Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **526,** 5118–5140 (2023) https://doi.org/10.1093/mnras/stad3008
Advance Access publication 2023 October 5

The Dark Energy Spectroscopic Instrument: one-dimensional power spectrum from first Ly α forest samples with Fast Fourier Transform

Corentin Ravoux [1,2★] Marie Lynn Abdul Karim,[2] Eric Armengaud [2] Michael Walther [3,4] Naim Göksel Karaçaylı [5,6,7,8] Paul Martini,[5,7,8] Julien Guy,[9] Jessica Nicole Aguilar,[9] Steven Ahlen [10] Stephen Bailey [9] Julian Bautista [1] Sergio Felipe Beltran [11] David Brooks,[12] Laura Cabayol-Garcia,[13] Solène Chabanier [9] Edmond Chaussidon [2] Jonás Chaves-Montero [13] Kyle Dawson,[14] Rodrigo de la Cruz [11] Axel de la Macorra [15] Peter Doel,[12] Kevin Fanning [8] Andreu Font-Ribera [13] Jaime Forero-Romero [16,17] Satya Gontcho A Gontcho [9] Alma X. Gonzalez-Morales [11,18] Calum Gordon,[13] Hiram K. Herrera-Alcantar [11] Klaus Honscheid,[5,6,8] Vid Iršič [19] Mustapha Ishak [20] Robert Kehoe,[20] Theodore Kisner [9] Anthony Kremin [9] Martin Landriau [9] Laurent Le Guillou [21] Michael Levi [9] Zarija Lukić,[9] Christophe Magneville,[2] Aaron Meisner [22] Ramon Miquel,[13,23] John Moustakas [24] Eva-Maria Mueller,[25] Andrea Muñoz-Gutiérrez,[15] Lucas Napolitano [26] Jundan Nie [27] Gustavo Niz [11,28] Nathalie Palanque-Delabrouille [2,9] Will Percival [29,30,31] Ignasi Pérez-Ràfols [13] Matthew Pieri,[32] Claire Poppett,[9,33,34] Francisco Prada [35] César Ramírez Pérez,[13] Graziano Rossi,[36] Eusebio Sanchez [37] David Schlegel,[9] Michael Schubnell,[38,39] Hee-Jong Seo [40] Francesco Sinigaglia,[41,42] Ting Tan,[21] Gregory Tarlé [39] Ben Wang [43] Benjamin Weaver,[22] Christophe Yèche [2] and Zhimin Zhou [27]

Affiliations are listed at the end of the paper

Accepted 2023 October 2. Received 2023 September 13; in original form 2023 June 13

ABSTRACT

We present the one-dimensional Ly α forest power spectrum measurement using the first data provided by the Dark Energy Spectroscopic Instrument (DESI). The data sample comprises 26 330 quasar spectra, at redshift $z > 2.1$, contained in the DESI Early Data Release and the first 2 months of the main survey. We employ a Fast Fourier Transform (FFT) estimator and compare the resulting power spectrum to an alternative likelihood-based method in a companion paper. We investigate methodological and instrumental contaminants associated with the new DESI instrument, applying techniques similar to previous Sloan Digital Sky Survey (SDSS) measurements. We use synthetic data based on lognormal approximation to validate and correct our measurement. We compare our resulting power spectrum with previous SDSS and high-resolution measurements. With relatively small number statistics, we successfully perform the FFT measurement, which is already competitive in terms of the scale range. At the end of the DESI survey, we expect a five times larger Ly α forest sample than SDSS, providing an unprecedented precise one-dimensional power spectrum measurement.

Key words: intergalactic medium – large-scale structure of Universe – cosmology: observations.

1 INTRODUCTION

The Ly α (Ly α) forest can be observed from the ground in the optical spectra of distant quasars at redshift between the end phase of reionization ($z \sim 6$) and the peak of galaxy formation $z \sim 2$. The Ly α forest consists of a series of Ly α absorption lines caused by intervening neutral hydrogen located at various redshifts between the quasar and the observer. The Ly α forest is a powerful probe of the underlying matter density field at redshift $z > 2$, together with the astrophysical state of the intergalactic medium (Gunn & Peterson 1965; Lynds 1971; Meiksin 2009; McQuinn 2016).

In particular, the small-scale distribution of neutral hydrogen (∼Mpc) is imprinted in the fluctuations of the Ly α forest along the line of sight that can be accessed by measuring the one-dimensional (1D) Ly α forest power spectrum (denoted $P_{1D,\alpha}$). This measurement is sensitive to the amplitude and slope of the matter power spectrum at redshift $z > 2$. The impact of cosmological parameters on $P_{1D,\alpha}$ can only be accurately predicted using hydrodynamical simulations (Borde et al. 2014; Bolton et al. 2017; Walther et al. 2021; Puchwein

★ E-mail: ravoux@cppm.in2p3.fr

et al. 2023). The realization of those simulations is made arduous by the large dynamic range needed to model the Ly α forest adequately (Lukić et al. 2015; Chabanier et al. 2022). Fitting data measurements with those simulation predictions provides constraints on the cosmological parameters σ_8, n_s, and Ω_m, as well as on parameters describing the thermal properties of the intergalactic medium. In particular, the simulations described in Walther et al. (2021) are able to predict $P_{1D,\alpha}$ with sufficient accuracy (at the 1 per cent level) when compared to expected uncertainties from the Dark Energy Spectroscopic Instrument (DESI) measurement.

Due to its sensitivity to the matter fluctuations at small scales, measurements of $P_{1D,\alpha}$ can constrain physics beyond the Standard Model, such as the mass of neutrinos, the mass of warm dark matter candidates, or a possible running of the spectral index due to primordial inflation physics. First, $P_{1D,\alpha}$ is well suited to constrain the sum of neutrino masses that damps the matter power spectrum at small scales (Lesgourgues & Pastor 2006, 2012). Stringent constraints are obtained by coupling $P_{1D,\alpha}$ with hydrodynamical simulations, and by combining it with the cosmic microwave background (Seljak, Slosar & McDonald 2006; Palanque-Delabrouille et al. 2015; Yèche et al. 2017; Palanque-Delabrouille et al. 2020). Secondly, several studies combined high- and moderate-resolution $P_{1D,\alpha}$ measurements to obtain constraints on the warm dark matter mass (Viel et al. 2005, 2008, 2013; Baur et al. 2016, 2017; Yèche et al. 2017; Palanque-Delabrouille et al. 2020). The hydrodynamical simulations used in those study either directly model neutrinos as particles or using a rescaling of the matter power spectrum to account for neutrinos (Pedersen et al. 2020a, b; Pedersen, Font-Ribera & Gnedin 2023). Finally, other exotic dark matter models such as fuzzy dark matter (Armengaud et al. 2017; Iršič et al. 2017) can also be constrained using $P_{1D,\alpha}$ measurement.

Between the first $P_{1D,\alpha}$ measurements (Croft et al. 1998; McDonald et al. 2000; Croft et al. 2002; Kim et al. 2004) and today, the large increase in observation capabilities brought numerous Ly α forest samples that can be split between moderate-resolution ($\lambda/\Delta\lambda \lesssim 5000$) and high-resolution ($\lambda/\Delta\lambda \gtrsim 20\ 000$) observations.

The computation of $P_{1D,\alpha}$ with high-resolution data sets such as SQUAD (Murphy et al. 2018), KODIAQ (O'Meara et al. 2015; O'Meara 2017), or XQ-100 (Lopez et al. 2016) are performed in Viel et al. (2013), Iršič et al. (2016), Walther et al. (2018), Boera et al. (2019), Day, Tytler & Kambalur (2019), Khaire et al. (2019), Gaikwad et al. (2021), and Karaçaylı et al. (2022). Those measurements use high signal-to-noise quasars to probe the intergalactic medium at very small scales (∼100 kpc) but does not provide sufficient statistics to accurately measure the large-scale clustering (>5 Mpc), needed for cosmological interpretation.

Moderate-resolution surveys provide large numbers of Ly α forests, which yield smaller statistical uncertainties on $P_{1D,\alpha}$ estimates. However, the resolution of such spectrographs limits the reach of very small scales. The $P_{1D,\alpha}$ measurement with moderate-resolution spectrograph was first performed on a small sample of the Sloan Digital Sky Survey (SDSS) data in McDonald et al. (2006). Subsequently, the increase of the Ly α forest statistic has largely improved this measurement with the Baryon Oscillation Spectroscopic Survey (BOSS) in Palanque-Delabrouille et al. (2013) and the extended Baryon Oscillation Spectroscopic Survey (eBOSS) in Chabanier et al. (2019) using 43 751 quasar spectra.

Several methods can be used to measure $P_{1D,\alpha}$ from Ly α forest samples. The most straightforward relies on the Fast Fourier Transform (FFT) and was applied in BOSS and eBOSS (Palanque-Delabrouille et al. 2013; Chabanier et al. 2019) analyses. The 1D power spectrum can also be measured with configuration space estimators such as the quadratic maximum likelihood estimator (QMLE). This method has already been applied to moderate resolution observations (McDonald et al. 2006; Palanque-Delabrouille et al. 2013) and more recently on high-resolution data in Karaçaylı et al. (2022). The QMLE method is applied to the same data than used in the present paper and presented in a companion paper (Karaçaylı et al. 2023a). The FFT method yields a straightforward calculation of $P_{1D,\alpha}$ and offers more control over the different calculation steps. Conversely, the more-complex QMLE estimation is not sensitive to gaps in the quasar spectra. The results between the two methods are presented in the companion paper and are in good agreement. FFT and QMLE results agree at 1 per cent level precision up to half the Nyquist frequency.

The purpose of this work is to compute $P_{1D,\alpha}$ from the first DESI data, following the same methodology as in the latest eBOSS measurement in Chabanier et al. (2019). Using the same method facilitate the comparison between eBOSS and DESI. The $P_{1D,\alpha}$ is sensitive to instrumental properties such as noise and spectral resolution. As the telescopes used and the data are very different, it is essential to characterize the DESI instrument. We improve the algorithms and methodology used in Chabanier et al. (2019) to account for systematic and instrumental differences between eBOSS and DESI. In particular, due to the spectral resolution improvement of DESI, our measurement allows accessing smaller scales than eBOSS.

The outline of this paper is as follows: Section 2 describes the DESI instrument and data processing used to perform this $P_{1D,\alpha}$ measurement. The $P_{1D,\alpha}$ pipeline is presented in Section 3, and the characterization of the DESI instrument in Section 4. We generate synthetic data to validate and correct our measurement in Section 5. The treatment of statistical and systematic uncertainties for the $P_{1D,\alpha}$ measurement is given in Section 6. Finally, we present our measurement on DESI data, as well as a comparison to previous measurements in Section 7, and conclude in Section 8.

2 INSTRUMENT AND DATA DESCRIPTION

DESI has as objective to measure the spectra of 40 million galaxies and quasars in a footprint of 14 000 deg^2 over 5 yr (Levi et al. 2013; DESI 2016a; Abareshi et al. 2022). This project aims to continue the cosmic mapping efforts started by SDSS, while drastically increasing its constraining power on the Lambda cold dark matter model and its possible extensions.

We first focus on the description of the data used for our measurement. In the following, we describe the data starting from the instrument (DESI), its associated spectroscopic pipeline, the different data acquisition phases, and the input catalogues of our study.

2.1 DESI instrument

The DESI instrument is mounted on the Mayall telescope, located on the Kitt Peak National Observatory (KPNO) in the Tohono O'odham Nation. The Mayall telescope is a reflective prime-focus telescope with a 4-m diameter primary mirror. The DESI instrument (DESI 2016b; Miller et al. 2023) receives photons through an optical corrector designed to increase the field of view to 7.5 deg^2 on the focal plane. The focal plane system, composed of 5000 robotically controlled fibres, can quickly modify its configuration to aim at the targeted objects on a specific footprint (Silber et al. 2022). An optical fibre system redirects the light of the observed objects to a separate climate-controlled enclosure containing 10 spectrographs. Each spectrograph comprises three CCD cameras, whose properties are given in the Table 1. In comparison to SDSS spectrographs,

Table 1. Spectral range and effective resolving power ($R = \Delta\lambda/\lambda$) for each channel of the DESI spectrographs (Abareshi et al. 2022).

Channel	Spectral range (Å)	Resolving power
Blue (B)	3600–5930	2000–3200
Red (R)	5600–7720	3200–4100
Near infrared (Z)	7470–9800	4100–5100

the effective resolving power ($\Delta\lambda/\lambda$) improved by at least a factor of 2.

2.2 DESI spectroscopic pipeline

The high complexity of the DESI survey induces the need for advanced software pipelines and products, including the imaging from the DESI Legacy Imaging Surveys (Zou et al. 2017; Dey et al. 2019; Schlegel et al., in preparation), a pipeline to select the targets to observe (Myers et al. 2022), a pipeline to assign fibres (Raichoor et al., in preparation), a pipeline to parse the survey and to optimize the observation strategy (Schlafly et al. 2023), and an extensive spectroscopic reduction pipeline (Guy et al. 2022).

This spectroscopic pipeline, called desispec https://github.com/desihub/desispec, transforms the raw CCD images into spectra, and is detailed in Guy et al. (2022). Before extracting the spectra, the images are subtracted by dark and bias calibration frames to remove expected background sources, and to estimate the associated readout noise (noise estimation details are given in Appendix C). The non-uniform CCD pixel response is corrected using a dedicated flat-field slit on the spectrograph and the CCD over-scan is removed. A dedicated software detects and flags cosmic rays or defective CCD pixels.

The spectral extraction is performed using the 'spectroperfectionism' method (Bolton & Schlegel 2010), an optimal spectroscopic extraction that correctly models complex 2D point spread functions (PDFs). This method provides the encoding for each fibre and each wavelength of the non-Gaussian instrument resolution into a resolution matrix (noted $\mathbf{R}$ in next sections) used to compute $P_{1D,\alpha}$.

All the spectra are defined on the same wavelength grid without additional resampling. Consequently, the extracted spectra are linearly binned in observed wavelength with a constant separation $\Delta\lambda_{pix} = 0.8$ Å. Conversely, the SDSS/BOSS spectrographs were logarithmically binned with a constant spectral pixel size of $\Delta v = 69$ km s^{-1}, which corresponds to $\Delta\lambda/\lambda = \Delta v/c = 2.3 \times 10^{-4}$, equivalent to $\Delta[\log(\lambda)] = 10^{-4}$ (Smee et al. 2013).

Once the spectrum of each fibre has been extracted, several post-processing steps allow removing a variety of further observational effects. The non-uniform response of individual fibres as a function of wavelength is corrected with flat-field frames by observing a white screen attached to the telescope dome and illuminated with a LED array. For all exposures, some fibres are dedicated to observing the sky. The so-called sky spectra associated with these fibres provide the sky level and the intensity of atmospheric emission lines and are subtracted to the spectra associated with targets (sky subtraction). The transmission defaults of the atmosphere and telescope as a whole are corrected by the observation of calibrated star spectra. This step converts CCD units (number of electrons) to observed flux units. Finally, the spectrum of an object is obtained by coadding its different exposures. The resulting spectrum is expressed separately into the three spectrographs bands described in Table 1.

All the software of the pipeline employed for the analysis of DESI data are listed in the repository desihub https://github.com/desihub. In particular, the spectra analysed in this article have been processed with the fuji https://github.com/desihub/desispec/releases/tag/0.51.13. version of the spectroscopic pipeline.

Table 2. Summary of the DESI data sets used in this study, their associated acronyms, a subjective description, and the total number of quasar spectra whose redshifts are between 2.0 and 5.0.

Data set	Acronym	Description	Quasar number
Target selection validation	SV1	Small sample with deeper exposures (up to 16) than the end of DESI	12 355
One-percent survey	SV3	Same density and depth as for the full 5 yr DESI main survey	12 686
First 2 months of the main survey	DESI-M2(M2)	Sparse data with small number of exposures	87 373

2.3 DESI data

The spectroscopic pipeline described previously was used to analyse data obtained over different periods. In this article, we use three data sets from the first observations of DESI described in Table 2, which includes a total of 112 414 quasar spectra whose redshifts are between 2.0 and 5.0.

The first two data sets 'Target Selection Validation' (SV1) (DESI 2023a) and 'One-Percent Survey' (SV3) are part of the Early Data Release (EDR) whose complete description is given in DESI (2023b). 'Target Selection Validation' was conducted from 2020 December to 2021 March and includes a large number of exposures (up to 16) for the same targets. The objective of this survey was among others, to study extensively the survey performance as a function of instrumental depth and to build visual inspection truth tables. The SV1 data set also includes 'Secondary Tiles' as detailed in DESI (2023b). Following the completion of 'Target Selection Validation' observations, the 'One-Percent Survey' phase was dedicated more specifically to the evaluation of the survey design. The number of exposures is similar to the main survey at its end (between 4 and 5 for each Ly α quasars) and the goal was to determine the best strategy to cover the sky while limiting fibre loss.

The main DESI survey started in June 2021 and in this article we also use the first 2 months of data, named DESI-M2 (and noted M2 in this paper for conciseness), which is not present in the EDR but will be included in the Data Release 1. In the M2 data set, most quasars have only one exposure. While all three data sets are studied in this article, in the end we removed SV1 due to its different noise properties (see Appendix A). The final measurement is computed on SV3 + M2 data set only.

2.4 Input catalogues

The input catalogues used to compute $P_{1D,\alpha}$ were obtained by applying specific procedures to the three data sets previously described.

