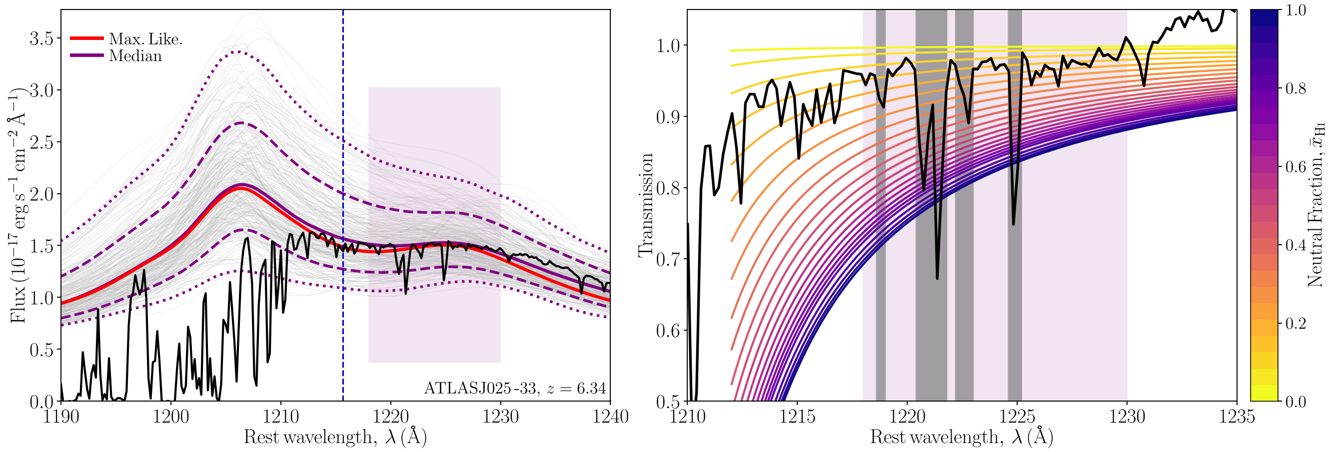

Figure 3. The same as Fig. 1 except after recalibrating our reconstruction pipeline using template spectra from two different QSO samples (see the text for further details).

and thus more likely to be the cause of the flux decrement than the broad Lyα component. Importantly, given that these components are sufficient for fitting QSOs within this spectral region (i.e. in our training set), it may also indicate that there are higher order correlations between our emission line parameters we are missing within our covariance matrix approach. In future, we aim to explore this further but for the purposes of this work, we will recalibrate our profiles to mitigate this feature.

It is worth pointing out that this feature was not identified in our previous study of $z > 7$ QSOs for several reasons. Other than for ULASJ1120 + 0641, there is no clear/prominent N V line in the data and further a very broad Lyα line is preferred. As a result, there was limited opportunity for this feature to manifest as it requires a notable N V line component. For ULASJ1120 + 0641, the relatively broad observed N V line profile extends sufficiently redward to benefit from our prior on the observed QSO flux. Therefore, this prior enables us to predict a relatively broad N V line component, limiting the potential impact of this feature. Thus, only with the exquisite quality of the X-shooter spectra coupled with the more prevalent N V line features has this feature become apparent.

It is important to note that this particular problem is intrinsic to our reconstruction method, as it relies upon the sum of three predicted Gaussian components to describe the Lyα + N V line complex. Other reconstruction pipelines in the literature provide flux predictions based off the sum of components following principal component analysis (PCA) or profiles drawn from machine learning approaches which are more agnostic to the direct emission line correlations (e.g. Davies et al. 2018b; Fathivavsari 2020; Reiman et al. 2020; Ďurovčíková et al. 2020; Bosman et al. 2021; Liu & Bordoloi 2021; Chen et al. 2022; Sun et al. 2023). As a result, these appear to be less susceptible to such systematic biases in the reconstructed profiles (see Greig et al. 2024, for more detailed discussions).

Crucially, and relevant for this work, the consistency of the flux ratios across the two samples implies that we should be able to recalibrate our reconstruction pipeline to correct for this systematic offset in our reconstruction profiles. To do so, we convolve each random draw from our covariance matrix reconstruction by the distribution of flux ratios shown in Fig. 2 (both from the X-shooter and SDSS sample). A consequence of this recalibration is that the distribution of reconstructed profiles will become broader, reflecting the increased scatter injected by sampling from these flux ratios and subsequently also broader uncertainties on the inferred IGM neutral fraction. Note, this convolution only occurs once we have corrected for the blueshift of the reconstructed QSO. That is, we shift these flux ratios to be centred on the location of the Lyα profile estimated from the reconstruction pipeline to ensure we are correctly recalibrating over the problematic region between Lyα and N V.

In Fig. 3, we demonstrate the performance of this recalibration on the QSO, ATLAS J025 − 33, (as shown in Fig. 1). In the left panel, we present our calibrated reconstruction and in the right panel, the corresponding Lyα transmission following this recalibration. Note, for the ML profile (red curve) we simply multiply it by the median flux ratio from Fig. 2 for visualization purposes. In the left panel, we can easily identify the two main aspects of this recalibration procedure. First, the shape of the median profile between 1210 and 1220 Å no longer possesses the parabolic dip observed in Fig. 1 (more readily visible in the right panel) and secondly, the broader distribution (and increased diversity) of the random profiles draws from our covariance matrix reconstruction. Note, that by performing our recalibration for this particular object, the predicted median flux is now below the observed spectrum at $\gtrsim 1230$ Å. However, this is not too concerning as it is at the redward edge of the N V line, whereas the purpose of the recalibration is to improve the reconstructions more so on the blue side of N V and the redward edge of Lyα (the joint contribution of the Lyα and N V line components). Nevertheless, this will be explored in greater detail in future work when working to improve on this QSO reconstruction method.

The right panel of Fig. 3 more readily demonstrates the improvements this recalibration brings to our reconstruction pipeline. Correcting this systematic offset in the shape of the reconstructed profiles produces corresponding Lyα transmission profiles which now better match the expected shape of the synthetic IGM damping wing profiles. That is, the observed Lyα transmission is now monotonically increasing with rest-frame wavelength, enabling more robust template matching against the synthetic sampling wing profiles drawn from the EoR simulations.

3.5 Joint fitting to obtain IGM neutral fraction constraints

We obtain our constraints on the IGM neutral fraction by jointly sampling our synthetic damping wing profiles from our EoR simulations convolved with our recalibrated intrinsic Lyα profile reconstructions. Our fitting pipeline is as follows:

(i) We draw $\sim 10^5$ intrinsic QSO profile estimates from the reconstructed Lyα and N V line profiles (see Section 3.1).

(ii) We convolve these $\sim 10^5$ reconstructed profiles by the blueshift corrected template flux ratios to recalibrate for a systematic $\sim$10 per cent underestimate of the intrinsic flux as outlined in Section 3.4 leading to $\sim 10^6$ corrected reconstruction profiles.

(iii) Each intrinsic profile is then multiplied by 10^5 synthetic damping wing opacities following Section 3.2. This results in $\sim 10^{11}$ mock spectra for each $\bar{x}_{\rm HI}$ snapshot from the EoR simulation.

(iv) Each $\sim 10^{11}$ mock spectra is then individually compared to the observed QSO spectrum over the $\lambda = 1218$–1230 Å region (consistent with Greig et al. 2017b; Greig et al. 2019; Greig et al. 2022). For each, we calculate a χ^2 relative to observed flux and the error spectrum. Where appropriate, absorption features or other problematic regions of the observed spectrum over this fitting range are masked out following visual inspection that could otherwise incorrectly bias our results.

(v) We then obtain a likelihood for the current $\bar{x}_{\rm HI}$ by averaging (marginalizing) over all $\sim 10^{11}$ mock spectra drawn from the corresponding $\bar{x}_{\rm HI}$ snapshot.

(vi) Steps (ii)–(v) are then repeated for all available $\bar{x}_{\rm HI}$ snapshots (24) to obtain a final 1D probability distribution function (PDF) of the $\bar{x}_{\rm HI}$ for our particular observed QSO.

Importantly, unlike in our previous works, in step (iv) we rebin the observed spectrum onto 1 Å bins over the entire $\lambda = 1218$–1230 Å region. Additionally, we were also required to enlarge the amplitude of the provided flux errors by a factor of 5.[13] These steps were necessitated owing to the considerably higher resolution and correspondingly small error spectrum provided by X-shooter. In the absence of this rebinning and enlarged errors, we found our simple analytic χ^2 estimate of the likelihood would encounter numerical difficulties owing to sampling large χ^2 values and correspondingly being too sensitive to narrow subsets of individual profiles who matched the observed data. Alternatively, we could have considered a narrow-wavelength range, however, doing so is more prone to biasing the results to particular features in the data. This highlights the necessity to improve our joint fitting pipeline which we will pursue in future work, where we will move away from using analytic likelihoods altogether in favour of likelihood-free or simulation-based inference (see e.g Cranmer, Brehmer & Louppe 2020, for a recent review on such methods and a preliminary exploration of simulation based inference for extracting constraints from QSO damping wings by Chen, Speagle & Rogers 2023).

