The Lyman-α forest catalog from the Dark Energy Spectroscopic Instrument Early Data Release

César Ramírez-Pérez, Ignasi Pérez-Ràfols, Andreu Font-Ribera, M Abdul Karim, E Armengaud, J Bautista, S.F Beltran, L Cabayol-Garcia, Z Cai, S Chabanier, et al.

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **528,** 6666–6679 (2024) https://doi.org/10.1093/mnras/stad3781
Advance Access publication 2023 December 7

The Lyman-α forest catalogue from the Dark Energy Spectroscopic Instrument Early Data Release

César Ramírez-Pérez [1★] Ignasi Pérez-Ràfols [2] Andreu Font-Ribera [1,3] M. Abdul Karim,[4] E. Armengaud [4] J. Bautista,[5] S. F. Beltran,[6] L. Cabayol-Garcia,[1] Z. Cai,[7,8,9] S. Chabanier [10] E. Chaussidon [10] J. Chaves-Montero [1] A. Cuceu [11,12,13] R. de la Cruz [6] J. García-Bellido [14] A. X. Gonzalez-Morales [6,15] C. Gordon,[1] H. K. Herrera-Alcantar [6] V. Iršič [16] M. Ishak [17] N. G. Karaçaylı [11,12,13,18] Zarija Lukić,[10] C. J. Manser [19,20] P. Montero-Camacho [9] L. Napolitano [21] G. Niz [6,22] M. M. Pieri,[23] C. Ravoux [5,4] F. Sinigaglia [24] T. Tan,[25] M. Walther [26,27] B. Wang [9,28] J. Aguilar,[10] S. Ahlen [29] S. Bailey [10] D. Brooks,[30] T. Claybaugh,[10] K. Dawson,[31] A. de la Macorra [32] G. Dhungana [33] P. Doel,[30] K. Fanning [13] J. E. Forero-Romero [34,35] S. Gontcho A Gontcho [10] J. Guy,[10] K. Honscheid,[11,12,13] R. Kehoe,[33] T. Kisner [10] M. Landriau [10] L. Le Guillou [25] M. E. Levi [10] C. Magneville,[4] P. Martini [11,13,18] A. Meisner [36] R. Miquel,[1,37] J. Moustakas [38] E. Mueller,[39] A. Muñoz-Gutiérrez,[32] J. Nie [40] N. Palanque-Delabrouille [4,10] W. J. Percival [41,42,43] G. Rossi,[44] E. Sanchez [45] E. F. Schlafly [46] D. Schlegel,[10] H. Seo [47] G. Tarlé [46,48] B. A. Weaver,[36] C. Yéche [4] and Z. Zhou[40]

Affiliations are listed at the end of the paper

Accepted 2023 December 4. Received 2023 October 30; in original form 2023 June 20

ABSTRACT

We present and validate the catalogue of Lyman-α forest fluctuations for 3D analyses using the Early Data Release (EDR) from the Dark Energy Spectroscopic Instrument (DESI) survey. We used 88 511 quasars collected from DESI Survey Validation (SV) data and the first two months of the main survey (M2). We present several improvements to the method used to extract the Lyman-α absorption fluctuations performed in previous analyses from the Sloan Digital Sky Survey (SDSS). In particular, we modify the weighting scheme and show that it can improve the precision of the correlation function measurement by more than 20 per cent. This catalogue can be downloaded from https://data.desi.lbl.gov/public/edr/vac/edr/lya/fuji/v0.3, and it will be used in the near future for the first DESI measurements of the 3D correlations in the Lyman-α forest.

Key words: catalogues – intergalactic medium – dark energy – large-scale structure of Universe.

1 INTRODUCTION

The Lyman-α forest is a pattern of absorption features caused by neutral hydrogen in the intergalactic medium (IGM), typically observed in the spectra of distant quasars ($z > 2$). Measurements of 1D correlations in a handful of high-resolution quasar spectra emerged as a powerful tool to study the large-scale distribution of matter (Croft et al. 1998; McDonald et al. 2000), opening a new field in the analysis of the high redshift universe and helping to constrain cosmological parameters.

Using data from the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), the 3D correlation function of absorption in the Lyman-α forest was measured for the first time in Slosar et al. (2011). Shortly after that, the first measurement of the baryonic acoustic oscillations (BAO) peak in the Lyman-α forest was presented (Busca et al. 2013; Slosar et al. 2013; Kirkby et al. 2013), using data from BOSS DR9 (Lee et al. 2013). These were followed by other BAO analyses using increasingly larger Lyman-α forest data sets from BOSS (Delubac et al. 2015; Bautista et al. 2017) and from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) (de Sainte Agathe et al. 2019). The precision of these BAO measurements was significantly improved with the measurement of the cross-correlation of quasars and the Lyman-α forest (Font-Ribera et al. 2014; du Mas des Bourboux et al. 2017; Blomqvist et al. 2019), and the final Lyman-α BAO measurement combining BOSS and eBOSS was presented in du Mas des Bourboux et al. (2020).

The Dark Energy Spectroscopic Instrument (DESI) is currently undergoing a 5-yr campaign to obtain close to a million quasar spectra with $z > 2$ (DESI Collaboration 2023a; Chaussidon et al. 2023). This data set will be four times larger than the state of the art (eBOSS

★ E-mail: cramirez@ifae.es

DR16 quasar sample Lyke et al. 2020) and will enable sub- per cent BAO measurements with the Lyman-α forest (Levi et al. 2013; DESI Collaboration 2016a).

In this publication, we present the first catalogue of Lyman-α forest fluctuations in DESI, including data from the Early Data Release [EDR; DESI Collaboration (2023b)] and from the first two months of the main survey (M2). This data set is used in a companion paper to measure the first 3D correlations in the Lyman-α forest from DESI (Gordon et al. 2023), and a comparison with synthetic data sets is presented in Herrera et al. (in preparation).

The methodology used here is similar to the one developed for eBOSS analyses, especially the most recent analysis by du Mas des Bourboux et al. (2020). This served as the basis for developing the data analysis pipeline of the DESI Lyman-α forest working group. In this publication, we provide a detailed description of our new pipeline, focusing on the changes with respect to the one used in eBOSS analyses. Some of these changes are motivated by changes in the input data: for instance, while SDSS spectra had pixels equispaced in the logarithm of the wavelength, DESI uses linearly spaced pixels. Other changes are motivated by studies that appeared after du Mas des Bourboux et al. (2020). For instance, following Ennesser et al. (2022), we now include in our analysis the spectra of Broad Absorption Line (BAL) quasars, after we mask the most contaminated regions. Finally, we also revisit the weighting scheme used to compute correlations in the Lyman-α forest, resulting in an improvement of more than 20 per cent in our precision.

All of the process followed in this paper, along with the updated changes, is executed using the publicly available code PICCA[1] and can be reproduced by the user using public DESI data. The PICCA package also includes modules for the computations of both auto- and cross-correlation with quasars, cosmological fits, and multiple useful tools for Lyman-α forest studies.

The catalogue described here is aimed at studies of 3D correlations in the Lyman-α forest. A similar data set, however, is also used in two companion publications that present the first measurements of 1D correlations with DESI data (Ravoux et al. 2023; Göksel Karaçaylı et al. 2023). Although analogous estimations of the unabsorbed quasar continuum are also needed in these studies, their methodology is somewhat different and these publications focus on studies of systematics that primarily affect the 1D correlations.

This paper is structured as follows. In Section 2, we present the data used for this work, including spectra and quasar catalogues. In Section 3, we explain how we obtain the flux-transmission field from spectra through the estimation of the expected flux of quasars in the continuum fitting process. This includes masking some wavelengths and applying corrections to the flux calibration and the reported uncertainties by the pipeline. In Section 4, we discuss aspects that require further clarification, such as modified weights and wavelength grid choices. Finally, in Section 5, we provide a summary of our findings and conclusions.

2 DATA

Data used in this publication come from two different DESI data sets. On the one hand, we have the EDR, which includes all Commissioning, Survey Validation (SV), and special survey data. On the other hand, we have EDR + M2, which includes all data from EDR as well as the first 2 months of the main survey.

[1]https://github.com/igmhub/picca/

Although this two samples are qualitatively similar, we are performing the analysis separately for each of them to achieve the two purposes of this paper:

(i) **Describe the Lyman-α forest Value Added Catalogue (VAC) in the context of EDR (DESI Collaboration 2023b):** VACs for a wide variety of tracers are being released using DESI early data. The present work provides the Lyman-α forest fluctuations catalogue.

(ii) **Describe the Lyman-α forest catalogue used in the context of early DESI data publications:** Multiple related publications are being published in this context within the Lyman-α working group. The objective of Gordon et al. (2023) is to obtain the first Lyman-α correlation measurements from DESI early data, testing the current pipeline and data quality, and compare its performance to previous eBOSS DR16 analyses. Herrera et al. (in preparation) provides details on the current status of the different procedures used to build mocks for Lyman-α forest analyses. Gontcho et al. (in preparation) characterizes the systematics caused by the DESI instrument on the 3D correlations of the Lyman-α forest. Bault et al. (in preparation) studies the impact of redshift errors on the 3D cross-correlation of quasars with the Lyman-α forest. P1D analyses are performed in two different papers: Ravoux et al. (2023) presents the 1 dimensional measurement using Fast Fourier Transform, while Göksel Karaçaylı et al. (2023) makes use of the Quadratic Maximum Likelihood Estimator.

Finally, the current publication provides the Lyman-α fluctuations catalogue, as well as its validation. The decision to use EDR + M2 data for these analyses was motivated by the need for a larger volume of data than the EDR could offer, which leads to better measurements of the correlation function, constraining power and estimation of systematics.

EDR data, including Lyman-α forest fluctuations is available now, including this VAC describing Lyman-α fluctuations. However, EDR + M2 will not be released as a separate piece of data and M2 will be released alongside Year 1 (Y1) data.

2.1 DESI spectroscopic data

DESI is the largest ongoing spectroscopic survey, operating on the Mayall 4-metre telescope at Kitt Peak National Observatory (DESI Collaboration 2022). DESI consists of 5000 fibers placed in the focal plane and distributed across 10 petals. Each fiber is controlled by a robotic positioner, allowing for quick and automatic positioning. The large number of fibers and the fast cadence provided by the automatic positioners allows DESI to measure up to 5000 spectra every 20 min over a ∼3° field (DESI Collaboration 2016b; Silber et al. 2023; Miller et al. 2023). Fibers carry light from the telescope into a separate room, where it is dispersed by ten spectrographs. Each of them with three cameras each (B, R, Z), covering different wavelength ranges.

While most of the fibers are assigned to science targets, some are used for calibration. Fibers assigned to the sky are essential for sky subtraction, as they enable the removal of the contribution to spectra caused by light from the background sky, particularly emission lines. Other fibers are assigned to standard stars in order to properly calibrate fluxes. This process is performed by comparing the measured counts in the CCD and the expected fluxes from the standard stars, allowing to estimate fluxes for all the other targets through a process called spectroperfectionism (described in Bolton & Schlegel 2010). Spectroperfectionism provides spectral fluxes and their variances. For the full description of the DESI spectroscopic reduction pipeline, see Guy et al. (2023).