The targeting of the quasars used in our study (Yèche et al. 2020; Chaussidon et al. 2022) was verified with visual inspection of subsets of early observations (Alexander et al. 2022). For quasars, the DESI pipeline categorizes the observed spectra and estimates their redshifts using the redrock spectral template fitting software https://github.com/desihub/redrock (Bailey et al., in preparation; Brodzeller et al. 2023). In order to optimize, the completeness of the quasar catalogue while keeping a high purity, additionally a broad Mg II line finder mgii_afterburner

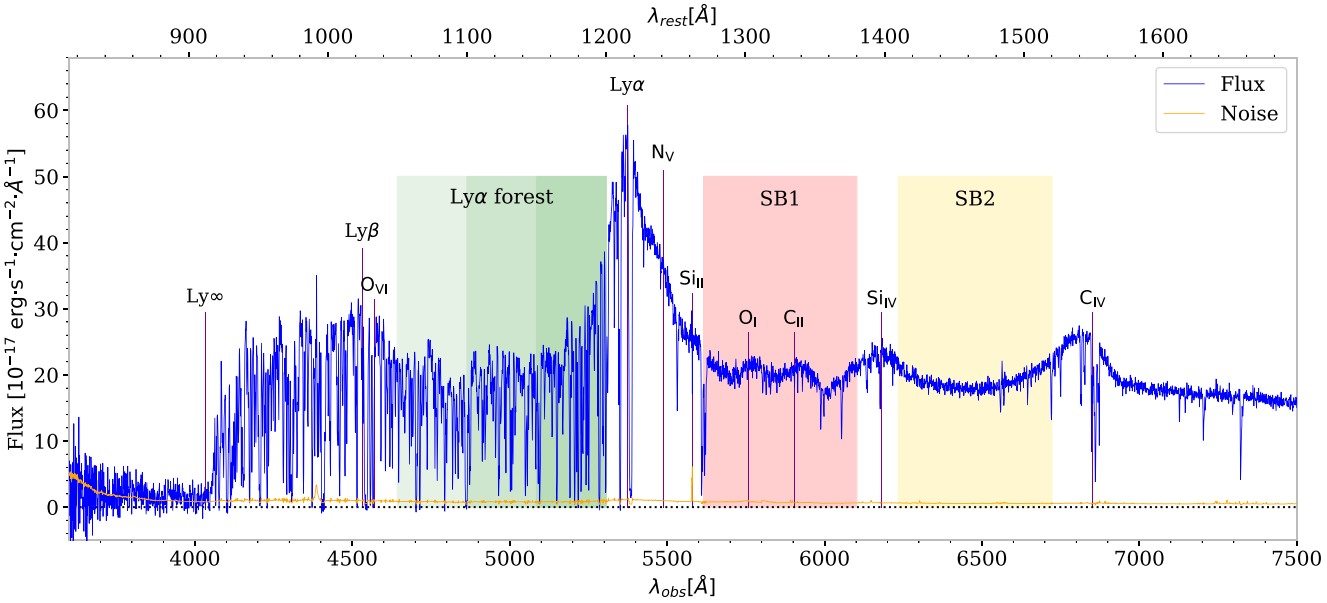

Figure 1. A particularly high-signal spectrum of a quasar located at a redshift $z = 3.42$ measured by DESI with an exposure time of 2300 s. This quasar was observed on 2021 April 12, in the SV3 programme, on DESI tile 221 (TARGETID $= 39627746095137037$, RA $= 217.263°$, Dec. $= -1.755°$). The quasar flux is represented in blue and its noise in orange. The Ly α forest is shown in green. The side-band regions 1 and 2 pictured in red and yellow are used to estimate the forest contamination by metals.

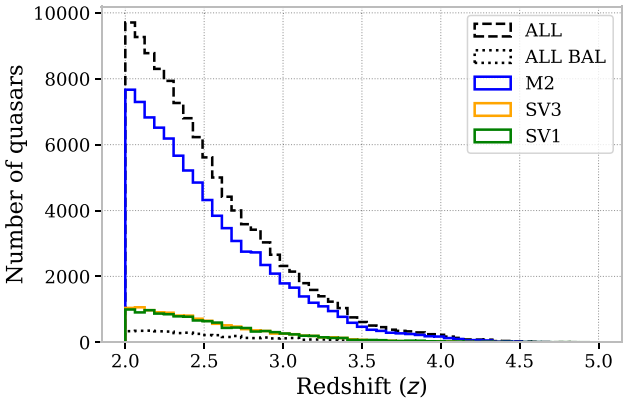

Figure 2. Histogram of the quasar redshift (with $z > 2.0$) whose spectra are observed in the SV1, SV3, and M2 data sets. The histogram of the sum of data sets is shown in dashed line. Finally, the number of BAL quasars characterized by a balnicity index higher than zero, and not used in the $P_{1D,\alpha}$ computation pipeline, is shown with a dotted line.

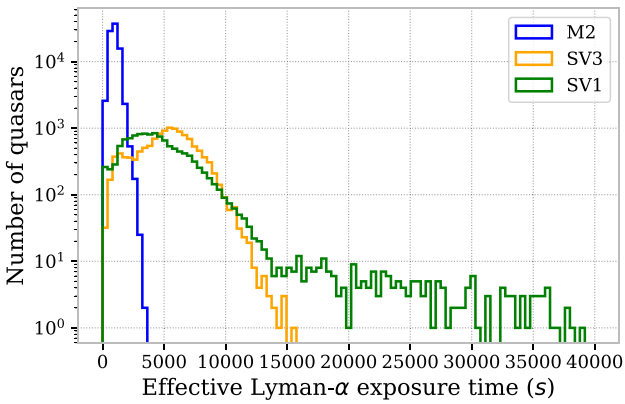

Figure 3. Normalized histogram of the effective exposure time in the Ly α forest region for the quasar spectra in the SV1, SV3, and M2 data sets. As mentioned in Section 2.3, there is a wide disparity of exposure time for the three data sets. As a reference, the nominal time of one DESI exposure is set to 1200 s.

https://github.com/desihub/desispec/blob/main/py/desispec/mgii_afterburner.py and a machine learning classifier applying deep convolutional neural networks QuasarNP https://github.com/desihub/QuasarNP (Busca & Balland 2018; Farr, Font-Ribera & Pontzen 2020) are run after redrock. Both post-processing programs are run on all objects targeted by DESI as detailed in Chaussidon et al. (2022).

An example of a quasar spectrum at redshift $z = 3.42$ is given in Fig. 1. This spectrum, with a particularly high flux, is part of the SV3 data set that contains spectra with exposure time equivalent to the end of the DESI survey. The redshifts and effective exposure times for each data set considered in this paper are shown in Figs 2 and 3, respectively. The effective exposure time accounts for nightly observing conditions by normalizing the real exposure time to a reference with airmass 1, zero galactic extinction, a 1.1 arcsec seeing (FWHM), and zenith dark sky (Guy et al. 2022). The nominal exposure time of one exposure is defined to 1000 s. The large differences in term of exposure time emphasize the need to treat the data sets differently, at least for noise properties. SV1 and SV3 contains a small number of forests but with heterogeneous exposure times, in opposition to M2, which contains many quasars only observed once.

Broad absorption line (BAL) quasars are specific quasars whose spectra exhibit consistent blueshifted absorptions associated with many spectral features. They are identified using the baltools https://github.com/paulmartini/baltools software. It consists of a χ^2 minimizer algorithm that looks for blueshifted C IV or Si IV absorptions in an unabsorbed quasar model. The fit is performed for rest-frame wavelengths between 1260 and 2400 Å. A quasar is considered a BAL type if its spectrum exhibits a region between C IV

and Si IV emission lines, with at least 10 per cent flux decrement below the continuum and a width greater than 2000 km s^{-1}. BAL quasars spectra (4.18 per cent in the total data set) are removed in the analysis performed in this work.

Damped Ly α absorbers (DLAs) are regions within a quasar spectrum that show oversaturated absorption with prominent Lorentzian wings as the quasar flux intersects the dense, circumgalactic medium of an intervening (proto-)galaxy. They are a subclass of high-column density systems and are a significant contaminant of the Ly α forest signal, particularly because of their wings and the additional contamination by circumgalactic metal absorption lines (McDonald et al. 2005). The correct modeling of such systems in simulations has been proven to be particularly complicated (Pontzen et al. 2008).

We use a catalogue resulting from the combination of a convolutional neural network (CNN) algorithm desi-dlas https://github.com/cosmodesi/desi-dlas (Parks et al. 2017; Zou et al., in preparation) and a Gaussian process (GP) algorithm https://github.com/jibanCat/gp_dla_detection_dr16q_public (Ho, Bird & Garnett 2021). desi-dlas is trained with SDSS spectra to identify candidate high-column density objects for rest-frame wavelengths between 900 and 1346 Å. It returns locations of high-column density systems in the spectra, as well as their H I column density and a confidence parameter. The GP finder provides similar output using the same training set and a Bayesian model selection. We only consider the high-column density objects with column density $N_{\rm HI} > 10^{20.3}$ cm^{-2} (DLAs). In accordance with recommendations from Parks et al. (2017) and Ho et al. (2021), we consider CNN confidence level higher than 0.2 as valid DLA detections when the ratio between the quasar continuum and the noise is higher than 3. We take a confidence level limit of 0.3 when this ratio is lower than 3. For the GP model, a 0.9 minimal confidence level is applied. In the case when absorbers are detected by the two models, the combined DLA catalogue uses $N_{\rm HI}$ values and DLA redshifts from GP model.

Although DLAs by themselves constitute tracers of the matter distribution, they have an extended impact on the observed spectra. They increase the correlations of neighbour spectrum pixels, thus artificially increasing $P_{\rm 1D,\alpha}$ level. Therefore, we choose to mask the core of DLA regions of the spectra by fixing the transmitted flux fraction to its mean value for spectrum pixels, where the DLA-induced absorption is larger than 20 per cent. In addition, the absorption in the Lorentzian damping wings that remain after the cut is corrected with a Voigt profile following Bautista et al. (2017) and Chabanier et al. (2021).

Finally, we use a catalogue of masks to account for atmospheric and Galactic emission lines, which has been adapted to DESI resolution https://github.com/corentinravoux/p1desi/blob/main/etc/skylines/list_mask_p1d_DESI_EDR.txt. The creation of this catalogue is detailed in Section 4.4.

3 1D POWER-SPECTRUM ESTIMATION

The $P_{\rm 1D,\alpha}$ estimator is build using the data product described previously in two phases. First, the fitting of the continuum of quasars is used to convert the absolute received flux to a normalized quantity δ_F. Secondly, $P_{\rm 1D,\alpha}$ is computed by employing an FFT and by averaging the product of this transformation for all the selected Ly α forests.

3.1 Continuum fitting

A standard normalized quantity used in the calculation of correlations and power spectra is the flux contrast δ_F of the Ly α forest, defined as

$$\delta_F(\lambda) = \frac{F(\lambda)}{\overline{F}(\lambda)} - 1 = \frac{f(\lambda)}{C_{\rm q}(\lambda, z_{\rm q})\overline{F}(\lambda)} - 1\,, \tag{1}$$

where F is the transmitted flux fraction, and $\overline{F}(\lambda)$ is its average value, the mean transmission of the intervening IGM. Note that for the purpose of this work, we do not need to know the individual quasar continua C_q, but only the product $C_{\rm q}(\lambda, z_{\rm q})\overline{F}(\lambda)$. Thus, similar to previous Ly α studies based on survey data (Chabanier et al. 2019; du Mas des Bourboux et al. 2020; Ramírez-Pérez et al. 2023), we directly measure this product in our continuum fitting process using the picca https://github.com/igmhub/picca. (du Mas des Bourboux et al. 2021) software package. This software also merges the quasar spectra over different bands to obtain f over the all wavelength range. The continuum of each quasar is modelled as the product of a universal continuum C common to all quasars, and a first-order polynomial term in wavelength:

$$C_{\rm q}(\lambda, z_{\rm q}) = \left(a_{\rm q} + b_{\rm q}\lambda\right) C\left(\lambda_{\rm rf} = \frac{\lambda}{(1+z_{\rm q})}\right)\,, \tag{2}$$

where $a_{\rm q}$ and $b_{\rm q}$ are quasar-dependent constants. In previous studies (Palanque-Delabrouille et al. 2013; Chabanier et al. 2019), $b_{\rm q} = 0$ was assumed for all quasars, i.e. only a wavelength-independent normalization factor was taken into account. We add an additional linear wavelength-dependent term to account for the diversity of quasars after verifying that this change does not impact the mean level of our $P_{\rm 1D,\alpha}$ measurement.

The $a_{\rm q}$ and $b_{\rm q}$ parameters for each quasar are determined along with C by maximizing the following log-likelihood:

$$\ln\mathcal{L} = -\frac{1}{2}\sum_i \frac{\left[f_i - \overline{F}(\lambda_i)C_{\rm q}\left(\lambda_i, z_{\rm q}, a_{\rm q}, b_{\rm q}\right)\right]^2}{\sigma_{\rm q}^2(\lambda_i)} - \ln\left[\sigma_{\rm q}^2(\lambda_i)\right]\,, \tag{3}$$

where the sum is run over all the spectrum pixels of the quasar q, and $\sigma_{\rm q}$ is the standard deviation estimator of the flux f.

In contrast to analyses of the large-scale 3D correlation function such as Bourboux et al. (2020), we want all spectrum pixels to contribute equally to the continuum fitting and the $P_{\rm 1D,\alpha}$ computation, as the opposite could bias $P_{\rm 1D,\alpha}$. Therefore, we impose noise-independent weights in the continuum fitting procedure:

$$\sigma_{\rm q}^2(\lambda) = \left(\overline{F}(\lambda)C_{\rm q}(\lambda)\right)^2 \tag{4}$$

This procedure is the same as in the previous $P_{\rm 1D,\alpha}$ analyses based on BOSS/eBOSS data (Palanque-Delabrouille et al. 2013; Chabanier et al. 2019). The standard deviation associated with δ_F at the end of the continuum fitting procedure is defined by

$$\sigma_{\delta_F}(\lambda) = \frac{\sigma_{\rm pip,q}(\lambda)}{\overline{F}(\lambda)C_{\rm q}(\lambda)}\,, \tag{5}$$

where $\sigma_{\rm pip,q}$ is the noise provided by the DESI spectroscopic pipeline detailed in Appendix C.

The universal continuum C and the $a_{\rm q}$ and $b_{\rm q}$ parameters are computed iteratively. In particular, C is estimated from the average of all spectra, i.e. in a non-parametric way. During the entire fitting procedure, spectrum pixels that are masked due to the presence of a DLA or an atmospheric line are not considered in the fit. We use 7 iterations and have verified that the continuum fits are converged at this point.

The noise level of a Ly α forest is characterized by defining the average signal-to-noise ratio $\overline{\mathrm{SNR}}$ in the Ly α forest region:

$$\overline{\mathrm{SNR}} = \left\langle \frac{f(\lambda)}{\sigma_{\mathrm{pip,q}}(\lambda)} \right\rangle_{\lambda} . \tag{6}$$

Only Ly α forests with a $\overline{\mathrm{SNR}}$ larger than 1 are used in the continuum fitting procedure. This procedure is also restricted to the observed redshift range 3600 Å $< \lambda <$ 7600 Å, to avoid the shorter wavelength range where a large fraction of the quasar spectra is absorbed by the atmosphere. We also select the rest-frame wavelength in the range 1050 Å $< \lambda_{\mathrm{rf}} <$ 1180 Å, so that the measured contrasts are dominated by the Ly α forest. In particular, we try to avoid the Ly β singlet and the OVIdoublet emission regions respectively located at $\lambda_{\mathrm{Ly}\,\beta} = 1025.72$ Å and $\lambda_{\mathrm{OVI}} = (1031.912, 1037.613)$ Å in the rest frame.

The cut $\lambda_{\mathrm{RF}} < 1180$ Å facilitates the continuum fitting procedure and mitigates most of the proximity effect: close to a quasar, the neutral hydrogen fraction is indeed influenced by the quasar's UV radiation in addition to the extragalactic UV radiation background (Bajtlik, Duncan & Ostriker 1988).

As detailed in Section 2.2, the quasar spectra are linearly binned in observed wavelength with $\Delta\lambda_{\mathrm{pix}} = 0.8$ Å. Note that when converting to rest-frame wavelength λ_{rf} the pixel size will be redshift dependent. For the continuum fitting process, we thus need to rebin our spectra to a uniform grid in λ_{rf}. As the quasar continuum is relatively smooth and to avoid noisy continuum fits for analyses of relatively small data sets, we chose a grid for the common continuum C that is 10 times coarser than the lowest redshift quasar pixels considered, i.e.:

$$\Delta\lambda_{\mathrm{pix,rf}} = 10\frac{\Delta\lambda_{\mathrm{pix}}}{1+z_{\mathrm{min}}} . \tag{7}$$

By taking $z_{\mathrm{min}} = 2.0$ as the lowest redshift, we obtain $\Delta\lambda_{\mathrm{pix,\,rf}} =$ 2.67 Å. With increasing size of the data set, such a rebinning could be relaxed in future DESI measurements. At the end of the continuum fitting procedure, the stacking of all the Lyαcontrasts is forced to be equal to zero to avoid introducing flux calibration errors.

3.2 FFT power spectrum estimator

Conceptually, the δ_F quantity can be separated into different contributions in the Ly α forest region:

$$\delta_F(\lambda) = \delta_{\mathrm{astro}}(\lambda) + \delta_{\mathrm{noise}}(\lambda) , \tag{8}$$

where δ_{astro} corresponds to the fluctuations caused by all the elements of the intergalactic medium (including Ly α), and δ_{noise} corresponds to noise fluctuations.

Considering instrumental effects, the true underlying flux contrast is modified on its way through the instrument in multiple ways. First, photons traverse the spectrograph leading to the output flux being convolved with the spectrograph line spread function $W(\lambda, \mathbf{R}, \Delta\lambda_{\mathrm{pix}})$ which depends on the resolution matrix presented in Section 2.2. The W term also account for the signal pixelization as the photons are counted into CCD pixels of size $\Delta\lambda_{\mathrm{pix}}$. Finally, we need to account for noise sourcing from the processes of photon counting and readout. In total, we can write the measured flux contrast δ_F as

$$\delta_F(\lambda) = \delta_{\mathrm{astro}}(\lambda) \circledast W(\lambda, \mathbf{R}, \Delta\lambda_{\mathrm{pix}}) + \delta_{\mathrm{noise}}(\lambda) . \tag{9}$$

We assume that the impact of the noise and the resolution term W are decorrelated, i.e. that the noise contrast is not affected by the instrumental effects of pixelization and resolution. Verifying this assumption is beyond the scope of this paper.

The δ_{astro} contrast can be decomposed between the Ly α signal and the one from all other elements of the intergalactic medium called metals. The contribution of those metal absorptions to δ_F can be decomposed into two parts. On the one hand, there are absorption lines with rest-frame wavelength $\lambda \gg \lambda_{\mathrm{Ly}\,\alpha} = 1215.67$ Å. Those lines can be independently observed redwards of the Ly α forest, using specific rest-frame spectral regions called side-bands SB1 (1270 $< \lambda_{\mathrm{rf}} <$ 1380 Å), and SB2 (1410 $< \lambda_{\mathrm{rf}} <$ 1520 Å), as shown in Fig. 1. We group the absorption of all those metals in a contrast noted δ_{metals}.

On the other hand, there are absorption lines with rest-frame wavelength $\lambda \lesssim \lambda_{\mathrm{Ly}\,\alpha}$, such as Si II and Si III elements ($\lambda_{\mathrm{SiII}} = 1, 190$ and 1193 Å, and $\lambda_{\mathrm{SiIII}} = 1, 206.50$ Å). They cannot be observed independently of the Ly α forest, but will show absorption that is correlated with the Ly α absorption and will lead to an oscillatory feature in the estimated $P_{\mathrm{1D},\alpha}$. For those lines, we adopt the same approach developed in McDonald et al. (2006) and subsequently used in other analyses. We leave the features inside our power spectrum estimates to be corrected by fitting an additional oscillation during parameter inference. We note $\delta_{\mathrm{Ly}\,\alpha}$ the contrast containing the Ly α forest and the effect of those latter metals. Finally, the flux contrast can be expressed:

$$\delta_F(\lambda) = \left(\delta_{\mathrm{Ly}\alpha}(\lambda) + \delta_{\mathrm{metals}}(\lambda)\right) \circledast W(\lambda, \mathbf{R}, \Delta\lambda_{\mathrm{pix}}) + \delta_{\mathrm{noise}}(\lambda) . \tag{10}$$

The 1D Ly α power spectrum ($P_{\mathrm{1D},\alpha}$) can be estimated from this decomposition by applying an FFT algorithm on δ_F of each Ly α forest. This method was applied in previous measurements (Croft et al. 1998; Palanque-Delabrouille et al. 2013; Chabanier et al. 2019).