4 RESULTS

4.1 Reconstruction of the XQR-30 sample

In Fig. 4, we provide the reconstructed intrinsic profiles for the 23 QSOs from the XQR-30 sample deemed suitable for IGM damping wing analysis (see Section 2.2). In each panel, we show the ML reconstruction profile (red curve), 300 random draws from our full posterior distribution (thin grey lines), the median (solid purple curve), and 68th (purple dashed) and 95th (purple dotted) percentile profiles obtained from the full posterior distribution. The purple-shaded box corresponds to the 1218–1230 Å region over which we fit for the damping wing imprint. In the inset panel, we provide the Lyα transmission profile over this 1218–1230 Å region obtained using the median profile, highlighting the presence (if any) of an IGM damping wing imprint. Grey-shaded regions correspond to features removed from our likelihood fitting.

Here, we will limit our discussions of these reconstructed QSOs to focus only on broad observations that can be made across the sample, and only focus on individual QSOs that warrant further discussion. First, for the most part, the Lyα transmission profile within the 1218–1230 Å region either appears flat, consistent with little to no IGM attenuation, or monotonically increasing with rest-frame wavelength. This indicates that our recalibration step to our reconstruction pipeline is performing well in mitigating our observed systematic offset in the predicted intrinsic flux between the Lyα and N V emission lines (see Sections 3.3 and 3.4).

However, potentially problematic QSO reconstructions remain, for example in the case of J060 + 24 and J359 − 06, which exhibit smoothly varying features in their transmission profiles. For the latter, this behaviour should not impact the results as the relative change in amplitude of the transmission profile is fairly modest. For the former, the shape appears to be driven more so by a stronger blueshift in the Lyα line profile leading to the uptick in transmission near 1218 Å. As a result this may result in a slight underestimate of the neutral fraction. ATLASJ029 − 36, VDESJ0224 − 47, and J1509 − 1749 also may exhibit similar features to those above but to a notably lesser extent. However, what is important is that since we draw from the full posterior distribution of reconstructed profiles that are convolved with the synthetic damping wing templates, profiles that better match the observed spectrum should result in a higher likelihood and thus dominate the constraining power. That is, using the full distribution of joint fits will be a better match than the median transmission profile that is provided for visualization purposes only.

4.2 Recovered IGM damping wing constraints

In Fig. 5, we present the marginalized 1D PDFs of the IGM neutral fraction for each of the 23 QSOs of the XQR-30 sample. For this, we separate the QSOs into redshift bins of $\Delta z = 0.1$. In each panel, the thin coloured lines correspond to an individual QSO. For the vast majority of QSOs in the sample we recover broad, one-sided distributions, consistent with no IGM attenuation. This, however, is not too surprising given the relatively small decrement in the Lyα transmission profiles shown in the inset panels of Fig. 4. Nevertheless, we still find several QSOs with marginalized 1D PDFs consistent with a detection of IGM attenuation.

We summarize the individual constraints/limits on the IGM neutral fraction from the XQR-30 sample in the fourth column of Table 1. For marginalized 1D PDFs consistent with a detection, we present the median and 68th percentiles of the posteriors. For the 1D PDFs consistent with no attenuation, we present the 68th percentile upper limits, which is the IGM neutral fraction which encloses 68 per cent of the total probability.

In the left panel of Fig. 6, we plot our individual QSO constraints on the IGM neutral fraction as a function of redshift. In Table 2, we summarize our joint constraints on the IGM neutral fraction after binning in redshift. In the right panel of Fig. 6, we provide the violin plots of the IGM neutral fraction after binning our individual QSOs into separate $\Delta z = 0.1$ redshift bins based loosely on the natural redshift sampling of the available QSOs in the XQR-30 sample. This choice of redshift sampling is fairly arbitrary but is adopted to ensure each bin contains multiple QSOs. We explored several alternative choices in binning (i.e. larger), but found the results to be generally

[13]We explored increasing this error by factors of 2, 5, and 10 and found no discernible difference in the inferred constraints other than a broadened PDF for increasing error. Thus we chose a factor of two to limit the amplitude of the increased error spectrum.

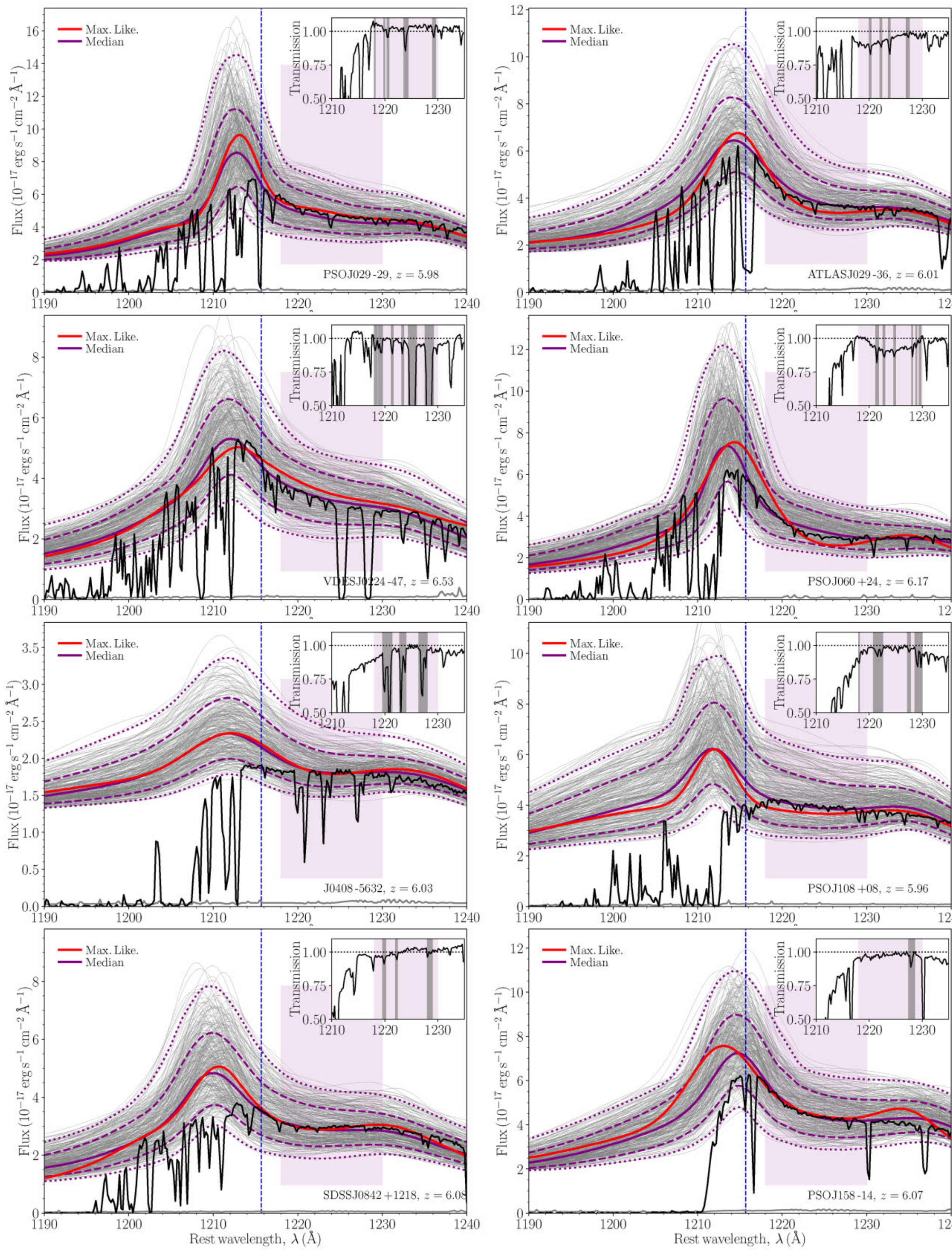

Figure 4. Intrinsic QSO reconstructions of the XQR-30 sample. We provide the ML reconstruction (red curve), the purple solid, dashed, and dotted curves correspond to the median, 68th and 95th percentile profiles and the thin grey curves represent 300 random draws from our full posterior distribution. The solid thick grey curve near zero corresponds to the QSO noise spectrum, amplified by a factor of 5 to improve the numerical performance of our neutral fraction inference approach (see the text for further details). The purple-shaded box demarcates our damping wing fitting region (1218–1230 Å region). Inset panel: the Lyα transmission profile highlighting the presence (if any) of an IGM damping wing imprint (i.e. observed transmission spectrum below unity) assuming the median (purple curve). Grey-shaded regions correspond to features in the observed spectrum removed from our likelihood fitting.

