Lawrence Berkeley National Laboratory

Title

The 3D Lyman-α forest power spectrum from eBOSS DR16

Permalink

Journal

Monthly Notices of the Royal Astronomical Society, 533(3)

ISSN

0035-8711

Authors

de Belsunce, Roger
Philcox, Oliver HE
Iršič, Vid
et al.

Publication Date

2024-08-29

DOI

10.1093/mnras/stae2035

Copyright Information

Peer reviewed

The 3D Lyman-α forest power spectrum from eBOSS DR16

Roger de Belsunce [1,2]★ Oliver H. E. Philcox [3,4] Vid Iršič,[5,6] Patrick McDonald,[1] Julien Guy[1] and Nathalie Palanque-Delabrouille[1]

[1]*Physics Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley CA 94720, USA*
[2]*Berkeley Center for Cosmological Physics, Department of Physics, University of California, Berkeley, CA 94720, USA*
[3]*Department of Physics, Columbia University, New York, NY 10027, USA*
[4]*Simons Society of Fellows, Simons Foundation, New York, NY 10010, USA*
[5]*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 OHA, UK*
[6]*Cavendish Laboratory, University of Cambridge, 19 J. J. Thomson Ave., Cambridge CB3 0HE, UK*

ABSTRACT

We measure the three-dimensional power spectrum (P3D) of the transmitted flux in the Lyman-α (Ly α) forest using the complete extended Baryon Oscillation Spectroscopic Survey data release 16 (eBOSS DR16). This sample consists of $\sim$205 000 quasar spectra in the redshift range $2 \leq z \leq 4$ at an effective redshift $z = 2.334$. We propose a pair-count spectral estimator in configuration space, weighting each pair by $\exp(i\mathbf{k} \cdot \mathbf{r})$, for wave vector $\mathbf{k}$ and pixel pair separation $\mathbf{r}$, effectively measuring the anisotropic power spectrum without the need for fast Fourier transforms. This accounts for the window matrix in a tractable way, avoiding artefacts found in Fourier-transform based power spectrum estimators due to the sparse sampling transverse to the line of sight of Ly α skewers. We extensively test our pipeline on two sets of mocks: (i) idealized Gaussian random fields with a sparse sampling of Ly α skewers, and (ii) log-normal `LyaCoLoRe` mocks including realistic noise levels, the eBOSS survey geometry and contaminants. On eBOSS DR16 data, the Kaiser formula with a non-linear correction term obtained from hydrodynamic simulations yields a good fit to the power spectrum data in the range $(0.02 \leq k \leq 0.35) \, h \, \mathrm{Mpc}^{-1}$ at the 1–2σ level with a covariance matrix derived from `LyaCoLoRe` mocks. We demonstrate a promising new approach for full-shape cosmological analyses of Ly α forest data from cosmological surveys such as eBOSS, the currently observing Dark Energy Spectroscopic Instrument and future surveys such as the Prime Focus Spectrograph, WEAVE-QSO, and 4MOST.

Key words: methods: statistical – galaxies: statistics – Cosmology: large-scale structure of Universe – Cosmology: theory.

1 INTRODUCTION

Light is absorbed by neutral hydrogen in the low density, highly ionized intergalactic medium (IGM), producing a series of characteristic absorption features in quasar spectra, called the Lyman-α (Ly α) forest. These absorption patterns are a measure of the distribution of neutral hydrogen along the line of sight. They represent a fundamental probe of large-scale structure at Mpc scales and below at high-redshifts ($2 \lesssim z \lesssim 5$) using ground-based observations. Given the sparseness of the Ly α forest[1], the two-point correlation function (Slosar et al. 2011) and the one-dimensional power spectrum (P1D) along the line of sight (McDonald et al. 2006) have been widely adopted in cosmological analyses of Ly α data (e.g. Busca et al. 2013; Palanque-Delabrouille et al. 2013; Slosar et al. 2013; du Mas des Bourboux et al. 2020).

The two-point correlation function (2PCF) can efficiently isolate effects at specific separations, e.g. the baryonic acoustic oscillation (BAO) feature, whilst the power spectrum is a more convenient statistic for measuring slowly varying effects down to small wavenumbers and provides less correlated errors (Font-Ribera, McDonald & Slosar 2018). In this work, we propose a weighted pair-count estimator, developed for the analysis of small-scale galaxy clustering (Philcox & Eisenstein 2020; Philcox 2021), to measure the three-dimensional Ly α forest power spectrum (P3D). By weighting each pair of Ly α pixels by $\exp(i\mathbf{k} \cdot \mathbf{r}_{ij})$ (for wave vector $\mathbf{k}$ and pair separation $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$), we directly measure the power spectrum without the need for grid-based fast Fourier transforms (FFTs) which are strongly affected by the highly non-trivial window function specific to Ly α surveys. This method also allows for almost exact computation of the window function using the same pixel pairs, albeit at a significant computational cost of $\mathcal{O}(N_{\mathrm{pix}}^2 R_0^3)$ (for number of spectral pixels N_{pix} and maximum pair separation R_0).

Over the past decades, Ly α surveys have improved in accuracy, size, and depth, resulting in large samples of medium-resolution spectra from the extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) and the currently observing Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration 2016). Whilst the eBOSS DR16 sample consists of 205 012 quasar spectra with absorber redshifts in the range $1.96 \leq z \leq 3.93$, DESI observes

★ E-mail: rbelsunce@lbl.gov

[1]The Ly α forest is sparsely sampled transverse to the line of sight but densely sampled along the line of sight.

the sky with increased spectral resolution of $2000 \lesssim R \lesssim 5500$ (Abareshi et al. 2022) and covers a much larger fraction of the sky (14 000 deg²). Over its survey duration it will provide ~840 000 quasar spectra at $z > 2.1$, a more than fourfold increase compared to BOSS/eBOSS (Adame et al. 2024). To complement the picture, smaller samples of high-resolution measurements from, e.g. the High Resolution Echelle Spectrometer (HIRES; Vogt et al. 1994; O'Meara et al. 2021) and the UV-Visual Echelle Spectrograph (UVES; Dekker et al. 2000; Murphy et al. 2019) allow for analyses deeper into the small-scale regime to almost $10x$ larger k_{max}. A number of future surveys will also capture spectra in both high and medium resolution modes, such as the William Herschel Telescope Enhanced Area Velocity Explorer survey (WEAVE-QSO; Pieri et al. 2016), the Prime Focus Spectrograph (PFS; Greene et al. 2022), and the 4-metre Multi-Object Spectroscopic Telescope (4MOST; de Jong et al. 2019).A fast and robust estimator for the analysis of Ly α forest at all scales is essential for accessing the full potential of the combination of these data sets.

One of the key advantages of the Ly α forest is that, since the density fields are only mildly non-linear at the respective redshifts, a much wider range of scales can be used to robustly probe cosmology than with most galaxy surveys. This makes the Ly α fluctuations a particularly powerful probe of early-Universe physics when combined with tracers that are sensitive to very large scales, e.g. cosmic microwave background anisotropies. The Ly α flux density contrast, δ_F, is used to derive constraints for cosmology and astrophysics, defined through

$$\delta_F(\lambda) = \frac{f_q(\lambda)}{\overline{F}(z_{Ly\alpha})q(\lambda)} - 1 \, , \qquad (1)$$

where $f_q(\lambda)$ is the observed transmitted flux as a function of τ, the optical depth, $\overline{F}(z_{Ly\alpha})$ is the mean transmitted flux at the H I absorber redshift, and $q(\lambda)$ is the unabsorbed continuum of the background quasar. The three-dimensional power spectrum, P_{3D}, of the fluctuations in the Ly α forest, given in equation (1), is principally sensitive to the amplitude of dark matter clustering σ_8, the shape of the matter power spectrum $\Gamma = \Omega_m h$, the spectral tilt n_s, and the sum of neutrino masses $\sum m_\nu$ (see e.g. Philcox & Ivanov 2022).

In this work, we present a measurement of the anisotropic three-dimensional Ly α forest power spectrum using eBOSS DR16 Ly α forest spectra. We compare our measurement to the best-fitting results from the 2PCF analysis in the range $(0.019 \leq k \leq 0.35)\,h\,\mathrm{Mpc}^{-1}$ (du Mas des Bourboux et al. 2020, hereafter dMdB20). It includes a Kaiser formula (Kaiser 1987) with a non-linear correction term obtained from hydrodynamical simulations (Arinyo-i-Prats et al. 2015) and modelling of contaminants (e.g. metals, distortion from continuum fitting, damped Ly α absorbers) specific to Ly α forest analyses. Predictions for the Ly α forest clustering have greatly improved over the last decade, using either hydrodynamical simulations (see e.g. Arinyo-i-Prats et al. 2015; Bolton et al. 2017; Doughty et al. 2023; Puchwein et al. 2023) or perturbation theory[2] (see e.g. Givans & Hirata 2020; Chen, Vlah & White 2021; Garny et al. 2021; Givans et al. 2022; Ivanov 2023).

In the context of simulations, different P3D methods have been proposed and tested, see McDonald (2003), Arinyo-i-Prats et al. (2015), Font-Ribera et al. (2018), and Horowitz, de Belsunce & Lukic (2024). The latter two methods provide close-to-optimal ways of

constructing a covariance matrix. Whilst we measure the anisotropic clustering in multipole space and as a function of $k = \sqrt{k_\parallel^2 + k_\perp^2}$, Font-Ribera et al. (2018) measure the P3D in the (more natural basis) $\{k_\parallel, \mathbf{k}_\perp\}$-basis (which is the Fourier space equivalent to the $\{r_\parallel, \mathbf{r}_\perp\}$-basis of the configuration space analysis in, e.g. dMdB20). Whilst the information content is (in principle) the same as in the present analysis, the $\{k_\parallel, \mathbf{k}_\perp\}$-basis allows to isolate (and marginalize out) modes that are, e.g. sensitive to distortions in the continuum fitting ($k_\parallel = 0$) (see Font-Ribera et al. 2018, for a discussion). In this work, we compute a simulation-based covariance matrix, whereas Font-Ribera et al. (2018) compute an approximate version of the global covariance matrix within the optimal quadratic estimator framework. A recent proof-of-principle presented in Karim et al. (2023) presents a small-scale measurement on simulations and eBOSS DR16 data which measures the cross-spectrum $P_\times(z, \theta, k_\parallel)$, originally developed in Hui, Stebbins & Burles (1999) and Font-Ribera et al. (2018).