The FFT estimator is implemented in picca (Bourboux et al. 2021). For each Ly α forest, the raw power spectrum[1] is defined from the Fourier transform of $\delta_F(r)$:

$$(2\pi)\delta_{\mathrm{D}}(k-k')P_{\mathrm{raw}}(k) = \delta_F(k)\delta_F(k') , \tag{11}$$

where δ_{D} is the 1D Dirac distribution.

Applying a Fourier transform on equation (10), the raw power spectrum is expressed by

$$P_{\mathrm{raw}}(k) = \left(P_{\mathrm{Ly}\,\alpha}(k) + P_{\mathrm{Ly}\,\alpha-\mathrm{Si\,II/Si\,III}}(k) + P_{\mathrm{metals}}(k)\right) \cdot W^2(k, \mathbf{R}, \Delta\lambda_{\mathrm{pix}}) + P_{\mathrm{noise}}(k) . \tag{12}$$

In this decomposition, $P_{\mathrm{Ly}\,\alpha}$, P_{metals}, and P_{noise} are the power spectra, respectively, associated with the contrasts $\delta_{\mathrm{Ly}\,\alpha}$, δ_{metals}, and δ_{noise} with the same definition as in equation (11). We assumed here that the SiIIand SiIIIpower spectra are negligible and that all cross-correlation terms between the contrasts are null. The only non-neglected cross-term is $P_{\mathrm{Ly}\,\alpha-\mathrm{Si\,II/Si\,III}}(k) = 2\left|\delta_{\mathrm{Ly}\alpha}(k)\delta_{\mathrm{Si\,II/Si\,III}}(k)\right|$, which corresponds to the correlated absorptions of Ly α with either Si II or Si III. The oscillations induced by this term have a wavenumber $2\pi/(\lambda_{\mathrm{Ly}\,\alpha} - \lambda_{\mathrm{Si\,II/Si\,III}})$ when 'k' is expressed in Å^{-1}.

The FFT estimator for the 1D power spectrum is computed as an average over all available Ly α forests in the measurement sample. It is designed to match the sum of Ly α and Ly $\alpha-$Si II/Si III power spectra in equation (12), so that we define

$$P_{\mathrm{1D},\alpha}(k) = \left\langle P_{\mathrm{Ly}\,\alpha}(k) + P_{\mathrm{Ly}\,\alpha-\mathrm{Si\,II/Si\,III}}(k) \right\rangle , \tag{13}$$

where $\langle . \rangle$ denotes the average over all the Ly α forests used for $P_{\mathrm{1D},\alpha}$ calculation. From equation (12), the estimator of $P_{\mathrm{1D},\alpha}$ is defined by

$$P_{\mathrm{1D},\alpha}(k) = \left\langle \frac{P_{\mathrm{raw}}(k) - P_{\mathrm{noise}}(k)}{W^2(k, \mathbf{R}, \Delta\lambda_{\mathrm{pix}})} \right\rangle - P_{\mathrm{metals}}(k) . \tag{14}$$

[1]This is not a proper power spectrum, strictly speaking, as it does not correspond to an average.

The $P_{1D,\alpha}$ measurement is split in different redshift bins to take into account its evolution. The Ly α forest is split into subforests that correspond to consecutive and non-overlapping subregions of equal length. This procedure also reduces the correlations between the different redshift bins. We chose to cut Ly α forests into three subforests whose rest-frame wavelength boundaries are $\lambda_{rf} = 1093.3$ and 1136.6 Å, so that the length of each subforest is $L_{sub} = 43.3$ Å. With this subforest separation, a single Ly α forest can contribute to up to three different redshift bins in the $P_{1D,\alpha}$ measurement. The subforest splitting constrains the minimal accessible observed wavenumber to $k_{min} = 2\pi/(L_{sub}(1 + z_{min})) = 0.0453$ Å^{-1}, by taking the minimal redshift used. Each subforest spans at most $\Delta z = 0.2$ and we choose the same Δz to define the redshift binning for $P_{1D,\alpha}$.

For observed wavelength $\lambda \lesssim 3,700$ Å, the noise level is high in comparison to the spectra because of atmospheric absorptions. To minimize the impact of this noise, we remove the spectrum pixels for which the observed wavelength is lower than 3750 Å, which corresponds to Ly α absorbers located at $z = 2.085$. In the future, with a dedicated study to control the noise at shorter wavelength, the $P_{1D,\alpha}$ analysis can be extended to redshift $z \sim 2$.

In accordance with the eBOSS study (Chabanier et al. 2019), we remove subforests shorter than 75 spectrum pixels due to a cut in the UV region or to the presence of a large DLA. We also do not consider the Ly α subforests with more than 120 masked spectrum pixels.

Unlike the analysis in Chabanier et al. (2019), we do not apply a second redshift-dependent $\overline{\text{SNR}}$ cut for the averaging of $P_{1D,\alpha}$. Instead, we develop and test a $\overline{\text{SNR}}$ weighting scheme, as detailed in the appendix B. This procedure is used for all the article except in Section 4.2.1 where the impact of the $\overline{\text{SNR}}$cut is investigated.

4 DESI INSTRUMENTAL CHARACTERIZATION

As the DESI instrument is new, we first focus the analysis on characterizing the instrumental effects on our $P_{1D,\alpha}$ measurement. In particular, we describe the impact of spectral resolution, instrumental noise, metal power spectrum, and atmospheric emission lines in the following.

4.1 Spectrograph resolution

4.1.1 Resolution correction modelling

As mentioned in Section 2.2, DESI spectra are linearly binned in observed wavelength, so that the natural unit for wavenumbers is Å^{-1}. The maximal measurable wavenumber follows the Nyquist-Shannon limit: $k_{max} = k_{Nyq} = \pi/\Delta\lambda_{pix} \simeq 3.92$ Å^{-1}.

For SDSS $P_{1D,\alpha}$ measurements (Croft et al. 1998; McDonald et al. 2006; Palanque-Delabrouille et al. 2013; Chabanier et al. 2019), the pixelization is logarithmically binned in observed wavelength, making it suitable to express the power spectrum in Hubble velocity unit v, because $\Delta v \propto \Delta\lambda/\lambda = \Delta\log(\lambda)$. In eBOSS (Chabanier et al. 2019), the resolution correction is modelled by a Gaussian function $W(k, \Delta\lambda) = \exp(-0.5(k[\text{s km}^{-1}]\Delta v)^2)$, where Δv is the spectral resolution in velocity units.

The spectroscopic resolution of DESI is improved with respect to the SDSS spectrographs. On the DESI blue band (see Table 1) where most of the Ly α forest are observed, the effective resolving power $R = \lambda/\Delta\lambda$ ranges from 2000 to 3200 (Abareshi et al. 2022). In comparison, SDSS spectrographs had a $1500 < R < 2300$ in its blue band ($3600 < \lambda < 6350$ Å). This improved resolution brings the opportunity to probe the clustering of matter at smaller scale by measuring the small fluctuations in the Ly α forest, as illustrated in Fig. 4.

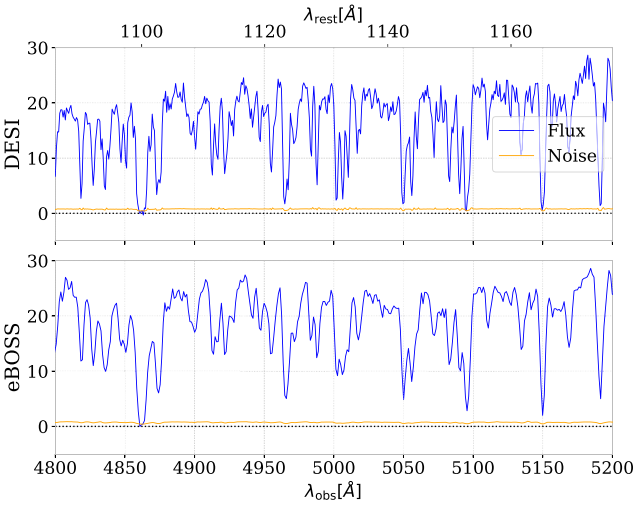

Figure 4. Illustration of the resolution improvement between eBOSS and DESI. The DESI spectrum of the quasar represented in Fig. 1 is zoomed on a region of the Ly α forest on the top panel. The spectrum of the same quasar obtained by the SDSS-IV (Gunn et al. 2006; Smee et al. 2013; Dawson et al. 2016; Blanton et al. 2017) in the 16th Data Release (DR16; Ahumada et al. 2020) of the eBOSS. The SDSS name associated with this quasar is 142 903.03–014 519.3. The DESI spectrum is obtained after three individual exposures for a total exposure time of 2300 s. The eBOSS spectrum have 11 exposures for a total of 6300 s. The large-scale absorption structures are similar but due to its improved spectroscopic resolution, the DESI spectrum clearly shows more details at small scales.

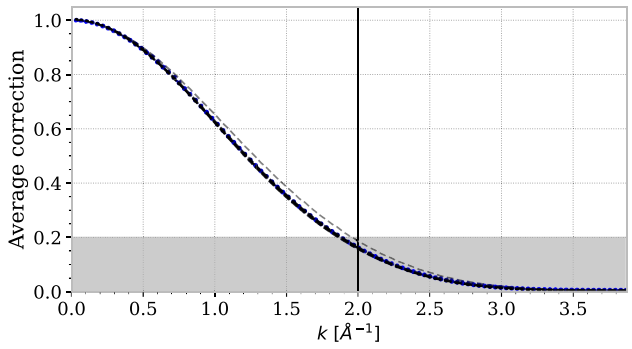

Figure 5. Average resolution correction $\langle W^2(k, \mathbf{R}, \Delta\lambda_{pix})\rangle = \langle \mathbf{R}^2(k) \cdot \rangle$ (blue points) using Ly α forest from SV3 + M2 data set. This resolution correction is weakly dependent on the redshift range (shown by shaded dashed black lines along the blue points). Only one shaded black line is above the blue points ($z = 3.8$). All the others are at the same level as the blue points. The mean value over all redshift bins is shown with the points. The shaded area represents the regime for which the associated impact of resolution and pixelization removes more than 80 per cent of the power spectrum. This criteria is chosen to define the maximal wavenumber of our $P_{1D,\alpha}$ measurement, shown with a vertical black line.

As described in Section 2.2, the DESI spectrograph resolution is entirely characterized by the resolution matrix $\mathbf{R}$ (Guy et al. 2022). In opposition to SDSS, the resolution matrix also accounts for the pixelization of the signal. Consequently, we choose to express the resolution correction function W in equation (12) directly as the Fourier transform of the resolution matrix that we note $\mathbf{R}(k)$.

Fig. 5 shows the average correction due to resolution and pixelization for the SV3 + M2 data set. This correction indicates that

resolution and pixelization suppress more than 95 per cent of the signal at $k > k_{\rm res,95} = 2.73$ Å^{-1}, and 98 per cent at $k > k_{\rm res,98} = 3.15$ Å^{-1}. This observation drives the maximal wavenumber that of the $P_{\rm 1D,\alpha}$ measurement. We choose the conservative value $k_{\rm max} = 2$ Å^{-1}, for which the average resolution corrections is equal to 80 per cent. We will extend this conservative limit in future studies after a full characterization of resolution on CCD pixel-level simulations.

4.1.2 *Validation with CCD image simulations*

We use CCD image simulations of the DESI instrument to verify our resolution modelling described previously. This method is also used for the QMLE estimation of $P_{\rm 1D,\alpha}$ in a companion paper (Karaçaylı et al. 2023a). Those simulations are built using the desisim https://github.com/desihub/desisim package developed to model and validate the spectroscopic extraction pipeline presented in Section 2.2 and detailed in Guy et al. (2022). We use the desisim package to produce realistic realizations of two-dimensional spectroscopic images. Those images simulate various instrumental effects such as the different sources of noise detailed in Appendix C, gain and bias of the CCD amplifiers, the throughput of each spectrograph, PSF of each fibre, and sky emission.

Using the desisim package, we transform Ly α transmissions that follows a given input $P_{\rm 1D,\alpha}$ into realistic Ly α forests with apparent magnitudes representative of DESI quasar targets and noise that is representative of a single, 1000 s exposure in nominal conditions. We simulate 45 000 Ly α forest spectra located over ten DESI tiles and process this simulated data set with the spectroscopic pipeline. Since we only want to see the impact of spectral resolution, the true imposed noise level is used to reduce the data.

Two sets of Ly α contrasts δ_F are generated from this simulated quasar sample. The first set, called *RAW*, is produced directly from Ly α transmissions. This type of realization does not contain the effect of spectral resolution. In parallel, we create Ly α contrasts from the full CCD image simulations, noted *CCD*, but using the true imposed continuum for each quasar to only see the impact of resolution. Finally, we run the $P_{\rm 1D,\alpha}$ FFT pipeline presented in Section 3.2 with our resolution modelling on both Ly α contrast sets.

The ratio between $P_{\rm 1D,\alpha}$ obtained from *RAW* and *CCD* sets is shown in Fig. 6. As expected, the main difference between those measurements resides in the smallest scales ($k > 1.0$ Å^{-1}). This ratio is not redshift dependent, as indicated by the light black curves in the background. We only consider the average overall redshifts. We checked that applying an additional pixelization correction $\mathrm{sinc}^2(0.5k\Delta\lambda_{\rm pix})$ similarly to eBOSS increases a lot the discrepancy between *RAW* and *CCD* power spectrum; thus, confirming that the resolution matrix accounts the pixelization, at least partly. We derive a correction by fitting a second-order polynomial function to the following averaged ratio:

$$A_{\rm res}(z,k) = \left\langle \frac{P_{\rm 1D,\alpha,RAW}(k,z)}{P_{\rm 1D,\alpha,CCD}(k,z)} \right\rangle_z . \quad (15)$$

This term is directly multiplied to our $P_{\rm 1D,\alpha}$ measurement to account for the miss-estimation of the resolution correction.

4.2 Noise power spectrum measurement

4.2.1 *Comparison of noise estimators with high-wavenumber data*

The $P_{\rm 1D,\alpha}$ measurement is significantly impacted by the noise power spectrum at small scales. Thus, it is necessary to obtain an accurate estimate of this component to correct for it.

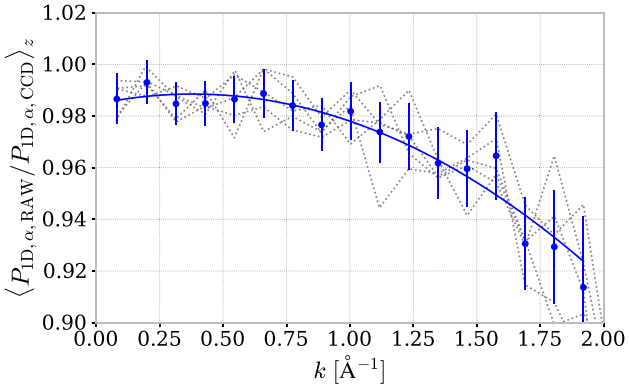

Figure 6. Ratios between the power spectrum obtained directly from Ly α transmissions (*RAW*) and the one derived from the CCD image simulations (*CCD*), for different redshift bins (light black) and averaged over all redshifts bins (blue points with error bars). Each power spectrum is re-binned by a factor of 3 to reduce error bars. A second-order polynomial is fitted to the ratio (blue continuous line). It is used to correct the miss estimation of resolution in the $P_{\rm 1D,\alpha}$ measurement.

The noise power spectrum is estimated either directly from the pipeline noise, or by using an exposure difference method. A detailed description of the obtained noise power spectrum estimators, respectively, noted $P_{\rm pipeline}$ and $P_{\rm diff}$, can be found in Appendix C.

Additionally, the noise power-spectrum level is determined by taking advantage of the combined effect of the resolution and the pixelization shown in Fig. 5, as well as the Ly α thermal broadening. Those effects erase essentially all 'signal' power and thus at large wavenumbers, equation (12) simplifies into $P_{\rm raw}(k) \simeq P_{\rm noise}(k)$. The difference (or ratio) of $P_{\rm raw}$ and $P_{\rm noise}$ on the largest $\boldsymbol{k}$-bins accessible can be used to validate the noise estimator and correct it empirically. We define the following asymptotic difference and ratio by averaging those quantities at large wavenumbers. We decide to use the criteria $k > k_{\rm res,98}$ where $k_{\rm res,98}$ is defined as the wavenumber for which the resolution and pixelization suppress more than 98 per cent of the signal:

$$\alpha = \langle P_{\rm raw}(k) - P_{\rm noise}(k)\rangle_{k>k_{\rm res,98}} ,$$
$$\beta = \left\langle \frac{P_{\rm noise}(k)}{P_{\rm raw}(k)} \right\rangle_{k>k_{\rm res,98}} . \quad (16)$$

Fig. 7 shows the measurement of α and β on SV3 data set with $\overline{\rm SNR} > 3$ (with $\overline{\rm SNR}$ defined by equation 6), using the pipeline noise to compute $P_{\rm noise}$. Considering the observed statistical fluctuations, we notice that the asymptotic behaviour of power spectra at high wavenumber enables a good measurement of α (respectively β), whose value is close to 0 (respectively 1). Additionally, we verified that the variation of α and β as a function of redshift is small in comparison to the statistical fluctuations of the ratio and difference.

4.2.2 *Characterization on DESI data sets*

We compute the α and β coefficients for the pipeline ($P_{\rm pipeline}$) and difference ($P_{\rm diff}$) noise estimators on SV1, SV3, and M2 data sets, while varying the applied $\overline{\rm SNR}$ cut.