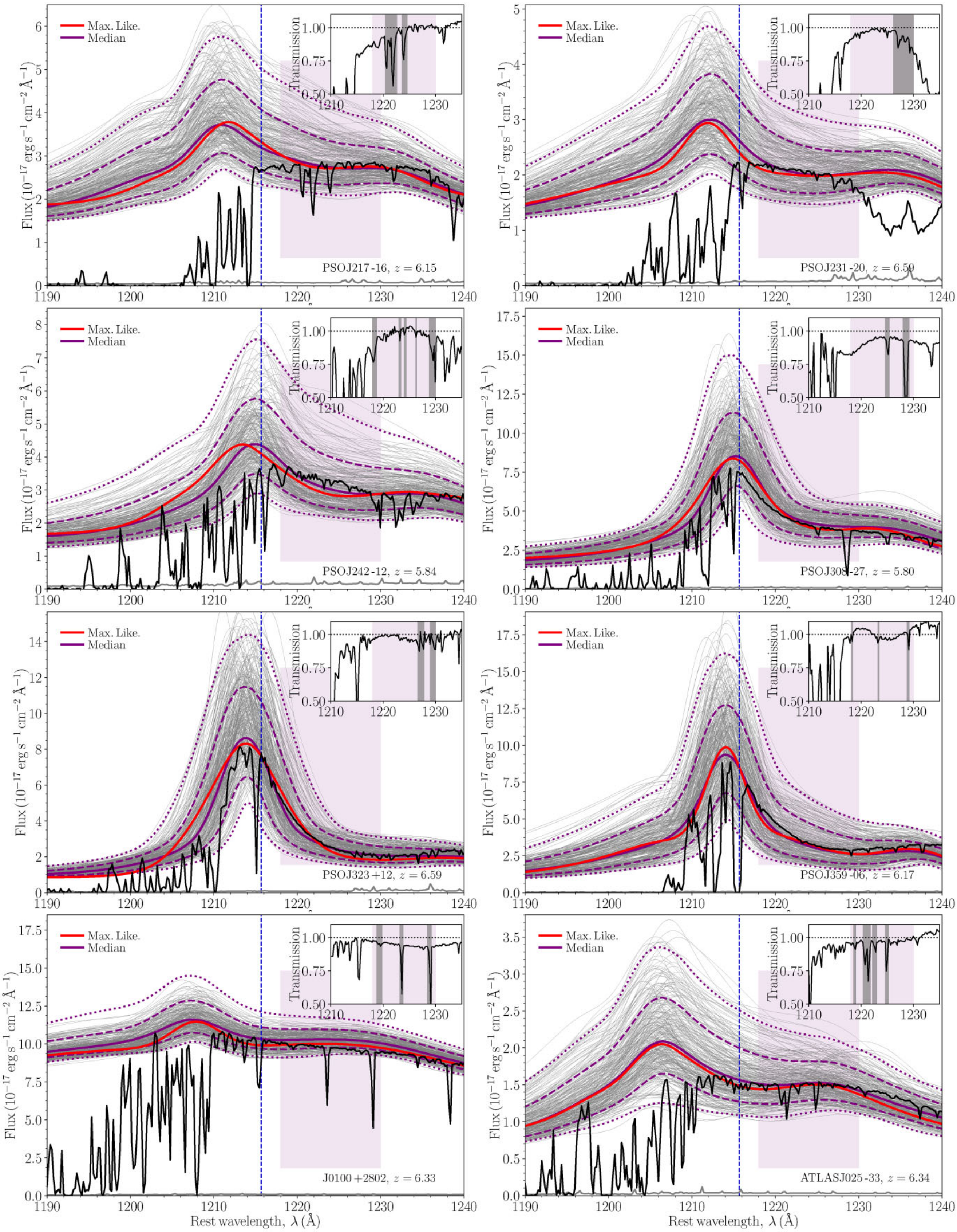

Figure 4. *Continued.*

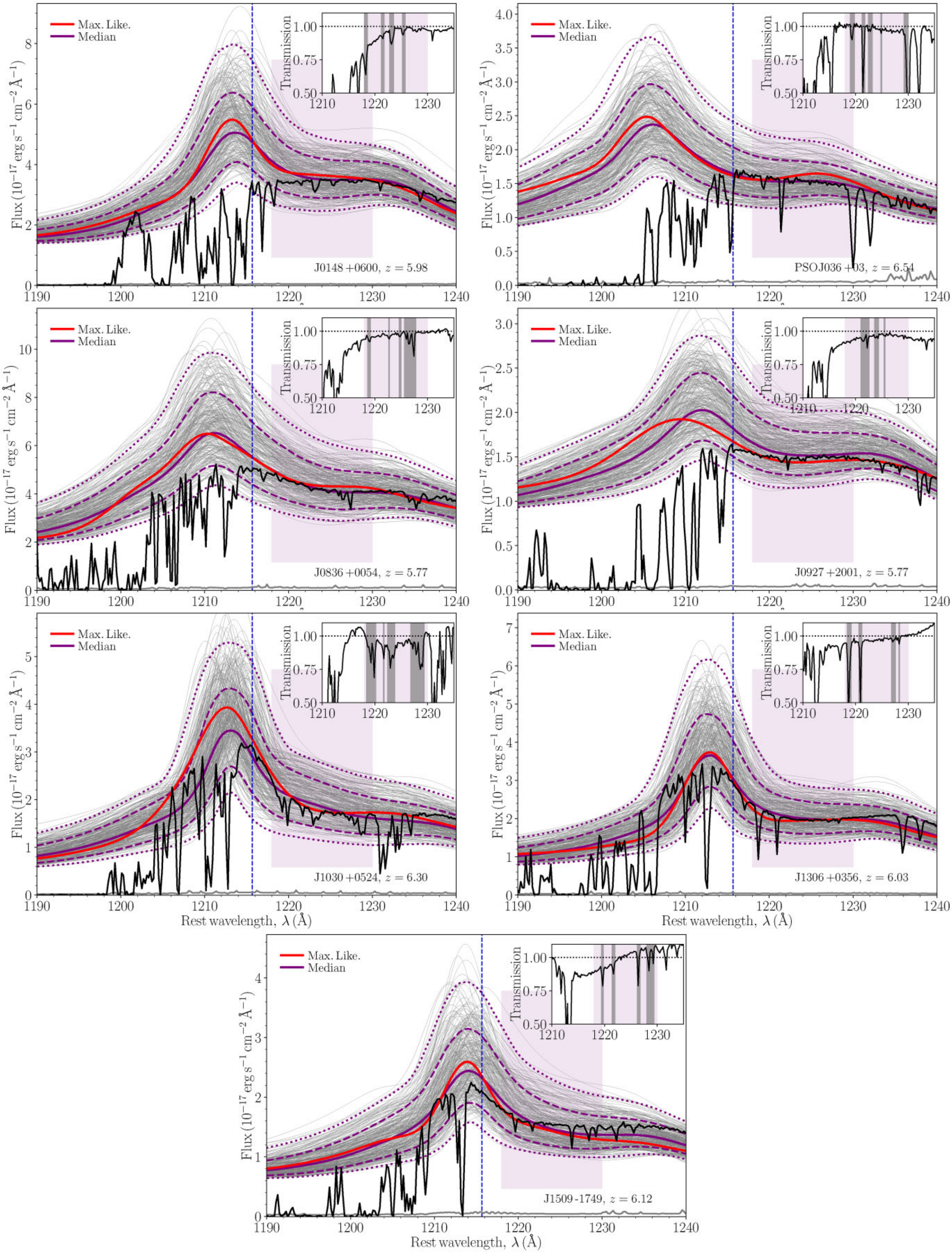

Figure 4. *Continued.*

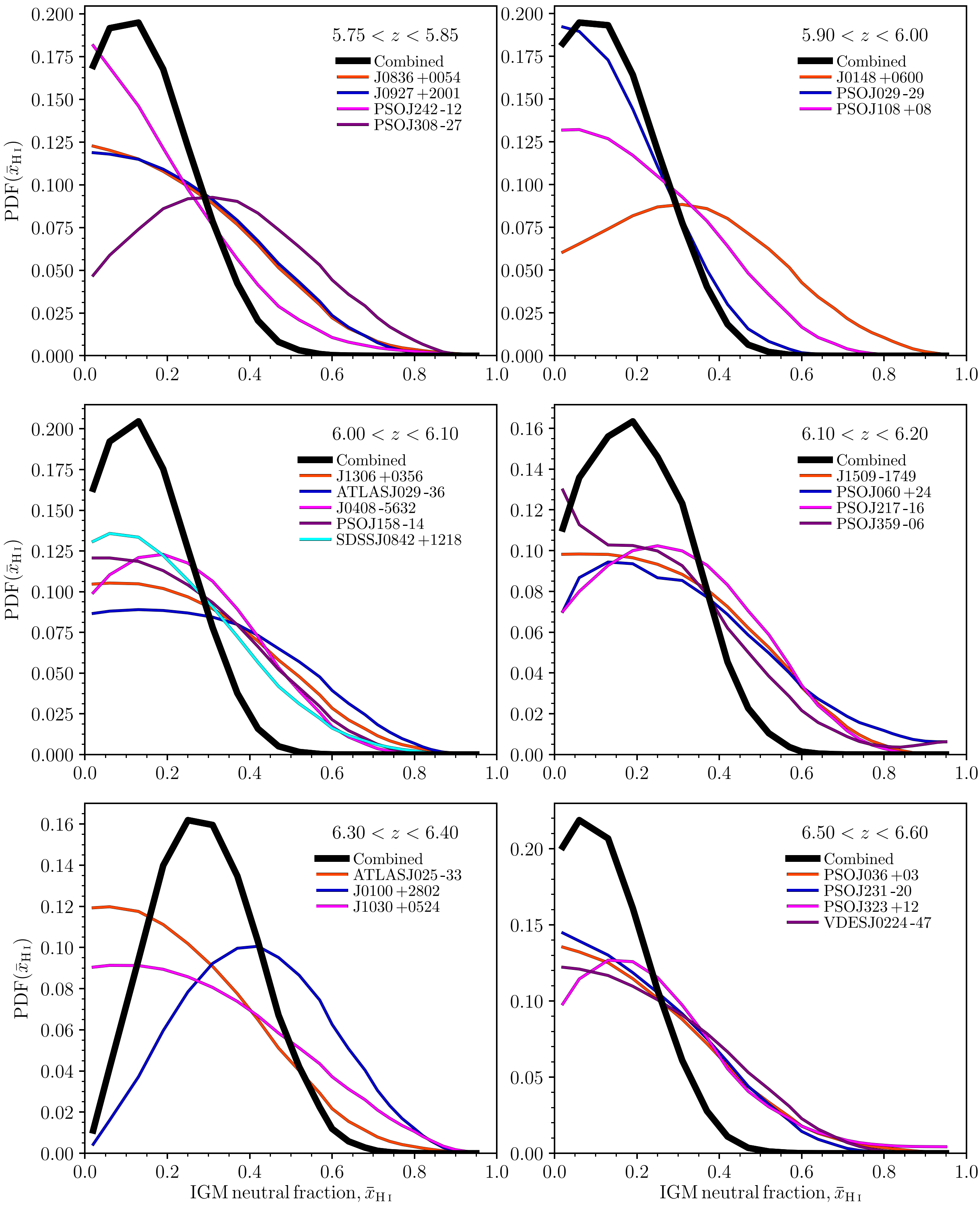

Figure 5. The marginalized 1D PDFs of the IGM neutral fraction for each of the XQR-30 QSOs, separated into redshift bins of $\Delta z = 0.1$. The thin coloured curves in each panel correspond to the recovered constraints for individual QSOs, whereas the thick black curve corresponds to the bin-averaged constraint on the IGM neutral fraction.

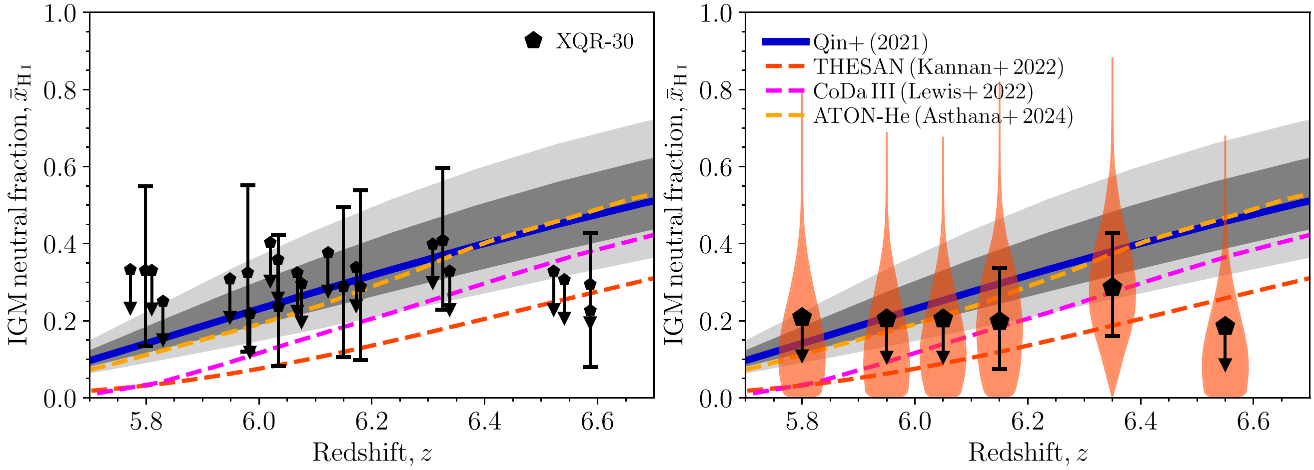

Figure 6. Our constraints on the IGM neutral fraction following our damping wing analysis of the XQR-30 sample. Left panel: the individual constraints on the IGM neutral fraction. We present the median and 68th percentiles for our constraints and 68th percentile upper limits for QSOs consistent with no IGM attenuation. Right panel: a violin plot demonstrating the constraints on the IGM neutral fraction after binning our QSOs into $\Delta z = 0.1$ redshift bins (whereby the individually recovered PDFs of all QSOs within each redshift bin are multiplied to obtain an averaged constraint). In both panels, the blue curve and the dark- and light-shaded regions correspond to the median, 68th and 95th percentile EoR histories obtained by Qin et al. (2021, see the text for further details). Further, the dashed red, magenta, and orange curves correspond to the fiducial reionization histories of state-of-the-art numerical simulations of the IGM from THESAN (Garaldi et al. 2022; Kannan et al. 2022), and CoDA III (Lewis et al. 2022) and Asthana et al. (2024).

Table 2. Summary of the redshift binned IGM neutral fraction constraints from our damping wing analysis. For IGM neutral fractions posteriors consistent with a detection, we provide the median and 68th percentiles. Remaining constraints are presented as 68th percentile upper limits.

Redshift range	$\bar{x}_{\rm H\,I}$	QSOs in bin
$5.85 \le z < 5.95$	<0.21	4
$5.90 \le z < 6.00$	<0.20	3
$6.00 \le z < 6.10$	<0.21	5
$6.10 \le z < 6.20$	$0.20^{+0.14}_{-0.12}$	4
$6.30 \le z < 6.40$	$0.29^{+0.14}_{-0.13}$	3
$6.50 \le z < 6.60$	<0.18	4

insensitive to the choice owing to the natural redshift sampling of the data.

In both panels, we also provide the constrained EoR histories from the state-of-the-art Monte-Carlo Markov Chain analysis of Qin et al. (2021). Here, these authors ran seminumerical simulations of the 21-cm signal during the EoR using 21CMFAST (Mesinger & Furlanetto 2007; Mesinger, Furlanetto & Cen 2011; Murray et al. 2020) coupled with a hybrid scheme to model the Lyα forest. Using these simulations, their astrophysical model describing the galaxies responsible for reionization are constrained only against existing observational constraints on the reionization epoch such as; the observed UV galaxy LFs at $z = 6$–10, the electron scattering optical depth, $\tau_{\rm e}$, measured by *Planck* (Planck Collaboration VI 2020), the dark pixel limits on the IGM neutral fraction (McGreer et al. 2015) and PDFs of the Lyα effective optical depth from the Lyα forest at $z =$ 5–6 (Bosman et al. 2018). Post-processing of the resultant posteriors on their inferred astrophysical model then yields a posterior on the allowed reionization history given the existing observational data. We overlay the median (blue curve) and the 68th (dark grey) and 95th (light grey) percentile regions of the constrained EoR histories.