The Ly α forest is a treasure trove of cosmological information, capable of measuring the expansion history of the Universe through BAO (McDonald & Eisenstein 2007; Busca et al. 2013; Slosar et al. 2013; dMdB20), the broad-band shape of the (large scale) three-dimensional Ly α correlation function (Slosar et al. 2013; Cuceu et al. 2021, 2023; Gordon et al. 2023a), neutrino masses (Seljak et al. 2005; Viel, Haehnelt & Springel 2010; Palanque-Delabrouille et al. 2013, 2020), and primordial black holes (Afshordi, McDonald & Spergel 2003; Murgia et al. 2019). The P3D is connected to the well-understood P1D by performing an integration perpendicular to the line of sight, which is sensitive to small scales (see e.g. Seljak et al. 2005; Viel et al. 2005; McDonald et al. 2006; Palanque-Delabrouille et al. 2013; Chabanier et al. 2019; Pedersen et al. 2021; Ravoux et al. 2023; Karaçaylı et al. 2024). At the smallest of scales, the Ly α absorption features provide means to test dark matter models (Viel et al. 2013; Baur et al. 2016; Armengaud et al. 2017; Iršič et al. 2017, 2023; Kobayashi et al. 2017; Murgia, Iršič & Viel 2018; Garzilli et al. 2019; Iršič, Xiao & McQuinn 2020; Rogers, Dvorkin & Peiris 2022; Villasenor et al. 2023), early dark energy models (Goldstein et al. 2023), and thermal properties of the ionized (cold) IGM (Zaldarriaga 2002; Viel et al. 2006; Bolton et al. 2008; Meiksin 2009; Garzilli et al. 2012; McQuinn 2016; Boera et al. 2019; Gaikwad et al. 2019, 2021; Walther et al. 2019; Villasenor et al. 2022; Wilson, Iršič & McQuinn 2022).

The remainder of this paper is organized as follows: we present the P3D estimator in Section 2, before discussing the theoretical modelling of the Ly α flux power spectrum in Section 3 and the forward modelling of the window function in Section 3.1. In Section 4, we describe the observational eBOSS DR16 Ly α forest data set as well as the employed synthetic Ly α spectra used to model covariances. In Section 5, we extensively test our pipeline on mocks before presenting the main result of our analysis in Section 6: the eBOSS DR16 P3D of the Ly α forest. Section 7 presents our conclusions. We make our HIPSTER-LYA implementation publicly available.[3]

2 THE THREE-DIMENSIONAL POWER SPECTRUM ESTIMATOR

In cosmological data analysis, one typically employs data compression, both to reduce the size of the data vector to a computationally manageable level and average over stochastic fluctuations. The

---

[2]On analyses of the connection between the physics of the Ly α forest and perturbative approaches, i.e. response function approaches, see Seljak (2012), Cieplak & Slosar (2016), and Iršič & McQuinn (2018).

[3]https://github.com/oliverphilcox/HIPSTER

particular compression statistic is specific to the problem at hand; for Ly α, one typically uses the 2PCF or the P1D along the line of sight (e.g. Busca et al. 2013; Palanque-Delabrouille et al. 2013; Slosar et al. 2013; Ravoux et al. 2023; Karaçaylı et al. 2024). Here, we compress the large number of Ly α forest spectra instead into the P3D,[4] paving the way for future full-shape cosmological analyses in Fourier space of eBOSS data (and beyond).[5]

Usually, Ly α summary statistics are computed using quadratic maximum likelihood (e.g. McDonald et al. 2006; Karaçaylı, Font-Ribera & Padmanabhan 2020) or FFT-based algorithms (e.g. Palanque-Delabrouille et al. 2013; Chabanier et al. 2019); in this work, we instead make use of configuration space pair-count estimators, following the HIPSTER algorithm of (Philcox & Eisenstein 2020; Philcox 2021). These were originally introduced in the context of the small-scale galaxy power spectrum and bispectrum, and translate well to our problem, given the sparse (dense) sampling of pixels transverse to (along) the line of sight, which typically leads to aliasing for FFT-based methods. As we discuss below, HIPSTER estimates spectra by summing pairs of points in configuration space, weighted by $\exp(i\mathbf{k} \cdot \mathbf{r}_{ij})$, for wave vector $\mathbf{k}$ and pair separation by $\mathbf{r}_{ij} \equiv \mathbf{r}_i - \mathbf{r}_j$.

Given some data field $D(\mathbf{x})$ (encoding Ly α fluctuations or galaxy overdensities, with some weighting), the (pseudo-)power spectrum is defined as a Fourier transform of the two-point function DD:

$$P(\mathbf{k}) \equiv \mathcal{F}[DD(\mathbf{r})] = \int \mathrm{d}^3\mathbf{r}\, DD(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} , \qquad (2)$$

where $\mathcal{F}$ is the Fourier transform operator, and we have assumed DD to be suitably normalized. Usually, one considers angularly averaged power spectra, defined by[6]

$$P^a = \frac{1}{V_a} \int_{\mathbf{k}} \Theta^a(|\mathbf{k}|)P(\mathbf{k}) , \qquad (3)$$

where we have restricted to a k-bin a, with volume $V_a = \int_{\mathbf{k}} \Theta^a(|\mathbf{k}|)$, for top-hat function $\Theta^a(k)$. Comparing to equation (2), we see that the $\mathbf{k}$-dependence is captured through the kernel

$$A^a(\mathbf{r}_i - \mathbf{r}_j) \equiv \frac{1}{V_a} \int_{\mathbf{k}} \Theta^a(k)e^{i\mathbf{k}\cdot(\mathbf{r}_i - \mathbf{r}_j)} \approx j_0(k_a|\mathbf{r}_i - \mathbf{r}_j|) , \qquad (4)$$

where j_0 is the 0th order spherical Bessel function. In the last step, we assumed the thin-shell limit and aligned the $|\hat{\mathbf{k}}|$-axis with $|\mathbf{r}_i - \mathbf{r}_j|$. Noting that $D(\mathbf{x}) = n(\mathbf{x})w(\mathbf{x})\delta(\mathbf{x})$ for background density n, weights w, and overdensity δ (which is the quantity of interest), the power spectrum can be written explicitly as

$$P^a = \frac{1}{V(nw)^2} \int_{\mathbf{r}_i} \int_{\mathbf{r}_j} n(\mathbf{r}_i)n(\mathbf{r}_j)w(\mathbf{r}_i)w(\mathbf{r}_j)\delta(\mathbf{r}_i)\delta(\mathbf{r}_j)$$
$$A^a(\mathbf{r}_i - \mathbf{r}_j) , \qquad (5)$$

adding a normalization term appropriate to an ideal uniform survey of volume V. In this approach, the power spectrum can be computed by explicitly evaluating equation (5), i.e. counting each pair of Ly α pixels, weighted by above kernel.

In practice, we also wish to estimate the anisotropy of the Ly α power spectrum. This is possible by a simple generalization for multipole ℓ:

$$P_\ell^a = \frac{1}{V(nw)^2} \int_{\mathbf{r}_i} \int_{\mathbf{r}_j} n(\mathbf{r}_i)n(\mathbf{r}_j)w(\mathbf{r}_i)w(\mathbf{r}_j)\delta(\mathbf{r}_i)\delta(\mathbf{r}_j) \times A_\ell^a(\mathbf{r}_i, \mathbf{r}_j)$$

$$A_\ell^a(\mathbf{r}_i, \mathbf{r}_j) \equiv \frac{2\ell + 1}{V_a} \int_{\mathbf{k}} \Theta^a(|\mathbf{k}|)e^{i\mathbf{k}\cdot(\mathbf{r}_i - \mathbf{r}_j)}\mathcal{L}_\ell(\angle[\mathbf{k}, \frac{1}{2}(\mathbf{r}_i + \mathbf{r}_j)]) , \qquad (6)$$

where $\angle[\mathbf{k}, \frac{1}{2}(\mathbf{r}_i + \mathbf{r}_j)]$ denotes the cosine of the angle between $\mathbf{k}$ and the local line of sight $(\mathbf{r}_i + \mathbf{r}_j)/2$. Equation (6) is the basis of the P3D estimator used herein.[7]

In practice, all pairs of Ly α pixels contribute to the measured power spectrum, which results in the naïve estimator of equation (6) being extremely expensive to compute. In practice however, well-separated pairs have negligible contributions to the power spectrum, thus we can truncate the pair counts at separations greater than $R_0 \geq 200\,h^{-1}\,\mathrm{Mpc}$, with minimal loss of information (Philcox & Eisenstein 2020). This is done via a smooth polynomial window function

$$W\left(\frac{|\mathbf{r}|}{R_0}\right) \equiv \begin{cases} 1, & 0 \leq \frac{r}{R_0} < \frac{1}{2} \\ 1 - 8(2\frac{r}{R_0} - 1)^3 + 8(2\frac{r}{R_0} - 1)^4, & \frac{1}{2} \leq \frac{r}{R_0} < \frac{3}{4} \\ -64(\frac{r}{R_0} - 1)^3 - 128(\frac{r}{R_0} - 1)^4, & \frac{3}{4} \leq \frac{r}{R_0} < 1 \\ 0, & \text{else,} \end{cases} \qquad (7)$$

which modifies the anisotropic power spectrum estimator in equation (6) by a factor $W((\mathbf{r}_i - \mathbf{r}_j)/R_0)$. Though this necessarily induces slight distortions to the measured spectra, its effect can be straight-forwardly included in the theoretical model for P_ℓ^a, thus avoiding any potential bias.

Finally, we must consider the window function of the data, denoted by $\Phi(\mathbf{r}) \equiv DD(\mathbf{r})/\xi(\mathbf{r})$, where $\xi(\mathbf{r}) = \langle\delta(\mathbf{x})\delta(\mathbf{x} + \mathbf{r})\rangle$, is the true 2PCF of the data. Due to the sampling of the Ly α skewers, this has a non-trivial form, and leads to our estimators returning a window-convolved power spectrum measurement. Explicitly, this effect can be included by forward modelling the window function in the theoretical model, using

$$\tilde{P}(\mathbf{k}) = \mathcal{F}[\xi(\mathbf{r})\Phi(\mathbf{r})W(r; R_0)](\mathbf{k}) , \qquad (8)$$

where $\xi(\mathbf{r})$ is the correlation function model, and $\tilde{P}(\mathbf{k})$ is the output window- (and pair-separation-) convolved power. This procedure will be discussed in more detail in Section 3.1.

In discrete form, the full estimator for the Ly α power spectrum used in this work is given by a sum over each pair of pixels, i, j:

$$\hat{P}_\ell^a = \frac{1}{V(nw)^2} \sum_i \sum_{j, i \neq j} w_i w_j \delta_{F,i}\delta_{F,j} A_\ell^a(\mathbf{r}_i, \mathbf{r}_j)W(|\mathbf{r}_i - \mathbf{r}_j|; R_0) , \qquad (9)$$

where we drop the (arbitrary) normalization factor and additionally remove the 'self-skewer' counts; an advantage of doing the measurement in configuration space to avoid (correlated) uncertainties stemming from continuum estimation. The weights w are obtained from the continuum fitting (dMdB20), and $\delta_{F,i}$ are the estimated Ly α

---

fluctuations. We use the same set of pixels to compute the window Φ (needed to forward model geometry effects):

$$\Phi_\ell^b = \frac{1}{V(nw)^2} \frac{2\ell+1}{V_b} \sum_i \sum_{j,\, i \neq j} w_i w_j \mathcal{L}_\ell(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{r}}_j) \Theta^b(|\mathbf{r}_i - \mathbf{r}_j|) \,, \quad (10)$$

restricting to radial bin b and Legendre multipole ℓ. The measurement of the multipoles of the power spectrum and corresponding window functions for a data set with $\sim 10^8$ pixels (similar to the data set used in this work) takes ~ 5 h on the Perlmutter computer at the National Energy Research Scientific Computing Center (NERSC) using one AMD Milan CPU.