Fig. 8 shows the measured α values on the Ly α forest regions for the different data sets and noise estimators. As previously stated, α should be equal to zero when the noise is perfectly estimated. The values of α are small compared to the absolute level of the noise power spectrum shown in Fig. 7. The SV1 data set exhibits a $\overline{\rm SNR}$ cut

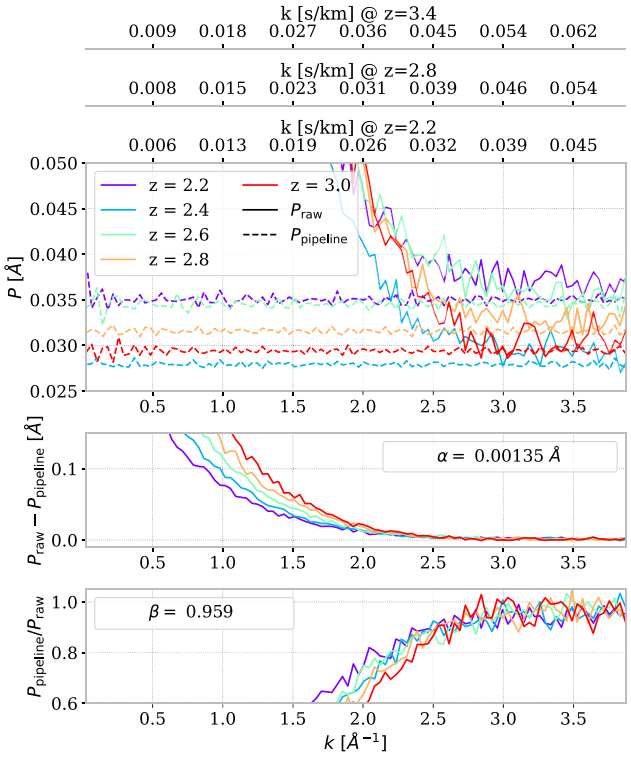

Figure 7. Measurement of the asymptotic difference (α) and ratio (β) between noise ($P_{\rm noise}$) and raw ($P_{\rm raw}$) power spectra. Those terms are defined by equation (16) and corresponds to the average of the difference and ratio of power spectra for large wavenumbers ($k > k_{\rm res,98}$). Here, α and β are measured for the pipeline noise ($P_{\rm pipeline}$), using the SV3 observations with a minimal $\overline{\rm SNR}$ cut of 3.

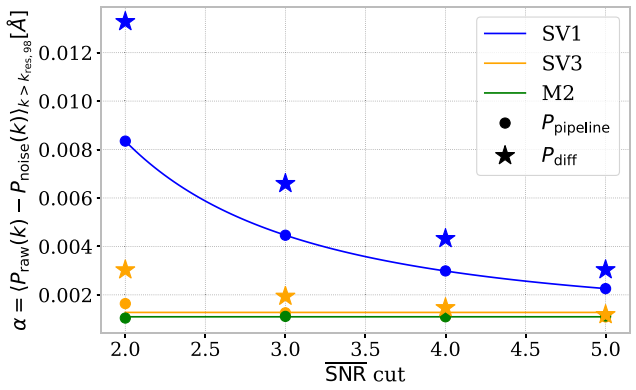

Figure 8. Asymptotic differences α between the noise and raw power spectra for SV1 (blue), SV3 (yellow), and M2 (green) data sets, as a function of the minimal $\overline{\rm SNR}$ cut. This difference is measured for both $P_{\rm noise}$ estimators from the pipeline ($P_{\rm pipeline}$,points) and from exposure differences ($P_{\rm diff}$,stars). The continuous lines are fits of the α values for pipeline noise, whose parameters are given in Table 3.

dependence which is not present for M2 and SV3. We choose to make a data set-dependent correction to remove this residual noise. The miss-estimated noise is higher for data sets with a larger number of exposures (such as SV1). We think this originates from unaccounted common sources of noise coming from the statistical uncertainties in the CCD calibration data (dark current, pixel flat field), which explain why this effect increases with the number of exposures. The

Table 3. Additive corrections applied to the pipeline noise for different spectral regions and data sets. An $\overline{\rm SNR}$ dependence is included in the case of SV1 only. The same parameters are used for both SB1 and SB2.

Band	Data	$P_{\rm noise,miss} = \alpha$ (Å)
Ly α	SV1	$0.026 \times \left(\overline{\rm SNR}\right)^{-1.77} + 0.00076$
	SV3	0.00127
	M2	0.00109
Side-bands	SV1	$0.018 \times \left(\overline{\rm SNR}\right)^{-1.52} + 0.000032$
	SV3	0.00048
	M2	0.00019

dependence in $\overline{\rm SNR}$ can be explained by the fact that this effect is amplified when we consider noisier spectra.

In Fig. 8, the exposure difference noise estimator $P_{\rm diff}$ is shown for SV1 and SV3. In M2, quasars are observed with one, or a small number of exposures; thus, $P_{\rm diff}$ cannot be reliably computed. For the SV1 and SV3 observations, the difference noise power spectrum exhibits the same trend as the pipeline noise estimate. However, $P_{\rm diff}$ consistently underestimates the noise level compared to $P_{\rm pipeline}$. We think that this underestimation is due to the fact that $P_{\rm diff}$ is not accounting for all the common sources of noise between exposures. As a consequence, we only consider the $P_{\rm pipeline}$ estimation from now on.

In Appendix C3, we perform additional studies to characterize the additivity of the miss-estimated noise and its behaviour for different spectra regions. Those regions, called side-bands, are used in the next section for metal power spectrum estimation. Table 3 summarizes the corrections to the pipeline noise we computed for the different data sets. For SV3 and M2, we choose to apply a constant additive correction (α). For SV1, given the observed dependence of α as a function of the minimal $\overline{\rm SNR}$ cut, we fit it with a power law.

4.3 Side-band power spectrum

Following previous $P_{\rm 1D,\alpha}$ studies (McDonald et al. 2006; Palanque-Delabrouille et al. 2013; Chabanier et al. 2019), special spectrum regions, called side-bands and that are devoid of Ly α absorption, are used to statistically estimate the power spectrum components $P_{\rm metals}$ caused by metal absorptions in the Ly α forest. The resulting signal from the side-bands contains information about the abundance, temperature and clustering of metals in the intergalactic medium. In our study, we aim at creating a model to closely reproduce the side-band power spectrum ($P_{\rm SB}$) so that we can statistically subtract it in the measurement of $P_{\rm 1D,\alpha}$ in equation (14).

We define the side-bands SB1 (1270 Å $< \lambda_{\rm rf} <$ 1380 Å) and SB2 (1410 Å $< \lambda_{\rm rf} <$ 1520 Å). In both side-bands, the fraction of transmitted flux contrast can be expressed similarly to equation (10):

$$\delta_F(\lambda)|_{\rm SB} = \delta_{\rm metals}(\lambda)|_{\rm SB} \circledast W(\lambda, \mathbf{R}, \Delta\lambda_{\rm pix}) + \delta_{\rm noise}(\lambda)|_{\rm SB} \,. \quad (17)$$

The $\delta_{\rm metals}(\lambda)|_{\rm SB}$ contrast contains all the fluctuations caused by metals with rest-frame absorption wavelength higher than 1380 Å for SB1 and 1520 Å for SB2. Similarly to the calculation of $P_{\rm 1D,\alpha}$, the side-band power spectrum writes:

$$P_{\rm SB}(k) = \left\langle \frac{P_{\rm raw}(k)|_{\rm SB} - P_{\rm noise}(k)|_{\rm SB}}{W^2(k, \mathbf{R}, \Delta\lambda_{\rm pix})} \right\rangle . \quad (18)$$

The main difference between both side-bands is that SB1 contains Si IV absorption, which is not present in SB2. Consequently, we use the side-band power spectrum of SB1 to estimate $P_{\rm metals}$ in

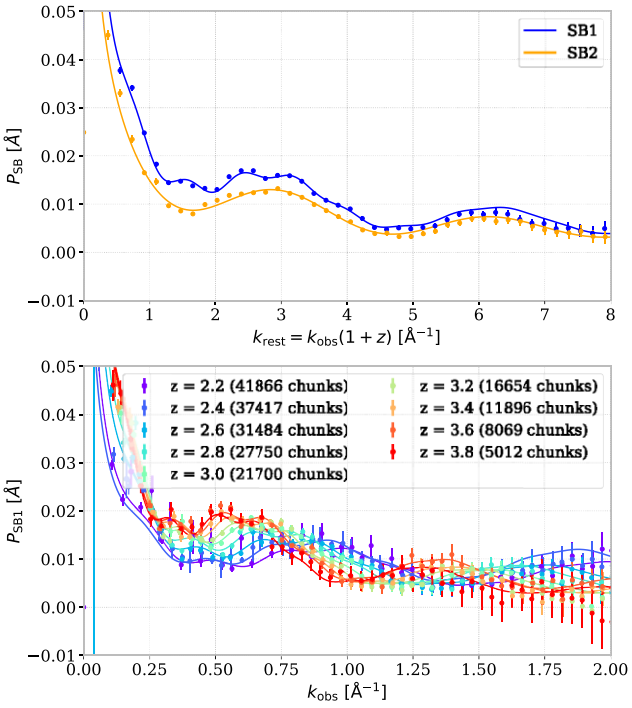

Figure 9. 1D power spectra measured in the side-band regions SB1 and SB2, using the SV3 + M2 data set after applying the data set-dependent noise correction. Top: Average of P_{SB} over all redshift bins in the rest wavenumber frame $k_{rest,\,i} = k_{obs,\,i} \times (1 + z)$. The fitted model represented in continuous line is given by equation (19). Bottom: P_{SB} on side-band SB1 as a function of redshift and observed wavenumber. Each redshift bin is fitted using the product between equation (19), with parameters fixed from previous fit, and a first-order polynomial function.

equation (14), and the SB2 power spectrum as a consistency check. The side-band power spectrum is computed at the same observed wavelength range as $P_{1D,\alpha}$. Because of the higher rest-frame absorption wavelength, the quasars employed to calculate the side-band power spectrum are at a lower redshift than the sample used for Ly α. In particular, quasars at $z < 2.0$ are employed to calculate P_{SB} in the lowest redshift bins of $P_{1D,\alpha}$. However, the metals in the intergalactic medium that produces the absorptions responsible for the side-band power spectrum (P_{SB}) are at the same redshift as those which produced the metal power spectrum (P_{metals}) in the $P_{1D,\alpha}$ calculation for Ly α. Consequently, the redshift dependence of metal absorptions is correctly taken into account.

We note that the method we use to remove metal contribution is not perfect. In particular, the blending of metals with Ly α emission pointed out in Day et al. (2019) is not fully accounted here. This second-order effect should be included in future studies for which the precision level will significantly improve.

The measurement of the side-band power spectrum using the SV3 + M2 data set is shown in Fig. 9. The top panel shows the stack of P_{SB} for all redshift bins as a function of wavenumber in rest-frame $k_{rest} = (1 + z) \times k_{obs}$. The bottom panel shows separated redshift bins as a function of observed wavenumber. A total number of 201 849 and 276 279 sub-forests were used for SB1 and SB2 respective measurements.

In top panel of Fig. 9, the average side-band power is lower for the SB2 than SB1, as expected by the addition of Si IV absorptions.

In the eBOSS measurement (Chabanier et al. 2019), a sixth-degree polynomial is used to fit the shape of the side-band power. For our measurement, we exploit the stacked P_{SB} profile to design a more physically motivated model.

A complete list of metals present in the Ly α forest and that impact $P_{1D,\alpha}$ is given in Pieri et al. (2014), Yang et al. (2022), and the strongest absorptions are from Si III, Si II, Si IV, and C IV.

The emission peaks, and consequently absorption peaks, of Si IV and C IV are actually two doublets. Their rest-frame wavelength given by NIST (Kramida et al. 2021) are $\lambda_{SiIV^a} = 1,393.76$ Å, $\lambda_{SiIV^b} = 1,402.77$ Å, $\lambda_{CIV^a} = 1,548.202$ Å, and $\lambda_{CIV^b} = 1,550.774$ Å. The presence of an absorption doublet in the side-band creates a peak in the 1D correlation function, which translates into an oscillatory pattern in the power spectrum, whose periodicity depends on the doublet separation. This effect is studied more in detail on the same data set in Karaçaylı et al. (2023b) to determine cosmic ion abundance. In the top panel of Fig. 9, both side-band power spectra display a large oscillation caused by the C IV doublet. As expected, the SB1 power spectrum shows an additional oscillation induced by Si IV doublet absorptions. These considerations lead us to model P_{SB} as the sum of a power law including all-metal contributions and oscillations due to Si IV and C IV doublets:

$$P_{SB,m} = A \times k^{-\epsilon} + \sum_i P_{doublet,i}(k, A_i, a_i, k_i, \psi_i)\,. \tag{19}$$

Oscillations induced by a doublet have a frequency characterized by the rest-frame wavenumber $k_{rest,\,i} = 2\pi/\delta\lambda_i$, where $\delta\lambda_i$ is the doublet separation in Å. We choose to use damped sinusoidal functions to model the doublet oscillations as follows:

$$P_{doublet,i}(k, A_i, a_i, k_i, \psi_i) = A_i e^{-a_i k} \sin\left(2\pi\left(\frac{k}{k_i}\right) + \psi_i\right)\,, \tag{20}$$

where k_i is a free parameter with a uniform prior centred around $k_{rest,\,i}$.

The result of the fit on the redshift-averaged P_{SB}, taking into account the oscillations of C IV and Si IV for SB1, and only C IV for SB2, is shown in the top panel of Fig. 9.

The SB1 fitted function is used to derive the redshift dependence of the side-band power spectrum, shown in the bottom panel of Fig. 9. For each redshift bin, we fit a product between the global SB1 fitted function (expressed in the observational wavelength frame), and a first-order polynomial. As P_{SB} may also include other uncorrelated contaminations besides metals, we do not seek to interpret the fitted values for each power spectrum. In particular, we note that the k_i are systematically shifted in comparison to their doublet oscillation frequency $k_{rest,i}$. For SB1, the fitted values are $k_{CIV} = 3.32$ Å^{-1} and $k_{SiIV} = 0.812$ Å^{-1}, whereas rest-frame wavenumber are $k_{rest,CIV} = 2.44$ Å^{-1} and $k_{rest,SiIV} = 0.697$ Å^{-1}. For SB2, we obtain $k_{CIV} = 3.23$ Å^{-1}. We think that this effect might be due to the blended impact of all metals present in the intergalactic medium. Thus, it is necessary to vary k_i parameters to closely fit our data.

This P_{SB} measurement already represents a clear improvement with respect to that of BOSS (Palanque-Delabrouille et al. 2013), and eBOSS (Chabanier et al. 2019): by eye, Si IV- and C IV-induced oscillatory patterns are seen even for individual redshift bins, and P_{SB} is essentially a decreasing function of wavenumber, even at high wavenumber. This indicates an improvement in the noise modelling.

4.4 Atmospheric and galactic emission lines

Atmospheric emission lines are corrected from DESI spectra by the spectral extraction pipeline as described in Section 2.2. The average of 15 000 sky spectra on exposures with optimal observing conditions, noted $\langle f_{sky} \rangle$, is shown in Fig. 10.

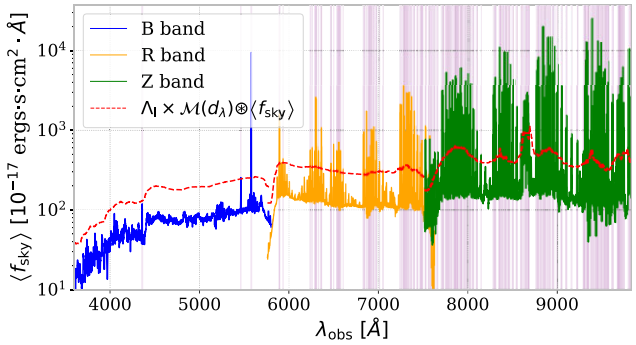

Figure 10. Average of the sky spectra of 15 000 sky fibres with optimal observation conditions (speed > 105, effective exposure time > 1, 100 s, seeing <1.05 deg, and airmass <1.3). These sky spectra originate from three exposures in the SV1 and SV3 data sets. The different spectral bands of DESI are represented (B in blue, R in orange, and Z in green). The median smoothing of this average sky spectrum multiplied by a threshold $\Lambda_l = 2.5$, shown as a dashed red line, is used to select atmospheric emission lines we want to mask (light purple lines). The line located at 4360 Å was added manually, considering its large impact on noise seen in Fig. 11.

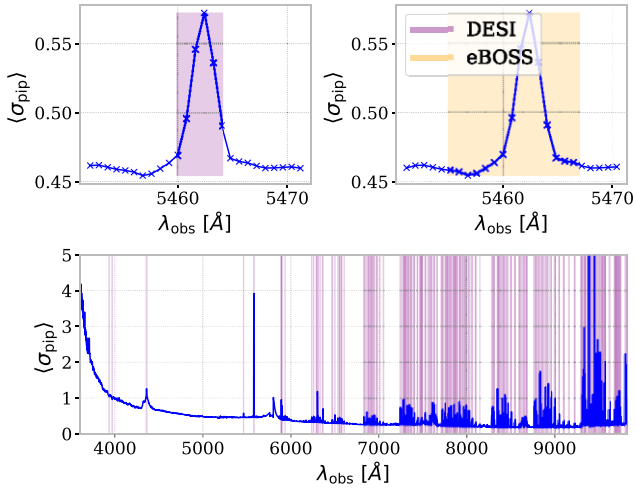

Figure 11. Comparison between the atmospheric emission line mask catalogues used for eBOSS and the new DESI catalogue we designed in our study. Bottom: Average of the DESI pipeline noise for 125 477 objects categorized by redrock as quasars or luminous red galaxies on SV3 observations (blue curve). The DESI ($\Lambda_l = 2.5$, $\Lambda_w = 1.2$) line catalogue is shown on top with purple vertical lines. Top: Zoom on a specific atmospheric emission line on the DESI stacked noise (blue curve). The mask used on this specific line for DESI is shown in the left panel in purple and for eBOSS in orange on the right panel. The DESI mask decreases the masked length in accordance with stacked DESI noise.

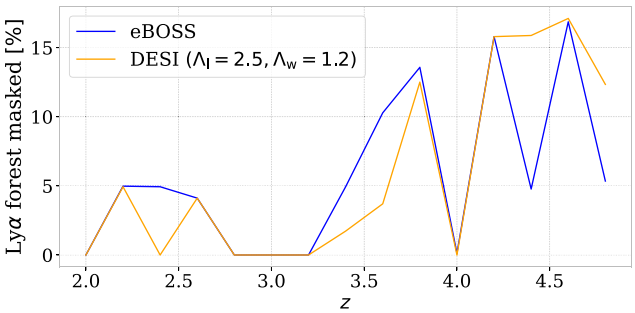

Figure 12. Percentage of spectral length masked by atmospheric lines as a function of redshift, for eBOSS and DESI ($\Lambda_l = 2.5$, $\Lambda_w = 1.2$) catalogues.