To help data processing, other pipelines are developed for multiple purposes: optimize the observation strategy (Schlafly et al. 2023), calculate the exposure time Kirkby et al. (in preparation), assign fibers to targets Raichoor et al. (in preparation), and target selection (Myers et al. 2023).

DESI target selection is based on the public Legacy Surveys (Zou et al. 2017; Dey et al. 2019; Schlegel et al. (in preparation)), with preliminary target selection details published for the MWS (Allende Prieto et al. 2020), BGS (Ruiz-Macias et al. 2020), LRGs (Zhou et al. 2020), ELGs (Raichoor et al. 2020), and QSOs (Yèche et al. 2020) samples. A series of five papers describe the target selection for these objects (MWS, Cooper et al. 2022; BGS, Hahn et al. 2023; LRG, Zhou et al. 2023; ELG, Raichoor et al. 2023; QSO, Chaussidon et al. 2023).

Quasars at redshift higher than 2.1 are identified by the DESI pipeline as high redshift quasars suitable for Lyman-α forest studies. These will be re-observed, allowing for a better measurement of their spectra. Re-observations of the same quasar are then coadded and stored in files grouped by healpix pixels (see Górski et al. 2005). The coaddition consists in a simple weighted average of fluxes for each wavelength in the spectra, using as weights their inverse-variance. Coadded files alongside the quasar catalogue are used by our pipeline to build the Lyman-α forest catalogue.

Each coadded file contains per-spectra information for the three arms (B, Z, R), as well as various metadata, including fiber positions, instrument configuration during observation and atmospheric conditions. The spectroscopic data include fluxes, estimated inverse-variance and a mask identifying invalid pixels.

Each of the three arms of the spectrograph covers a different wavelength region:

(i) B arm: [3600, 5800] Å
(ii) R arm: [5760, 7620] Å
(iii) Z arm: [7520, 9824] Å

with some overlap between each of the arms. The Lyman-α forest is predominantly observed in the blue arm (B arm) of the spectrograph, and only partially in the red arm (R arm). DESI follows a linear wavelength grid with 0.8 Å steps, `picca` is adapted to work with this resolution.

In Fig. 1, we show the distribution of wavelength and pipeline-reported error on the flux measurements for the EDR + M2 data set, where $\sigma_{\rm pip}$ the variance reported by the DESI pipeline based on photon statistics and readout noise. The image highlights the main differences between EDR and M2 samples. EDR data corresponds to longer observations, with fewer quasars observed. M2 data, on the other hand, includes a large number of objects, but with shorter observations, leading to a larger value for $\sigma_{\rm pip}$. The same panel shows an increase in variance around 4400 Å, caused by the collimator mirror reflectivity (Guy et al. 2023). For reference, the distribution of fluxes is centred around $5 \cdot 10^{-18}$erg s^{-1}cm^{-2}Å^{-1} in the blue end of the spectra, and $5 \cdot 10^{-19}$erg s^{-1}cm^{-2}Å^{-1} in the red end.

To match the accuracy of our most reliable mock data (Herrera et al. (in preparation)) and due to the limited amount of Lyman-α forest data available at longer wavelengths, we have limited our analysis to fluctuations fulfilling $z < 3.79$. Given that Lyman-α fluctuations can be related to a specific location in the wavelength grid ($z = \lambda/\lambda_{\rm Ly\alpha} - 1$), this redshift cut corresponds to a maximum wavelength of 5772 Å entering our analysis.

2.2 Description of the quasar catalogue

Objects in the spectroscopic data are identified using the template-fitting code Redrock (Bailey et al. (in preparation), and for the special case of QSOs, Brodzeller et al. 2023). Redrock employs a set of templates as representatives of the main object classes observed by DESI: quasars, galaxies, and stars. For each observed spectrum, Redrock determines the best-fitting redshift and template by comparing it to the set of templates. This process allows Redrock to accurately identify the objects in the DESI spectroscopic data.

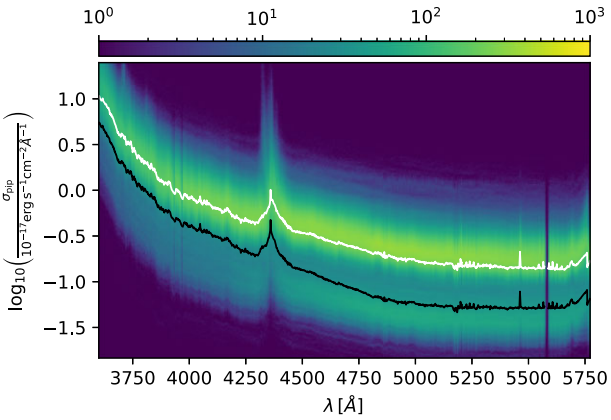

Figure 1. Distribution of wavelength and pipeline-reported error on the flux measurement for the larger EDR + M2 sample. This measurement is performed in the C III region of the spectra, as defined in Table 2. The solid lines mark the mean value of the error for a given wavelength, black for the EDR sample and white for the EDR + M2. Apart from the clear distinction in these two lines, a subtle division into two bands at larger wavelengths can also be observed. This is caused by the better signal-to-noise in the EDR sample, thanks to multiple re-observations of the same targets. We can also observe how the variance relatively increases in the two ends of the spectrograph and in the area affected by the collimator mirror reflectivity around 4400 Å.

However, Redrock sometimes misidentifies quasars as galaxies. To address this issue, a set of independent quasar-identification approaches, referred to as 'afterburners' are also run. These afterburners can identify features missed by Redrock and improve the accuracy of quasar identification. In our case, the most relevant afterburners are QuasarNet (Busca & Balland 2018; Farr, Font-Ribera & Pontzen 2020), the Mg II afterburner, and SQUEzE (Pérez-Ràfols et al. 2020). While only the first two methods were used to build the final catalogue, all of them were tested during the SV phase (Alexander et al. (2023), also see Lan et al. (2023) for SV results on galaxies).

For the case of QuasarNET, its main role is to correct for cases where Redrock identifies an object as a low redshift galaxy when it is actually a high redshift quasar. In these cases, QuasarNet is used to re-identify the object as a quasar based on its spectral features using Machine Learning techniques and trained using visually inspected data sets. If QuasarNet identifies an object as a high redshift quasar, Redrock is run again with a high redshift prior. This will likely change the object identification to a high redshift quasar, and if confirmed, the object would be included in the final catalogue.

The other relevant afterburner is the Mg II afterburner. Its main role is to correct for cases where Redrock misidentifies a quasar as a galaxy. For every object identified as a galaxy by Redrock, the Mg II afterburner checks the width of the Mg II emission line. If the line is wide enough, the object is re-identified as a quasar and included in the final catalogue.

For this release, only objects targeted and confirmed as quasars are used. The resulting catalogues include a total of 68 750 quasars for the EDR sample and 318 691 quasars for the EDR + M2 (see Table 1). This value includes all quasars in the input sample, being the actual value of quasars used for each region detailed in Table 2. The redshift and spatial distributions of the quasars is shown in Fig. 2,

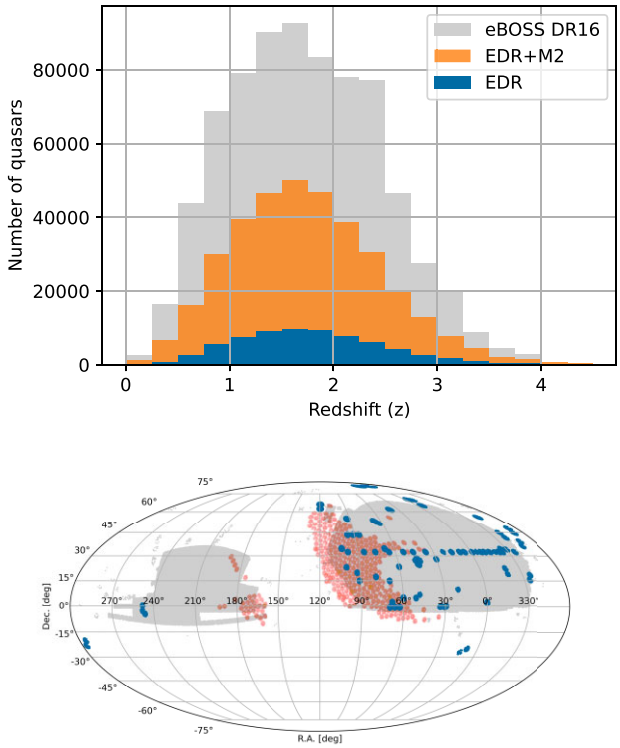

Figure 2. Distribution of objects in the two samples used in this publication, compared with the final eBOSS DR16 sample presented in du Mas des Bourboux et al. (2020). Top: Redshift distribution of the three catalogues. Bottom: Sky distribution of the same catalogues. We observe that M2 makes a significant portion of the EDR+M2 sample. When compared to the eBOSS data, the EDR + M2 sample is approximately halfway to matching the number of objects in the former.

compared to the larger catalogue used in eBOSS DR16 analysis (du Mas des Bourboux et al. 2020). The number of quasars in the EDR + M2 is clearly dominated by the first 2 months of main survey (M2), containing almost half of the objects in the eBOSS sample.

2.3 BAL and DLA information

Damped Lyman-α Absorption (DLA) systems caused by H I-rich galaxies affect the Lyman-α forest by generating absorption features that can interfere with the continuum fitting process. DLA information is provided using a DLA finder based on a convolutional neural network and a Gaussian process, for a full description of the DLA catalogue see Zou et al. (in preparation). We selected from the full sample the objects where both the convolutional neural network and the Gaussian process detected a DLA with a confidence larger than 50 per cent.

Apart from DLAs, the Lyman-α forest can also be affected by BALs believed to be caused by the existence of ionized plasma outflows from the accretion disc. These BAL quasars can be added to the analysis, although they have to be appropriately masked (see Section 3.1). The algorithm for BAL identification was presented in Guo & Martini (2019), and the detailed description of the BAL catalogue for DESI data is presented in Filbert et al. (2023).

The number of DLAs and BALs for both EDR and EDR + M2 samples is shown in Table 1. Given the better signal to noise in the EDR sample, the identification of DLA and BAL objects in this sample is higher in this smaller data set.

Table 1. Number of Lyman-α quasars in the two samples, number of quasars showing BAL features and number of DLAs affecting forests from the Lyman-α quasars. The better signal to noise in the EDR sample allows for the detection of a higher number of DLA and BAL features.