3 THEORY POWER SPECTRUM MODELLING

One of the key advantages of the Lyα forest is that it probes the Universe at redshifts $2 \leq z \leq 5$, i.e. with many more linear modes than seen in galaxy surveys. For our theory modelling, we use the same procedure to the one outlined in dMdB20 and briefly summarize it below.[8] The real-space quasi-linear power spectrum[9], $P_{QL}(k, \mu, z)$, is modulated by a simulation-based non-linear fitting function, F_{NL}, and connected to the redshift-space flux power through the well-known Kaiser formula (Kaiser 1987)

$$P_F(k, \mu, z) = b_F(z)^2 \left(1 + \beta_F \mu^2\right)^2 F_{NL}(k, \mu) P_{QL}(k, z) \,, \quad (11)$$

where b_F denotes the redshift dependent linear bias parameter, β_F is the redshift space distortion (RSD) parameter (which is assumed to be redshift independent) and μ is the angle of $\mathbf{k} = \{k_\parallel, \mathbf{k}_\perp\}$ to the line of sight, $\mu \equiv k_\parallel/k$. The redshift evolution of the Lyα forest bias parameter enters the weights, w, which are taken from equation (7) in dMdB20, and enters the power spectrum measurement through equations (9) and (10). Following McDonald et al. (2006), dMdB20, and Gordon et al. 2023b, the product of Lyα forest bias and growth factor is assumed to have a redshift dependence described by $(1 + z)^{\gamma_{Ly\alpha}-1}$ where $\gamma_{Ly\alpha} = 2.9$ is the redshift evolution parameter of the Lyα forest bias. Over sufficiently large scales the Lyα forest fluctuations are, to linear order, given by $\delta_F = b_{\delta,F}\delta + b_{\eta,F}\eta$ where δ are mass fluctuations and $\eta = -(\partial v_p/\partial x_p)/aH$ is the gradient of the peculiar velocity v_p over the comoving line-of-sight coordinate x_p with the scale factor a and the Hubble constant H. As such, the bias factors are the partial derivatives with respect to δ and η (see e.g. Arinyo-i-Prats et al. 2015). The RSD parameter for Lyα forest surveys is

$$\beta_F = \frac{f b_{\eta,F}}{b_{\delta,F}} \,, \quad (12)$$

for our fiducial cosmology we obtain $f = 0.9704$. We can estimate b_η and b_δ from simulations, by either fitting the (i) autopower spectrum using the Kaiser formula (e.g. Arinyo-i-Prats et al. 2015); (ii) cross-power of Lyα and matter (e.g. Givans et al. 2022); or (iii) from separate Universe simulations (e.g. Cieplak & Slosar 2016). For the Lyα forest this results in an RSD parameter greater than unity, yielding a quadrupole that is larger than the monopole (in contrast to the case for galaxies). To account for the broadening in $k_\parallel$ stemming from high column density (HCD) systems (see e.g. Rogers et al.

2018, and references therein), the Lyα bias parameters (b, β) are remapped to

$$b_F' = b_F + b_{HCD} F_{HCD}(k_\parallel) \,, \quad (13)$$

$$b_F' \beta_F' = b_F \beta_F + b_{HCD} \beta_{HCD} F_{HCD}(k_\parallel) \,, \quad (14)$$

where $F_{HCD}(k_\parallel) = \exp(-L_{HCD} k_\parallel)$ is fitted to hydrodynamical simulations and L_{HCD} is the typical scale for unmasked HCDs, set to $10\, h^{-1}$ Mpc in dMdB20.

The non-linear (NL) correction to the power spectrum is obtained from fits to hydrodynamical simulations (Arinyo-i-Prats et al. 2015)

$$F_{NL}(k, \mu) = \exp\left(q_1 \Delta^2(k) \left[1 - \left(\frac{k}{k_\nu}\right)^{\alpha_\nu} \mu^{b_\nu}\right] - \left(\frac{k}{k_p}\right)^2\right) \,, \quad (15)$$

with the usual dimensionless power spectrum $\Delta^2(k) \equiv k^3 P_{lin}(k)/(2\pi^2)$. The parameter q_1 is a dimensionless parameter and controls the importance of the non-linear enhancement.[10] While q_1 is only (very) weakly dependent on redshift, it is, together with α_ν and $k_\nu^{\alpha_\nu}$, inversely proportional to the amplitude of the linear power spectrum. The remaining parameters model the Jeans smoothing and line-of-sight broadening stemming from non-linear peculiar velocities and thermal broadening (see McDonald 2003; Arinyo-i-Prats et al. 2015). We set the parameters to $q_1 = 0.8558$, $k_\nu = 1.11454\, h^{-1}$ Mpc , $\alpha_\nu = 0.5378$, $b_\nu = 1.607$, and $k_p = 19.47\, h^{-1}$ Mpc , following dMdB20.

The quasi-linear power spectrum is given by

$$P_{QL}(k, z) = P_{sm}(k, z) + \exp\left[-\frac{k_\parallel^2 \Sigma_\parallel^2 + k_\perp^2 \Sigma_\perp^2}{2}\right] P_{peak}(k, z) \,, \quad (16)$$

which decomposes the power spectrum into a smooth (no BAO feature) and a peak (isolating the BAO feature) component. This description includes a correction for the non-linear broadening of the BAO peak denoted by $\Sigma_\parallel/\Sigma_\perp = 1 + f$ where f is the (linear) growth rate (see e.g. Eisenstein, Seo & White 2007). Analogously to dMdB20, we set the smoothing parameters to $\Sigma_\perp = 3.26\, h^{-1}$ Mpc and $\Sigma_\perp = 6.42\, h^{-1}$ Mpc for a growth rate of $f \approx 0.97$.

Contaminants in the Lyα forest, such as absorption by metals and correlations due to the sky–subtraction procedure of the eBOSS data pipeline, are modelled as additive terms in the Lyα autocorrelation function (see section 4 in dMdB20, for a detailed discussion). The computation of comoving separations is based on the (erroneous) assumption that all absorption stems solely from the H I Lyα transition. Transitions from other elements (e.g. silicon and carbon often denoted as metals) contribute to the measured Lyα forest fluctuations as well. Thus, the measured absorption at a given wavelength is a mixture of absorption at different redshifts (from different transitions). For each pair of possible transitions the offset between the true and assumed redshift is computed. The resulting re-mapping matrix between the true and assumed comoving separation is called the metal matrix, see e.g., Blomqvist et al. (2018) and dMdB20, relating the measured to the 'true' summary statistic in configuration space, e.g. the correlation function.

To estimate the window-convolved power spectra, we will require the correlation function multipoles: these are defined as an angular

---

[8]Note that we do not include the $G(\mathbf{k})$ in equation (27) of dMdB20 since we do not grid the data for our power spectrum computation.

[9]The linear input power spectrum is computed using the Boltzmann solver CAMB (https://camb.info/) as part of publicly available VEGA package (https://github.com/andreicuceu/vega/).

[10]The parameter q_1 determines the impact of non-linearities, and is related to the parameter k_{NL} in McDonald (2003).

integral over the full correlation function ξ,

$$\xi_\ell(r) \equiv \frac{2\ell+1}{2} \int\limits_{\mu=-1}^{+1} d\mu\, \xi(r,\mu)\mathcal{L}_\ell(\mu)\,. \tag{17}$$

As described below, these will then be used to compute the mode mixing introduced by the window matrix in equation (23), allowing accurate comparison of theory and data.

3.1 Forward modelling the window function

In the present analysis, we forward model the effect of the survey geometry, captured by the configuration-space window matrix $\Phi(\mathbf{r}) \equiv DD(\mathbf{r})/\xi(\mathbf{r})$, on the theoretically expected power spectrum instead of removing the window function from the data. This factor is independent of the Ly α fluctuations, δ_F, and can be explicitly computed by counting pairs of Ly α pixels, as in equation (10). This is analogous to the approach used in galaxy surveys e.g. Beutler et al. (2017), Castorina & White (2018), and Beutler & McDonald (2021), whence one writes

$$\Phi(\mathbf{r}) \equiv \frac{RR(\mathbf{r})}{RR_{\text{model}}(\mathbf{r})} = \frac{RR(\mathbf{r})}{V(nw)^2}\,, \tag{18}$$

for (weighted) background density $R(\mathbf{x}) = n(\mathbf{x})w(\mathbf{x})$, and an ideal bin volume $RR_{\text{model}}(\mathbf{r}) = \frac{4\pi}{3}(r_{\text{max}}^3 - r_{\text{min}}^3)$, given some set of (thin) bins in r_{ij}. In the galaxy case, this is estimated using catalogues of random particles; for us, the procedure simply involves counting the unweighted Ly α pixels.[11]

Expressed in terms of (even) Legendre multipoles, the pair-count estimator of equation (9) measures the power spectrum convolved with the functions

$$\Gamma_\ell^2(r) \equiv \Phi_\ell(r)W(r;R_0)\,, \tag{19}$$

where $W(r;R_0)$ is the pair-truncation function given in equation (7). The convolved power spectrum multipoles can be obtained by first considering the distortion induced to the 2PCF (see e.g. Beutler et al. 2017):

$$\tilde{\xi}_0(r) = \xi_0\Gamma_0^2 + \frac{1}{5}\xi_2\Gamma_2^2 + \frac{1}{9}\xi_4\Gamma_4^2 + \ldots \tag{20}$$

$$\tilde{\xi}_2(r) = \xi_0\Gamma_2^2 + \xi_2\left[\Gamma_0^2 + \frac{2}{7}\Gamma_2^2 + \frac{2}{7}\Gamma_4^2\right]$$
$$+\xi_4\left[\frac{2}{7}\Gamma_2^2 + \frac{100}{693}\Gamma_4^2 + \frac{25}{143}\Gamma_6^2\right] + \ldots \tag{21}$$

$$\tilde{\xi}_4(r) = \xi_0\Gamma_4^2 + \xi_2\left[\frac{18}{35}\Gamma_2^2 + \frac{20}{77}\Gamma_4^2 + \frac{45}{143}\Gamma_6^2\right]$$
$$+\xi_4\left[\Gamma_0^2 + \frac{20}{77}\Gamma_2^2 + \frac{162}{1001}\Gamma_4^2\right.$$
$$\left.+ \frac{20}{143}\Gamma_6^2 + \frac{490}{2431}\Gamma_8^2\right] + \ldots \tag{22}$$

From equation (8), the power spectra can then be expressed as a Hankel transform:

$$\tilde{P}_\ell^a = 4\pi i^\ell \int_0^{R_0} dr\, r^2 \tilde{\xi}_\ell(r)K_\ell^a(r)\,, \tag{23}$$

where $K_\ell^a(r)$ is the usual spherical Bessel function $j_\ell(kr)$ integrated over a k-bin to reduce oscillations:

$$K_\ell^a(r) = \frac{\int_{k_{\text{min}}^a}^{k_{\text{max}}^a} k^2\, dk\, j_\ell(kr)}{\int_{k_{\text{min}}^a}^{k_{\text{max}}^a} k^2\, dk}\,, \tag{24}$$

as in equation (3).[12] For a thin k-bin centred at k_a, $K_\ell^a(r) \approx j_\ell(k_a r)$. In this work, we truncate our expansion of the window-convolved correlation function, $\tilde{\xi}_\ell$, at $\ell_{\text{max}} = 4$ for the 2PCF. Note that we include *all* corresponding window matrix multipoles for each multipole in $\tilde{\xi}_\ell$.[13] In Appendix A, we discuss the contribution of each correlation function multipole to the final result.