The noise of spectrum pixels associated with intense atmospheric lines is strongly increased. It induces additional oscillations in the Ly α contrasts and increases the level of $P_{1D,\alpha}$. We need to correct this effect as those atmospheric lines are not linked to IGM physics. We choose to mask the major atmospheric lines as in previous measurements (Palanque-Delabrouille et al. 2013; Chabanier et al. 2019). The catalogue of lines in these studies https://github.com/igmhub/picca/blob/master/etc/list_veto_line_Pk1D.txt was adapted to the spectral resolution of the SDSS instrument. The improved resolution of DESI makes it possible to reduce the masking size for narrow atmospheric lines, decreasing the impact of masking on the $P_{1D,\alpha}$ measurement

We develop an algorithm similar to Lee et al. (2013) to compute an atmospheric line catalog adapted to the DESI instrument. A median smoothing $\mathcal{M}(d_\lambda)$ of spectral width $d_\lambda = 160$ Å is applied on the average sky spectrum. Atmospheric lines are selected when the average sky spectrum is larger than the product of the smoothed sky flux, $\mathcal{M}(d_\lambda) \circledast \langle f_{sky} \rangle$, by a threshold Λ_l. In Fig. 10, the dashed red line represents this product for $\Lambda_l = 2.5$ and $d_\lambda = 160$ Å.

A second threshold $\Lambda_w = 1.2$ defines the width of atmospheric lines. The upper and lower wavelength limits of an atmospheric line are defined as the first wavelengths on each side whose average sky spectrum is lower than $\Lambda_w \times \mathcal{M}(d_\lambda) \circledast \langle f_{sky} \rangle$. To remain conservative and prevent numerical effects potentially caused by masking at a spectrum pixel position, the line widths are increased by 1 Å on each side.

In this atmospheric line catalogue, we also add the galactic absorption lines, which correspond to relatively broad absorptions made by dust in the Milky Way. We take the same lines as eBOSS: Ca II H and K lines at 3968 and 3933 Å, and the Na D doublet at 5893 Å. The DESI atmospheric emission line catalogue built from this procedure is available online.

We verify that the produced atmospheric line catalogue correctly masks the DESI noise. For this purpose, we compute an average noise by stacking the pipeline noise σ_{pip} of 125 477 objects categorized by redrock as quasars or luminous red galaxies on SV3 observations. This average noise with the DESI atmospheric line catalogue such that ($\Lambda_l = 2.5$, $\Lambda_w = 1.2$) is shown in the bottom panel of Fig. 11. A zoom on an atmospheric line is shown in the top panels of Fig. 11 with eBOSS and DESI masks. The eBOSS mask is too wide for DESI stacked noise, which highlights the benefit of creating a new catalogue. After a visual inspection of most atmospheric lines, we validate that all the spectrum pixels showing an increase in noise are masked by setting the width threshold to $\Lambda_w = 1.2$.

Comparing the average noise (Fig. 11) to the average sky flux (Fig. 10), there is a consistency between atmospheric emission lines and observed peaks in the pipeline noise. The feature at 4360 Å is an exception, as it appears wide and relatively high in the DESI noise and not in the average sky flux. Its wavelength is inside a known transmission dip around 4400 Å due to an issue with DESI's spectrograph collimator coating (Guy et al. 2022). For this specific line, we take the same value as the eBOSS catalogue and force the algorithm to consider it as an atmospheric line even if it does not pass the Λ_l requirement.

A comparison of the percentage of Ly α forest masked for eBOSS and DESI catalogue, as a function of redshift, is given in Fig. 12. To remain conservative, we chose the value of $\Lambda_l = 2.5$ to obtain a catalogue of atmospheric emission lines with a spectral length

masked similar to the eBOSS catalogue. A complete study on synthetic data, out of this paper's scope, will be done to decrease the length masked, and consequently the impact on $P_{1D,\alpha}$.

5 SYNTHETIC DATA CORRECTIONS

Synthetic data (otherwise called mocks) are generated to characterize the impact on $P_{1D,\alpha}$ of continuum fitting, spectral resolution, noise modelling, and spectrum pixel masking. From this, we derive empirical corrections of these effects and apply them to the data measurement.

5.1 Synthetic data sets

We generated a set of DESI-Lite (Karaçaylı, Font-Ribera & Padmanabhan 2020) mocks, specifically designed for $P_{1D,\alpha}$. The full description of these mocks is given in Karaçaylı et al. (2023a). The DESI-LITE software produces uncorrelated Ly α forests that mimic the redshift and noise distribution of the SV3 + M2 data set. Ten independent realizations are generated with different initial conditions. For each realization, a random catalogue of DLAs is created to follow the redshift and column density distribution of the latest eBOSS catalogue (Chabanier et al. 2021).

The quickquasars software https://github.com/desihub/desisim/blob/main/py/desisim/scripts/quickquasars.py (Herrera-Alcantar et al., in preparation), included in the desisim package, transforms Ly α transmission into spectra with observational and astrophysical contaminants. For each realization, two sets of spectra are generated by imprinting DLAs according to the catalog aforementioned or not.

We run five different $P_{1D,\alpha}$ variations to study the impact of different Ly α contaminants:

(i) *TRUECONT*: The true quasar continuum imposed by quickquasars is applied instead of the continuum fitting procedure described in Section 3.1. In comparison to *RAW* mocks, this realization is impacted by finite noise and resolution.
(ii) *CONT*: The Ly α contrasts are calculated using the pipeline detailed in Section 3.1. This type of mocks includes the impact of continuum fitting.
(iii) *DLAm*: Realization for which the DLAs are not added to forests at the quickquasars stage, though we mask spectrum pixels as if they were present. The objective of this kind of mocks is to characterize the impact of DLA masking.
(iv) *LINEm*: Similarly to *DLAm* but masking the atmospheric emission lines catalogue built in Section 4.4.
(v) *DLA*: Realization for which the DLA are applied to the spectra without masking them. The objective of this mocks is to measure the impact of DLA to compute a DLA completeness systematic error in Section 6.

For all the mocks, we take the same procedure as for the FFT calculation on observational data. In particular, the same $\overline{\text{SNR}}$-weighting is applied, and the number of sub-forest for each realization is around 81 500 with a small statistical variation between realizations. This is slightly larger than the number of sub-forest of the data sample given in Section 7.

In the next sections, all the results are shown for the combination of ten independent realizations. To decrease the error bars, we also performed a linear rebin that provides a wavenumber binning three times coarser than what we used for observational data. The error bars of the presented ratios are computed in quadrature.

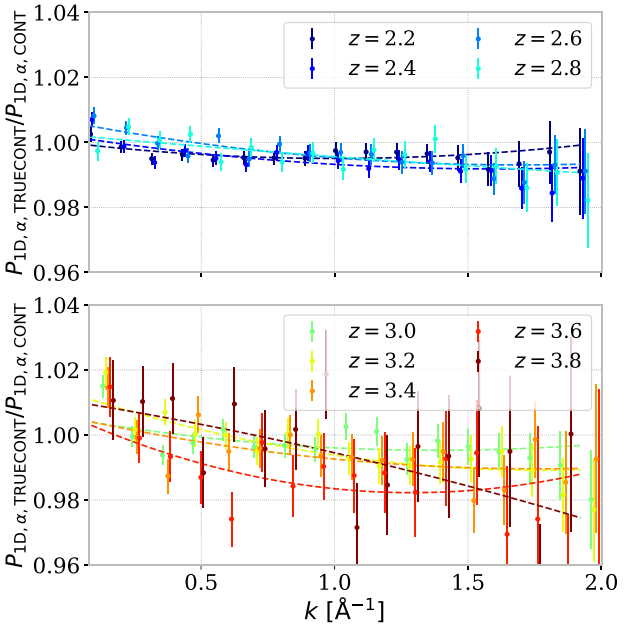

Figure 13. Ratios between the power spectrum obtained using true continuum (*TRUECONT*) and the one derived with our pipeline (*CONT*) on the combination of 10 mocks. Each power spectrum is re-binned by a factor of 3 to reduce error bars. Fitting functions are represented by continuous lines, and used to correct the $P_{1D,\alpha}$ measurement. For clarity, we artificially offset the points corresponding to different redshifts.

5.2 Continuum-fitting correction

The continuum-fitting procedure defined in Section 3.1 systematically distort the measured $C_q(\lambda, z_q)\overline{F}(\lambda)$ term by suppressing large-scale modes, and may bias the $P_{1D,\alpha}$ measurement. This is a well-known effect in BAO measurements (Bourboux et al. 2020). To create a correction which contains this effect, we compare the mock computed using the true continuum (*TRUECONT*) with the one which follows the standard continuum fitting procedure (*CONT*):

$$A_{\text{cont}}(z,k) = \frac{P_{1D,\alpha,\text{TRUECONT}}(k,z)}{P_{1D,\alpha,\text{CONT}}(k,z)}\,. \tag{21}$$

This correction is shown for the combination of ten DESI-Lite mocks in Fig. 13. We use a second-order polynomial function to fit this correction and apply it to the $P_{1D,\alpha}$ measurement.

This correction differs in amplitude compared to the eBOSS measurement (Chabanier et al. 2019). As for eBOSS, the 1D power spectrum with continuum fitting is higher than that measured with the true continuum. However, in our case, the impact is much smaller than eBOSS, for which this ratio was near 4 per cent (without using a first-order polynomial function in the continuum fitting). Furthermore, we do not have a large-scale impact as significant as eBOSS.

5.3 Spectrum pixel masking

For both DLA and atmospheric line masking, we remove some data points from the measured spectra. This does not impact studies performed on real-space spectra, such as the continuum fitting or the QMLE for $P_{1D,\alpha}$ (Karaçaylı et al. 2022). On the other hand, the FFT calculation requires that spectrum pixels are equally spaced. Consequently, when computing the Fourier transform, we impose a value of $\delta_F = 0$ (equivalent to mean transmitted flux fraction value for

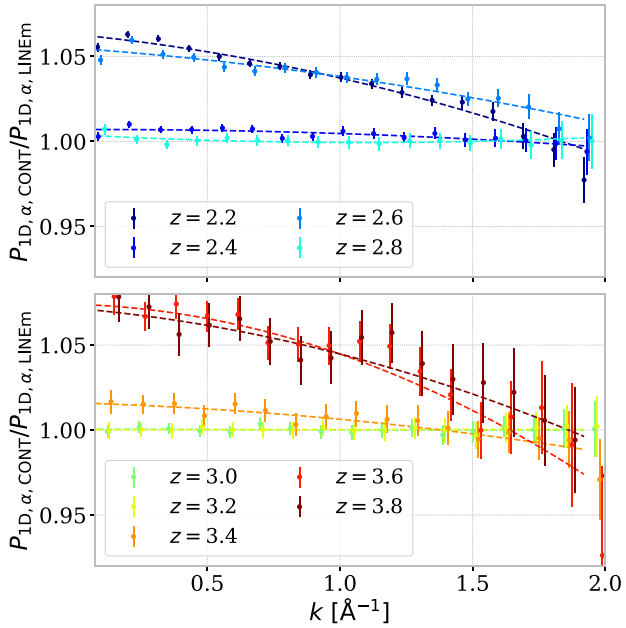

Figure 14. Ratio between the unmasked (*CONT*) and masked (*LINEm*) power spectra (equation 22) for atmospheric line masking on the combination of the 10 mocks. The DESI ($\Lambda_1 = 2.5$, $\Lambda_w = 1.2$) atmospheric line catalogue is used. Each power spectrum is re-binned by a factor of 3 to reduce error bars. Second-order polynomial functions are employed to fit the corrections in each redshift bin. For clarity, we artificially offset the points corresponding to different redshifts.

F) and infinite standard deviation to the masked spectrum pixels. This masking introduces a *k*-dependent bias, which we need to quantify.

In order to determine and correct this bias in our $P_{1D,\alpha}$ measurement, we compare mocks for which DLAs or atmospheric lines are masked (respectively *DLAm* and *LINEm*) with mocks where no masking is applied (*CONT*). On those two mocks, the DLAs and atmospheric emission lines are not imposed on spectra. We want to derive only the corrections of masking in order to apply them on data. The coefficients used for both masking corrections are defined as the ratio between the unmasked and the masked power spectra:

$$A_{\rm line}(k,z) = \frac{P_{1D,\alpha,\rm CONT}(k,z)}{P_{1D,\alpha,\rm LINEm}(k,z)}\,. \tag{22}$$

$$A_{\rm dla}(k,z) = \frac{P_{1D,\alpha,\rm CONT}(k,z)}{P_{1D,\alpha,\rm DLAm}(k,z)}\,. \tag{23}$$

5.3.1 Atmospheric emission lines

The correction induced by the DESI atmospheric line mask, as defined in Section 4.4, is shown in Fig. 14. We verified that, for all redshifts where the masked Ly α forest length of DESI is close to eBOSS, the impact of masking is lower in the DESI case. It indicates that applying thinner masks to our measurement mitigates the impact of masking.

As expected, the correction roughly scales with the number of masked spectrum pixels. The effect of masking is a relatively smooth function of wavenumber, and its main impact is at low wavenumber. The most impacted redshift bins are $z = 2.2$ (CaIIgalactic absorption lines), $z = 2.6$ (lines at 4360 Å in the transmission dip), and at high redshift for which many atmospheric lines need to be masked. As shown in Fig. 12, redshifts $z = 2.4$, 2.8, 3.0, 3.2 have no masks applied, and only a few for $z = 3.4$. It is also in agreement with the level of corrections. The impact of atmospheric line masking is qualitatively in agreement with the eBOSS results in Chabanier et al. (2019). We choose to model $A_{\rm line}(k,z)$ by a second-order polynomial fit and use this correction in the final calculation of $P_{1D,\alpha}$.

5.3.2 DLA masking

DLAs are added at random locations in the Ly α forest during the creation of the mocks. For this study, we do not attempt to characterize the completeness of the DLA finder applied to the data, and we use a 'truth' DLA catalogue for masking.

We mask the 'truth' catalog with the same parameters as the DLA data catalogue, i.e. for $N_{\rm HI} > 10^{20.3}$ cm^{-2}. The correction induced by the masking, $A_{\rm dla}(k,z)$, is represented in Fig. 15. As it was already seen in the eBOSS measurement (Chabanier et al. 2019), the DLA masking has a small impact compare to atmospheric emission lines. This is due to the random distribution of DLAs and the smaller masking in terms of Ly α forest length. As the impact is very similar for all wavenumbers, we apply a *k*-independent correction $A_{\rm dla}(k, z) = A_{\rm dla}(z)$, whose amplitude is 0.5 per cent on average.

6 UNCERTAINTY ESTIMATION

The statistical uncertainty of our averaged $P_{1D,\alpha}$ measurement, noted $\sigma_{\rm stat}$, is obtained during the $\overline{\rm SNR}$weighting scheme presented in Appendix B. For each (k, z) bin, a binned histogram of standard deviation as a function of $\overline{\rm SNR}$ is derived. Fitting this histogram provides a function $\sigma_{k,z}(\overline{\rm SNR})$ that is used to define the statistical uncertainty:

$$\sigma_{\rm stat}(k,z) = \sqrt{\frac{1}{\sum_i \left(\sigma_{k,z}(\overline{\rm SNR}_i)\right)^{-2}}}\,, \tag{24}$$

where σ is the fitted function, and the *i* index runs over the $\overline{\rm SNR}$ bins chosen.

The obtained statistical uncertainties are shown in Fig. 16. Despite using a $\overline{\rm SNR}$-dependent weighting in the $P_{1D,\alpha}$ calculation compared to a redshift-dependent $\overline{\rm SNR}$ cut as eBOSS, we find similar trends as in Chabanier et al. (2019). The statistical uncertainty depends mainly on the number of sub-forests for each redshift bin, so that $\sigma_{\rm stat}$ is an increasing function of redshift. This error bar depends also on the power spectrum level, as pointed out by the right panel of Fig. 16, for which low-redshift bins are not separated. In the case of $z \sim 2.2-2.4$, at small scales ($k \gtrsim 1.5$ Å^{-1}), the statistical uncertainties are crossing each other. It is caused by the large noise increase in the blue spectral band due to atmospheric absorptions.

Looking at wavenumber dependence, $\sigma_{\rm stat}$ increases as a function of k for small scales ($k \gtrsim 1.5$ Å^{-1}). This is due to the resolution correction, which effectively increases the rms of individual $P_{1D,\alpha}$. At large scales ($k \lesssim 1$ Å^{-1}), $\sigma_{\rm stat}$ is a decreasing function of k, mainly due to the decrease in Fourier modes available to compute $P_{1D,\alpha}$.

In our study, we characterized the impact of several instrumental and astrophysical contaminants. From this extensive study, we associate systematic errors, noted $\sigma_{\rm syst}$, to our $P_{1D,\alpha}$ measurement. Fig. 17 shows the systematic uncertainties for different redshift bins and their relative values with respect to statistical errors. Similarly to Chabanier et al. (2019), we made conservative choices to define those uncertainties:

(i) *Noise estimation*: As presented in Section 4.2, the pipeline noise is corrected using the α corrective term, which depends on

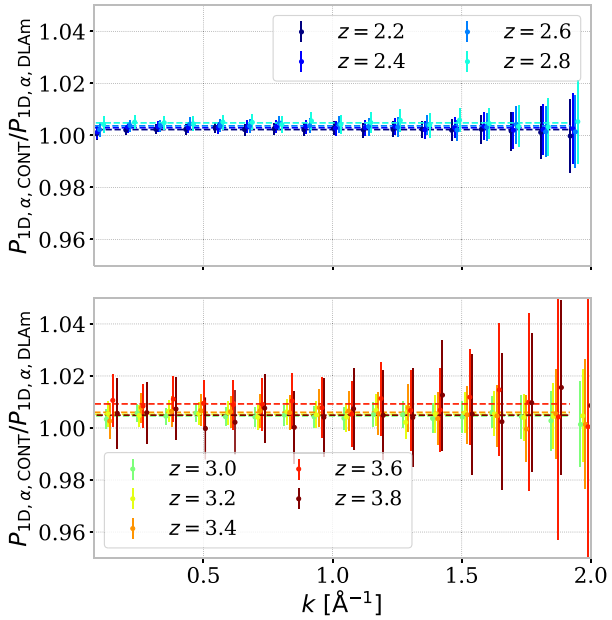

Figure 15. Ratio between the unmasked (*CONT*) and masked (*DLAm*) power spectra (equation 23) for DLA masking on the combination of the 10 mocks. Each power spectrum is re-binned by a factor of 3 to reduce error bars. Continuous lines shows constant fits are used for correction. For clarity, we artificially offset the points corresponding to different redshifts.

the data set considered. We assign a systematic uncertainty equal to 30 per cent of the average α for each redshift bin.

(ii) *Resolution*: We fit the average resolution correction given in Fig. 5 by a simplified model $\exp(-0.5(k\Delta\lambda)^2) \cdot \mathrm{sinc}(0.5k\Delta\lambda_{\rm pix})$ with $\Delta\lambda_{\rm pix}$ fixed to the DESI spectral pixel separation. This procedure allows to determine the effective spectral resolution $\Delta\lambda$ in Å. Measurements of DESI PSF stability, shown in Abareshi et al. (2022, fig. 30), indicate that its fractional change is less than 1 per cent over all spectrographs. We therefore assign a conservative systematic uncertainty $\sigma_{\Delta\lambda} = 1$ per cent$\Delta\lambda$. Using the above-mentioned simplified resolution model, this translates into a $P_{1D,\alpha}$ uncertainty equal to $2k^2\Delta\lambda\sigma_{\Delta\lambda} \cdot P_{1D,\alpha}(k)$.