	EDR	EDR + M2
Quasars	68 750	318 691
DLAs	5006	17 375
BALs	9294	28 185

3 SPECTRAL REDUCTION

The continuum fitting procedure estimates the expected flux for each of the quasars in the catalogue, this process is essential for computing Lyman-α fluctuations. The flux-transmission field can be defined as:

$$\delta_q(\lambda) \equiv \frac{f_q(\lambda)}{\overline{F}(\lambda) C_q(\lambda)} - 1, \quad (1)$$

where $\overline{F}$ is the mean transmitted flux at a specific wavelength, C_q the unknown unabsorbed quasar continuum for quasar q, and f_q its observed flux. The combination $\overline{F}C_q(\lambda)$ is the mean expected flux of the quasars and is the quantity that we fit for the spectra.

Continuum fitting is the procedure to compute the flux-transmission field. It can be split in an initial clean-up phase and a second phase where the quasar continuum is actually fitted. The clean-up phase involves two sequential procedures: masking multiple unmodelled features and re-calibrating the spectra to eliminate residuals from the DESI calibration procedure.

In this section, we first explain these two procedures and then provide details on the core of the continuum fitting process.

3.1 Masks (sky lines, galactic absorption, DLA, and BAL)

There are multiple features in the spectra that are not caused by Lyman-α fluctuations but are still present in our spectra due to contamination. In order to simplify the cosmological modeling when using Lyman-α fluctuations for correlation measurements, we mask these features and remove them from our analysis.

The three type of masks applied in our analysis are: DESI pipeline; BAL and DLA, applied to absorption features in the forest[2] region; and galactic absorption and sky emission lines that appear in the spectrograph when observing Lyman-α quasars. We note that the DESI pipeline mask is applied before the individual observations are co-added (see Section 2.1). The other masks are applied after the coaddition.

3.1.1 DESI pipeline masking

A masking process is applied within the DESI pipeline to identify bad pixels in the spectra. These are typically caused by CCD defects or cosmic rays hitting the spectrograph CCD. This mask is found to only affect about 0.1 per cent of the DESI pixels used for the Lyman-α analysis. Given the small fraction of pixels discarded, we decided to remove these pixels from the analysis.

[2]We will refer to each individual line-of-sight as forest, following the convention from previous Lyman-α publications.

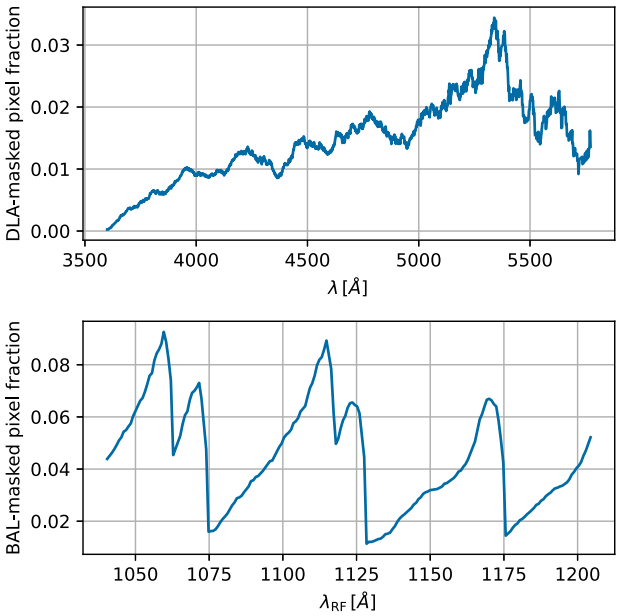

Figure 3. Fraction of pixels masked due to DLAs (top) and BAL (bottom) features. As expected, the number of detected DLAs increase with redshift (and therefore with λ), yielding a fraction of masked pixels always below the 5 per cent. For the BAL case, we show the masked fraction as a function of λ_{RF}, which allow us to observe the strong wavelength dependence of the masking, associated with emission lines for different elements. This figure used the full EDR + M2 data set.

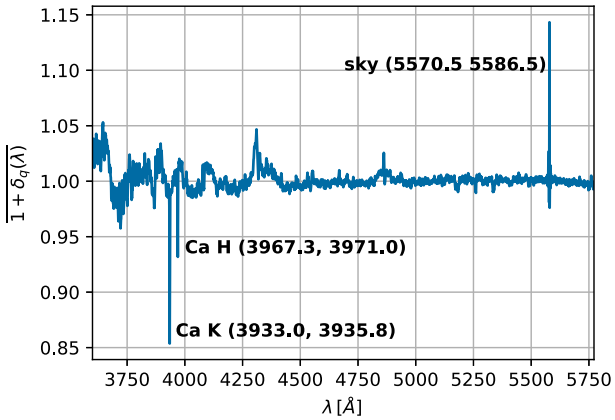

Figure 4. Weighted average of the flux-transmission field measured in the C III region. This measurement has been performed without masking sharp features, and they can be clearly observed at different positions in the spectrograph. The three masked regions are specified in the plot, showing the wavelength range affected by the mask. Smooth features in this measurement can also be observed. Their impact can be corrected through the flux re-calibration (see Section 3.2). For this figure, we used the full EDR + M2 data set.

3.1.2 DLA and BAL masking

DLA absorption affects our spectra by imprinting itself in the spectra of quasars, resulting in zero flux around the true redshift of the DLA and damping wings further away. A secondary effect of this is the reduction in the mean flux of the affected spectra, affecting the overall mean absorption. This biases our estimate of $\overline{F}C_q$, potentially affecting all the quasars in the sample (see Section 3.3).

Although it is possible to include the absorption features from DLAs into our models for the correlation functions, for simplicity, and following the prescription in previous analyses (du Mas des Bourboux et al. 2020), we mask the regions of the Lyman-α forest that are affected by identified DLAs when the DLA reduces the transmission by more than 20 per cent. We then correct the absorption in the wings using a Voigt profile as suggested by Noterdaeme et al. (2012), being able to include the pixels affected by these wings without affecting the mean absorption.

The top panel of Fig. 3 shows the fraction of pixels that have been masked due to DLAs as a function of observed wavelength. Although there is an increase in the number of detected DLAs with redshift, the number of masked pixels is always smaller than 5 per cent.

Regarding BALs, in previous analyses quasars exhibiting BAL features were directly excluded from the analysis, although they were later used as tracers for cross-correlations between quasars and the Lyman-α forest. However, Ennesser et al. (2022) showed that quasars with BAL features could be safely included in the analysis if all expected locations of these absorption features are masked. Here, we follow their proposed approach.

In the bottom panel of Fig. 3, we observe the fraction of pixels masked due to BALs as a function of rest-frame wavelength. In this case, we can see a pattern of absorption that matches the expected absorption features described in Ennesser et al. (2022). Since not all the BAL quasars suffer from the same absorption profile, we expect the BAL-masked pixel fraction to be maximal around the central absorption and decrease as we go away from it. It is worth noting that for the quasars used in the Lyman-α region continuum fitting, the percentage of BAL quasars is 16 per cent for the EDR + M2 sample and 23 per cent for the EDR sample.

3.1.3 Galactic absorption and sky emission masking

There are certain absorption and emission features in our spectra that differ from Lyman-α fluctuations or other absorbers in the IGM. These absorption and emission features can be easily identified because they affect specific wavelengths, resulting in sharp features when the spectra of multiple quasars is combined. We can remove them by simply masking out the corresponding wavelengths from our analysis.

The two most significant features in this regard are galactic absorption and sky emission. Galactic absorption is caused by material in the interstellar medium (ISM) absorbing at specific wavelengths. The ISM absorption for some of these lines cannot be easily separated from the intrinsic absorption in the atmospheres of the stars used for flux calibration, and thus they are not properly accounted for. Here, we are affected by the Ca K and H transitions, and we mask the corresponding wavelengths (see Fig. 4).

Sky emission comprises emission lines generated by atmospheric effects and are mostly corrected in the modelling of sky lines. However, inaccuracies in this modelling result in spurious features in our spectra. Here, we also mask the affected wavelengths (see Fig. 4).

In Fig. 4, we present the measurement of the estimated $\overline{1+\delta_q(\lambda)}$ in the C III region (see Table 2). Two kind of features can be observed in this plot: first, we observe the sharp features corresponding to the mentioned galactic absorption and sky emission; then we observe smooth features caused by inaccuracies in the DESI calibration process. To correct for the former, we mask the pixels in the wavelength intervals shown in the plot. The later can be corrected through re-calibration (see Section 3.2).

Table 2. Statistics for the regions considered during the analysis, the region span is defined in the quasar rest-frame wavelength (λ_{RF}). The number of forests corresponds to the number of quasars whose spectra can be observed in the spectrograph. A minor number of forests are rejected due to low SNR or due to them being too short.

Region	$\lambda_{RF,\,min}$ Å	$\lambda_{RF,\,max}$ Å	# forests EDR	# forests EDR + M2
C III	1600	1850	49 810	233 310
C IV	1410	1520	41 445	189 984
Mg II-R	2900	3120	7290	33 936
Mg II	2600	2760	13 373	62 628
S IV	1260	1375	34 044	152 979
Lyman-α	1040	1205	20 281	88 511

Here we use the C III region to build our masks as it lacks the Lyman-α absorption features. This makes it easier to estimate absorption and emission effects.

3.2 Re-calibration

DESI calibrates fluxes using standard stars (Guy et al. 2023). However, inaccuracies in the modelling of these calibration stars introduce features in the measured spectra when fluxes are predicted for other objects (such as quasars). These features can be observed by examining the mean $\delta_q(\lambda)$ in any region of the spectra, particularly in the regions without Lyman-α absorption, where the variance of the measured flux is expected to be lower. This is possible because the features caused by calibration defects are function of observed wavelength, while the continuum estimate is a function of rest-frame wavelength. Looking again at Fig. 4, we can observe smooth features in the stack of fluctuations ($\overline{1+\delta_q(\lambda)}$) alongside the masked sharp features.

These features are correlated between different forests, and could therefore bias the measurement of the correlation function. To avoid this, on top of the DESI pipeline calibration, we re-calibrate our spectra. As stated above, these features affect the same region of observed wavelength regardless of the rest-frame position where they act. Therefore, one could in principle use the entire quasar spectra for calibration purposes. However, the larger spectral diversity near quasar emission lines can unnecessarily complicate this. In consequence, we search for a featureless region redwards of the Lyman-α emission line. We will refer to all the candidate regions as re-calibration regions.

In the absence of Lyman-α fluctuations, flux fluctuations are ideally only caused by noise. In any of the re-calibration regions, the measurement of $\overline{1+\delta_q(\lambda)}$ is expected to be consistent with 1; and its measurement can be considered a null test that is not fulfilled in general given the issues described above.

In the bottom panel of Fig. 5, we show the measurement of $\overline{1+\delta_q(\lambda)}$ for multiple different candidate re-calibration regions. The fact that all of them show similar features, in spite of being at different regions of the spectra, suggests that all fluxes can be corrected consistently. Furthermore, we see comparable features in the white dwarf (WD) average residuals in the top panel of the same figure. This similarity can be seen in some of the regions of the spectrum, especially for $\lambda \in [3600, 4200]$ Å, and it further justifies the re-calibration process.