4 LYMAN-α FOREST DATA FROM EBOSS DR16

In the present analysis, we compute the P3D from Ly α forest spectra from the 16th data release (DR16Q; Lyke et al. 2020) of the completed eBOSS (Dawson et al. 2016) of the fourth generation of the Sloan Digital Sky Survey (SDSS-IV; York et al. 2000). The primary scientific goal of eBOSS was to constrain cosmological parameters using BAO and RSDs (see e.g. dMdB20; Alam et al. 2021). Briefly summarized, the DR16 data consists of the complete 5-yr BOSS and 5-yr eBOSS survey. The Ly α spectra were observed with a double spectrograph mounted on the 2.5m Apache Point telescope to map 10 000 deg^2 of the sky. The observations are conducted in the observed wavelength range of $3600 < \lambda_{\text{obs}} < 10,000$ Å with a spectral resolution of $R \sim 1500$–2500 (Dawson et al. 2016). The quasar selection and algorithms of the pipeline are explained in detail in Myers et al. (2015) as well as technical details related to the survey itself in Dawson et al. (2016).

4.1 Data selection

The full sample consists of 205 012 quasar spectra with absorber redshifts in the range $1.96 \leq z \leq 3.93$. The quasar sample is split into two disjoint regions on the sky; the northern (NGC) and southern (SGC) galactic caps, with 147 392 and 57 620 sightlines, respectively. The number of spectral pixels are $34.3 \cdot 10^6$ for the entire data set with an average signal-to-noise ratio of 2.56 per pixel calculated in the Ly α forest. The forest used for cosmological analysis is defined to be $1040 \leq \lambda_{\text{rf}} \leq 1200$ Å. For ease of comparison to dMdB20, we use the same set of flux decrements, δ_F, introduced in the following section. We note that BOSS/eBOSS observed spectra in spectral pixels of with $\Delta \log_{10}(\lambda) \sim 10^{-4}$ which have been rebinned for the purpose of the final analysis on to a grid of $\Delta \log_{10}(\lambda) \sim 3 \times 10^{-4}$. Damped Ly α absorbers (DLAs), defined as regions where the transmission is reduced by more than 20 per cent in the flux decrements, are masked out. A Voigt profile is fitted to each DLA to correct for the absorption in the wings (Noterdaeme et al. 2012). Quasars with broad absorption lines (BALs) have been removed for the present analyses (see Lyke et al. 2020, for more details). Both are the main contaminants affecting the data selection.

---

[11]In full, the Ly α case can be imagined as a set of (correlated) pencil-beam surveys, whose sampling is known precisely.

[12]These integrals of the spherical Bessel function are analytic (see e.g. https://dlmf.nist.gov/).

[13]We tested that including 2PCF contributions up to $\ell_{\text{max}} = 6$ with all corresponding window matrix multipoles up to $\ell_{\text{max}} = 10$ did not affect our results.

4.2 Continuum fitting

To measure clustering statistics such as the P3D, we use the flux decrement given in equation (1). Extracting the unabsorbed flux, i.e. the quasar continuum, is a daunting task in general. In this paper, we use the publicly available continuum-fitted spectra from eBOSS DR16.[14] In the following, we briefly outline the continuum-fitting procedure from dMdB20: to each Ly α forest a slope and an amplitude is fitted to the stacked spectra in the quasar sample using the PICCA package.[15] The product in the denominator of equation (1) assumes a common continuum for all quasars in rest frame, $\overline{q}(\lambda_{\rm rf})$, corrected by a first-order polynomial for each quasar q in $\Lambda \equiv \log \lambda$ which accounts for the diversity of the quasars

$$\overline{F}(\lambda)q(\lambda) = \overline{q}(\lambda) \left(a_q + b_q \frac{\Lambda - \Lambda_{\rm min}}{\Lambda_{\rm max} - \Lambda_{\rm min}} \right), \tag{25}$$

which results in a biased mean and spectral slope of each forest flux decrement $\hat{\delta}(\lambda) = F(\lambda)/ \left[(a_q + b_q \Lambda(\lambda))\overline{q}(\lambda) \right] - 1$. The fluctuations are then given by

$$\tilde{\delta}(\lambda) = \hat{\delta}(\lambda) - \frac{1}{W_q} \sum_{\lambda'} w(\lambda')\hat{\delta}(\lambda') \left[1 + \frac{\Lambda(\lambda)\Lambda(\lambda')}{\overline{\Lambda^2(\lambda)}} \right], \tag{26}$$

with $W_q = \sum_{\lambda'} w(\lambda')$ and $\Lambda(\lambda) = \log \lambda - \overline{\log \lambda}$. We use this quantity to measure the power spectrum. Estimating the continuum directly from the forest distorts the field, requiring correction at the level of the flux decrement (see e.g. dMdB20). The continuum fitting and projection to centre the mean flux decrement at zero suppresses modes along the line of sight in the forest. A good approximation is to treat these distortions as linear and model them in the theoretically expected correlation function through a distortion matrix (DM; dMdB20)

$$\xi_{\rm A}^{\rm DM} = \sum_{\rm A'} D_{\rm AA'}\xi_{\rm A'}, \tag{27}$$

where A and A' denote two bins of the correlation function

$$D_{\rm AA'} = \frac{1}{\sum_{\lambda} w(\lambda)} \sum_{(i,j)\in A} w_i w_j \sum_{(i',j')\in A'} \eta_{ii'}\eta_{jj'}, \tag{28}$$

with

$$\eta_{ii'} \equiv \delta_{ii'}^{\rm K} - \frac{w_j}{\sum_k w_k} - \frac{w_j \Lambda_i \Lambda_j}{\sum_k w_k \Lambda_k^2}, \tag{29}$$

for forest pixel pairs (i, j) and (i', j') in sightlines A and A'. Whilst this approach marginalizes out large-scale modes, it lends itself well to measure the small-scale continuum for the P1D. In the current implementation, we apply the distortion matrix to the theory correlation function in $\{r_{\parallel}, \mathbf{r}_{\perp}\}$−space.[16] (We denote the theory correlation function to be 'distorted' after applying the distortion matrix.) The distorted multipoles are then used to compute the window convolved theory power spectrum in equation (23). See Appendix B for the effect of the distortion matrix on theory power spectra and correlation function multipoles.

5 CONSISTENCY TESTS ON SIMULATIONS

Before presenting the main results of this paper, we first test our pipeline by applying the power spectrum estimator, presented in

Section 2, to idealized Gaussian realizations, with results displayed in Section 5.1. In Section 5.2, we apply our pipeline to realistic eBOSS DR16 mocks including contaminants (such as HCD absorbers and metals). In particular, we discuss the effect of continuum fitting and the extraction of the unabsorbed flux, on the resulting power spectrum measurements.

5.1 Gaussian random field simulations

The first series of tests of the power spectrum estimator presented in Section 2 is on anisotropic Gaussian random fields (GRFs), generated with a known linear matter power spectrum $P_{\rm lin}(k)$. The aim is to investigate biases as well as the range of validity and robustness of the estimator to non-trivial survey geometries with a known input matter power spectrum.[17] The GRFs are realized in a periodic box of size $L = 1380\,h^{-1}$ Mpc with $N = 512$ cells at redshift $z = 2.4$ with a Λ cold dark matter (CDM) linear power spectrum as theory input (Planck Collaboration VI 2020). The fundamental mode is $k_F = 2\pi/L = 0.0045\,h\,{\rm Mpc}^{-1}$ with the Nyquist frequency being $k_{\rm Ny} = 1.17\,h\,{\rm Mpc}^{-1}$.

We sample $N_s = 900$ sightlines from the box, corresponding to a (very sparse) quasar density of ∼2–3 qso deg^{-2}, and 460 800 Ly α pixels.[18] For the Ly α forest fluctuations, δ, we use the pixel value, i.e. the modes from the GRF, and model the selection function using weights of unity. For the power spectrum measurement, we do not include pixels in the same skewer, as in the main Ly α analysis. Additionally, we apply a non-trivial selection function, i.e. with edges and sparse sampling transverse to the line of sight to measure the effect of the window function on the measured power spectrum. To this end, we remove the periodicity of the box and add a 10 per cent root-mean-squared error, relative to the mean, Gaussian, white (uncorrelated), noise to each of the pixels in the box, corresponding to somewhat realistic Ly α noise levels for continuum fitting and pipeline noise.

We measure the monopole, quadrupole, and hexadecapole power spectra from the GRFs and compute the 2PCF contributions up to $\ell_{\rm max} = 4$ with all corresponding window matrix multipoles. We use 30 equidistant k-bins in the range $(0.01 \leq k \leq 0.50)\,h\,{\rm Mpc}^{-1}$ and $N_r = 1,000$ linearly spaced bins in the range $(0 \leq r \leq R_0 = 200)\,h^{-1}$ Mpc for the window matrix, respectively.[19] We compare the measured power spectrum multipoles to the theory input power spectrum in Fig. 1 and recover the input spectrum well within the 1σ error bars up to $k \lesssim 0.5\,h\,{\rm Mpc}^{-1}$ with the measured anisotropic power spectra. Note that the quadrupole is larger than the monopole due to the large RSD parameter $\beta_F = 1.5$, typical for Ly α surveys. Note that the pair separation window $W(r; R_0)$ suppresses power on scales larger than the ones defined by $k = 2\pi/R_0$ which is at $k = 0.031\,h\,{\rm Mpc}^{-1}$ for our chosen value of R_0. Thus, we discard k-values below the cut-off in our analysis pipeline. In Appendix A, we discuss the effect of the number of skewers on the multipoles of the window matrix itself as well as their convolution with the window convolved theory power spectra.[20]

---

[14]Publicly available at https://data.sdss.org/sas/dr16/.

[15]Publicly available at https://github.com/igmhub/picca.

[16]Analogous to dMdB20 we measure the distortion matrix in bins of $4\,h^{-1}$ Mpc in the range $0 \leq \{r_{\parallel}, r_{\perp}\} \leq 200\,h^{-1}$ Mpc.

[17]In the following section, we will discuss a series of tests on more realistic mocks.

[18]We performed the same test using a DESI-like quasar density of ∼ 30 qso deg^{-2} and recover the input power spectrum to per-cent-level precision over the same k-range.

[19]We tested that our results are robust to the choice of N_r and the $\ell_{\rm max}$ of the 2PCF (and the corresponding window matrix multipole).