(iii) Resolution correction: We apply a correction to the resolution modeling as presented in Section 4.1, by multiplying $A_{\rm res}$ to $P_{1D,\alpha}$. We add an associated systematic error defined as 30 per cent of this correction.

(iv) Side-band: The fitted side-band power spectrum $P_{\rm SB1,m}$ measured in Section 4.3 is subtracted to $P_{1D,\alpha}$ to account essentially for metal absorptions in the Ly α forest region. We associate with this correction a systematic uncertainty equal to the statistical errors of the measured SB1 power spectrum. This is a conservative choice, as the modelling performed in Section 4.3 closely reproduces $P_{\rm SB}$.

(v) Spectrum pixel masking: The impact of masking DLAs and atmospheric emission lines on the $P_{1D,\alpha}$ measurement was determined with synthetic data in Section 5.3. Spectrum pixel masking is corrected by multiplying $A_{\rm line}(z,k) \cdot A_{\rm dla}(z,k)$ to the $P_{1D,\alpha}$ estimator. We define the systematic error associated with each masking as 30 per cent of this correction.

(vi) Continuum fitting: Similarly, we assign a systematic error of 30 per cent times the $A_{\rm cont}(z,k)$ correction computed in Section 5.2.

(vii) DLA completeness: Using the synthetic data described in Section 5, we derived the impact of DLA on the 1D power spectrum as the ratio between mock with DLA (*DLA*) and without (*CONT*). We fit this ratio with an adapted function provided by Rogers et al. (2017). As detailed previously in Section 2.4, our DLA catalogue of data results from the combination of two finders. The trend of this ratio is reported on the penultimate panel of Fig. 17. In Chabanier et al. (2021), the authors perform a full study on eBOSS data and provide the completeness of the CNN finder. The completeness of this finder is higher than 85 per cent for $\log(N_{\rm HI}) > 20.3$. To be conservative, we choose to associate an uncertainty of 15 per cent of the total impact of DLAs on $P_{1D,\alpha}$ to the incompleteness of our catalogue. We stress that this uncertainty is overestimated, as the CNN finder has a higher completeness for DLAs with higher column density and since we are using an additional GP algorithm.

For all the corrections we applied on our $P_{1D,\alpha}$ measurement, the choice of 30 per cent in the associated systematic uncertainties is motivated by the fact that we consider a shift randomly ranging between no correction and 100 per cent of the correction. It is

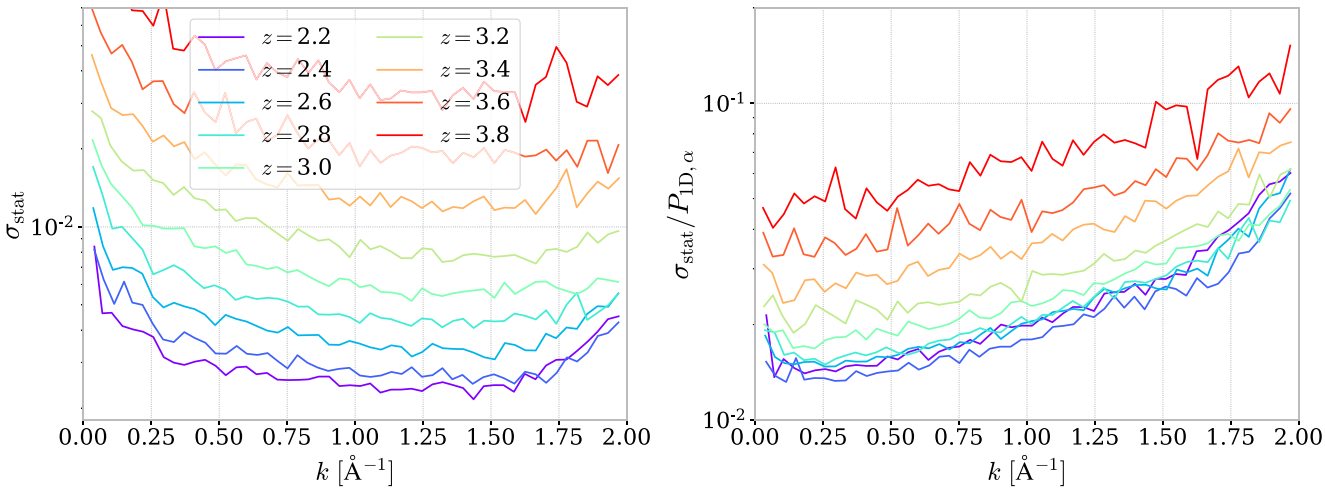

Figure 16. Statistical uncertainties (left) and its relative value with respect to $P_{1D,\alpha}$ (right) of the DESI SV3 + M2 measurement in Å, as a function of wavenumber.

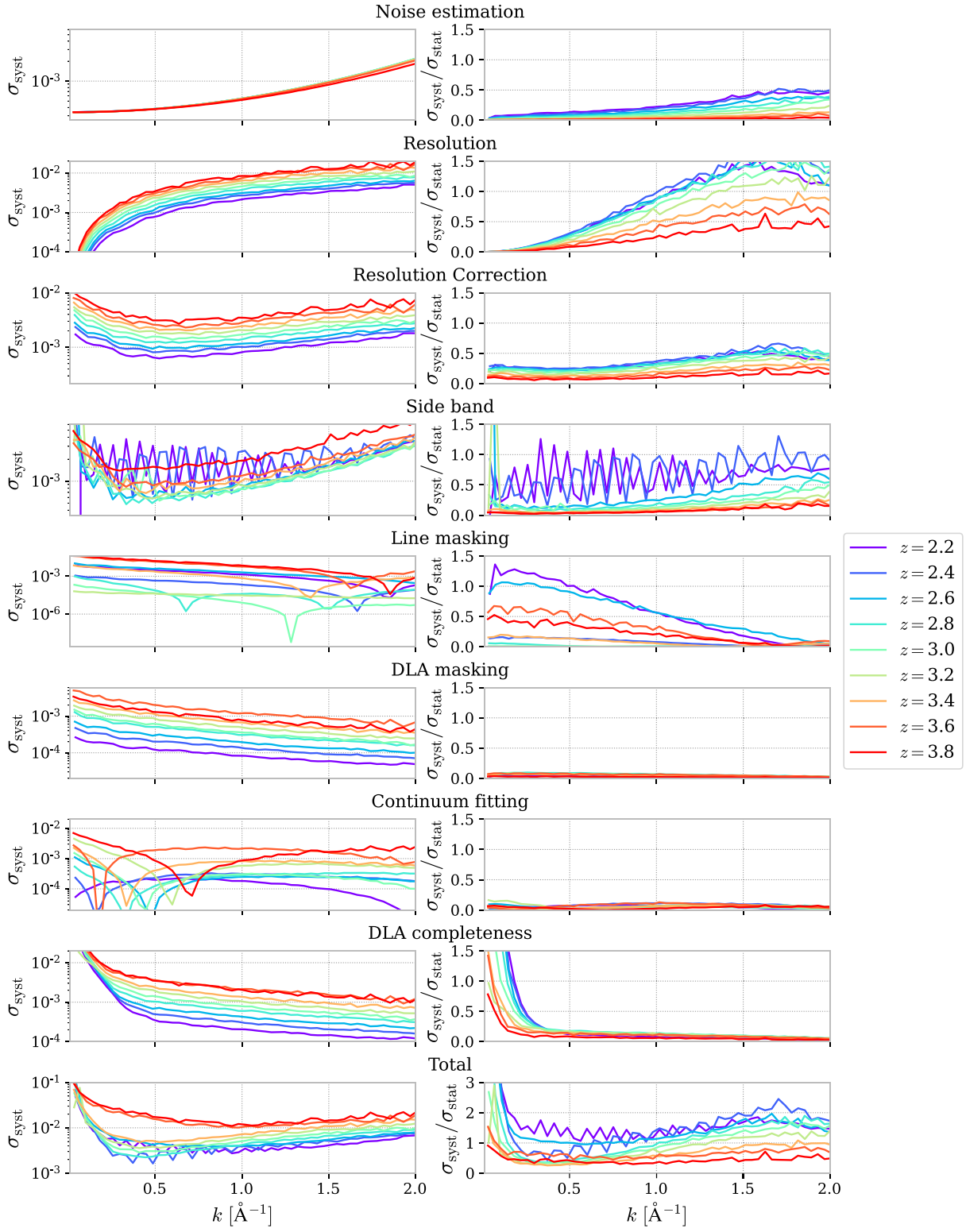

Figure 17. Value of the systematic uncertainties $\sigma_{\rm syst}$ in Å, for different redshift bins, on the DESI SV3 + M2$P_{\rm 1D,\alpha}$ measurement. Each line is associated to a systematic considered in this article. The left panels show the absolute uncertainties, and the right panels, their relative values with respect to statistical uncertainties showed in Fig. 16.

described by a uniform distribution between 0 and 1. The standard deviation of the distribution, equal to 0.30, quantifies the spread among the possible values, leading to a systematic uncertainty equal to 30 per cent of the correction.

Opposite to eBOSS (Chabanier et al. 2019), we chose to not account for the incompleteness of the BAL catalog in our analysis, as it is one of the weaker contaminants. The study of BAL catalogue completeness will be performed on further studies.

The general trends in Fig. 17 are similar to those of the eBOSS measurement (Chabanier et al. 2019). Given our limited statistics, most of the systematic errors are smaller than the statistical uncertainties for all redshift bins and all scales. However, this is expected to change for future DESI measurements, which will offer unprecedented statistics, thus reducing the statistical errors.

There is room for improvement for the major source of systematic uncertainties presented above. The noise modelling can be improved by understanding and correcting the source of unaccounted noise. Regarding the resolution modelling, the mathematical model and its verification with the relatively new CCD mocks presented in Section 4.1 can be improved with additional tests and larger data sets. Decreasing statistical error on the side-band power spectrum will directly reduce the associated systematic error. Concerning the spectrum pixel masking, especially for the atmospheric lines, thinner masks can be applied considering the improvement of atmospheric emission line subtraction in DESI (see Guy et al. 2022). Furthermore, a more complex but analytical mathematical correction could be derived for this regular masking. For the DLA completeness, a more advanced study as the one performed in Chabanier et al. (2019) is needed to reduce the associated systematic uncertainties. Additionally, as shown in Rogers et al. (2017), the impact and correction of objects with lower column density (sub-DLAs, Lyman limit systems, etc.) should be accounted for in a future, more developed study. Finally, the systematics are defined from a simplified assumption of uncertainty propagation, along with conservative choices. We plan to improve the modeling of each systematic and obtain more reliable uncertainties by modeling them directly into large samples of synthetic data.

7 DESI MEASUREMENT

We apply the methodology and corrections described in previous sections to the combination of SV3, and M2 data sets, noted SV3 + M2. We choose to remove the SV1 data set from this measurement due to considerations on the noise power spectrum shown in Section 4.2 and measurements performed in the Appendix A.

The $P_{1D,\alpha}$ measurement is done using the pipeline and parameters presented in Section 3. Considering all the corrections defined in the previous section, the final $P_{1D,\alpha}$ estimator is defined by

$$P_{1D,\alpha}(k) = A_{\rm line}(z,k)\cdot A_{\rm dla}(z,k)\cdot A_{\rm cont}(z,k)\cdot A_{\rm res}(k)\cdot \left(\left\langle \left[P_{\rm raw}(k) - P_{\rm pipeline}(k) - \alpha\right]\cdot \mathbf{R}^{-2}(k)\right\rangle - P_{\rm SB1,m}(k)\right) . \quad (25)$$

Fig. 18 presents the normalized $P_{1D,\alpha}$ measurement such that $\Delta^2_{1D,\alpha} = kP_{1D,\alpha}/\pi$. This observable is shown for 9 redshift bins ranging from 2.2 to 3.8, for wavenumbers $0.145 \leq k \leq 2$ Å^{-1}, and using a total of 73 839 subforests extracted from 26 330 quasar spectra. The represented error bars are the statistical and systematic uncertainties added in quadrature. The details of subforest properties for each redshift bin are given in Table 4.

7.1 Comparison with other measurements

We perform a comparison with past measurements. The comparison of the DESI $P_{1D,\alpha}$ with the last moderate-resolution measurement (eBOSS survey Chabanier et al. 2019) is shown in Fig. 19 (left). We also compare our measurement with the last high-resolution measurement obtained using the combination of KODIAQ, SQUAD, and XQ-100 surveys (Karaçaylı et al. 2022) in Fig. 19 (right).

Both moderate- and high-resolution measurements were expressed in velocity units (s km^{-1}). The conversion between this unit system and the one used for DESI (Å^{-1}) is defined by

$$k\left[\text{s}\cdot\text{km}^{-1}\right] = k\left[\text{Å}^{-1}\right]\times \lambda_\alpha(1+z)/c\,. \quad (26)$$

This conversion is performed at the stage of the FFT estimation by converting the terms in equation (14) in velocity units before performing the ensemble average to compute $P_{1D,\alpha}$. We found that doing this conversion on the averaged $P_{1D,\alpha}$ significantly shift $P_{1D,\alpha}$ because of the redshift term in the equation (26). This is caused by the non-uniform redshift distribution in each z-bins, pointed out by Fig. 2. Similarly, we converted all corrections terms appearing in equation (25) in velocity units. For the comparison with eBOSS measurement, we compute $P_{1D,\alpha}$ in velocity units with the same binning. For the high-resolution measurement, we rebin our DESI measurement to the same wavenumber binning, and account this rebinning in the calculation of error bars, to obtain a fair comparison.

The comparison with the eBOSS measurement in Fig. 19 (left) yields a ~ 15 per cent difference at small wavenumber ($k < 0.01$ km s^{-1}). To investigate this discrepancy, we performed a detailed investigation by varying most of the parameters of our analysis, which are susceptible to impact small scales. The discussion concerning those tests is detailed in Appendix D. On the other scales (0.01 km s$^{-1} < k < 0.02$ km s^{-1}), our measurement agrees with eBOSS considering the error bars.

The major improvement lays at large wavenumber ($k > 0.015$ km s^{-1}) where the improved DESI resolution and noise modelling allows us reaching much smaller scales than eBOSS, especially for high redshifts. We are able to conservatively reach the wavenumber $k_{\rm max} = 2.0$ Å^{-1} for all redshift bins. In comparison, the eBOSS measurement (Chabanier et al. 2019) achieved maximal wavenumber of $k_{\rm max} = 1.54$ Å^{-1} for $z = 2.2$ and $k_{\rm max} = 1.03$ Å^{-1} for $z = 3.8$. At large wavenumbers, the eBOSS measurement is highly contaminated by noise and resolution. We consider that our measurement is more suitable to probe those scales, considering the improvement in resolution and noise estimation.

We compare our measurement with the high-resolution measurements in Fig. 19 (right). This measurement (Karaçaylı et al. 2022) is performed with a statistically smaller sample of quasars but with a very high spectral resolution ($5000 \lesssim R \lesssim 60\,000$) and $\overline{\rm SNR}$. The high-resolution $P_{1D,\alpha}$ measurements thus allow reaching very small scales with large error bars. Our agreement shows a 20 per cent disagreement at small scales ($k > 0.02$ km s^{-1}), mostly for the lower redshifts measured. It indicates that there is still room for improvement in our measurement's noise and resolution modeling. We note that considering error bars, both measurements agree on intermediate scales.

8 CONCLUSION AND PROSPECTS

We performed the first measurement of the 1D power spectrum ($P_{1D,\alpha}$) with DESI data. The main objective of this paper is to carefully characterize the different contaminants of $P_{1D,\alpha}$ with regard to DESI instrument. In particular, we modelled the noise and spectral resolution of DESI. In comparison to the previous eBOSS measurement (Chabanier et al. 2019), we improved the analysis of side-band power spectrum and atmospheric emission lines.

We used adapted synthetic data to correct the impact of spectrum pixel masking, continuum fitting and spectral resolution modelling. We performed a complete review of the systematic uncertainties linked to the $P_{1D,\alpha}$ pipeline and compared the DESI measurement with previous moderate- and high-resolution measurements. We find a relatively good agreement, except for a slight difference at

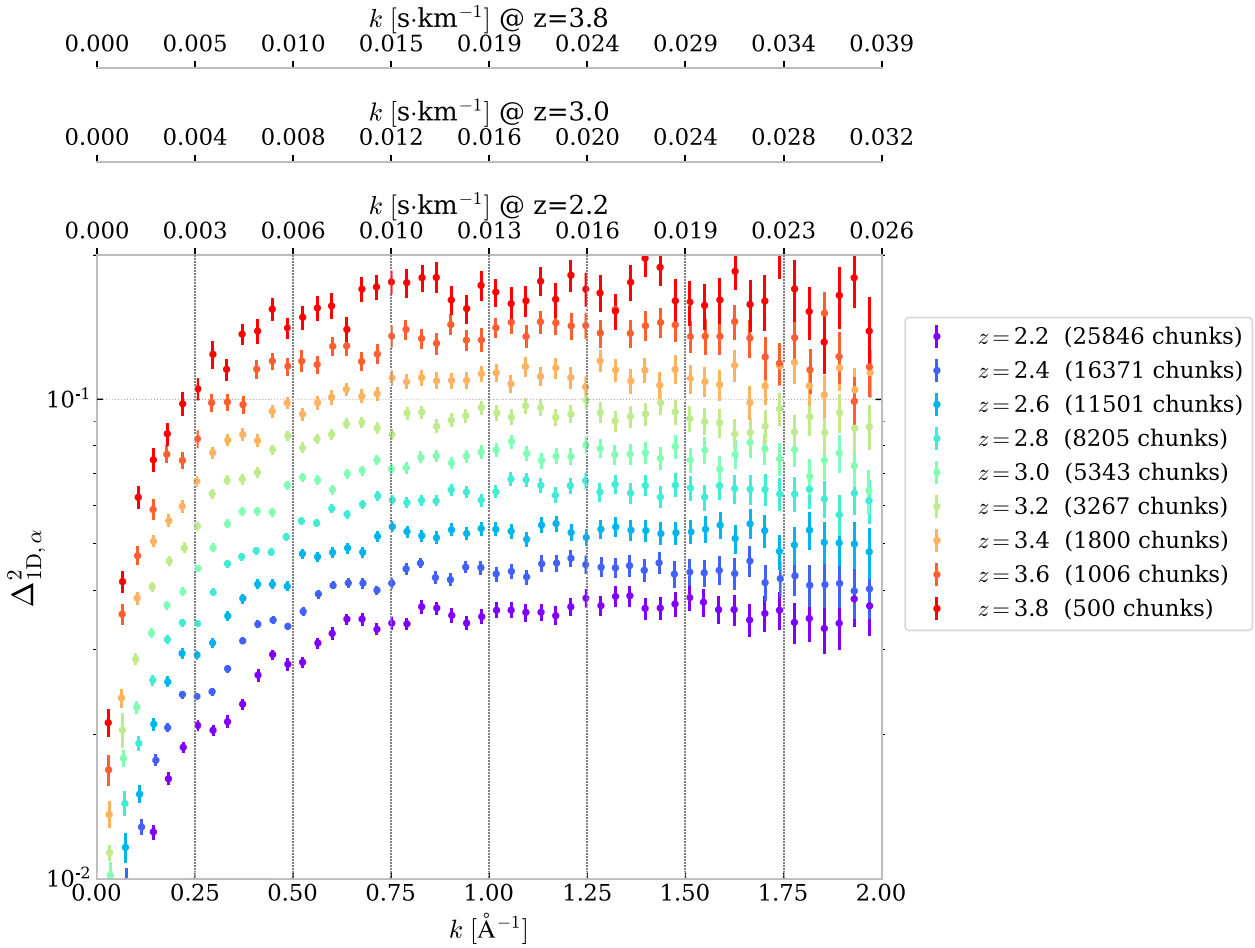

Figure 18. Normalized 1D Ly α forest power spectrum ($\Delta_{1D,\alpha}(k)$) using the SV3 + M2 data set, for redshift bins from $z = 2.2$ to $z = 3.8$. All the corrections given in equation (25) are applied to perform this measurement. As an illustration, wavenumbers in velocity space for different redshifts are represented at the top of the figure. Error bars are systematic and statistical uncertainties added in quadrature.