In previous analyses, a region to the right of the Mg II emission line (Mg II-R) was selected to re-calibrate fluxes, partly because it is located further to the right of the spectra, reducing potential contamination from other absorption lines. However, in this work, we

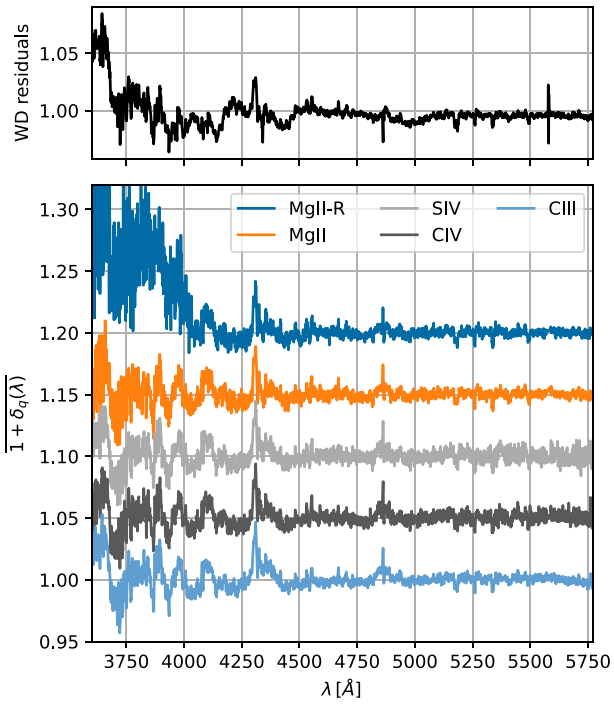

Figure 5. Bottom: Weighted average of the flux-transmission field measured at different regions of the spectra. Its value has been shifted to better distinguish the different regions. As opposed to the results shown in Fig. 4, sharp features are not present here because these samples have already been masked. The smooth features in the spectra are similar for all the measured regions, being the Mg II-R an outlier in this tendency, likely to be caused due to the reduced number of pixels available for this region at low wavelengths (see Fig. 6). Results are computed from the full EDR + M2 sample. Top: Average residuals to WD spectra in the blue arm of the DESI spectra, as seen in Guy et al. (2023). A similar trend can also be observed here between these residuals and our measured average of the flux-transmission field, especially at smaller wavelengths, further justifying our re-calibration process.

selected the C III region ($\lambda \in [1600, 1850]$ Å) due to the larger number of pixels available and given the similar behaviour compared to the other regions. Fig. 6 shows the number of pixels at each wavelength bin of size $\Delta\lambda = 55.58$ Å for the different regions, and the number of forests available for each region can be found in Table 2. In both cases, we see that C III has the largest number of pixels available for the analysis.

The choice of a re-calibration region at a larger wavelength than the Lyman-α forest also leads to an increase in the number of quasars that can be used for this process. The Lyman-α forest analysis includes

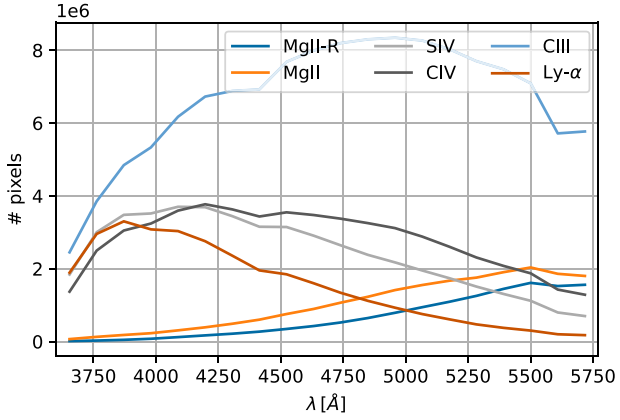

Figure 6. Number of pixels available for the different regions measured in bins of $\Delta\lambda = 55.58$ Å. Given the location of the different regions at different positions in the quasar spectra, the number of available pixels is different for each of them. This is because our spectral coverage lies in the range (3600, 5772) Å. Thanks to the larger number of pixels available for the C III region, it was selected as the region to be used for the re-calibration process. This figure used the full EDR + M2 sample.

quasars with redshifts in the range $z \in (2.1, 3.7)$, while for the C III region, quasars in the range $z \in (0.9, 2.6)$ are included. By looking at the quasar distribution (Fig. 2), we see that the number of objects in the second range will be larger, and hence a larger number of pixels available in the re-calibration region.

We use a similar procedure as du Mas des Bourboux et al. (2020) for this re-calibration, correcting all the fluxes by using the mean flux in the re-calibration region:

$$f_{\rm new}(\lambda_i) = f_{\rm orig}(\lambda_i)/\overline{f_{\rm calib,orig}}(\lambda_i)\ . \tag{2}$$

The error reported by the pipeline for these fluxes also needs to be corrected through:

$$\sigma_{\rm pip,new}(\lambda_i) = \sigma_{\rm pip,orig}(\lambda_i)/\overline{f_{\rm calib,orig}}. \tag{3}$$

After this correction, smooth features in the spectra will be alleviated, which prevents biased results. In previous analyses, a second correction was applied to correct the estimation of the flux variance provided by the pipeline (see Appendix A). However, the better performance of DESI in this aspect allows us to skip this step. Evidence in this direction is the better behaviour of the estimated correction to the pipeline error (see Fig. 7).

3.3 Continuum fitting

The core of the continuum fitting process consists of obtaining the expected flux $\overline{F}C_q$ for each quasar in the sample. This allows for the determination of the flux-transmission field (equation 1). The process is performed iteratively, with several quantities fitted simultaneously, and performed on each of the quasar regions independently: in the case of a re-calibrated Lyman-α analysis, it is firstly performed in the re-calibration region (C III in this case), and afterwards in the Lyman-α region.

In order to simplify the process, the expected flux $\overline{F}(\lambda)C_q(\lambda)$ is assumed to be a universal function of rest-frame wavelength, $\bar{C}(\lambda_{\rm RF})$, corrected by a first degree polynomial in log λ.:

$$\overline{F}(\lambda)C_q(\lambda) = \overline{C}(\lambda_{\rm RF})\left(a_q + b_q\frac{\Lambda - \Lambda_{\rm min}}{\Lambda_{\rm max} - \Lambda_{\rm min}}\right), \tag{4}$$

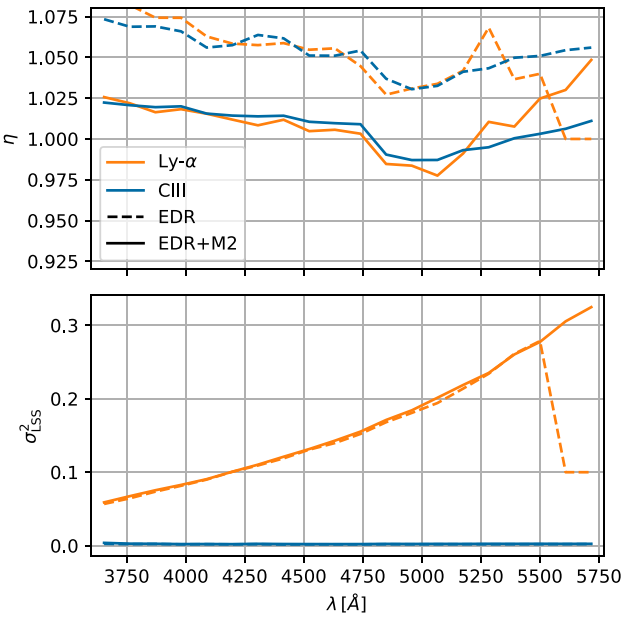

Figure 7. Wavelength evolution of the fitted parameters η and $\sigma^2_{\rm LSS}$, measured in both the C III and Lyman-α regions for the EDR (dashed) and EDR + M2 (solid) samples. Top: The pipeline error correction η is found to be larger for the EDR sample, caused by the worse estimation of the pipeline-reported variance for the first part of the SV program. Bottom: $\sigma^2_{\rm LSS}$, in this case it is consistent between the two samples. As expected the C III region shows a value close to 0 for all wavelengths, while for the Lyman-α region follows the expected increase in its intrinsic variance with redshift. In both η and $\sigma^2_{\rm LSS}$, measurements for the EDR sample, the fitted parameters could not be obtained for the larger wavelength value due to the reduced number of pixels available, falling to the default values $\eta = 0.1$ and $\sigma^2_{\rm LSS} = 0.1$.

where $\Lambda \equiv \log\lambda$ and $\Lambda_{\rm min,\ max}$ identify its minimum and maximum values inside the region to be fitted. That is, the spectra of quasars in our sample are assumed to have the same underlying shape or continuum for all objects, allowing for variations in the amplitude and tilt. In the analysis, the transmitted flux $\overline{F}$ and the quasar continuum C_q cannot be fitted independently, although they are not needed separately in our analysis. For this reason, we choose to directly estimate the expected flux of quasars.

The parameters $(a_q,\ b_q)$ are fitted by maximizing the likelihood function

$$2\ln L = -\sum_i \frac{\left[f_i - \overline{F}C_q(\lambda_i, a_q, b_q)\right]^2}{\sigma_q^2(\lambda_i)} - \sum_i \ln\left[\sigma_q^2(\lambda_i)\right], \tag{5}$$

where $\sigma_q^2(\lambda)$ is the variance of the flux f_i. This value has to be estimated, and since it depends on $(a_q,\ b_q)$, we include this dependence in the likelihood function.

The full variance of the flux, $\sigma_q^2(\lambda)$, includes not only the obvious contribution from the noise estimated by the DESI pipeline but also the intrinsic variance of the Lyman-α forest.[3] We account for the intrinsic variance in the following way:

$$\frac{\sigma_q^2(\lambda)}{\left(\overline{F}C_q(\lambda)\right)^2} = \eta(\lambda)\tilde{\sigma}^2_{\rm pip,q}(\lambda) + \sigma^2_{\rm LSS}(\lambda). \tag{6}$$

where $\sigma^2_{\rm LSS}$ is the mentioned intrinsic variance of the Lyman-α forest, and $\tilde{\sigma}_{\rm pip,q} = \sigma_{\rm pip,q}(\lambda)/\overline{F}C_q(\lambda)$ where $\sigma_{\rm pip,q}$ is flux variance

[3]This quantity is expected to be very small outside the Lyman-α forest region.

as estimated by the DESI pipeline. In this expression, we also include a correction $\eta(\lambda)$ to account for inaccuracies in the pipeline noise estimation. We discard here an extra term in this expression accounting for quasar variability at high signal-to-noise ratio (SNR), which was previously used in du Mas des Bourboux et al. (2020) (see A2 for details). It is worth noting that σ_q^2 here is the variance of the flux and therefore has flux units, whether all quantities at the right-hand side are dimensionless.

In the iterative continuum fitting process, we estimate the quantities η, $\sigma^2_{\rm LSS}$, a_q, b_q, $\overline{C}(\lambda_{\rm RF})$. The sequential steps of the process are:

(i) **Assume wavelength-independent values for unknown quantities:** A flat assumption is assumed for all the quantities to be fitted and measured. In practice, this means setting $\overline{C}(\lambda_{\rm RF}) = 1$, $\eta(\lambda) = 1$ and $\sigma^2_{\rm LSS}(\lambda) = 0.1$.