[20]Note that we use the terms 'window function' and 'window matrix' interchangeably.
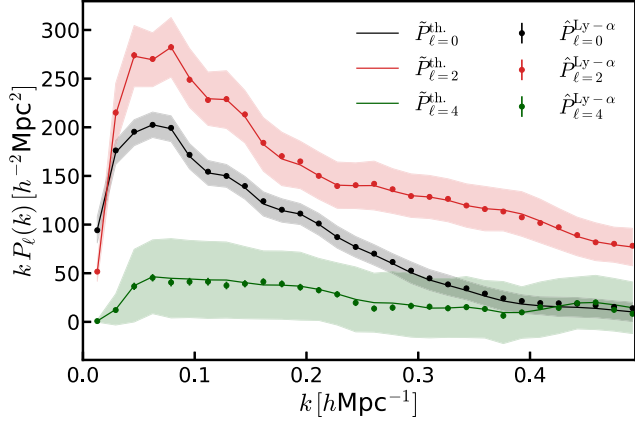
Figure 1. Consistency test of the power spectrum estimator, comparing the window-convolved linear power spectrum at $z = 2.4$ (monopole in black, quadrupole in red, hexadecapole in green, respectively) to the mean power spectrum of $N = 100$ Gaussian random field simulations (dots), computed using pair counts truncated $R_0 = 200\,h^{-1}$ Mpc . The shaded regions give the square root of the diagonal of the covariance matrix and illustrate the variance between the $N = 100$ realizations; this is also shown by the error-bars, which include the factor of $1/\sqrt{N} = 1/10$. The 2PCF contributions are computed up to $\ell_{\max} = 4$ with all corresponding window matrix multipoles, as described in Section 3.

The error bars shown in Fig. 1 capture the variation between differing GRF realizations and are obtained from the diagonal of the covariance matrix computed from $N = 100$ realizations of the GRFs. For Legendre multipoles ℓ_1, ℓ_2 of the anisotropic power in k-bins m, n, this is given by

$$\mathbf{C}_{\ell_1 \ell_2}^{mn} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\hat{P}_{\ell_1,i}^m - \bar{P}_{\ell_1}^m \right) \left(\hat{P}_{\ell_2,i}^n - \bar{P}_{\ell_2}^n \right), \tag{30}$$

with the mean of the power spectra given by

$$\bar{P}_\ell^m = \frac{1}{N} \sum_{i=1}^{N} \hat{P}_{\ell,i}^m , \tag{31}$$

where i indexes the N simulations. The correlation matrix is defined as

$$\mathbf{R}_{\ell_1 \ell_2}^{mn} \equiv \frac{\mathbf{C}_{\ell_1 \ell_2}^{mn}}{\sqrt{\mathbf{C}_{\ell_1 \ell_1}^{mm} \mathbf{C}_{\ell_2 \ell_2}^{nn}}} ; \tag{32}$$

by construction, this is unity along the diagonal. We show all combinations of the correlation matrix moments in Fig. 2: the first column displays the correlation between k-bins in the monopole with itself, the quadruple, and the hexadecapole (from bottom to top, respectively). Analogously, the second column displays the correlations of the quadrupole with the same power spectrum multipoles and the third column shows the correlations of the hexadecapole with said moments. The anisotropies sourced by the imposed survey geometry and intrinsic correlations from the generated anisotropies are (weakly) visible as correlations along the diagonal between the monopole and quadrupole (and the quadrupole with the hexadecapole, respectively.). We find negligible correlation between the monopole and hexadecapole (the latter being mostly dominated by noise). For the $\ell_1 = \ell_2$ correlations, we observe a correlation length of up to two k-bins, resulting in an approximately bi-diagonal correlation matrix. Note that for the quadrupole and the hexadecapole, the lowest three k-bins are highly correlated. This can
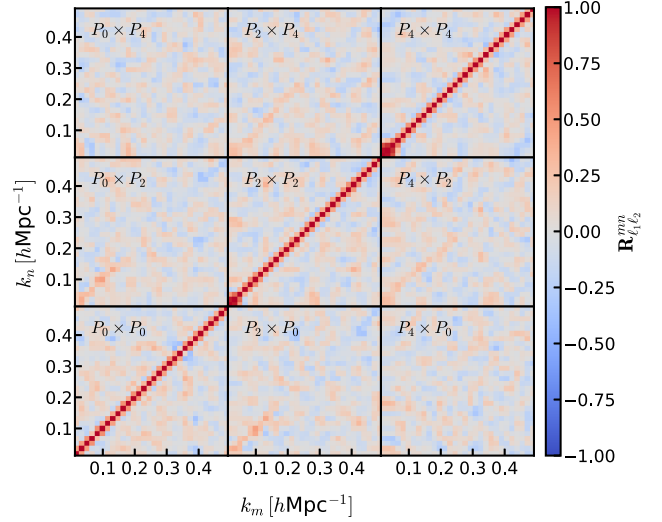
Figure 2. Correlation matrix of the power spectrum multipoles $P_{\ell_1} \times P_{\ell_2}$ measured from $N = 100$ Gaussian random field realizations, as in Fig. 1. The correlation matrix is defined in equation (32); red denotes fully correlated and blue fully anticorrelated power spectrum bins.

be removed by increasing R_0 in the pair separation function, though we caution that computation time-scales as $\mathcal{O}(N_{\text{pix}}^2 R_0^3)$.

Whilst the above tests on GRFs give us confidence that the P3D estimator recovers the input power spectrum to high precision in the presence of survey geometries typical for Ly α data analysis, we stress that these simplistic simulations ignore crucial systematic and instrumental effects as well as details of the data reduction pipeline present in real data, e.g. the co-addition of individual spectra, per spectral pixel noise estimates, sky residuals, continuum fitting, metal contamination, damped Ly α absorbers, and broad absorption lines. Whilst resolution mostly affects high-k modes, the key systematic arising along the line of sight, i.e. at large scales, is the uncertainty in the continuum measurement.[21] In the following section, we address (some of) these issues using more realistic Ly α forest synthetic data.

5.2 Synthetic Ly α spectra: `LyaCoLoRe` simulations

In this section, we present tests of our estimator on realistic Ly α forest simulations (for a more complete discussion of the `LyaCoLoRe` mocks, see Farr et al. 2020; dMdB20). In the following, we will briefly summarize the key steps to generate these simulations: Each realization is based on a GRF of length $10\,h^{-1}$ Gpc and 4096^3 particles yielding a resolution of $\sim 2.4\,h^{-1}$ Mpc generated using `CoLoRe`'s log-normal density model (Ramírez-Pérez et al. 2022).[22] The input cosmology is based on the best-fitting ΛCDM-parameters obtained by the *Planck* satellite (see column 1 of table 3 in Planck Collaboration XIII 2016). The Gaussian field is subsequently transformed to a log-normal density field which is Poisson sampled assuming a quasar density of 59 quasars per deg² and a functional form for the quasar bias redshift evolution (see e.g. dMdB20). The line-of-sight skewers are computed by interpolating from the GRF

---

[21]See Font-Ribera et al. (2018) and Horowitz et al. (2024) for approaches on how to marginalize over continuum uncertainties in the power spectrum computation.

[22]Publicly available at https://github.com/damonge/CoLoRe.

on to the 'observed' pixel positions and the radial velocity field to the centre of the box. The obtained skewers are then post-processed using the LyaCoLoRe package (Farr et al. 2020; Herrera-Alcantar et al. 2023).[23] The sparse sampling of the Ly α forest transverse to the line of sight (also denoted by δ-sampling) results in power on small scales contributing to the error on large scales (see e.g. McDonald & Eisenstein 2007) which is quantified by the P1D. Thus, additional power is added to each skewer by sampling from a Gaussian with variance set by the P1D. From the log-normal transformation of the final skewer the optical depth field, τ, is computed using the fluctuating Gunn–Peterson approximation (FGPA; Croft et al. 1998). RSDs are obtained by convolving the τ field with the peculiar velocity field. The observed flux is then related to τ via $F = \exp(-\tau)$.

The LyaCoLoRe mocks mimic noise and instrumental systematics present in the eBOSS data.[24] In particular, the effect of continuum fitting, instrumental noise, HCD absorbers[25], Ly α absorption and metals. Additionally, random redshift errors for the spectra are included by drawing from a Gaussian with mean zero and $\sigma_z = 400 \, \mathrm{km \, s^{-1}}$. The eBOSS analysis pipeline and spectral resolution are simulated using the SPECSIM package (Kirkby et al. 2021).[26] The co-addition of the resulting skewers with instrumental noise and a realistic quasar continuum is done using the DESISIM package.[27]

In the following, we present tests of our power spectrum estimator on N realizations of mocks with increasing levels of systematics. We compare four different set of mocks (and use a similar nomenclature as the one adopted in table 4 of dMdB20):

(i) Ly α raw mocks (eboss-raw): We estimate the P3D directly from the simulated Ly α forest fluctuations. This analysis is similar to the case of the GRFs but using the eBOSS Ly α survey geometry with the corresponding effective volume. We use $N = 10$ realizations.

(ii) Distorted Ly α raw mocks (eboss-raw-dist): We include a variation of the 'raw' mocks of which we remove the mean of each skewer and denote them by 'raw-dist'. We use $N = 10$ realizations.

(iii) +continuum + noise (eboss-0.0): We add the effect of measuring quasar continua of the spectra and include instrumental noise. The performed continuum fitting, introduced in Section 4.2, requires forward modelling the suppression of modes along the line of sight through a so-called distortion matrix given in equation (27). We use $N = 10$ realizations.

(iv) + metals + HCDs + σ_ν (eboss-0.2): We estimate P3D from realistic Ly α mock spectra including the effects stemming from continuum fitting and instrumental noise (from step ii) and adding metals, HCD absorbers, and random redshift errors. Analogous to the analysis on real data, HCDs are masked based on the cut $\log N_{\mathrm{Hi}} > 20.3 \, \mathrm{cm^{-2}}$. We use $N = 100$ realizations.

In Fig. 3, we compare the power spectra of the LyaCoLoRe mocks for the four different mock configurations, alongside their errors. As in the main analysis, we measure the power spectra up to
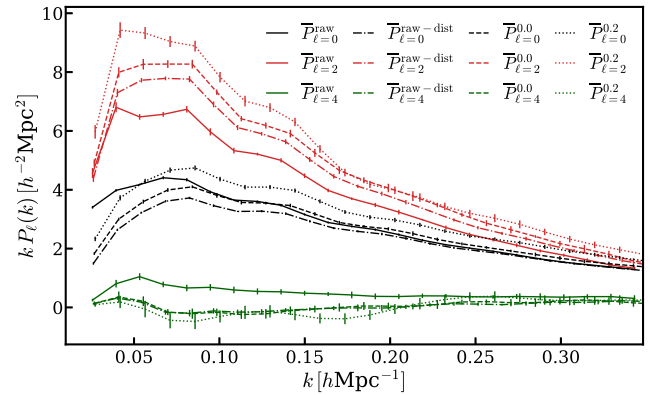
[23]Publicly available at https://github.com/igmhub/LyaCoLoRe.

[24]Note that while this set of synthetic spectra have been tuned to BAO-scale analyses and have also been tested in the context of full-shape analyses of the Ly α forest with the 2PCF (see e.g. Cuceu et al. 2023), these have not been explicitly developed for the purpose of P3D analyses.

[25]Following dMdB20, we denote systems with neutral hydrogen exceeding $10^{17.2}$ atoms $\mathrm{cm^{-2}}$ as HCDs.

[26]Publicly available at https://github.com/desihub/specsim.

[27]Publicly available at https://github.com/desihub/desisim.