Table 4. Number of subforest, average redshift, and signal-to-noise ratio for each redshift bins in the final data set sample used in this measurement.

z bin	2.2	2.4	2.6	2.8	3.0	3.2	3.4	3.6	3.8
#	25 846	16 371	11 501	8205	5343	3267	1800	1006	500
$\langle z \rangle$	2.2	2.4	2.6	2.8	2.99	3.19	3.39	3.59	3.79
$\overline{\text{SNR}}$	2.79	3.0	3.08	3.13	3.28	3.27	3.14	3.07	3.21

large scales with the eBOSS measurement, partially due to the different residual correction we apply. Our estimation of $P_{1D,\alpha}$ is also compared with the QMLE method in a companion paper Karaçaylı et al. (2023a). Measurements with FFT and QMLE methods agree at 1 per cent level precision up to half the Nyquist frequency.

The DESI spectral resolution is approximately two times better than SDSS. Consequently, our $P_{1D,\alpha}$ measurement is of high scientific interest to probe the small scales of the intergalactic medium. However, the data sets we exploited remains inferior to eBOSS in terms of statistics. If we apply the same $\overline{\text{SNR}}$ cut as in Chabanier et al. (2019), our sample contains 17 333 subforests compared to 94 558 for eBOSS. However, we expect future DESI data sets to provide high increase of statistics (up to 1 million Ly α forest, thus almost 3 million subforests). This unprecedented data set will allow obtaining a subper cent precision measurement. The resulting $P_{1D,\alpha}$ measurement will provide stringent constraints on the sum of neutrinos masses, warm dark matter models, and on the parameters of the intergalactic medium (DESI 2016a; Valluri et al. 2022).

We plan to improve our analysis to keep the level of systematic error close to the statistical one. First, applying stricter constraints on a larger Ly α forest sample will be beneficial to reduce systematical uncertainties. Furthermore, we also plan to improve the treatment of contaminants presented in Section 6. In particular, we plan to test extensively the resolution and noise estimations on pixel-level simulations of the CCD camera.

High-resolution hydrodynamical simulations, associated with Gaussian processes emulator such as in Pedersen et al. (2020b)

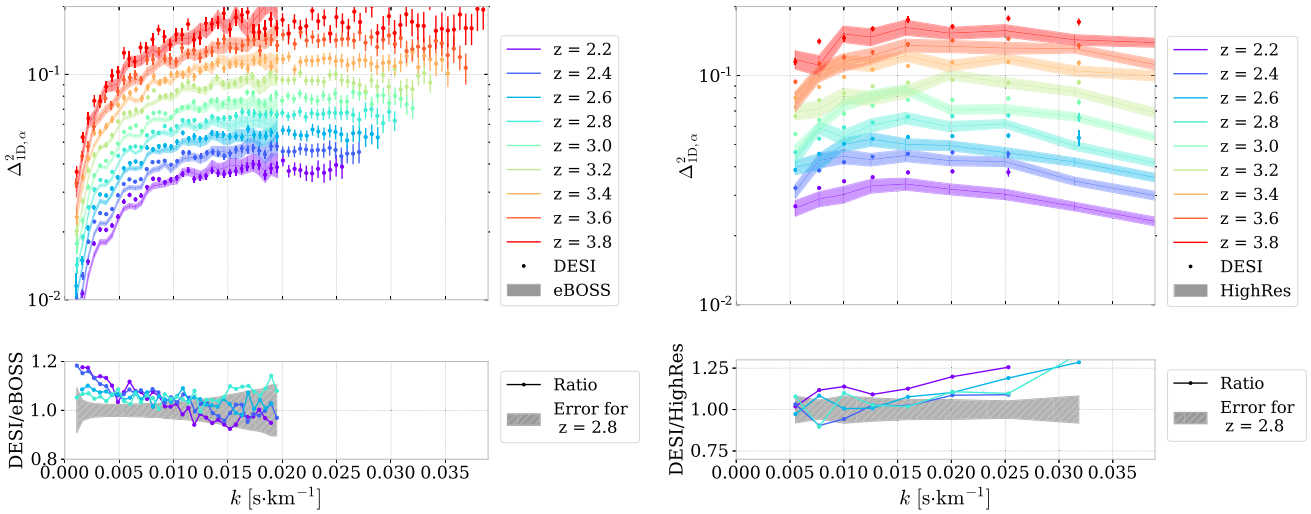

Figure 19. Left: Comparison between the measurement performed here and that of the eBOSS data (Chabanier et al. 2019). For this comparison, all the analysis was performed in velocity units using the conversion given by equation (26). Both normalized $P_{1D,\alpha}$ are shown in the top panel (our measurement with points and the eBOSS measurement with shaded coloured areas). The ratio between DESI and eBOSS measurements is shown in the bottom panel. The striped grey area in the bottom panel shows the centered error bar of the ratio averaged over all shown redshift bins. We remove high-redshift bins for clarity. A description of the tests used to explain the difference between DESI and eBOSS can be found in Appendix D. (right) Same comparison with the high-resolution measurement obtained using the combination of KODIAQ, SQUAD, and XQ-100 surveys (Karaçaylı et al. 2022). In this case, our $P_{1D,\alpha}$ measurement and the error bars associated are rebinned to the wavenumber binning of the high-resolution measurement. For clarity, only the four first redshift bins are shown in the bottom panel of each figure, along with the error on the ratio for $z = 2.8$ centred to unity.

and Walther et al. (2021) will be employed with next DESI $P_{1D,\alpha}$ measurement to obtain constraints on cosmological and intergalactic medium parameters.

ACKNOWLEDGEMENTS

The authors acknowledge support from Agence National de la Recherche (grant ANR-16-CE31-0021). The project leading to this publication has received funding from Excellence Initiative of Aix-Marseille University – A*MIDEX, a French 'Investissements d'Avenir' programme (grant AMX-20-CE-02–DARKUNI).

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 U.S. 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 Technology 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 DESI spectra from SV1 and SV3 will be publicly available in the Early Data Release (EDR) in 2023. We added the first 2 months of the main survey to improve our statistics. The main survey spectra will be made publicly available as part of Year 1 data release in the future. All the data points of the figures in this article are made available according to the data management policy of DESI (Ravoux 2023). All the outputs and plots of this article are generated with picca (8.1.2) and p1desi https://github.com/corentinravoux/p1desi (1.0.0).

REFERENCES

Abareshi B. et al., 2022, Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2205.10939
Ahumada R. et al., 2020, ApJS, 249, 3
Alexander D. M. et al., 2022, The DESI Survey Validation: Results from Visual Inspection of the Quasar Survey Spectra. Available at: http://arxiv.org/abs/2208.08517
Armengaud E., Palanque-Delabrouille N., Yèche C., Marsh D. J. E., Baur J., 2017, MNRAS, 471, 4606
Bajtlik S., Duncan R. C., Ostriker J. P., 1988, ApJ, 327, 570
Baur J., Palanque-Delabrouille N., Yèche C., Boyarsky A., Ruchayskiy O., Armengaud É., Lesgourgues J., 2017, J. Cosmol. Astropart. Phys., 2017, 013
Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, J. Cosmol. Astropart. Phys., 2016, 012
Bautista J. E. et al., 2017, Astron. Astrophys., 603, A12
Blanton M. R. et al., 2017, AJ, 154, 28
Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101
Bolton A. S., Schlegel D. J., 2010, Publ. Astron. Soc. Pac., 122, 248
Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897
Borde A., Palanque-Delabrouille N., Rossi G., Viel M., Bolton J., Yèche C., LeGoff J.-M., Rich J., 2014, J. Cosmol. Astropart. Phys., 2014, 005

Brodzeller A. et al., 2023, Performance of the Quasar Spectral Templates for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2305.10426
Busca N., Balland C., 2018, QuasarNET: Human-level spectral classification and redshifting with Deep Neural Networks. Available at: http://arxiv.org/abs/1808.09955
Chabanier S. et al., 2019, J. Cosmol. Astropart. Phys., 2019, 017
Chabanier S. et al., 2021, ApJS, 258, 18
Chabanier S. et al., 2022, Modeling the Lyman-α forest with Eulerian and SPH hydrodynamical methods. Available at: http://arxiv.org/abs/2207.05023
Chaussidon E. et al., 2022, Target Selection and Validation of DESI Quasars. Available at: http://arxiv.org/abs/2208.08511
Croft R. A. C., Weinberg D. H., Bolte M., Burles S., Hernquist L., Katz N., Kirkman D., Tytler D., 2002, ApJ, 581, 20
Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44
Dawson K. S. et al., 2016, AJ, 151, 44
Day A., Tytler D., Kambalur B., 2019, MNRAS, 489, 2536
DESI, 2016a, The DESI Experiment Part I: Science,Targeting, and Survey Design. Available at: http://arxiv.org/abs/1611.00036
DESI, 2016b, The DESI Experiment Part II: Instrument Design. Available at: http://arxiv.org/abs/1611.00037
DESI, 2023a, Validation of the Scientific Program for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2306.06307
DESI, 2023b, The Early Data Release of the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2306.06308
Dey A. et al., 2019, ApJ, 157, 168
du Mas des Bourboux H. et al., 2020, ApJ, 901, 153
du Mas des Bourboux H. et al., 2021, Astrophysics Source Code Library, record ascl:2106.018
Farr J., Font-Ribera A., Pontzen A., 2020, JCAP, 2020, 015
Gaikwad P., Srianand R., Haehnelt M. G., Choudhury T. R., 2021, MNRAS, 506, 4389
Gunn J. E. et al., 2006, AJ, 131, 2332
Gunn J. E., Peterson B. A., 1965, ApJ, 142, 1633
Guy J. et al., 2022, The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2209.14482
Ho M.-F., Bird S., Garnett R., 2021, MNRAS, 507, 704
Iršič V. et al., 2016, MNRAS, 466, 4332
Iršič V., Viel M., Haehnelt M. G., Bolton J. S., Becker G. D., 2017, Phys. Rev. Lett., 119, 031302
Karaçaylı N. G. et al., 2022, MNRAS, 509, 2842
Karaçaylı N. G. et al., 2023a, Optimal 1D Ly α Forest Power Spectrum Estimation – III. DESI Early Data. Available at: http://arxiv.org/abs/2306.06316
Karaçaylı N. G. et al., 2023b, MNRAS, 522, 5980
Karaçaylı N. G., Font-Ribera A., Padmanabhan N., 2020, MNRAS, 497, 4742
Khaire V. et al., 2019, MNRAS, 486, 769
Kim T.-S., Viel M., Haehnelt M. G., Carswell R. F., Cristiani S., 2004, MNRAS, 347, 355
Kramida A., Ralchenko Yu., Reader J., ASD NIST, 2021, NIST Atomic Spectra Database (ver. 5.9), National Institute of Standards and Technology. Available at: https://physics.nist.gov/asd
Lee K.-G. et al., 2013, AJ, 145, 69
Lesgourgues J., Pastor S., 2006, Phys. Rep., 429, 307
Lesgourgues J., Pastor S., 2012, Adv. High Energy Phys., 2012, 1
Levi M. et al., 2013, The DESI Experiment, a whitepaper for Snowmass 2013. Available at: https://arxiv.org/abs/1308.0847
Lopez S. et al., 2016, A&A, 594, A91
Lukić Z., Stark C., Nugent P., White M., Meiksin A., Almgren A., 2015, MNRAS, 446, 3697
Lynds R., 1971, ApJ, 164, L73
McDonald P. et al., 2006, ApJS, 163, 80
McDonald P., Miralda-Escude J., Rauch M., Sargent W. L. W., Barlow T. A., Cen R., Ostriker J. P., 2000, ApJ, 543, 1
McDonald P., Seljak U., Cen R., Bode P., Ostriker J. P., 2005, MNRAS, 360, 1471
McQuinn M., 2016, Annu. Rev. Astron. Astrophys., 54, 313
Meiksin A. A., 2009, Rev. Mod. Phys., 81, 1405
Miller T. N. et al., 2023, The Optical Corrector for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2306.06310
Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2018, MNRAS, 482, 3458
Myers A. D. et al., 2022, The Target Selection Pipeline for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2208.08518
O'Meara J. M. et al., 2015, AJ, 150, 111
O'Meara J. M., Lehner N., Howk J. C., Prochaska J. X., Fox A. J., Peeples M. S., Tumlinson J., O'Shea B. W., 2017, AJ, 154, 114
Palanque-Delabrouille N. et al., 2013, A&A, 559, A85
Palanque-Delabrouille N. et al., 2015, J. Cosmol. Astropart. Phys., 2015, 011
Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 2020, 038
Parks D., Prochaska J. X., Dong S., Cai Z., 2017, MNRAS, 476, 1151
Pedersen C., Font-Ribera A., Gnedin N. Y., 2023, ApJ, 944, 223
Pedersen C., Font-Ribera A., Kitching T. D., McDonald P., Bird S., Slosar A., Rogers K. K., Pontzen A., 2020a, JCAP, 2020, 025
Pedersen C., Font-Ribera A., Rogers K. K., McDonald P., Peiris H. V., Pontzen A., Slosar A., 2020b, JCAP, 2021, 033
Pieri M. M. et al., 2014, MNRAS, 441, 1718
Pontzen A. et al., 2008, MNRAS
Puchwein E. et al., 2023, MNRAS, 519, 6162
Ramírez-Pérez C. et al., 2023, The Lyman-alpha forest catalog from the Dark Energy Spectroscopic Instrument Early Data Release. Available at: http://arxiv.org/abs/2306.06312
Ravoux C., 2023, The Dark Energy Spectroscopic Instrument: One-Dimensional Power Spectrum from First Lyman-alpha Forest Samples with Fast Fourier Transform. preprint (arXiv:2306.06311)
Rogers K. K., Bird S., Peiris H. V., Pontzen A., Font-Ribera A., Leistedt B., 2017, MNRAS, 474, 3032
Schlafly E. F. et al., 2023, Survey Operations for the Dark Energy Spectroscopic Instrument. Available at: http://arxiv.org/abs/2306.06309
Seljak U., Slosar A., McDonald P., 2006, J. Cosmol. Astropart. Phys., 2006, 014
Silber J. H. et al., 2022, The Robotic Multi-Object Focal Plane System of the Dark Energy Spectroscopic Instrument (DESI). Available at: http://arxiv.org/abs/2205.09014
Smee S. A. et al., 2013, AJ, 146, 32
Valluri M. et al., 2022, Snowmass2021 Cosmic Frontier White Paper: Prospects for obtaining Dark Matter Constraints with DESI. Available at: https://arxiv.org/abs/2203.07491
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., Rauch M., Sargent W. L. W., 2008, Phys. Rev. Lett., 100, 041304
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
Walther M., Armengaud E., Ravoux C., Palanque-Delabrouille N., Yèche C., Lukić Z., 2021, J. Cosmol. Astropart. Phys., 2021, 059
Walther M., Hennawi J. F., Hiss H., Oñorbe J., Lee K.-G., Rorai A., O'Meara J., 2018, ApJ, 852, 22
Yang L., Zheng Z., Bourboux H. d. M. d., Dawson K., Pieri M. M., Rossi G., Schneider D. P., de la Macorra A., 2022, ApJ, 935, 121
Yèche C. et al., 2020, Res. Notes Am. Astron. Soc., 4, 179
Yèche C., Palanque-Delabrouille N., Baur J., BourBoux H. d. M. d., 2017, J. Cosmol. Astropart. Phys., 2017, 047
Zou H. et al., 2017, PASP, 129, 064101

APPENDIX A: DESI DATA SET COMPARISON

SV1, SV3, and M2 are very different data sets: they have different target selections and exposure times and were collected for different states of the DESI instrument. From a $P_{1D,\alpha}$ point of view, the noise properties of SV1 are a potential issue. While our initial goal was to analyse the full data sample available, we choose first to compare

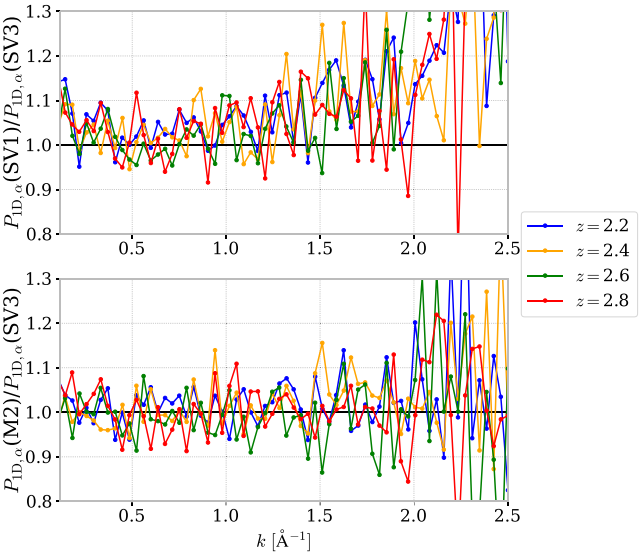

Figure A1. Ratio of the $P_{1D,\alpha}$, measured with the same parameters, between SV1 and SV3 (top), and M2 and SV3 (bottom), for four redshift bins.

the measured $P_{1D,\alpha}$ on the separate SV1, SV3, and M2 data sets. Fig. A1 shows their respective ratios, on the four redshift bins with largest statistics. It appears that the measurement of $P_{1D,\alpha}$ on SV1 is biased compared to the other two data sets. In particular, we believe that the difference at $k \gtrsim 1.0$ Å^{-1} is due to an imperfection in the noise correction presented in Section 4.2. Consequently, we decide to remove the SV1 data set in this study to remain conservative.

