photo-heating rate of all the UVB models predict systematically larger  $u_0$  than consistent with our temperature measurements. Not that the  $u_0$  evolution accounts for the heating but not the cooling of the IGM. Hence to match the  $T_0 - \gamma$  evolution, the whole of the  $u_0$  evolution is important and not just matching  $u_0$  at some redshift. A sharp increase in the  $u_0$  inferred from our temperature measurements is evident at  $3.0 \leq z \leq 3.6$ .