The vast majority of our inferred IGM damping wing constraints from the XQR-30 sample strongly align with the inferred EoR histories from Qin et al. (2021). Further, in general our XQR-30 data points monotonically increase in IGM neutral fraction for increasing redshift, consistent with ongoing reionization. The main discrepancy with our results and those from Qin et al. (2021) are the inferred IGM neutral fraction limits at $z \gtrsim 6.5$. Our IGM damping wing limits at $z \gtrsim 6.5$ are systematically below the inferred EoR histories from Qin et al. (2021). Our one constraint within this bin, PSOJ323 + 12, returns an IGM neutral fraction of $\bar{x}_{\rm H\,I} = 0.23^{+0.19}_{-0.15}$, which owing to the broad 68th percentiles is consistent with the inferred EoR history. For the remaining three QSOs, we recover upper limits at $z \gtrsim 6.5$ that are inconsistant at between 68th and 95th percentiles of the joint distribution. In Appendix B, we explored the role of the assumed prior on our minimum H II bubble size, $R_{\rm min}$. For our available four choices of $R_{\rm min}$, we observe a shift of $\sim$0.02 in the inferred IGM neutral fraction from our lowest to highest choice of $R_{\rm min}$ therefore our adopted choice of averaging over $R_{\rm min}$ is not likely a contributing factor to the lower IGM neutral fractions. Within the literature there are a broad range of approaches to both extract the damping wing imprint and to predict the intrinsic QSO profile, thus it would be prudent to perform a re-analyses of these QSOs using these alternative reconstruction/damping wing fitting pipelines to add further confidence to the results presented in this work. However, this is a significant undertaking and we leave such a re-analysis to future work.

To place these constraints into a broader context, we additionally provide the reionization histories from recent state-of-the-art radiation-hydrodynamic simulations from THESAN (Garaldi et al. 2022; Kannan et al. 2022) and Cosmic Dawn III (CoDa III, Lewis et al. 2022). We also provide the reionization history from Asthana et al. (2024) who post-processed the Sherwood simulation suite (Bolton et al. 2017) using radiative transfer. Note that unlike the Qin et al. (2021) posteriors, which are constrained specifically by the Bosman et al. (2018) Lyα forest data, the THESAN and CoDA III simulations are single simulations designed to match a considerably broader range of existing observational data. As a result, their reionization histories will differ to those from Qin et al. (2021). On the other hand, the Asthana et al. (2024) simulations are specifically designed to match the XQR-30 Lyα forest data of Bosman et al.