Figure 3. Comparison of Ly α power spectra measured from LyaCoLoRe mocks for four different configurations: 'raw' are the raw Ly α forest fluctuations (solid lines); 'raw-dist' are the distorted fluctuations where the mean and the slope along each sight line have been removed (dashed–dotted lines); '0.0' include continua and instrumental noise (dashed lines); '0.2' include metals, HCDs, and random redshift errors (dotted lines). The coloured lines (black for the monopole, red for the quadrupole, and green for the hexadecapole, respectively) are the mean power spectra and the error bars are the variance between the different realizations.

$\ell_{\max} = 4$ and compute the window function multipoles up to $\ell_{\max} = 8$. Therefore, we use 24 equidistant k-bins in the range ($0.02 \leq k \leq 0.35$) $h \, \mathrm{Mpc^{-1}}$ and $N_r = 1,000$ linearly spaced bins in the range ($0 \leq r \leq R_0 = 200$) $h^{-1} \, \mathrm{Mpc}$ for the window matrix, respectively. We treat the mean power spectra measured from the 'raw' Ly α forest fluctuations mocks (solid coloured lines in the plot), as ground truth to assess the effect of contaminants on the power spectra.[28]

First, we consider the effects of removing the mean and the slope of each line of sight to mimic the effect of the distortion matrix (dashed–dotted lines in Fig. 3). This isolates the effect of the distortion matrix modelling and can be seen to reduce the monopole power on all scales with most pronounced impact at large scales where it reduces the power up to a factor of two. The quadrupole is also affected at all scales, leading to an increase in power by approximately 20 per cent. The hexadecapole is also affected at all scales, and exhibits a sign change at large scales. Note that the effect becomes increasingly small above $k \geq 0.3 \, h \, \mathrm{Mpc^{-1}}$ for all spectra.

Second, we assess the impact of the continuum and instrumental noise (dashed lines in Fig. 3), denoted by 'eboss-0.0', discussed in Appendix B. This affects all multipoles at large scales. For the monopole it removes, as expected, power at $k \lesssim 0.10 \, h \, \mathrm{Mpc^{-1}}$, since the large-scale modes are projected out of the continuum. For the quadrupole (red dashed line) it enhances power at all scales with a further enhancement of 10–15 per cent up to $k \lesssim 0.15 \, h \, \mathrm{Mpc^{-1}}$. The hexadecapole is strongly affected and switches sign below $k \lesssim 0.20 \, h \, \mathrm{Mpc^{-1}}$. Note that we do not expect the 'raw-dist' mocks to agree exactly with the '0.0' mocks since removing the mean from each line of sight is not exactly the same as introducing a continuum in each spectrum (the latter also affecting the weights). These differences will propagate into the distortions and yield the observed qualitative agreement with a small offset.

[28]Note that we do not compare the measured power spectrum to theoretically expected power spectra since the true power in those mocks is not known, except that at large scales where it follows the Kaiser formula given in equation (11) with the best-fitting parameters from table 1 in Farr et al. (2020) for the parameters $\{\alpha_\parallel, \alpha_\perp, b_{\eta,F}, \beta_F, b_{\delta,F}\}$, shown in Fig. B1.
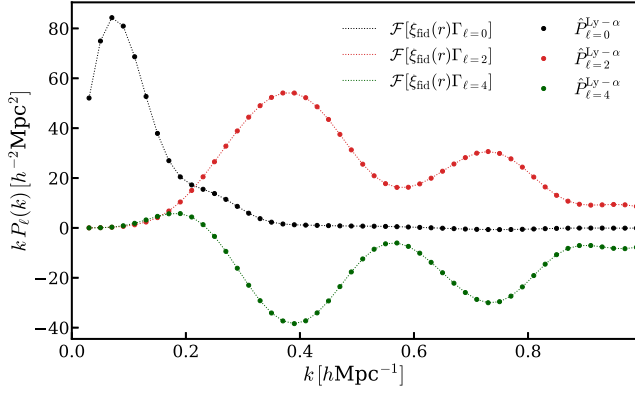
Figure 4. Consistency test for anisotropies induced by the (eBOSS-like) survey geometry. This shows the power spectra computed from an isotropic fiducial correlation function (without stochasticity), estimated from skewers taken from a LyaCoLoRe mock using pairs up to $R_0 = 200 \, h^{-1} \, \mathrm{Mpc}$. The dots represent the measurement of the P3D estimator (black, red, and green for $\ell = 0, 2, 4$, respectively). The dashed lines are the fiducial correlation function, $\xi(r)$ multiplied by the window function and the pair count separation window, $\Phi_\ell W(r; R_0)$ (same colour coding). For ease of comparison, we divide the monopole by a factor of 10.

Third (dotted lines in Fig. 3), we include metals, HCDs, and random redshift errors. For the monopole (dashed to dotted) this adds ∼10 per cent power at all scales, creating a small offset. For the quadrupole this mostly affects scales up to $k \lesssim 0.15 \, h \, \mathrm{Mpc}^{-1}$. The hexadecapole, however, is barely affected. We see that the measurements are noisier and that the power is further decreased in two regions around $k \sim 0.05 \, h \, \mathrm{Mpc}^{-1}$ and $k \sim 0.15 \, h \, \mathrm{Mpc}^{-1}$.

5.3 Anisotropies from the survey geometry

Finally, we illustrate the effects of window function-induced anisotropy by constructing a simple test built upon a known isotropic correlation function, $\xi_{\mathrm{fid}}(\mathbf{r}) = \exp(-|\mathbf{r}|/10)$. To remove the stochasticity of the data set, we replace the $w_i w_j \delta_i \delta_j$ term in our pair count estimator (given in equation 9) with the explicit correlation function ξ_{fid}, evaluated at the pair separation $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. We apply the resulting estimator to the LyaCoLoRe mocks as before, which tests the sensitivity of the present approach to recovering anisotropies purely sourced by the survey geometry, i.e. our measurement of the window matrix. Although the input correlation function is isotropic (with $\xi_{\ell>0} = 0$), mode-mixing is expected to occur due to the window; thus $\hat{P}_\ell(k)$ will be non-zero in practice. To compare data and theory, we use equations (20)–(22), as before, setting $\xi_{\ell>0} = 0$. In this limit, $\xi_\ell(r) = \xi_0(r)\Gamma^2 \ell(r).[29]$

In Fig. 4, we show the results using all the skewers from our 'eboss-0.2' mocks which mimic the realistic eBOSS survey geometry.[30] The δ-fields of these mocks have HCDs that are masked out, i.e. this introduces an additional level of realism in reconstructing the window function. We recover the monopole (black), quadrupole (red), and hexadecapole (green) to better than 0.1 per cent precision over the entire k-range of $(0.0 < k \leq 1.0) \, h \, \mathrm{Mpc}^{-1}$, indicating that our modelling of the window function is highly accurate. This gives

us confidence in the present approach to measure the P3D at all scales from Ly α data.

6 RESULTS

In this section, we use the pair-count estimator described in Section 2 to measure the P3D from 205 012 eBOSS DR16 quasar spectra. In Section 6.1 and Fig. 5, we present the key result of the paper: the P3D measurement on eBOSS data in a single redshift bin $2 \leq z \leq 4$ for both data sets (NGC and SGC). We compare the measurements to window convolved theory power spectra described in Section 3. Throughout this section we adopt a *Planck* 2016 best-fitting ΛCDM cosmology: $\Omega_c h^2 = 0.1197$, $\Omega_b h^2 = 0.02222$, $\Omega_\nu h^2 = 0.0006$, $H_0 = 67.31 \, \mathrm{km \, s^{-1} \, Mpc^{-1}}$, $\sigma_8 = 0.8299$, $n_s = 0.9655$ (Planck Collaboration XIII 2016). For the Ly α forest, we use an effective mean redshift of $z = 2.334$ and for theory power spectra, given in equation (11), we use the eBOSS DR16 best-fitting values from the Ly α autocorrelation presented in table 6 in dMdB20. The BAO parameters are defined as usual

$$\alpha_\parallel \equiv \frac{D_H(z_{\mathrm{eff}})/r_d}{[D_H(z_{\mathrm{eff}})/r_d]_{\mathrm{fid}}} = 1.047 \, ,$$

$$\alpha_\perp \equiv \frac{D_M(z_{\mathrm{eff}})/r_d}{[D_M(z_{\mathrm{eff}})/r_d]_{\mathrm{fid}}} = 0.980 \, , \tag{33}$$

where D_H and D_M are the Hubble and comoving angular diameter distances, respectively, evaluated at some effective redshift, z_{eff}. The denominator is evaluated at a fiducial ΛCDM cosmology. The bias and RSD best-fitting parameters are (dMdB20)

$$b_{\eta, \mathrm{Ly}\alpha} = -0.201, \, \beta_{\mathrm{Ly}\alpha} = 1.657 \, , \tag{34}$$

the metal biases with the associated rest-frame wavelength in parentheses are

$$b_{\eta, \mathrm{Si\,ii\,(1190)}} = -0.0029 \, , \, b_{\eta, \mathrm{Si\,ii\,(1193)}} = -0.0021 \, ,$$

$$b_{\eta, \mathrm{Si\,iii\,(1207)}} = -0.0045 \, , \, b_{\eta, \mathrm{Si\,ii\,(1260)}} = -0.0022 \, , \tag{35}$$

and the bias parameters for the HCD systems (Rogers et al. 2018)

$$b_{\mathrm{HCD}} = -0.0522 \, , \, \beta_{\mathrm{HCD}} = 0.610 \, . \tag{36}$$

In Section 6.2, we discuss our error bars and covariance matrix for the P3D measurement obtained from $N = 100$ realistic Ly α simulations. Increasing the pair separation in the pair truncation window function $W(r; R_0)$ to $R_0 = 400 \, h^{-1} \, \mathrm{Mpc}$ did not improve the agreement between theory and measured power spectrum, thus we do not show the corresponding results in this work. For our main analysis, we measure the power spectra up to $\ell_{\mathrm{max}} = 4$ in 24 equidistant k-bins in the range $(0.02 \leq k \leq 0.35) \, h \, \mathrm{Mpc}^{-1}$ and the window function multipoles up to $\ell_{\mathrm{max}} = 8$ with $N_r = 1,000$ linearly spaced bins in the range $(0 \leq r \leq R_0 = 200) \, h \, \mathrm{Mpc}^{-1}$, respectively.