APPENDIX B: SNR WEIGHTING

For this $P_{1D,\alpha}$ measurement, we keep all Ly α subforests available in the measurement sample, independently of their $\overline{\text{SNR}}$, unlike what was done in eBOSS (Chabanier et al. 2019), reminding that only one $\overline{\text{SNR}}$ cut is applied at earlier stages of the analysis, during the continuum fitting procedure as described in Section 3.1.

Individual Ly α power spectra, falling into the same wavenumber bin, do not have the same dispersion which varies as function of the $\overline{\text{SNR}}$. In our analysis, we account for this effect at the last step of our FFT estimator pipeline, by weighting each of the Ly α subforests by a $\overline{\text{SNR}}$ dependent factor, while averaging over the full measurement sample.

First, for each Ly α subforest (i), the variance of individual $P_{1D,\alpha,i}$ is fitted according to the following:

$$\sigma^2(P_{1D,\alpha,i}) = \frac{a}{(\overline{\text{SNR}_i} - 1)^2} + b\,, \tag{B1}$$

where $\overline{\text{SNR}_i}$ is the mean signal-to-noise ratio of the Ly α sub-forest (i), defined in equation (6).

Hence, the weighting factor is

$$W_i = \frac{1}{\sigma^2(P_{1D,\alpha,i})}\,. \tag{B2}$$

The employed fitting model works well empirically with the measured σ^2 ($P_{1D,\alpha,i}$). Also according to equation (B1), W_i tends to 0 as $\overline{\text{SNR}_i}$ tends to 1, consistently with the applied $\overline{\text{SNR}}$ cut $= 1$.

Tests for this $\overline{\text{SNR}}$ weighting method were done on DESI-LITE mocks that mimic the SV1 + SV3 data set described in Section 2, and are specifically designed for $P_{1D,\alpha}$ measurement. A comparison

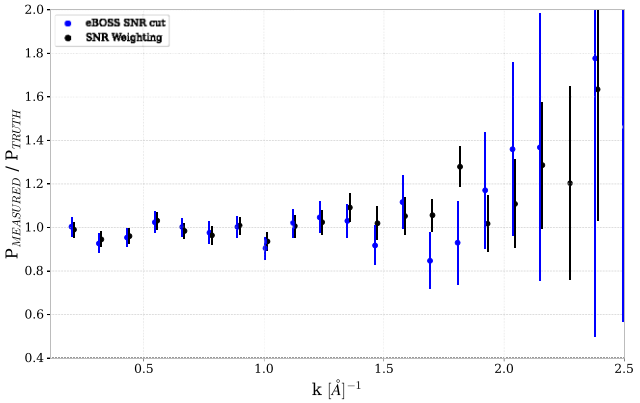

Figure B1. Comparison between measured $P_{1D,\alpha}$ and DESI-LITE mocks truth power spectrum, for the redshift bin z $= 3.8$, for both $P_{1D,\alpha}$ measured with eBOSS $\overline{\text{SNR}}$ cut method, and the $\overline{\text{SNR}}$ weighting method.

between measured $P_{1D,\alpha}$ and mocks truth power spectrum is represented in Fig. B1, for both $P_{1D,\alpha}$ measured with eBOSS $\overline{\text{SNR}}$ cut method, and our $\overline{\text{SNR}}$ weighting method.

Fig. B1 shows that we have an improvement compared to the eBOSS method, especially at large wavenumber and redshift values, where we are mostly limited by the statistics, as well as at low wavenumber and redshift values, while for the eBOSS measurement, there was no possible optimization at both small and large wavenumber ranges at the same time.

APPENDIX C: DETAILS ON NOISE ESTIMATION

C1 Pipeline noise

The noise associated to each spectrum is computed with the spectroscopic pipeline presented in Section 2.2. It is modeled as the addition of effects at the CCD scale. Each contribution is calculated by measuring the spatial variance on the CCD image from the associated noise source. It is assumed that this noise comes from the following four sources:

(i) **Poisson noise**: Measuring photons with a CCD is a statistical process. It creates a noise source which is directly linked to the input flux, and particularly dominant for low fluxes. For DESI, this noise is estimated by modelling the CCD.

(ii) **Overscan**: The overscan measures the bulk offset, i.e. the average level of all CCD pixels. It is used to remove small variations in the bias. Overscan suppression introduces noise.

(iii) **Bias**: Noise due to the response of the CCD to a minimal exposition time. It emerges from parasite electron or CCD pixel defects. The master bias estimates this noise.

(iv) **Dark current**: Readout noise due to the thermal motion of the atoms composing CCD material which induces charge deposit. Dark current is estimated using the master dark. In DESI, the modeling of noise is improved to account for Poisson noise in the dark frames.

All those noise sources have all been corrected for their dependence on the CCD position. By adding these four terms, we obtain a CCD noise estimator which is propagated to the spectra by the 'spectroperfectionism' formalism. A pipeline noise estimator σ_{pip} is then obtained.

The noise power spectrum estimated from pipeline ($P_{\rm pipeline}$) is computed from the standard deviation σ_{δ_F} linked to $\sigma_{\rm pip}$ by equation (5). For each unmasked spectrum pixel, a contrast $\delta_{\rm pipeline}$ is generated following a normal probability distribution such that:

$$\delta_{\rm pipeline}(\lambda) \hookrightarrow \mathcal{N}\left(0, \sigma_{\delta_F}(\lambda)\right) . \tag{C1}$$

This procedure is repeated $N_{\rm G}$ times ($N_{\rm G} = 2500$) to obtain a converged noise power spectrum. For each quasar, the associated noise power spectrum is the average of the $N_{\rm G}$ noise contrasts after Fourier transformation:

$$P_{\rm pipeline}(k) = \left\langle \left|\delta_{\rm pipeline}\right|^2 \right\rangle_{N_{\rm G}} . \tag{C2}$$

C2 Exposure difference noise

Another noise estimation can be done using the difference between exposures of the same quasar, when several exposures are available for the same object. The difference between exposures removes the physical signal, leaving only the fluctuations due to noise. We implemented a noise power spectrum estimator using this principle. We define the difference coadd of a quasar of index j by separating half of its exposures in the even category ($N_{\rm even}$ exposures) and the other half in the odd category ($N_{\rm odd}$ exposures) such that

$$\Delta f_j = \frac{1}{2}\left(\frac{\sum_{k=1}^{N_{\rm even}} \left(\sigma_{{\rm pip},k}\right)^{-2} f_k}{\sum_{k=1}^{N_{\rm even}} \left(\sigma_{{\rm pip},k}\right)^{-2}} - \frac{\sum_{k=1}^{N_{\rm odd}} \left(\sigma_{{\rm pip},k}\right)^{-2} f_k}{\sum_{k=1}^{N_{\rm odd}} \left(\sigma_{{\rm pip},k}\right)^{-2}} \right) , \tag{C3}$$

where $\sigma^2_{\rm pip,k}$ is the pipeline noise of the exposure k for quasar j. In the case where the total number of exposures is even, $N_{\rm even} = N_{\rm odd}$. The standard deviation of Δf_j can be calculated from the variances of individual exposures:

$$\sigma_{\Delta f_j} = \frac{1}{2}\sqrt{\frac{1}{\sum_{k=1}^{N_{\rm even}} \left(\sigma_{{\rm pip},k}\right)^{-2}} + \frac{1}{\sum_{k=1}^{N_{\rm odd}} \left(\sigma_{{\rm pip},k}\right)^{-2}}} . \tag{C4}$$

This difference coadd is unbiased, i.e. of zero average, whatever the values of the sum of the inverse variance for both exposure populations. Finally, this estimator does not necessarily need an even total number of exposures.

To derive an estimator of $P_{\rm noise}$, the variance of Δf_j must be equal to that of the coadded flux defined by

$$f_j = \frac{\sum_k \left(\sigma_{{\rm pip},k}\right)^{-2} f_k}{\sum_k \left(\sigma_{{\rm pip},k}\right)^{-2}} , \tag{C5}$$

To obtain the same variance, we multiply Δf_j by the ratio $\sigma_{f_j}/\sigma_{\Delta f_j}$:

$$\Delta f_j^{\rm corr} = 2\frac{\frac{1}{\sqrt{\sum_{k=1}^{N_{\rm tot}} \left(\sigma_{{\rm pip},k}\right)^{-2}}}}{\sqrt{\frac{1}{\sum_{k=1}^{N_{\rm even}} \left(\sigma_{{\rm pip},k}\right)^{-2}} + \frac{1}{\sum_{k=1}^{N_{\rm odd}} \left(\sigma_{{\rm pip},k}\right)^{-2}}}} \Delta f_j , \tag{C6}$$

where $N_{\rm tot}$ is the total number of exposure for the quasar j ($N_{\rm tot} = N_{\rm even} + N_{\rm odd}$).

In SDSS analysis (McDonald et al. 2006; Palanque-Delabrouille et al. 2013; Chabanier et al. 2021), the variance of all exposures for a given object was considered equal. In this case, $\sigma_{\Delta f_j}$ can be simplified, and a correction was applied only in the case of an odd number of exposures. In the case of a constant exposure variance, the difference coadd in equation (C6) is equal to the one derived in Palanque-Delabrouille et al. (2013) and Chabanier et al. (2019). Our new estimator corrects the variance for any exposure time. It is essential in the case of DESI first data, for which the exposure times can be very variable compared to SDSS.

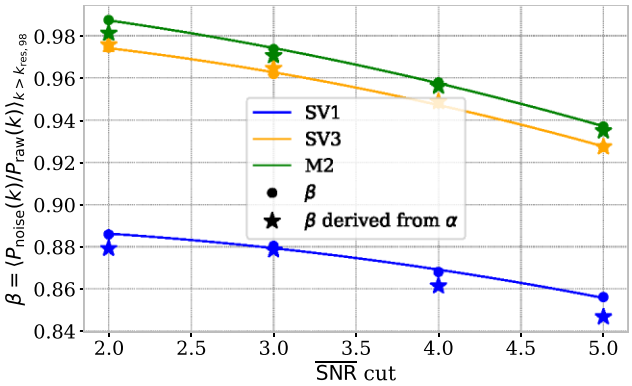

Figure C1. Asymptotic ratios β between the noise and raw power spectra for SV1 (blue), SV3 (yellow), and M2 (green) data sets, as a function of the minimal $\overline{\rm SNR}$ cut, for the pipeline noise. Points give the direct estimation of β. Stars represent β as derived from the asymptotic differences in Fig. 8. Second-order polynomial fits are shown only for representation.

The exposure difference coadd is computed in picca (Bourboux et al. 2021). We obtain an estimator of $P_{\rm noise}$ called difference power spectrum and noted $P_{\rm diff}$, such that

$$P_{\rm diff}(k) = \left| \mathcal{F}\left[\frac{\Delta f_j^{\rm corr}(\lambda)}{\overline{F}(\lambda) C_{\rm q}(\lambda)} - 1 \right] \right|^2 . \tag{C7}$$

C3 Additional considerations on noise estimation

Fig. C1 shows the asymptotic ratios β for the three data sets. We remark that for SV3 and M2, the $\overline{\rm SNR}$dependence of β is much more pronounced than that of α. The absolute noise level is the main parameter which varies when changing the minimal $\overline{\rm SNR}$ cut. It indicates that the residual noise source is additive rather than multiplicative. To support this hypothesis, we computed the β values derived from the α of Fig. 8, using the mean value of $P_{\rm noise}$ for all redshift. They are shown as stars in Fig. C1, and exhibit similar trends to the direct β computation, which corroborates that the missing noise is additive. Consequently, we decide to correct $P_{\rm pipeline}$ using an additive term α ($P_{\rm noise} = P_{\rm pipeline} + \alpha$).

The same noise study is performed on side-band regions SB1 and SB2 for which the astrophysical signal, i.e. absorption from intergalactic elements, is much lower than the Ly α band. The α values are shown in Fig. C2. For side-bands, the overall missing noise level exhibits similar trends as a function of the minimal $\overline{\rm SNR}$ cut, but is lower than for the Ly α band. This is likely due to the use of different quasar populations employed for side-band and Ly α measurements. Indeed, the DESI observation strategy is different for low redshift quasars (used for side-band study) and Ly α quasars. On average, the number of exposures is larger for Ly α than low-redshift quasars. Consequently, and in accordance with our previous interpretations, the misestimation of noise is larger for Ly α measurement than side-bands. The difference might also be due to the much lower astrophysical signal in the side-bands, allowing an improved estimation of asymptotic noise level. As the results for SB1 and SB2 are very similar, we decide to apply the same correction for these two bands.

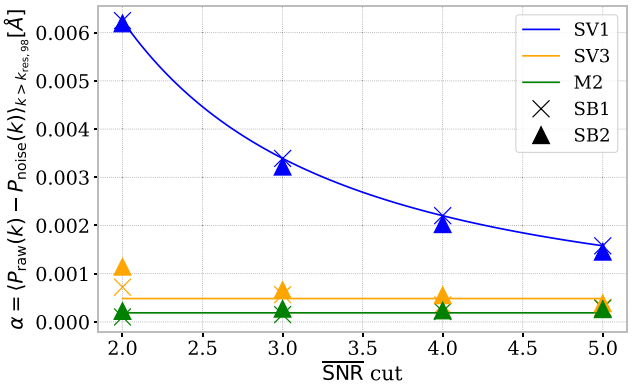

Figure C2. Asymptotic differences α for SV1(blue), SV3(yellow), and M2(green) data sets in the side-band regions SB1 (crosses) and SB2 (triangles), for the pipeline noise. The continuous lines are fits of the α values, whose parameters are given in Table 3.

APPENDIX D: COMPARISON WITH EBOSS MEASUREMENT ON LARGE SCALES

We performed a series of tests to investigate the discrepancy at large scales ($k < 0.01$ km s^{-1}) between our measurement and the eBOSS measurement in Chabanier et al. (2019) as seen in Fig. 19 (left).

We first focused on reproducing the eBOSS measurement with the picca software used in our analysis and the eBOSS parameters. Starting from SDSS spectra, we applied the pipeline used in Chabanier et al. (2019), and successively replaced each step (continuum fitting in Section 3.1, Fourier transform and averaging in Section 3.2) by the new picca software. We assessed that the version used in our analysis could reproduce the eBOSS measurement without noticeable bias at all scales.

Compared to the eBOSS measurement, some parameters are changed for the continuum fitting presented in Section 3.1. For eBOSS, this pipeline step was performed separately on subforests instead of the total Ly α forest. We checked that performing our continuum fitting on subforest does not modify the large-scales level of $P_{1D,\alpha}$. Additionally, we have performed the following changes in the continuum analysis. We removed the smoothing of the common continuum in equation (7), which was not used in Chabanier et al. (2019). We changed the polynomial order in equation (2) to zero as in eBOSS. Additionally, we tested to modify parameters of the continuum fitting procedure, which were used as eBOSS but could potentially change the level of $P_{1D,\alpha}$ at large scales. We also applied non-constant weights in equation (4), removed the forcing to zero of the Ly α contrast stack, or changed the observed wavelength to a smaller range. The conclusion of those continuum fitting tests is that none of those effects could be responsible for the ~ 15 per cent discrepancy visible on the eBOSS comparison. Furthermore, the corrections we derived from the mocks in Section 5 are different than the eBOSS corrections and these differences could explain the disagreement between the two measurements. To check the impact of corrections, we computed an eBOSS measurement without any corrections, and we realized the comparison in Fig. 19 (left), adding the corrections successively. This test yields that part of the disagreement (between 4 and 6 per cent) is due to the continuum fitting correction, which is different from eBOSS (see Fig. 7 in Chabanier et al. 2019).

Another effect that could impact the largest scale is the possible impurity and incompleteness of the DLA catalog. In order to eliminate effects due to differences in the DLA catalogue, we created a common set of quasars and DLAs by merging eBOSS and DESI catalogues. We found that the results from this common set of quasars are not significantly different, which indicates that missing DLAs cannot fully explain the disagreement. Finally, we varied the DLA and BAL catalogue used by varying confidence levels, or column density N_{HI} to include sub-DLAs. Masking and correcting the damping wings for $\log(N_{HI}) < 20.3$ systems decreases the discrepancy only for the very first wavenumber bins ($k < 0.004$ km s^{-1}).

To conclude, the large-scale disagreement between our measurement and eBOSS (Chabanier et al. 2019) cannot be fully explained for the moment. The continuum correction is responsible for a portion of this discrepancy. A complete study on DLA completeness or a comparison at the spectrum level between eBOSS and DESI will be performed in future studies to investigate in detail this discrepancy.

[1]*Aix Marseille Univ, CNRS/IN2P3, CPPM, F-13288 Marseille, France*
[2]*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[3]*Excellence Cluster ORIGINS, Boltzmannstrasse 2, D-85748 Garching, Germany*
[4]*University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, D-81677 München, Germany*
[5]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[6]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[7]*Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, USA*
[8]*The Ohio State University, Columbus, 43210 OH, USA*
[9]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[10]*Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA*
[11]*Departamento de Física, Universidad de Guanajuato – DCI, C.P. 37150, Leon, Guanajuato, México*
[12]*Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK*
[13]*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, E-08193 Bellaterra (Barcelona), Spain*
[14]*Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA*
[15]*Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México*
[16]*Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia*
[17]*Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia*
[18]*Consejo Nacional de Ciencia y Tecnología, Av. Insurgentes Sur 1582. Colonia Crédito Constructor, Del. Benito Juárez C.P. 03940, México D.F. México*
[19]*Kavli Institute for Cosmology, Department of Physics, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*
[20]*Department of Physics, Southern Methodist University, 3215 Daniel Ave., Dallas, TX 75205, USA*
[21]*Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France*
[22]*NSF's NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[23]*Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, E-08010 Barcelona, Spain*
[24]*Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA*
[25]*Department of Physics & Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[26]*Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA*

[27]*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing 100012, P.R. China*
[28]*Instituto Avanzado de Cosmología A. C. San Marcos 11 – Atenas 202. Magdalena Contreras, 10720. Ciudad de México, México*
[29]*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[30]*Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
[31]*Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[32]*Aix Marseille Univ, CNRS, CNES, LAM, F-13013 Marseille, France*
[33]*Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA*
[34]*University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA*
[35]*Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*
[36]*Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea*
[37]*CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*
[38]*Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA*
[39]*University of Michigan, Ann Arbor, MI 48109, USA*
[40]*Department of Physics & Astronomy, Ohio University, Athens, OH 45701, USA*
[41]*Instituto de Astrofísica de Canarias, C/ Vía Láctea, s/n, E-38205 La Laguna, Tenerife, Spain*
[42]*Universidad de La Laguna, Dept. de Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[43]*Department of Astronomy, Tsinghua University, 30 Shuangqing Road, Haidian District, Beijing, 100190, China*

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