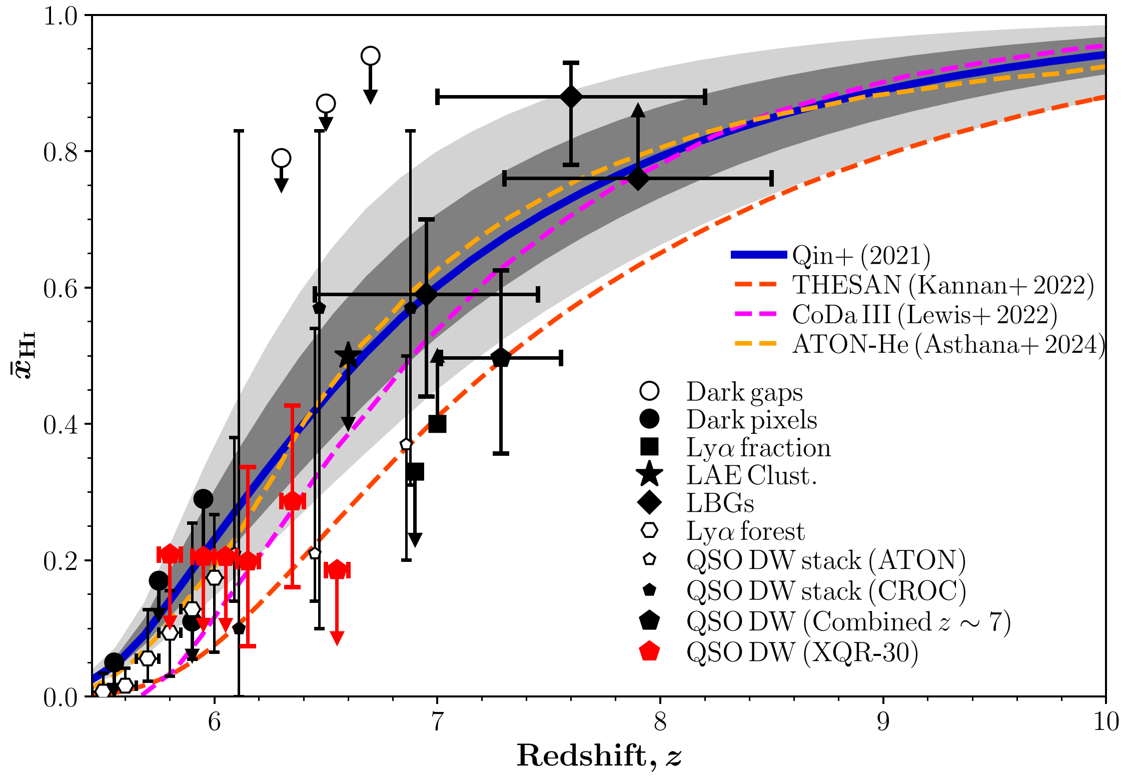

Figure 7. Our binned IGM neutral fraction constraints as a function of redshift in the context of other EoR observables. Red pentagons: the XQR-30 sample (this work). Large black pentagon: the combined IGM damping wing constraints of all 4 $z > 7$ QSOs (Greig et al. 2022). Small pentagons: IGM damping wing constraints from stacking QSO reconstruction profiles (Ďurovčíková et al. 2024). Open hexagon: the inferred constraints from the Lyα forest using numerical simulations at $z = 5.5$–6.0 (Gaikwad et al. 2023). Open circles: dark gaps at $z = 6.3$, 6.5, and 6.7 (Zhu et al. 2022, obtained with XQR-30). Circles: dark pixels at $z = 5.9$ (McGreer, Mesinger & D'Odorico 2015) and $z = 5.55$, 5.75, and 5.95 (Jin et al. 2023). Squares: the Lyα fraction at $z = 6.9$ (Wold et al. 2022) and $z = 7$ (Mesinger et al. 2015), stars: LAE clustering at $z = 6.6$ (Sobacchi & Mesinger 2015), diamonds: LBGs at $z = 7$ (Mason et al. 2018), $z = 7.6$ (Hoag et al. 2019), and $z = 8$ (Mason et al. 2019). The blue curve and the dark- and light-shaded regions corresponds to the median, 1σ and 2σ constraints from (Qin et al. 2021, see the text for further details) and the dashed red, magenta, and orange curves correspond to the fiducial reionization histories of state-of-the-art numerical simulations of the IGM from THESAN (Garaldi et al. 2022; Kannan et al. 2022), and CoDA III (Lewis et al. 2022) and Asthana et al. (2024).

(2022) and as such the reionization history is very similar to the median history recovered by Qin et al. (2021).

Our individual inferred IGM damping wing constraints form the XQR-30 sample remain broadly consistent with the reionization histories from these state-of-the-art numerical simulations. Although the THESAN and CoDa III reionization histories are $\Delta\bar{x}_{\rm H\,I} \sim 0.1 - 0.15$ lower than that expected explicitly from the Lyα forest, since the vast majority of our results are only upper limits these remain statistically consistent. For the handful of constraints that we recover, the CoDa III simulation is consistent with all but one constraint at $z \sim 5.8$ whereas the THESAN simulations are inconsistent with several of our constraints. However, as noted above these simulations are not explicitly constrained to match the Lyα forest data, but instead to broadly match a large range of existing observations. Thus, any inconsistencies are not too concerning. Importantly, at $z \gtrsim 6.5$ where our XQR-30 damping wing results begin to disagree with the inferred constraints from the Lyα forest data, the THESAN and CoDA III simulations are perfectly consistent within the 68th percentile uncertainties of our constraints and upper limits. Therefore, the stronger inconsistencies with our damping wing results and the inferred constraints on the reionization history using the Lyα forest data at $z \sim 6.5$ can likely be softened once additional observational constraints are folded into such an analysis. Nevertheless, once we bin the individual constraints (right panel of Fig. 6) our inferred limits at $z \gtrsim 6.5$ are still inconsistent at more than the 68th percentiles with all these numerical simulations.

4.3 Compilation of reionization constraints

To place our constraints in a broader reionization context, in Fig. 7 we compare our binned IGM neutral fraction constraints against a compilation of other existing constraints on the IGM neutral fraction. Here, we consider constraints and limits obtained from: (i) dark gaps (Zhu et al. 2022) (ii) dark pixels (McGreer et al. 2015; Jin et al. 2023), (iii) the Lyα fraction at $z = 6.9$ (Wold et al. 2022) and at $z = 7$ (Mesinger et al. 2015), (iv) the clustering of Lyα emitters (LAEs) at $z = 6.6$ (Sobacchi & Mesinger 2015), (v) Lyman-break galaxies (LBGs) at $z = 7$ (Mason et al. 2018), $z = 7.6$ (Hoag et al. 2019) and at $z = 8$ (Mason et al. 2019), (vi) the inferred constraints comparing numerical simulations to the Lyα forest at $z = 5.5$–6.0 (Gaikwad et al. 2023), (vii) the joint QSO IGM damping constraint combining all 4 $z > 7$ QSOs (Greig et al. 2022), and (viii) IGM damping constraints after stacking continuum normalized QSO spectra (Ďurovčíková et al. 2024). Note, this is not intended to be an exhaustive list, but instead provide a broad selection of approaches and redshift ranges. We also include the reionization histories of the various state-of-the-art numerical simulations as introduced previously.

Our constraints on the IGM neutral fraction after binning in redshift intervals of $\Delta z = 0.1$ remain consistent with existing constraints on the EoR in the literature. Except for our limit at $z \sim 6.5$, which is now inconsistent at more than 95 per cent. However, it is consistent with the recent QSO damping wing analysis of Ďurovčíková et al. (2024), although their posteriors are extremely broad owing to the parameter space binning of their simulations. The reason for our increased inconsistency is due to the multiplication of the individual posteriors within this redshift bin. Since three of the four QSOs are consistent with no IGM attenuation, the multiplication of these strengthens the preference for a posterior consistent with no IGM attenuation. In order to confirm the results of this work, independent re-analysis of the XQR-30 sample using alternative reconstruction and/or damping wing fitting are required.

4.4 Discussion

It is important to point out that the IGM damping wing constraints presented in this work use only a single QSO reconstruction methodology (covariance matrix). As highlighted previously, there are numerous approaches throughout the literature based on different underlying assumptions. For example, rather than explicitly using the correlations amongst emission lines to reconstruct the intrinsic flux, one can instead use correlations amongst the PCA components (e.g. Davies et al. 2018b; Bosman et al. 2021; Chen et al. 2022). Instead of directly mapping these correlations one can learn them via neural networks (Ďurovčíková et al. 2020) or deep learning (Liu & Bordoloi 2021) and project the reconstruction uncertainties using normalizing flows (Reiman et al. 2020). Alternatively, one can use factor analysis to produce a more flexible basis set of components rather than those from PCA (Sun et al. 2023). Finally, one can also perform deep learning to predict the QSO flux per pixel based on a training set of QSOs (Fathivavsari 2020).

In Greig et al. (2024), we perform a detailed comparison of these reconstruction pipelines on a unified set of QSOs from two separate instruments (X-shooter and SDSS). Qualitatively speaking, these reconstruction pipelines tend to be consistent with one another within their associated modelling uncertainties. However, quantitatively, there can be some notable differences from object to object and these differences would filter through to attempts to recover constraints on the IGM neutral fraction through the damping wing. In the case of this data set, with the high-quality X-shooter spectra, these numerical differences in the predicted QSO profiles could result in considerable differences in the inferred IGM neutral fractions from various reconstruction pipelines.