(ii) **Fit the per-quasar parameters (a_q, b_q):** Using the current values of $\overline{C}(\lambda_{\rm RF})$, $\eta(\lambda)$ and $\sigma^2_{\rm LSS}(\lambda)$, the pair of values (a_q, b_q) is estimated for each individual forest through the minimization of the likelihood defined in equation (5).

(iii) **Estimate the flux-transmission field $\delta_q(\lambda)$:** Using the current value $\overline{C}(\lambda_{\rm RF})$ and the best-fitting parameters for (a_q, b_q), we compute the expected flux for each quasar following equation (4). This allows for the computation of the fluctuations around the expected flux ($\delta_q(\lambda)$) as defined in equation (1).

(iv) **Fit the variance functions:** The functions $\eta(\lambda)$ and $\sigma^2_{\rm LSS}(\lambda)$, defined in equation (6), are now fitted using the estimated $\delta_q(\lambda)$ from the previous step. In order to do so, different values of the flux-transmission field are grouped by wavelength and $\tilde{\sigma}_{\rm pip}$. We take 20 bins for the wavelength split and 100 for the split on $\tilde{\sigma}_{\rm pip}$, generating a grid of 2000 points. We compute the variance $\sigma^2(\lambda, \tilde{\sigma}_{\rm pip}) = \overline{\delta_q^2(\lambda, \tilde{\sigma}_{\rm pip})}$ for each point in the grid.[4] The functions $\eta(\lambda)$ and $\sigma^2_{\rm LSS}$ are fitted for each of the 20 wavelengths independently using the following likelihood function:

$$2 \ln L = -\sum_{\tilde{\sigma}_{\rm pip}} \frac{\left[\delta_q^2(\lambda, \tilde{\sigma}_{\rm pip}) - \eta(\lambda)\tilde{\sigma}^2_{\rm pip} - \sigma^2_{\rm LSS}(\lambda)\right]^2}{\delta_q^4(\lambda, \tilde{\sigma}_{\rm pip})}\,, \tag{7}$$

where the sum in $\tilde{\sigma}_{\rm pip}$ include all the valid bins in $\tilde{\sigma}_{\rm pip}$ (at most 100). The points in the grid computed using less than 100 pixels are considered unreliable and discarded from the fits. We note that in those cases where the fit does not converge (for example, if there are too few quasars), then the default values from step (ii) are kept.

(v) **Recompute the mean expected flux:** At this stage we update the value of $\overline{C}(\lambda_{\rm RF})$. This is performed by computing the weighted average of all quasars expected flux sharing the same $\lambda_{\rm RF}$ value. Here, we use the optimal weights as defined in equation (8; see Section 4.1).

(vi) **Compute and save relevant statistics:** Relevant statistics are computed at each iteration and stored in `delta_attributes_iteration{i}.fits.gz` files (where i is the iteration number). The saved statistics include the stack of fluctuations ($\overline{1 + \delta_q(\lambda)}$), the fitted variance functions (η, $\sigma^2_{\rm LSS}$), the mean continuum ($\overline{C}(\lambda_{\rm RF})$) and fit metadata, including the tilt and slope values (a_q, b_q), the number of pixels used for the fit and the χ^2 of the fit.

(vii) **Continue next iteration starting in step (ii):** The next iteration starts using the updated values of $\overline{C}(\lambda_{\rm RF})$, η and $\sigma^2_{\rm LSS}$.

[4]Here, we assume $\overline{\delta_q(\lambda, \tilde{\sigma}_{\rm pip})} = 0$ as an approximation. Fluctuations around this assumption are small (see Fig. 8)

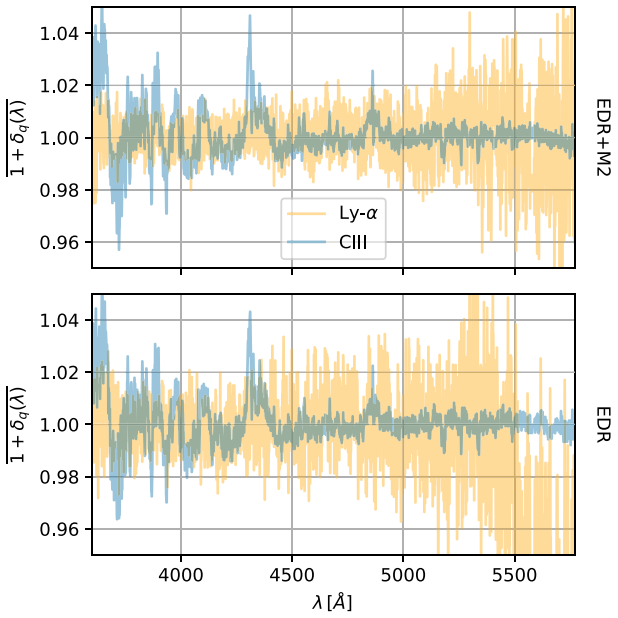

Figure 8. Measurement of the weighted mean of the flux-transmission field for both the Lyman-α and C III regions. The Lyman-α region shows an expected higher variance at all wavelengths compared to the C III region, although it lacks smooth features thanks to the re-calibration process. Similarly, the EDR sample shows a higher variance than EDR + M2. In both cases, this is caused by the decrease in number of pixels available. Given the low number of pixels available at larger wavelengths for the EDR sample for the Lyman-α region, it departs from the expected unity behaviour.

This process is performed five times, resulting in stable estimates of all the defined quantities.

In Fig. 7, we show the final estimates for the two fitted function η and $\sigma^2_{\rm LSS}$ for the two regions considered in this analysis (Lyman-α and C III) and for the two different samples (EDR and EDR+M2). The correction to the pipeline estimated variance, η, shows a ∼2 per cent deviation from the ideal value of 1 in the case of EDR + M2 and up to ∼7.5 per cent deviation in the case of EDR. This difference is due to the worse estimation of the pipeline-reported variance for the first part of the SV program (see Ravoux et al. 2023 for details). For $\sigma^2_{\rm LSS}$, its estimation is similar for both samples, showing a clear dependence on wavelength (and hence redshift) for the Lyman-α region and an expected (due to the lack of Lyman-α fluctuations) nearly zero value for the C III region. The fit fails at the higher wavelength values in the case of the EDR sample due to the small number of pixels available at these wavelengths, and then the initial value is kept.

Fig. 8 shows the mean flux-transmitted field for the same two regions and samples. In this case, differences between the two samples are smaller, only showing a higher variance in the case of the smaller (EDR) sample. The variances are particularly worse at the largest wavelengths in the Lyman-α region due to the reduced number of pixels available. The C III region, for which no re-calibration has been performed, still shows the smooth features already discussed in Section 3.2. They do not appear in the already re-calibrated Lyman-α region, although there is a higher variance caused by the less number of pixels (as shown in Fig. 6).

3.3.1 *Example of quasar fit*

An example of a high SNR quasar spectrum is shown in Fig. 9. We can see the Lyman-α forest at the left side of the Lyman-α emission line

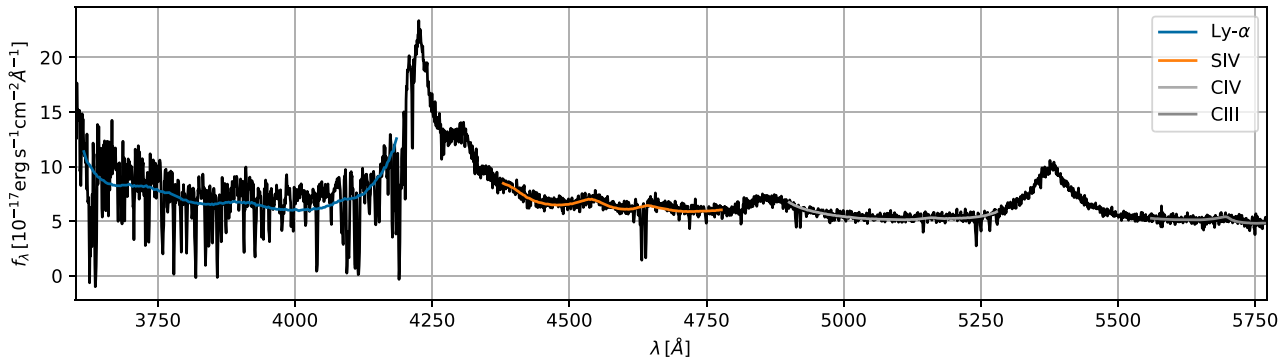

Figure 9. Example of a high signal-to-noise quasar spectrum. The mean expected flux $\overline{C}(\lambda_{\rm RF})$ as in equation (4) is shown for the Lyman-α, S IV, C IV, and C III regions. The quasar has a redshift $z_q = 2.495$ and is identified as a DESI object with TARGETID = 39628443918272474. As in this example, we occasionally observe metal absorption in the calibration regions.

($\lambda_{\rm RF} = 1215.7$ Å), with the characteristic absorption features. The solid lines show the mean expected flux for that particular forest at different spectral regions. Metal absorption features can be observed in this example for the S IV and C IV re-calibration regions.

4 DISCUSSION

The main science driver of the Lyman-α catalogue presented in this publication is the measurement of 3D correlations that allows us to constraint the BAO scale (Gordon et al. 2023). In this section, we discuss two methodological novelties in the construction of the Lyman-α catalogue with respect to previous eBOSS analyses (du Mas des Bourboux et al. 2020), and we do this by looking at the impact of these changes to the precision with which we will be able to measure the 3D correlations. In particular, in Section 4.1 we discuss the optimization of the weights assigned to each Lyman-α fluctuation, while in Section 4.2, we will choose the rest-frame wavelength range based on the precision of the correlation function measurements.

We conclude in Section 4.3 with a discussion on the distribution of per-quasar parameters that capture the diversity of quasar continua in the data set.

4.1 Optimal weights

Correlations in the Lyman-α forest are estimated as weighted averages of products of fluctuations $\delta_q(\lambda)$ in pixel pairs at a given separation. An optimal quadratic estimator would use the inverse of the pixel covariance as the weight matrix, but the large number of pixels in current Lyman-α data sets makes this inversion not feasible.

Ignoring the small correlation between pixels in different quasar spectra, one can approximate the covariance as block-diagonal, and make the inversion tractable. This approximation has been used in several measurements of the 1D power spectrum (McDonald et al. 2006; Karaçaylı, Font-Ribera & Padmanabhan 2020; Karaçaylı et al. 2022; Göksel Karaçaylı et al. 2023), it has been proposed to measure the 3D power spectrum (Font-Ribera, McDonald & Slosar 2018) and it was used in one of the first BAO measurements with the Lyman-α forest (Slosar et al. 2013).