6.1 Power spectrum multipoles from eBOSS DR16

In Fig. 5, we show the P3D multipoles measured on eBOSS DR16 data for the NGC (SGC) in the left (right) panel and compare them to the window convolved quasi non-linear theory power spectrum multipoles. We show the monopole, quadrupole, and hexadecapole from the data and compute the window function multipoles up to $\ell_{\mathrm{max}} = 8$, as described in Section 3. The error bars are given by the square root of the diagonal of the covariance matrix, introduced in Section 6.2. To account for the employed continuum-fitting procedure that projects
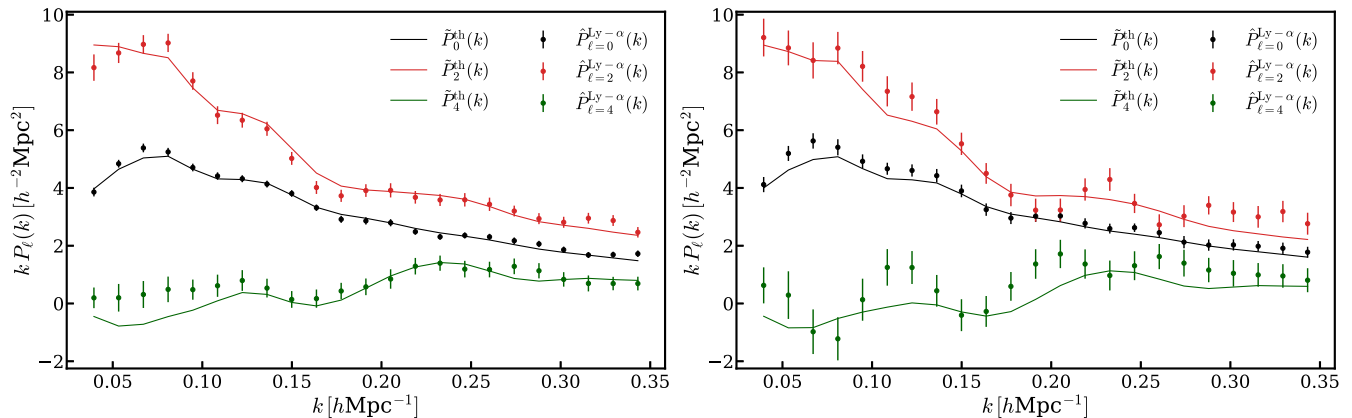
---

[29]In the limit of a fully sampled and infinite box the recovered signal would be isotropic with $\hat{P}_{\ell=2} = \hat{P}_{\ell=4} = 0$.

[30]We performed the same test using the eBOSS DR16 skewers and recover the anisotropies to 0.1 per cent precision over the same k-range.

Figure 5. Power spectrum multipoles measured from eBOSS DR16 data for the NGC (left) and SGC (right), showing the monopole in black, quadrupole in red, hexadecapole in green. The dots give the measurements and the solid lines show the quasi-linear theory predictions using eBOSS DR16 best-fitting values (dMdB20). The error bars are the square root of the diagonal of the covariance matrix obtained from $N = 100$ realistic LyaCoLoRe mocks. The window function contributions are computed up to $\ell_{\mathrm{max}} = 8$ and convolved with the distorted theory power spectrum. All measurements use the pair-count estimator of equation (9), truncated at $R_0 = 200\,h^{-1}\,\mathrm{Mpc}$ via a continuous window function $W(r; R_0)$.

out large-scale modes, we apply a distortion matrix, further altering the large-scale shape.[31] The estimator uses a pair separation window truncating the pair count at $R_0 \geq 200\,h^{-1}\,\mathrm{Mpc}$, as before.

For the monopole, we find excellent agreement with the theory prediction over the entire k-range. The quadrupole is also in good agreement with the theory prediction, bar the last few measurement points in the $k \gtrsim 0.20\,h\,\mathrm{Mpc}^{-1}$ range. The quadrupole and hexadecapole both have correlated features at large scales: We conjecture that this stems from the mode mixing introduced by the distortion matrix. In particular, the $\sim$1–2σ discrepancy at large scales in the hexadecapole has to be dealt with cautiously when extracting cosmology from the Lyα forest. In Appendix B and Fig. B1, we justify this by showing that the forward modelling of the DM does not capture exactly the hexadecapole on large scales: first, the distortion matrix is computed out to separations of $\gtrsim 400\,h^{-1}\,\mathrm{Mpc}$ but should, in principle, be computed up to the scales of the survey. Second, it assumes a correlation function measurement in bins of size $\{4\,h^{-1}\,\mathrm{Mpc} \times r_\parallel, 4\,h^{-1}\,\mathrm{Mpc} \times r_\perp\}$. In the range $k \gtrsim 0.2\,h\,\mathrm{Mpc}^{-1}$, the quadrupole and hexadecapole measurements and the theory predictions are in $\sim 1\sigma$ agreement with each other. Note that the error bars from SGC are larger than the ones from the NGC which stems from the smaller statistics, i.e. NGC approximately has three times the number of spectra than the SGC data set.

In Fig. 6, we show the ratio of the quadrupole and hexadecapole to the monopole to obtain an estimate for the RSD parameter, β_F. Note that although the measurements are quite noisy, we get good agreement for the measured power spectra to the non-linear theory predictions at scales $k \gtrsim 0.1\,h\,\mathrm{Mpc}^{-1}$. For comparison, we include the linear theory prediction (horizontal dashed lines) to guide the eye which we compute by integrating equation (17) in Fourier space. Assuming non-linear effects are negligible, the Kaiser effect gives the analytic pre-factors $Q_{\ell=0}(\beta) = 1 + 2\beta/3 + \beta^2/5$, $Q_{\ell=2}(\beta) = 4\beta/3 + 4\beta^2/7$, and $Q_{\ell=4}(\beta) = 8\beta^2/35$ for the monopole, quadrupole, and hexadecapole, respectively. In this limit, the multipoles of our theory power spectrum are given by $P_\ell(k) = Q_\ell(\beta)b^2 P_{\mathrm{lin}}(k)$; inserting the value of $\beta = 1.627$ gives

1.39 and 0.23 for the ratios of quadrupole to monopole and hexadecapole to monopole, respectively. The errorbars are taken from the diagonals of the covariance matrices, re-scaled by the monopole power, and should be taken indicatively rather than at face value.

6.2 Covariance matrix from mocks catalogues

To capture the variance between spectra and account for instrumental and systematic noise in the data, we compute a covariance matrix from $N = 100$ realizations of 'eboss-0.2' LyaCoLoRe simulations, introduced in Section 5.2. The error bars in Fig. 5 are the square root of the diagonal of the covariance matrix with the corresponding correlation matrix shown in Fig. 7 for which we vary both Legendre multipoles and k-bins, as in Section 5.1. The resulting correlation matrix has a diagonal structure with non-negligible off-diagonal terms and blocks: it is interesting to note that the eBOSS DR16 survey geometry mixes modes between 'neighbouring multipoles', e.g. between monopole and quadrupole as well as between quadrupole and hexadecapole. For the $\ell_1 = \ell_2$ correlations, we observe a correlation length of up to two k-bins. Note that we do not include any scales with k^{-1} larger than the pair separation window R_0, thus (in contrast to the GRF correlation matrix) the lowest k-bins are not artificially correlated. The resulting covariance matrix is positive definite and can thus be used for cosmological inference. We note that the covariance matrix of the monopole, $P_{\ell=0} \times P_{\ell=0}$, shows correlations between k-bins of order 10 per cent on larger and up to 30 per cent on smaller scales. This may stem from the application of the distortion matrix, which effectively mixes δ-modes. Thus, we expect a less correlated correlation matrix when using an alternative continuum-fitting method that does not project out large-scale modes at the expense of increased noise in the forest.

7 SUMMARY AND CONCLUSIONS

The Lyα forest is a treasure trove of cosmological information on the expansion history of the Universe and beyond. Given its high-redshift range ($2 \leq z \leq 4$) and sensitivity to Mpc scales and below, it is an ideal tracer to probe a wide range of scales: early-Universe physics from combinations with a large-scale tracer, e.g. the cosmic

---

[31]We remind the reader that we use the same data selection and Lyα flux decrement obtained from eBOSS DR16 data for our analysis as dMdB20.
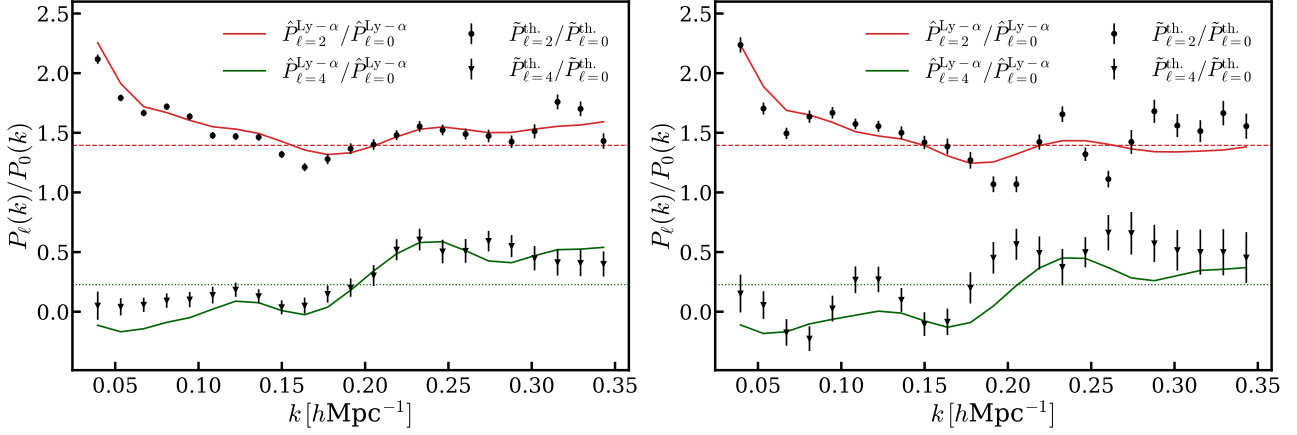
Figure 6. Ratio of measured quadrupole (black dots) and hexadecapole (black inverted triangles) power spectra to monopole power for eBOSS DR16 data (left panel: NGC, right panel: SGC), with linear expectation for $\beta_F = 1.627$ given by the horizontal dashed lines and ratio of the (non-linear) theory power spectra for the ratio of the quadrupole (red line) and hexadecapole (green line) to the theory monopole. The error bars are taken from the power spectrum, i.e. the square root of the diagonal of the covariance matrix obtained from LyaCoLoRe simulations, and are rescaled by the monopole power.
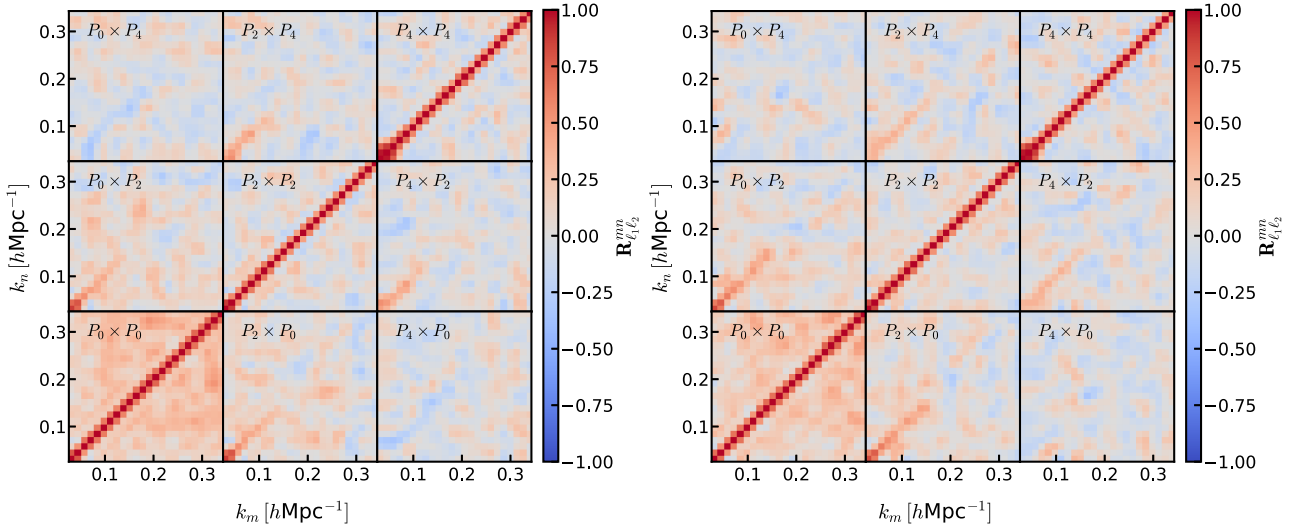
Figure 7. Correlation matrix of the power spectrum multipoles $P_{\ell_1} \times P_{\ell_2}$ obtained from $N = 100$ realizations of the LyaCoLoRe simulations, emulating that of the eBOSS data. The left (right) panel shows the NGC (SGC) measurements. For each multipole, we display the k-modes used in the analysis: $0.02 \le k \le 0.35 \, h \, \mathrm{Mpc}^{-1}$. Pair counts are truncated above $R_0 > 200 \, h^{-1} \, \mathrm{Mpc}$. The correlation matrix is defined in equation (32); red denotes fully correlated and blue fully anticorrelated power spectrum bins. In general, the correlation matrix shows little off-diagonal terms except for the anisotropies sourced by the sparse Ly α sampling which are visible in the 'cross' terms $P_\ell \times P_{\ell'}$ for the combinations $\{\ell = 0, \ell' = 2\}$ and $\{\ell = 2, \ell' = 4\}$ (and their inverses). The autocorrelation of the monopole is at the level of 10–30 per cent for the off-diagonal terms. Cross-terms of the hexadecapole with the monopole are uncorrelated and with the quadrupole only slightly correlated.