Not limited to just the reconstruction method, but the methodology to infer the IGM damping wing signature also differs in the literature. Here, we restrict our approach to fitting the smoothed IGM imprint using large-volume EoR simulations to generate synthetic damping wing profiles in the range $\lambda = 1218$–1230 Å. Our choice is based on a preference to avoid having to model and marginalize over the uncertainties of the host QSO environment. In taking into account the QSO host environment, one can better deal with uncertainties in the QSO lifetime and affords a larger region over which to search for the damping wing imprint (e.g. $\lambda \sim 1190$–1230 Å in Davies et al. 2018b). Alternatively, one can simply use an analytic model of the red damping wing imprint (Miralda-Escudé 1998) in order to gain an estimate of the IGM neutral fraction (e.g. Bañados et al. 2018; Reiman et al. 2020).

All of the approaches mentioned previously have their own underlying assumptions and uncertainties. Therefore, to be able to obtain robust estimates of the IGM neutral fraction from the smooth imprint of the IGM damping wing, realistically, one could average (marginalize) over the vast array of different approaches in the literature. In doing so, one would average over the differences in methodology and thus the various systematics and uncertainties. At the very least, independent confirmation using any of the existing other methods would add credence to our recovered constraints.

Another important point worth reiterating is that to be able to obtain our IGM damping wing constraints we had to (i) recalibrate our reconstruction profile and (ii) degrade the observed X-shooter spectra. In the case of the former, this implies further improvements are necessary to our reconstruction pipeline to robustly account for the predicted Lyα and N V line profiles. In the case of the latter, we rebinned the spectra onto 1 Å bins and further had to increase the associated error in the spectra owing to the inflexibility of the analytic (χ^2) expression used in our joint likelihood fitting. In addition to it being inflexible, it also does not fully take advantage of covariances in the observational data. Presently, the observed flux in each wavelength bin is take to be an independent measurement, ignoring the known correlations with neighbouring bins. In fact, we use these correlations to predict our reconstructed profiles (Gaussian components). Given the wealth of theoretical data at hand, one could employ a form of simulation based inference (see e.g Cranmer et al. 2020, for a recent review) to learn the likelihood that more directly connects our reconstructed QSO profiles to the underlying IGM neutral fraction of our synthetic damping wing profiles (see e.g. Chen et al. 2023, for a preliminary demonstration).

Recently there have also been observations of Lyα emission from galaxies extending into the EoR (e.g. Endsley & Stark 2022; Jung et al. 2022; Hayes & Scarlata 2023; Saxena et al. 2023; Umeda et al. 2023; Whitler et al. 2024; Witstok et al. 2024). Although galaxies are much fainter than QSOs, in principle a similar analysis can be performed to attempt to infer the presence of a Lyα damping wing in the individual galaxy spectra (see e.g. Keating et al. 2023). However, this is extremely difficult as the local interstellar medium and H I in the local circumgalactic medium make it much more challenging to infer the unabsorbed emission compared to that of QSOs. Improving the S/N or reducing the galaxy to galaxy variance through stacking (e.g. Umeda et al. 2023) is fundamentally limited by the fact that the damping wing imprint is non-linear and cannot be described by a mean profile convolved with a mean galaxy spectral energy distribution (i.e. the average of a product is not the same as the product of averages).

5 CONCLUSION

We performed an IGM damping wing analysis on the enlarged XQR-30 sample, consisting of 42 high-quality X-shooter spectra spanning $5.8 \lesssim z \lesssim 6.6$. Following careful selection cuts removing QSOs demonstrating BAL features or possible pDLA absorption systems along the line of sight, we are left with 23 of the original 42 QSOs for our IGM damping wing analysis. Nevertheless, this is a factor of 5 improvement over the number of individual QSOs that have been explored for signs of IGM damping wing attenuation due to ongoing reionization.

Our IGM damping wing analysis utilizes a covariance matrix reconstruction approach to predict the intrinsic QSO profile near Lyα (Greig et al. 2017b). The fundamental assumption of this approach is that emission lines can be accurately modelled as Gaussian profiles and that the Lyα and N V emission lines can be reconstructed from a covariance matrix of their correlations with other high-ionization emission lines. These lines, observed redward of Lyα (e.g. C IV, Si IV + O IV], and C III]), are easily measurable and unaffected by IGM attenuation or other line-of-sight contamination. In our approach we fit the observed QSO spectrum over $\lambda = 1275$–2300 Å and draw reconstructed profiles spanning $\lambda = 1180$–1260 Å from our resultant marginalized covariance matrix describing the properties of Lyα and N V.

In the process of analysing the XQR-30 sample, we identified a systematic offset in the predicted QSO flux between rest-frame Lyα and N V owing to our methodology of modelling emission lines as Gaussian profiles. Within this region, slight differences between the wavelength separation or in the widths of either the broad component of Lyα or N V can result in a very modest flux decrement. Using two distinct samples of QSOs from X-shooter and SDSS between $3.5 < z < 4.5$ (unaffected by IGM attenuation), we quantitatively explored whether this was specific to the higher quality X-shooter spectra, or a general feature of our pipeline. We found a consistent median flux decrement of $\sim$10 per cent across both samples, indicative of it being intrinsic to our reconstruction pipeline. Importantly, the consistency in the shape and amplitude of this flux decrement allowed us to recalibrate our covariance matrix reconstruction pipeline by drawing from templates of our original reconstructions compared to the known truth (unattenuated QSO spectrum).

After recalibrating our reconstructed QSO profiles we jointly sample these with synthetic IGM damping wing profiles drawn from large volume EoR simulations (1.6 Gpc on a side with an EoR morphology driven by galaxies residing in $M_h \gtrsim 10^9\,\mathrm{M}_\odot$ haloes). Within our Bayesian framework we fit our reconstructed QSOs profiles multiplied by the synthetic IGM damping wing profiles against the observed QSO spectra of the XQR-30 sample. Specifically, we fit for the smooth component of the IGM damping wing imprint only redward of Lyα ($\lambda = 1218$–1230 Å). Following this pipeline, we recover 1D marginalized posteriors on the IGM neutral fraction from each individual QSO spectrum.

Across the available sample of 23 high-z QSOs, we find 7 with constraints on the IGM neutral fraction during the EoR while for the remaining 16 we recover upper limits. Making use of the large number of QSOs at our disposal, we binned our results in redshift intervals of $\Delta z = 0.1$. Following this, we obtain our final results on the IGM neutral fraction (median and 68th percentiles) during the tail end of reionization of $0.20^{+0.14}_{-0.12}$ and $0.29^{+0.14}_{-0.13}$ at $z = 6.15$ and 6.35. Further, we report 68th percentile upper-limits on the IGM neutral fraction of $\bar{x}_{\mathrm{H\,I}} < 0.21$, 0.20, 0.21, and 0.18 at $z = 5.8$, 5.95, 6.05, and 6.55.

These constraints on the IGM neutral fraction are consistent with those obtained from alternative methods in the literature. In particular, all but one of our constraints are within the joint 68th percentile constraints on the IGM neutral fraction obtained by the state-of-the-art forward modelling of recent Lyα forest data by Qin et al. (2021). The only discrepancy occurs in the highest redshift bin, $6.5 \le z < 6.6$, which is inconsistent at 95 per cent certainty. Of the four QSOs from the XQR-30 sample in this redshift range, three are best modelled by upper limits consistent with no IGM attenuation. The fourth (PSOJ323 + 12) indicates an IGM neutral fraction of $\bar{x}_{\mathrm{H\,I}} = 0.23^{+0.19}_{-0.15}$. Taken at face value this implies that these 4 QSOs may reside in a patch of the Universe that was reionized earlier. We also explored whether our choice of averaging over the adopted minimum local H II bubble size, R_{min} containing the QSO plus galaxy contribution impacted these results. We found that increasing R_{min} had a fairly modest impact on the inferred value of $\bar{x}_{\mathrm{H\,I}}$, smaller than the corresponding 68th percentile uncertainties. Given the broad range of approaches in the literature designed to reconstruct the intrinsic properties of high-z QSOs, along with differing methodologies for extracting the IGM damping wing imprint, it would be valuable to repeat this analysis across multiple methods, averaging over the different pipeline systematics and uncertainties.

ACKNOWLEDGEMENTS

We would like to thank the anonymous referees whose comments improved this manuscript. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project no. CE170100013. AM acknowledges support from the Ministry of Universities and Research (MUR) through the PRIN project 'Optimal inference from radio images of the epoch of reionization' as well as the PNRR project 'Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing'. GDB was supported by the National Science Foundation (NSF) through grant AST-1751404. SEIB is supported by the Deutsche Forschungsgemeinschaft (DFG) under Emmy Noether grant no. BO 5771/1-1. HC thanks the support by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference no. DIS-2022-568580. MGH has been supported by STFC grants ST/N000927/1 and ST/Y004191/1. YZ was supported by the NSF through award SOSPADA-029 from the NRAO. This work was performed on the OzSTAR national facility at Swinburne University of Technology. The OzSTAR program receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government, and from the Victorian Higher Education State Investment Fund (VHESIF) provided by the Victorian Government. Based on observations collected at the European Southern Observatory under ESO programmes: 60.A-9024, 084.A-0360, 084.A-0390, 085.A-0299, 086.A-0162, 086.A-0574, 088.A-0897, 091.C-0934, 294.A-5031, 096.A-0095, 096.A-0418, 097.B-1070, 098.B-0537, 0100.A-0625, 0101.B-0272, 0102.A-0154, 1103.A-0817.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss4.org.

SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, Center for Astrophysics | Harvard & Smithsonian (CfA), the Chilean Participation Group, the French Participation Group, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Software: CYTHON (Behnel et al. 2011), NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), and COSMOHAMMER (Akeret et al. 2013).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abdurashidova Z. et al., 2022a, ApJ, 924, 51
Abdurashidova Z. et al., 2022b, ApJ, 925, 221
Akeret J., Seehars S., Amara A., Refregier A., Csillaghy A., 2013, Astron. Comput., 2, 27
Alam S. et al., 2015, ApJS, 219, 12
Asthana S., Haehnelt M. G., Kulkarni G., Aubert D., Bolton J. S., Keating L. C., 2024, preprint (arXiv:2404.06548)
Bañados E. et al., 2019, ApJ, 885, 59
Bañados E. et al., 2018, Nature, 553, 473
Becker G. D., Bolton J. S., Madau P., Pettini M., Ryan-Weber E. V., Venemans B. P., 2015, MNRAS, 447, 3402
Becker G. D. et al., 2019, ApJ, 883, 163
Behnel S., Bradshaw R., Citro C., Dalcin L., Seljebotn D., Smith K., 2011, Comput. Sci. Eng., 13, 31
Bischetti M. et al., 2022, Nature, 605, 244
Bischetti M. et al., 2023, ApJ, 952, 44
Bolton J. S., Haehnelt M. G., Warren S. J., Hewett P. C., Mortlock D. J., Venemans B. P., McMahon R. G., Simpson C., 2011, MNRAS, 416, L70
Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897
Bosman S. E. I., Fan X., Jiang L., Reed S., Matsuoka Y., Becker G., Haehnelt M., 2018, MNRAS, 479, 1055
Bosman S. E. I., Ďurovčíková D., Davies F. B., Eilers A.-C., 2021, MNRAS, 503, 2077
Bosman S. E. I. et al., 2022, MNRAS, 514, 55
Chen H. et al., 2022, ApJ, 931, 29
Chen H., Speagle J., Rogers K. K., 2023, preprint (arXiv:2311.16238)
Cranmer K., Brehmer J., Louppe G., 2020, Proc. Natl. Acad. Sci., 117, 30055
D'Aloisio A., McQuinn M., Trac H., 2015, ApJ, 813, L38
D'Odorico V. et al., 2023, MNRAS, 523, 1399
Davies F. B. et al., 2018a, ApJ, 864, 142
Davies F. B. et al., 2018b, ApJ, 864, 143
Davies F. B., Bañados E., Hennawi J. F., Bosman S. E. I., 2023, preprint (arXiv:2312.06747)
Dawson K. S. et al., 2013, AJ, 145, 10
Ďurovčíková D., Katz H., Bosman S. E. I., Davies F. B., Devriendt J., Slyz A., 2020, MNRAS, 493, 4256
Ďurovčíková D. et al., 2024, preprint (arXiv:2401.10328)
Endsley R., Stark D. P., 2022, MNRAS, 511, 6042
Fan X. et al., 2006, AJ, 132, 117
Fathivavsari H., 2020, ApJ, 898, 114
Gaikwad P. et al., 2023, MNRAS, 525, 4093
Garaldi E., Kannan R., Smith A., Springel V., Pakmor R., Vogelsberger M., Hernquist L., 2022, MNRAS, 512, 4909
Ghara R. et al., 2020, MNRAS, 493, 4728
Ghara R., Giri S. K., Ciardi B., Mellema G., Zaroubi S., 2021, MNRAS, 503, 4551
Greig B., Mesinger A., McGreer I. D., Gallerani S., Haiman Z., 2017a, MNRAS, 466, 1814
Greig B., Mesinger A., Haiman Z., Simcoe R. A., 2017b, MNRAS, 466, 4239
Greig B., Mesinger A., Bañados E., 2019, MNRAS, 484, 5094
Greig B., Mesinger A., Koopmans L. V. E., 2020, MNRAS, 491, 1398
Greig B., Trott C. M., Barry N., Mutch S. J., Pindor B., Webster R. L., Wyithe J. S. B., 2021, MNRAS, 500, 5322
Greig B., Mesinger A., Davies F. B., Wang F., Yang J., Hennawi J. F., 2022, MNRAS, 512, 5390
Greig B. et al., 2024, preprint (arXiv:2404.01556)
HERA Collaboration, 2023, ApJ, 945, 124
Harris C. R. et al., 2020, Nature, 585, 357
Hayes M. J., Scarlata C., 2023, ApJ, 954, L14
Heintz K. E. et al., 2023, preprint (arXiv:2306.00647)
Hoag A. et al., 2019, ApJ, 878, 12
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jin X. et al., 2023, ApJ, 942, 59
Jung I. et al., 2022, preprint (arXiv:2212.09850)
Kannan R., Garaldi E., Smith A., Pakmor R., Springel V., Vogelsberger M., Hernquist L., 2022, MNRAS, 511, 4005
Keating L. C., Puchwein E., Haehnelt M. G., 2018, MNRAS, 477, 5501
Keating L. C., Bolton J. S., Cullen F., Haehnelt M. G., Puchwein E., Kulkarni G., 2023, preprint (arXiv:2308.05800)
Keating L. C., Puchwein E., Bolton J. S., Haehnelt M. G., Kulkarni G., 2024, MNRAS, 531, L34
Lewis J. S. W. et al., 2022, MNRAS, 516, 3389
Liu B., Bordoloi R., 2021, MNRAS, 502, 3510
López S. et al., 2016, A&A, 594, A91
Lyke B. W. et al., 2020, ApJS, 250, 8
Mason C. A., Treu T., Dijkstra M., Mesinger A., Trenti M., Pentericci L., de Barros S., Vanzella E., 2018, ApJ, 856, 2
Mason C. A. et al., 2019, MNRAS, 485, 3947
Mazzucchelli C. et al., 2023, A&A, 676, A71
McGreer I. D., Mesinger A., D'Odorico V., 2015, MNRAS, 447, 499
Mertens F. G. et al., 2020, MNRAS, 493, 1662
Mesinger A., Furlanetto S., 2007, ApJ, 669, 663
Mesinger A., Haiman Z., 2007, ApJ, 660, 923
Mesinger A., Furlanetto S., Cen R., 2011, MNRAS, 411, 955
Mesinger A., Aykutalp A., Vanzella E., Pentericci L., Ferrara A., Dijkstra M., 2015, MNRAS, 446, 566
Mesinger A., Greig B., Sobacchi E., 2016, MNRAS, 459, 2342
Miralda-Escudé J., 1998, ApJ, 501, 15
Mondal R. et al., 2020, MNRAS, 498, 4178
Mortlock D. J. et al., 2011, Nature, 474, 616
Murray S., Greig B., Mesinger A., Muñoz J., Qin Y., Park J., Watkinson C., 2020, J. Open Source Softw., 5, 2582
Planck Collaboration XIII, 2016, A&A, 594, A13
Planck Collaboration VI, 2020, A&A, 641, A6
Prochaska J. X., Hennawi J. F., Herbert-Fort S., 2008, ApJ, 675, 1002
Qin Y., Mesinger A., Bosman S. E. I., Viel M., 2021, MNRAS, 506, 2390
Reiman D. M., Tamanas J., Prochaska J. X., Ďurovčíková D., 2020, preprint (arXiv:2006.00615)
Rybicki G. B., Lightman A. P., 1979, Radiative Processes in Astrophysics. Wiley-Interscience, New York
Satyavolu S. et al., 2023, MNRAS, 522, 4918
Saxena A. et al., 2023, A&A, 678, A68
Sobacchi E., Mesinger A., 2015, MNRAS, 453, 1843
Sun Z., Ting Y.-S., Cai Z., 2023, ApJS, 269, 4
Trott C. M. et al., 2020, MNRAS, 493, 4711
Umeda H., Ouchi M., Nakajima K., Harikane Y., Ono Y., Xu Y., Isobe Y., Zhang Y., 2023, preprint (arXiv:2306.00487)
Vernet J. et al., 2011, A&A, 536, A105
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Wang F. et al., 2020, ApJ, 896, 23
Whitler L., Stark D. P., Endsley R., Chen Z., Mason C., Topping M. W., Charlot S., 2024, MNRAS, 529, 855
Witstok J. et al., 2024, A&A, 682, A40
Wold I. G. B. et al., 2022, ApJ, 927, 36
Yang J. et al., 2020, ApJ, 897, L14
Yang J. et al., 2021, ApJ, 923, 262
Zhu Y. et al., 2022, ApJ, 932, 76

APPENDIX A: UPDATED COVARIANCE MATRIX

For this work, we have substantially increased the number of QSOs that are included within our training set for constructing our emission line covariance matrix, from 1673 to 30 166 QSOs. In Fig. A1, we provide an updated correlation coefficient matrix for this new training set to demonstrate the available emission line parameter correlations. Each emission line is separated by solid vertical and horizontal dashed lines, while dashed lines denote the separation between broad and narrow components, respectively. Along the top, each emission line is identified.

This correlation coefficient matrix, $\mathbf{R}_{ij}$ is determined by computing,

$$\mathbf{R}_{ij} = \frac{\mathbf{C}_{ij}}{\sqrt{\mathbf{C}_{ii}\mathbf{C}_{jj}}}, \tag{A1}$$

where the ith and jth subscripts denote the different emission line parameters and $\mathbf{C}_{ij}$ is covariance matrix of the full training set computed using,

$$\mathbf{C}_{ij} = \frac{1}{N-1}\sum_{i}^{N}(\mathbf{X}_i - \boldsymbol{\mu}_i)(\mathbf{X}_j - \boldsymbol{\mu}_j). \tag{A2}$$

Here, $\mathbf{X}_i$ is the data vector containing all ith emission line parameters (21) from the full QSO sample, $\boldsymbol{\mu}$ is its mean and N is the size of our training set.