Recent measurements of 3D correlations in the Lyman-α forest (Delubac et al. 2015; Bautista et al. 2017; de Sainte Agathe et al. 2019; du Mas des Bourboux et al. 2020), on the other hand, have ignored all correlations between pixels and have used instead a diagonal weight matrix. These studies weighted each pixel with the inverse of its variance, including the instrumental noise and the intrinsic fluctuations [see equation (6)], effectively approximating the inverse covariance matrix with the inverse of its diagonal elements. This approximation is simple and easy to implement, but it is not the optimal diagonal weight matrix.

In this section, we study a simple modification of our diagonal weight matrix, where we add an extra free parameter $\sigma^2_{\rm mod}$ that modulates the contribution of the intrinsic fluctuations $\sigma_{\rm LSS}$ to the weights:

$$w_q(\lambda) = \frac{1}{\eta(\lambda)\tilde{\sigma}^2_{{\rm pip},q}(\lambda) + \sigma^2_{\rm mod}\sigma^2_{\rm LSS}(\lambda)} \,. \tag{8}$$

A small value of $\sigma^2_{\rm mod}$ is equivalent to weighting the pixels based solely on the noise variance, while a large value gives the same weight to all pixels at a given redshift, regardless of instrumental noise.

In Fig. 10, we show that a value of $\sigma^2_{\rm mod}$ around 7-8 can improve the precision of the auto-correlation measurement by 25 per cent, at no additional cost. As expected, the gain in precision in the cross-correlation is roughly half of that, and we find a 10 per cent improvement for values of $\sigma^2_{\rm mod}$ around 6–7.

It is important to note that the actual gain will depend on the properties of the data set. For instance, $\sigma_{\rm pip}$ and $\sigma_{\rm LSS}$ change differently with the width of the pixels used, and we have tested that when using pixels of 2.4 Å (similar to the ones used in the BAO measurement of du Mas des Bourboux et al. 2020) the optimal value is smaller, $\sigma^2_{\rm mod} = 3.1$, and the gain in the auto-correlation is only 8 per cent.

For this publication and the associated Value Added Catalogue, we decide to use a value of $\sigma^2_{\rm mod} = 7.5$, and we leave for future work the implementation of a block-diagonal weighting.

4.2 Selection of rest-frame wavelength range

Due to diversity in quasar spectra near the emission lines, the rest-frame wavelength range that can be used for analysing the Lyman-α forest is limited. This happens at the red end of the spectra due to the Lyman-α emission line ($\lambda_{\rm RF} = 1215.67$ Å) and at the blue end limited due to the Lyman-β line ($\lambda_{\rm RF} = 1025.72$ Å).

Extending the analysis towards longer wavelengths closer to the Lyman-α emission line allows for the incorporation of more data. By including these extra pixels in our analysis, we can improve our measurements of the correlation function if the improvement due to

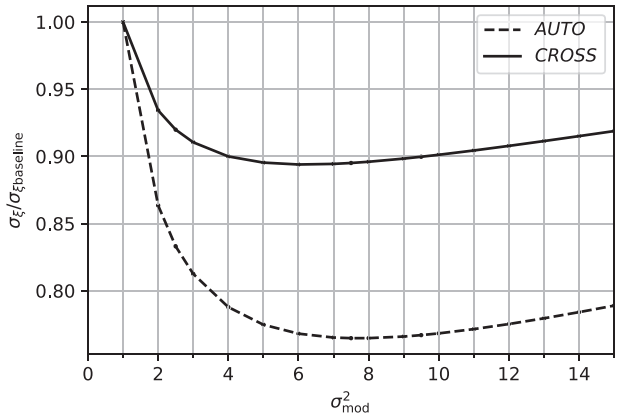

Figure 10. Measurement of the Lyman-α auto- (solid) and cross- (dashed) correlation errorbars for different choices of the σ^2_{mod} parameter, scaled to a reference value at $\sigma^2_{\mathrm{mod}} = 1$. The value of the errorbars was averaged over all scales, given the small scale dependency. The σ^2_{mod} parameter modified the inverse-variance weighting scheme as defined in equation (8), the use of a $\sigma^2_{\mathrm{mod}} \neq 1$ allows us for the optimization of the weighting scheme. In both auto- and cross-correlations, we observe a reduction in the size of the errorbars when we approach the optimal value of the σ^2_{mod} parameter. This optimal value is slightly different, but in both cases is found around 7–8. In the optimal value, the improvement in the measurement of the auto-correlation is about 20 per cent and 10 per cent for the case of the cross-correlation. Both measurements have been performed with the full EDR + M2 data set.

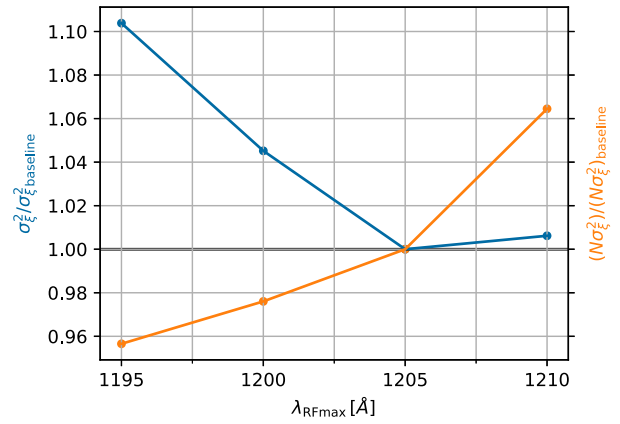

Figure 11. Comparison of the size of the errorbars in the Lyman-α auto-correlation for different choices of $\lambda_{\mathrm{RF,\,max}}$. The blue points show the errorbars size after averaging over all scales (given the small scale dependence of this quantity), and scaled to a reference value at $\lambda_{\mathrm{RF,\,max}} = 1205$ Å. The orange points show the equivalent measurement but in this case for the product of errorbar sizes times the number of pixel pairs, removing the dependence on the number of pairs used in the measurement. Following the blue points, we observe an improvement in the precision of our auto-correlation measurements when increasing the $\lambda_{\mathrm{RF,\,max}}$ parameter, having the optimal value at 1205 Å. The orange points show that the quality of these added points actually decrease for higher wavelengths, revealing that the improved performance is only driven by the inclusion of more information in our sample. We selected 1205 Å as our default value as a compromise between these two features. We used data from the whole EDR + M2 data set for these measurements.

the large number of pixels is not counterbalanced by of the larger pixel variance.

In Fig. 11, the blue points show the size of the errorbars in the auto-correlation measurement when we extend the analysis to higher

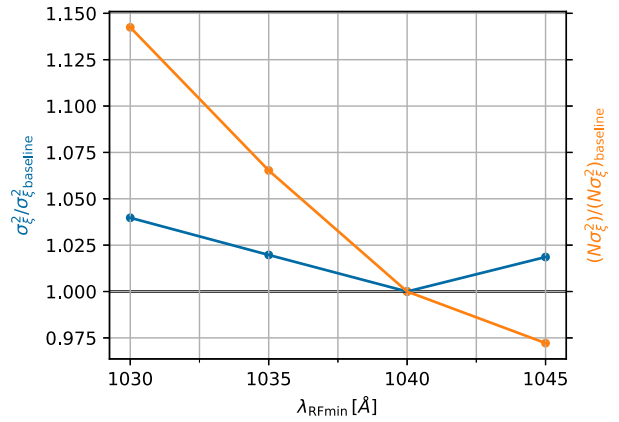

Figure 12. Identical measurements as the ones displayed in Fig. 11, in this case for values of $\lambda_{\mathrm{RF,\,min}}$. The optimal wavelength is found at $\lambda_{\mathrm{RF,\,min}} = 1040$ Å, while having similar features for the orange points: a decrease in the quality of the points added when approaching the emission line. Following the same arguments as in the case of $\lambda_{\mathrm{RF,\,max}}$, we decided to keep the limit at 1040 Å. Again, these measurements were performed using the EDR + M2 sample.

$\lambda_{\mathrm{RF,\,max}}$. The value at $\lambda_{\mathrm{RF,\,max}} = 1205$ Å yields the smallest errorbars. To separate the two effects of increased number of pixels and increased pixel variance, we also plot $\sigma^2_\xi N$, where N is the number of pairs in the correlation measurement, since we expect $\sigma^2 \propto 1/N$. This will show how valuable the added points are when extending the wavelength range. Its evolution in Fig. 11 shows that extending $\lambda_{\mathrm{RF,\,max}}$ to higher wavelengths adds less valuable information, and therefore the decrease in the size of the errorbars is only driven by the increase in the sample size.

Given this result, we decided to set $\lambda_{\mathrm{RF}} = 1205$ Å because we found that further increasing the limit did not add constraining power. The increased variance near the emission line is likely the reason behind this, as it is not accounted for in our calculations and could potentially affect our measurements if we approached it too closely. A more detailed study of quasar continuum variance, using a larger data set than the one presented here, is necessary to fully understand its effects. We leave this for future releases of DESI data.

The same exercise was performed for the blue side of the quasar spectrum, and the results are shown in Fig. 12. When compared to the prescription of previous analyses ($\lambda_{\mathrm{RF,\,min}} = 1040$ Å in du Mas des Bourboux et al. 2020), we observe a degradation of the errorbars either going to smaller $\lambda_{\mathrm{RF,\,min}}$ due to the decrease of pixel count; or to larger $\lambda_{\mathrm{RF,\,min}}$ approaching the emission line. For this reason we retain this prescription of $\lambda_{\mathrm{RF,\,min}} = 1040$ Å.

4.3 Per-quasar parameters (a,b)

As mentioned in Section 3.3, the quasar continua is assumed to follow a universal function of rest-frame wavelength, with a correction by a first degree polynomial parametrized by a_q and b_q. Quasar variability can be larger at both ends of the Lyman-α forest, due to the Lyman-β and Lyman-α emission lines, but variability in the weak quasar emission lines at 1017 and 1123 Å (Suzuki 2006) could also impact the results.

In Fig. 13, we examine the distribution of these two parameters to test this assumption. The plot does not show special features apart from the long tail at large values of $|b_q/a_q|$. This feature is caused by spectra where only a small part of the forest appears in the

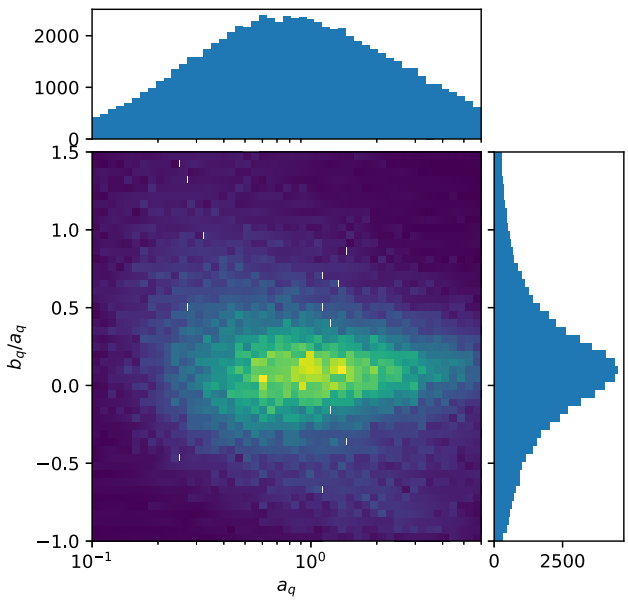

Figure 13. Distribution of a_q and b_q/a_q parameters defined in equation (4) for the EDR + M2 sample. The a_q parameter modifies the amplitude of the mean continuum $\overline{C}(\lambda_{\rm RF})$, while the b_q/a_q term introduces a modification that tilts it as a function of rest-frame wavelength. This last parameter relates to the intrinsic width of the spectral index. The faint long tail at large values of $|b_q/a_q|$ is caused by short forests in the sample, where the continuum fit can be problematic.

spectrograph. In these cases, fitting the real continuum of the quasar is difficult, leading to poor fits. For this reason, we require forests to have a minimum length of 150 pixels of 0.8 Å; pixels.