Generally speaking, we find only very small differences between the line correlation strengths from our updated training set compared to those shown in Greig et al. (2022).

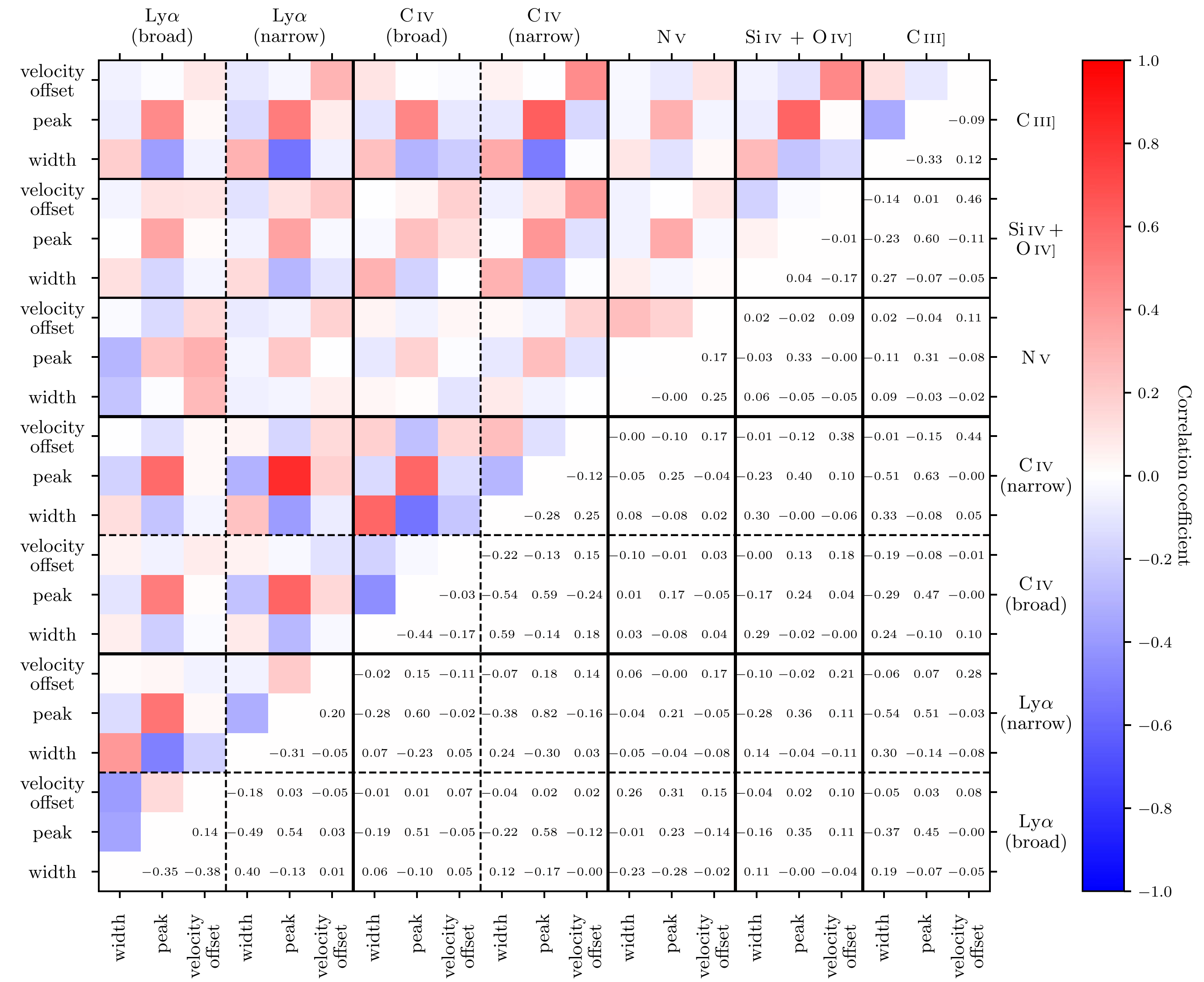

Figure A1. The updated correlation coefficient matrix following the expansion of the training set to 30 166 QSOs with S/N > 6.5 and spanning 2.08 < z < 4.0 from BOSS DR16Q (Lyke et al. 2020). This 21D matrix highlights the correlation coefficient strengths between the various emission line parameters (separated by solid black vertical and horizontal lines), with the Lyα and C IV lines modelled by a double-component Gaussian and the N V, Si IV + O IV], and C III] lines modelled as a single-component Gaussian. Each emission line component is fully described by three parameters, the peak width, height, and velocity offset from the systemic line centre. The data set used in this work corresponds to an order of magnitude increase over previous work (Greig et al. 2022).

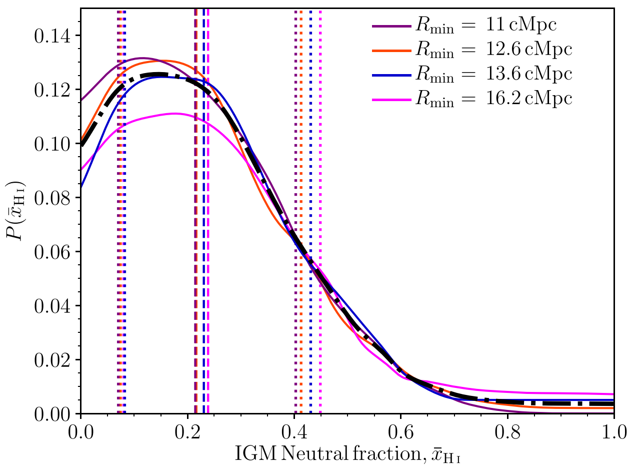

Figure B1. A comparison of the recovered IGM neutral fraction from PSOJ323 + 12 (z = 6.5872) after averaging over different minimum H II bubble sizes, $R_{\rm min}$ (black dotted–dashed curve) and when considering each choice of $R_{\rm min}$ individually. In total, we have data available for four choices of $R_{\rm min}$, distinguished by the coloured curves. The vertical dashed lines correspond to the median inferred constraint on the IGM neutral fraction for each individual choice of $R_{\rm min}$, while the dotted lines correspond to the 68th percentile region.

APPENDIX B: IMPACT OF PROXIMITY ZONE SIZE PRIOR

Our constraints on the IGM neutral fraction are determined by averaging over four different minimum local H II bubble sizes, $R_{\rm min}$ (see e.g. Section 3.2 for further details) based on the available data used for our previously analysed $z > 7$ QSOs. Here, we investigate this assumption of averaging over this $R_{\rm min}$ by determining the inferred IGM neutral fraction for a single QSO, PSOJ323 + 12 (z = 6.5872) for each individual available $R_{\rm min}$ compared to that obtained when averaging over all four.

In Fig. B1, we provide the inferred IGM neutral fraction PDF following our pipeline outlined in Section 3.5 for an $R_{\rm min}$ of 11 (purple), 12.6 (red), 13.6 (blue), and 16.2 cMpc (magenta). The black dot dashed curve corresponds to averaging over all four choices while the vertical dashed lines correspond to the median of the recovered PDF, whereas the dotted lines denote the 68th percentile uncertainties.

For these different available choices of $R_{\rm min}$, we obtain constraints of $\bar{x}_{\rm H\,I} = 0.21^{+0.18}_{-0.14}$, $0.22^{+0.19}_{-0.15}$, $0.23^{+0.20}_{-0.15}$, and $0.24^{+0.21}_{-0.16}$, respectively. For reference, after averaging we recover $\bar{x}_{\rm H\,I} = 0.23^{+0.19}_{-0.15}$. This implies that for an increasing choice of $R_{\rm min}$, we infer both an increasing IGM neutral fraction and also a broader PDF, which is to be expected. Increasing $R_{\rm min}$ effectively amounts to a marginal leftward shift in the synthetic damping wing profiles in the left panel of Fig. 1. As the distance to the first patch of neutral IGM is larger (for the joint QSO + galaxy contribution), for the same fixed amplitude attenuation (determined by the reconstruction pipeline), we require higher IGM neutral fractions. Equally, due to the increased $R_{\rm min}$, we are sampling further into the tail of the Lyα scattering cross-section, resulting in a lower amplitude variation for fixed wavelength in the synthetic damping wing profiles (reduced scatter in the mean profiles as a function of neutral fraction). This lower width between the damping wing profiles leads to an inferred broadening of the IGM posteriors as the width of our intrinsic profile distribution remains unchanged.

Overall, we recover a shift in the inferred IGM neutral fractions of $\Delta\bar{x}_{\rm H\,I} = 0.022$ for a shift in $R_{\rm min}$ of $\sim$5 cMpc. Additionally, we observe an increase of 0.02 to the inferred 68th percentile region. Therefore, averaging over $R_{\rm min}$ does not have a strong impact on our inferred IGM neutral fraction constraints, especially given the overall width of the 68th percentiles in the posterior.

This paper has been typeset from a TEX/LATEX file prepared by the author.