5 SUMMARY AND CONCLUSIONS

In this publication, we present the first measurement of Lyman-α fluctuations using DESI data. We use two data samples: EDR, which mainly consists of SV data; and EDR+M2, which also includes the first 2 months of the main survey (M2) data. The EDR sample contains 68 750 quasars, of which 20 281 have valid forests for Lyman-α studies, i.e. with the Lyman-α region visible and identifiable. The EDR + M2 sample contains 318 691 quasars, of which 88 511 have valid forests. We release the Lyman-α fluctuations catalogue for the EDR sample as part of the first DESI data release.

To achieve this measurement, we adapted the methodology proposed in previous analyses, especially the one outlined in du Mas des Bourboux et al. (2020), to suit the unique characteristics of the DESI early data sample. We discussed and applied several important modifications.

(i) We adjusted our pipeline to work with linear-spaced bins in wavelength to match the DESI scheme, preserving its format and precision.

(ii) We adapted the re-calibration process by simplifying it, removing unneeded steps that did not improve the performance of the whole re-calibration. We also switched the re-calibration region to allow for more pixels to be used in this process.

(iii) We simplified the weighting scheme by removing terms that added unnecessary complexity and did not contribute to the accuracy of the results, specially for this small data release.

(iv) By optimizing our diagonal weight matrix, we improved measurements of the auto-correlation by about 20 per cent, and of the cross-correlation by about 10 per cent

The presented flux-transmission field catalogue could be used for other studies apart from the standard BAO analysis. As in Font-Ribera et al. (2012) and Pérez-Ràfols et al. (2018a, 2018b), its cross-correlation with DLAs can be measured. A similar analysis could be performed with Strong Blended Lyman-α absorption systems, as in Pérez-Ràfols et al. (2022). The full-shape analysis of the Lyman-α forest 3D correlation function would allow for the measurement of the Alcock-Paczynski effect, as performed in Cuceu et al. (2023) using eBOSS DR16 data. Furthermore, IGM tomography is also possible with this sample.

As data sets get larger and larger, we will need to be more careful with the analysis. Weights used to measure clustering could be improved in two ways. As discussed in Section 4.1, the implementation of a block-diagonal weighting scheme will account for correlations between pixels within the same forest. Alternatively, taking into account that quasar diversity makes the estimation of the continua more difficult, one possible way of improving the weights is adding to the weighting scheme proposed in equation (6) an extra term accounting for errors in the estimation in the continuum. This term would be dependent on $\lambda_{\rm RF}$ and would be larger around the emission lines, where quasar diversity is expected to be higher. Finally, the analysis could be extended into the Lyman-β region.

This study can serve as a solid foundation for future research using DESI data. We are excited about the potential for further investigations in this area, as more comprehensive analyses become possible with the availability of future DESI releases.

ACKNOWLEDGEMENTS

CRP is partially supported by the Spanish Ministry of Science and Innovation (MICINN) under grants PGC-2018-094773-B-C31 and SEV-2016-0588. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. IPR is partially funded by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 754510. AFR acknowledges financial support from the Spanish Ministry of Science and Innovation under the Ramon y Cajal program (RYC-2018-025210) and the PGC2021-123012NB-C41 project, and from the European Union's Horizon Europe research and innovation programme (COSMO-LYA, grant agreement 101044612). IFAE is partially funded by the CERCA program of the Generalitat de Catalunya.

This material is based upon work supported by the US Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the US National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF's National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: https://www.desi.lbl.gov/collaborating-institutions. Any opinions, findings, and conclusions or recommendations expressed

in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

The authors are honoured to be permitted to conduct scientific research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

DATA AVAILABILITY

The Lyman-α forest catalogue for the EDR dataset is available in the DESI public database, and can be accessed at https://data.desi.lbl.gov/public/edr/vac/edr/lya/fuji/v0.3. For the case of the EDR + data set, it will be released alongside DESI Y1 data. Data points for the figures in this publication ara available in the zenodo repositor, at https://doi.org/10.5281/zenodo.8009535.

REFERENCES

Alexander D. M. et al., 2023, AJ, 165, 124
Allende Prieto C. et al., 2020, Res. Notes Am. Astron. Soc., 4, 188
Bautista J. E. et al., 2017, A&A, 603, A12
Blomqvist M. et al., 2019, A&A, 629, A86
Bolton A. S., Schlegel D. J., 2010, PASP, 122, 248
Brodzeller A. et al., 2023, AJ, 166, 66
Busca N., Balland C., 2018, preprint (arXiv:1808.09955)
Busca N. G. et al., 2013, A&A, 552, A96
Chaussidon E. et al., 2023, ApJ, 944, 107
Cooper A. P., et al., 2023, ApJ, 947, 37
Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44
Cuceu A., Font-Ribera A., Nadathur S., Joachimi B., Martini P., 2023, Phys. Rev. Lett., 130, 191003
DESI Collaboration, 2016a, preprint (arXiv:1611.00036)
DESI Collaboration, 2016b, preprint (arXiv:1611.00037)
DESI Collaboration, 2022, AJ, 164, 207
DESI Collaboration, 2023a, preprint (arXiv:2306.06307)
DESI Collaboration, 2023b, preprint (arXiv:2306.06308)
Dawson K. S. et al., 2013, AJ, 145, 10
Dawson K. S. et al., 2016, AJ, 151, 44
Delubac T. et al., 2015, A&A, 574, A59
Dey A. et al., 2019, AJ, 157, 168
de Sainte Agathe V. et al., 2019, A&A, 629, A85
du Mas des Bourboux H. et al., 2017, A&A, 608, A130
du Mas des Bourboux H. et al., 2020, ApJ, 901, 153
Ennesser L., Martini P., Font-Ribera A., Pérez-Ràfols I., 2022, MNRAS, 511, 3514
Farr J., Font-Ribera A., Pontzen A., 2020, J. Cosmol. Astropart. Phys., 2020, 015
Filbert S. et al., 2023, preprint (arXiv:2309.03434)
Font-Ribera A. et al., 2012, J. Cosmol. Astropart. Phys., 2012, 059
Font-Ribera A. et al., 2014, J. Cosmol. Astropart. Phys., 2014, 027
Font-Ribera A., McDonald P., Slosar A., 2018, J. Cosmol. Astropart. Phys., 2018, 003
Göksel Karaçaylı N. et al., 2023, preprint (arXiv:2306.06316)
Gordon C., et al., 2023, J. Cosmol. Astropart. Phys., 2023, 045
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
Guo Z., Martini P., 2019, ApJ, 879, 72
Guy J. et al., 2023, AJ, 165, 144
Hahn C., et al., 2023, AJ, 165, 253
Karaçaylı N. G., Font-Ribera A., Padmanabhan N., 2020, MNRAS, 497, 4742
Karaçaylı N. G. et al., 2022, MNRAS, 509, 2842
Kirkby D. et al., 2013, J. Cosmol. Astropart. Phys., 2013, 024
Lan T.-W. et al., 2023, ApJ, 943, 68
Lee K.-G. et al., 2013, AJ, 145, 69
Levi M. et al., 2013, preprint (arXiv:1308.0847)
Lyke B. W. et al., 2020, ApJS, 250, 8
McDonald P., Miralda-Escudé J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P. et al., 2006, ApJS, 163, 80
Miller T. N. et al., 2023, preprint (arXiv:2306.06310)
Myers A. D. et al., 2023, AJ, 165, 50
Noterdaeme P. et al., 2012, A&A, 547, L1
Pérez-Ràfols I. et al., 2018a, MNRAS, 473, 3019
Pérez-Ràfols I., Miralda-Escudé J., Arinyo-i-Prats A., Font-Ribera A., Mas-Ribas L., 2018b, MNRAS, 480, 4702
Pérez-Ràfols I., Pieri M. M., Blomqvist M., Morrison S., Som D., 2020, MNRAS, 496, 4931
Pérez-Ràfols I., Pieri M. M., Blomqvist M., Morrison S., Som D., Cuceu A., 2022, MNRAS , 524, 1464
Raichoor A. et al., 2020, Res. Notes Am. Astron. Soc., 4, 180
Raichoor A. et al., 2023, AJ, 165, 126
Ravoux C., et al., 2023, MNRAS, 526, 5118
Ruiz-Macias O. et al., 2020, Res. Notes Am. Astron. Soc., 4, 187
Schlafly E. F., et al., 2023, 166, 166, 259
Silber J. H. et al., 2023, AJ, 165, 9
Slosar A. et al., 2011, J. Cosmol. Astropart. Phys., 2011, 001
Slosar A. et al., 2013, J. Cosmol. Astropart. Phys., 2013, 026
Suzuki N., 2006, ApJS, 163, 110
Yèche C. et al., 2020, Res. Notes Am. Astron. Soc., 4, 179
Zhou R. et al., 2020, Res. Notes Am. Astron. Soc., 4, 181
Zhou R. et al., 2023, AJ, 165, 58
Zou H. et al., 2017, PASP, 129, 064101

APPENDIX A: CHANGES WITH RESPECT TO PREVIOUS ANALYSES

In this publication, we have presented some changes compared to the last Lyman-α forest analysis conducted using SDSS data (du Mas des Bourboux et al. 2020). These changes were necessary due to the differences in the way data are structured in DESI as compared to SDSS. One of the most significant is the linear pixelization of the wavelength grid.

There are three relevant changes in the way the analysis is performed. First, the re-calibration has been updated to account for the improvements offered by the new instrument. Secondly, the variance estimation process has been simplified. Thirdly, the re-calibration region has been switched to C III. We have already discussed the third change in Section 3.2, now we will describe the first two changes in detail.

A1 Re-calibration of spectra in SDSS

The improved estimation of pipeline noise ($\sigma_{\rm pip}$) in DESI has eliminated the need for multiple re-calibration steps as it was performed in the SDSS analysis.

For SDSS analyses, re-calibration was performed in two steps, both of which were conducted in the same re-calibration region. The first step was equivalent to the one used in this analysis (Section 3.2). After the first step, a second re-calibration step was added to further correct the reported variance of the pipeline. In this second step, the continuum fitting process was run again to obtain a new estimation of η.

Then, this new value of η is applied in the main Lyman-α fluctuations run to correct for the values of $\sigma_{\rm pip}$:

$$\sigma^2_{\rm pip\ \ (Ly\alpha)}(\lambda) = \eta^{\rm (calib2)}(\lambda)\,\sigma^2_{\rm pip\ \ (calib2)}(\lambda). \tag{A1}$$

The extra re-calibration step in SDSS aimed to make the final estimation of η closer to 1, but it also added an unnecessary layer of complexity to the process with the sole purpose of doing so. However,

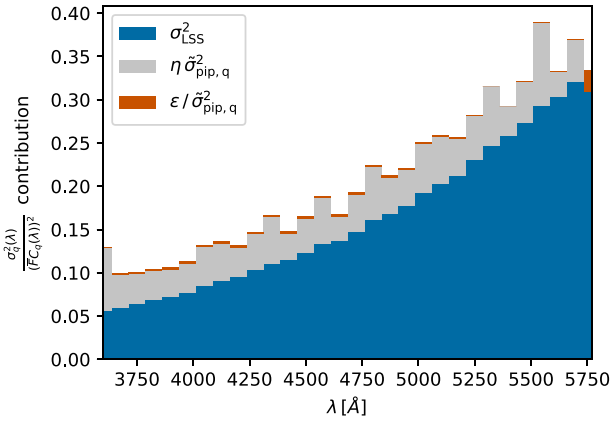

Figure A1. Contribution of each term in equation (A2) to the full variance, for multiple wavelength bins. For this measurement we only used the 5 per cent of quasars with the highest SNR in the EDR + M2 sample. This analysis reveals that the effect of the ϵ term in equation (A2) is minimal, even for the data subset where it is supposed to be most significant.

it did not modify the product $\eta \cdot \sigma^2_{\rm pip,q}$, and as a result, the overall variance assigned to pipeline noise was unchanged.

A2 Variance estimator

In equation (6), we have defined the model for our variance of the flux $\sigma_q^2(\lambda)$. Its expression has been simplified from previous SDSS analysis, where this variance was modeled as:

$$\frac{\sigma_q^2(\lambda)}{\left(\overline{F}C_q(\lambda)\right)^2} = \eta(\lambda)\tilde{\sigma}^2_{\rm pip,q}(\lambda) + \sigma^2_{\rm LSS}(\lambda) + \frac{\epsilon(\lambda)}{\tilde{\sigma}^2_{\rm pip,q}(\lambda)}. \quad \text{(A2)}$$

The additional term was added to account for the observed increase in variance at high SNR, which is likely caused by the diversity of quasar spectra.

In Fig. A1, we present the weight of each term in equation (A2) contributing to the total variance, selecting only the highest SNR quasars. This shows that the effect of this added term is small, and for simplicity we decided to not include it in our analysis. The plot displays only the top 5 per cent higher SNR quasars since this subset is expected to have the largest contribution (the third term in equation (A2) becomes larger). It is worth noting that the contribution is too small to be discernible when all the objects in the sample are considered.

However, it is important to take into account quasar diversity in future analyses. We plan to incorporate a term in equation (6) that considers the variance in the quasar's continuum. This term will be dependent on rest-frame wavelength, $\sigma_C^2(\lambda_{\rm RF})$, and will also help us select the region of the quasar suitable for Lyman-α analyses, showing us how close the emission lines we can get.

Nevertheless, for this DESI EDR analysis, we decided to omit this modification since its inclusion lacked sufficient justification, and its influence on the analysis is negligible.

APPENDIX B: CATALOGUE DESCRIPTION

The catalogue is released in multiple compressed FITS files, one file per healpix pixel. Some additional files, with statistical information of the sample and of each of the continuum fitting steps are also included. Finally, the PICCA configuration file generated at the moment of building the catalogue is also provided, this will allow for the reproduction of the catalogue for any user having access to DESI data.

Each of the delta files with Lyman-α fluctuations are stored under `Deltas/` and have the following internal structure:

```
Filename: Delta/delta-16.fits.gz
No. Name Ver Type Cards Dimensions Format
0 PRIMARY 1 PrimaryHDU 6 ()
1 LAMBDA 1 ImageHDU 13 (2376,) float64
2 METADATA 1 BinTable-
HDU 33 54R x 9C [K, D, D, D, D, K, 12A, 12A,
12A]
3 DELTA 1 ImageHDU 12 (2376, 54) float64
4 WEIGHT 1 ImageHDU 11 (2376, 54) float64
5 CONT 1 ImageHDU 11 (2376, 54) float64
```

The `METADATA` card contains metadata information for each of the forests that belong to the given healpix pixel:

```
XTENSION =  'BINTABLE' / binary table ex-
tension
BITPIX  =  8 / 8-bit bytes
NAXIS  =  2 / 2-dimensional binary table
NAXIS1  =  84 / width of table in bytes
NAXIS2  =  315 / number of forests
PCOUNT  =  0 / size of special data area
GCOUNT  =  1 / one data group (re-
quired keyword)
TFIELDS  =  9 / num-
ber of fields in each row
TTYPE1  =  'LOS_ID ' / picca line-of-
sight id
TFORM1  =  'K ' / data format of field: 8-
byte INTEGER
TTYPE2  =  'RA ' / right ascension
TFORM2  =  'D ' / data format of field: 8-
byte DOUBLE
TUNIT2  =  'rad ' / physical unit of field
TTYPE3  =  'DEC ' / declination
TFORM3  =  'D ' / data format of field: 8-
byte DOUBLE
TUNIT3  =  'rad ' / physical unit of field
TTYPE4  =  'Z ' / redshift
TFORM4  =  'D ' / data format of field: 8-
byte DOUBLE
TTYPE5  =  'MEANSNR ' / mean signal-to-
noise ratio
TFORM5  =  'D ' / data format of field: 8-
byte DOUBLE
TTYPE6  =  'TARGETID' / object identifica-
tion
TFORM6  =  'K ' / data format of field: 8-
byte INTEGER
TTYPE7  =  'NIGHT ' / observation night(s)
TFORM7  =  '12A ' / data for-
mat of field: ASCII Character
TTYPE8  =  'PETAL ' / observation petal(s)
TFORM8  =  '12A ' / data for-
mat of field: ASCII Character
TTYPE9  =  'TILE ' / observation tile(s)
TFORM9  =  '12A ' / data for-
mat of field: ASCII Character
EXTNAME  =  'METADATA' / name of this bi-
nary table extension
BLINDING =  'none ' / blinding scheme used
```

```
   COMMENT Per-forest metadata
   CHECKSUM =  'Q93lR70iQ70iQ70i' / HDU check-
sum updated 2023-05-20T00:19:40
   DATASUM  =  '1813998756' / data unit check-
sum updated 2023-05-20T00:19:40
```

The rest of information inside the file comes in ImageHDU format. In order to read it correctly, one has to use the card `LAMBDA` as observed wavelength. The rest of features: `DELTA`, `WEIGHT`, and `CONT` consist of a 2D array of the values for delta of matter fluctuations, weight assigned to each delta and quasar expected flux. The first index in this 2D array designates the quasar, and the second the wavelength (the ones given in the `LAMBDA` image).

The auxiliary files are stored under the `Log/` directory. The file `delta_attributes.fits.gz` contains statistical properties of the sample and fit values. Its structure is the following:

```
   Filename: Log/delta_attributes.fits.gz
   No. Name Ver Type Cards Dimensions Format
   0 PRIMARY 1 PrimaryHDU 6 ()
   1 STACK_DELTAS 1 BinTable-
HDU 20 2376R x 3C [D, D, D]
   2 VAR_FUNC 1 BinTable-
HDU 25 20R x 6C [D, D, D, D, J, L]
   3 CONT 1 BinTableHDU 20 200R x 3C [D, D, D]
   4 FIT_METADATA 1 BinTable-
HDU 24 60R x 6C [K, D, D, D, K, L]
```

The `STACK_DELTAS` card contains the measurement of $\overline{1+\delta_q(\lambda)}$. `VAR_FUNC` stores the fitted variance function, alongside information from the fits. `CONT` includes the mean expected flux. Finally, the card `FIT_METADATA` holds information of the quasar continuum fit for each of the quasars.

Finally, the file `Log/rejection_log.fits.gz` includes a single card with metadata for all the objects in the quasar catalogue (even the ones rejected) alongside the rejection status.

The full data documentation can be found within the catalogue files.

[1]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra Barcelona, Spain*
[2]*Departament de Física Quàntica i Astrofísica, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain*
[3]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[4]*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[5]*Aix Marseille Univ, CNRS/IN2P3, CPPM, F-13288, Marseille, France*
[6]*Departamento de Física, Universidad de Guanajuato - DCI, C.P. 37150, Leon, Guanajuato, México*
[7]*Department of Astronomy and Astrophysics, UCO/Lick Observatory, University of California, 1156 High Street, Santa Cruz, CA 95065, USA*
[8]*Department of Astronomy and Astrophysics, University of California, 1156 High Street, Santa Cruz, CA 95065, USA*
[9]*Department of Astronomy, Tsinghua University, 30 Shuangqing Road, Haidian District, Beijing, 100190, China*
[10]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[11]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[12]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[13]*The Ohio State University, Columbus, 43210 OH, USA*
[14]*Instituto de Física Teórica (IFT) UAM/CSIC, Universidad Autónoma de Madrid, Cantoblanco, E-28049, Madrid, Spain*
[15]*Consejo Nacional de Ciencia y Tecnología, Av. Insurgentes Sur 1582. Colonia Crédito Constructor, Del. Benito Juárez C.P. 03940, México D.F. México*
[16]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[17]*Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA*
[18]*Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, USA*
[19]*Astrophysics Group, Department of Physics, Imperial College London, Prince Consort Rd, London, SW7 2AZ, UK*
[20]*Department of Physics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK*
[21]*Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA*
[22]*Instituto Avanzado de Cosmología A. C., San Marcos 11 - Atenas 202. Magdalena Contreras, 10720. Ciudad de México, México*
[23]*Aix Marseille Univ, CNRS, CNES, LAM, F-13388, Marseille, France*
[24]*Instituto de Astrofísica de Canarias, C/ Vía Láctea, s/n, E-38205 La Laguna, Tenerife, Spain Universidad de La Laguna, Dept. de Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[25]*Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France*
[26]*Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany*
[27]*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81677 München, Germany*
[28]*Beihang University, Beijing 100191, China*
[29]*Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA*
[30]*Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[31]*Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA*
[32]*Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México CP 04510, México*
[33]*Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA*
[34]*Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia*
[35]*Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia*
[36]*NSF's NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[37]*Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, E-08010 Barcelona, Spain*
[38]*Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA*
[39]*Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK*
[40]*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, PR China*
[41]*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[42] *Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
[43]*Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[44]*Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea*
[45]*CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*
[46]*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218 , USA*
[47]*Department of Physics & Astronomy, Ohio University, Athens, OH 475701, USA*
[48]*University of Michigan , Ann Arbor, MI 48109, USA*

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