microwave background anisotropies, from large scales or dark matter models as well as thermal properties of the ionized (cold) IGM from small scales. Previous analyses have focused on the 2PCF (see e.g. dMdB20) or the small-scale P1D from Fourier transforms along each line of sight (see e.g. Ravoux et al. 2023; Karaçaylı et al. 2024, in the context of DESI); however, a pioneering study on simulations (Font-Ribera et al. 2018) and a recent small-scale measurement (Karim et al. 2023) paired with advancements in the theoretical modelling of the Ly α forest (see e.g. Ivanov 2023) have now paved the way for P3D analyses. To achieve this goal, we require accurate estimates of the three-dimensional Ly α power spectrum: this is challenging, given the sparse sampling transverse to and dense sampling along the line of sight. In this paper, we attack this problem by presenting a pair count estimator to measure the P3D from the Ly α forest. Our

approach is based on the pair counting estimator HIPSTER (Philcox & Eisenstein 2020; Philcox 2021) and weighs each pair by $\exp(i\mathbf{k} \cdot \mathbf{r}_{ij})$, for wavevector $\mathbf{k}$ and particle pair separation $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. This directly measures the power spectrum without the need for grid-based Fourier transforms which are affected by the sparse δ−function type sampling of the Ly α forest.

We have presented the first large-scale P3D measurement of Ly α forest spectra on 205 012 medium-resolution Ly α forest spectra from the eeBOSS DR16 in two disjoint regions of the sky: the NGC and SGC. Furthermore, we have compared our P3D measurement to the best-fitting quasi-linear theory prediction from the 2PCF presented in dMdB20. We extensively test the estimator on Gaussian random fields, i.e. realizations of a linear input power spectrum, with Gaussian noise and realistic continuum error levels of 10 per cent. We

recover the input power spectrum at the $\sim (1)$ 3σ level for (dense) sparse configurations, respectively. Further, our pipeline has been extensively tested on synthetic Ly α spectra (LyaCoLoRe) with increasing levels of realism to probe for the effects of distorted continua, instrumental noise, HCDs, metals and random redshift errors. We demonstrated that we can forward model the effects of anisotropies sourced by the survey geometry and the distortions introduced by the continuum fitting (modes along the line of sight, $k_\parallel = 0$) accurately for the multipoles.

We present a covariance matrix derived from 100 realistic Lya-CoLoRe simulations for further cosmological analysis of the Ly α forest data set. We obtain a fairly diagonal covariance matrix with correlations between adjacent multipoles stemming from the survey geometry and, for the autocorrelation of the monopole, contaminants and the mode-mixing distortion matrix. We present the anisotropic clustering of eBOSS DR16 Ly α spectra up to the hexadecapole and obtain good agreement to the Kaiser formula with a non-linear correction term obtained from hydrodynamical simulations (Arinyo-i-Prats et al. 2015) up to $k \leq 0.35\,h\,\mathrm{Mpc}^{-1}$. The best-fitting theory power spectrum obtained from dMdB20 shows deviations at the $\sim 2\sigma$ level on the largest scales, in particular for the hexadecapole. The quadrupole and hexadecapole mix a large range of scales when convolving the window matrix with the correlation function and Hankel transforming it to Fourier space. This effect is even more pronounced if a distortion matrix (an artefact from the present continuum fitting approach) is applied. We leave for future work to explore alternative continuum fitting methods as well as the accuracy of the presented covariance matrix.

Our main conclusion is that the novel estimator is well-suited to measure clustering statistics from the Ly α forest and can deal with non-trivial survey geometries and masked data vectors. This will facilitate robust measurements of P3D for the currently observing DESI and future surveys such as the PFS, WEAVE-QSO, and 4MOST, opening the door to a wide variety of novel cosmological analyses.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request. The eBOSS DR16 data set is publicly available at https://data.sdss.org/sas/dr16.

REFERENCES

Abareshi B. et al., 2022, AJ, 164, 207

Adame A. G. et al., 2024, AJ, 167, 62

Afshordi N., McDonald P., Spergel D. N., 2003, ApJ, 594, L71

Alam S. et al., 2021, Phys. Rev. D, 103, 083533

Arinyo-i-Prats A., Miralda-Escudé J., Viel M., Cen R., 2015, J. Cosmol. Astropart. Phys., 2015, 017

Armengaud E., Palanque-Delabrouille N., Yèche C., Marsh D. J. E., Baur J., 2017, MNRAS, 471, 4606

Baur J., Palanque-Delabrouille N., Yèche C., Magneville C., Viel M., 2016, J. Cosmol. Astropart. Phys., 2016, 012

Beutler F. et al., 2017, MNRAS, 466, 2242

Beutler F., McDonald P., 2021, J. Cosmol. Astropart. Phys., 2021, 031

Blomqvist M. et al., 2018, J. Cosmol. Astropart. Phys., 2018, 029

Boera E., Becker G. D., Bolton J. S., Nasir F., 2019, ApJ, 872, 101

Bolton J. S., Puchwein E., Sijacki D., Haehnelt M. G., Kim T.-S., Meiksin A., Regan J. A., Viel M., 2017, MNRAS, 464, 897

Bolton J. S., Viel M., Kim T. S., Haehnelt M. G., Carswell R. F., 2008, MNRAS, 386, 1131

Busca N. G. et al., 2013, A&A, 552, A96

Castorina E., White M., 2018, MNRAS, 476, 4403

Chabanier S. et al., 2019, J. Cosmol. Astropart. Phys., 2019, 017

Chen S.-F., Vlah Z., White M., 2021, J. Cosmol. Astropart. Phys., 2021, 053

Cieplak A. M., Slosar A., 2016, J. Cosmol. Astropart. Phys., 2016, 016

Croft R. A. C., Weinberg D. H., Katz N., Hernquist L., 1998, ApJ, 495, 44

Cuceu A., Font-Ribera A., Joachimi B., Nadathur S., 2021, MNRAS, 506, 5439

Cuceu A., Font-Ribera A., Nadathur S., Joachimi B., Martini P., 2023, Phys. Rev. Lett., 130, 191003

Dawson K. S. et al., 2016, AJ, 151, 44

de Jong R. S. et al., 2019, The Messenger, 175, 3

Dekker H., D'Odorico S., Kaufer A., Delabre B., Kotzlowski H., 2000, in Iye M., Moorwood A. F., eds, Proc. SPIE Conf. Ser. Vol. 4008, Optical and IR Telescope Instrumentation and Detectors. SPIE, Bellingham, p. 534

DESI Collaboration, 2016, preprint (arXiv:1611.00036)

Doughty C. C., Hennawi J. F., Oñorbe J., Davies F. B., Lukić Z., 2023, preprint (arXiv:2312.01480)

du Mas des Bourboux H. et al., 2020, ApJ, 901, 153, dMdB20

Eisenstein D. J., Seo H.-J., White M., 2007, ApJ, 664, 660

Farr J. et al., 2020, J. Cosmol. Astropart. Phys., 2020, 068
Font-Ribera A., McDonald P., Slosar A., 2018, J. Cosmol. Astropart. Phys., 2018, 003
Gaikwad P., Srianand R., Haehnelt M. G., Choudhury T. R., 2021, MNRAS, 506, 4389
Gaikwad P., Srianand R., Khaire V., Choudhury T. R., 2019, MNRAS, 490, 1588
Garny M., Konstandin T., Sagunski L., Viel M., 2021, J. Cosmol. Astropart. Phys., 2021, 049
Garzilli A., Bolton J. S., Kim T. S., Leach S., Viel M., 2012, MNRAS, 424, 1723
Garzilli A., Magalich A., Theuns T., Frenk C. S., Weniger C., Ruchayskiy O., Boyarsky A., 2019, MNRAS, 489, 3456
Givans J. J. et al., 2022, J. Cosmol. Astropart. Phys., 2022, 070
Givans J. J., Hirata C. M., 2020, Phys. Rev. D, 102, 023515
Goldstein S., Hill J. C., Iršič V., Sherwin B. D., 2023, Phys. Rev. Lett., 131, 201001
Gordon (2023) Journal of Cosmology and Astroparticle Physics, 045 , 1475-7516
Gordon C. et al., 2023a, preprint (arXiv:2308.10950)
Gordon C. et al., 2023b, J. Cosmol. Astropart. Phys., 2023, 045
Greene J., Bezanson R., Ouchi M., Silverman J., the PFS Galaxy Evolution Working Group, 2022, preprint (arXiv:2206.14908)
Herrera-Alcantar H. K. et al., 2023, preprint (arXiv:2401.00303)
Horowitz B., de Belsunce R., Lukic Z., 2024, preprint (arXiv:2403.17294)
Hui L., Stebbins A., Burles S., 1999, ApJ, 511, L5
Iršič V. et al., 2017, Phys. Rev. D, 96, 023522
Iršič V. et al., 2023, Phys. Rev. D, 109, 043511
Iršič V., McQuinn M., 2018, J. Cosmol. Astropart. Phys., 2018, 026
Iršič V., Xiao H., McQuinn M., 2020, Phys. Rev. D, 101, 123518
Ivanov M. M., 2023, Phys. Rev. D, 109, 023507
Kaiser N., 1987, MNRAS, 227, 1
Karaçaylı N. G. et al., 2024, MNRAS, 528, 3941
Karaçaylı N. G., Font-Ribera A., Padmanabhan N., 2020, MNRAS, 497, 4742
Karim M. L. A., Armengaud E., Mention G., Chabanier S., Ravoux C., Lukić Z., 2023, J. Cosmol. Astropart. Phys., 2024, 088
Kirkby D. et al., 2021, The DESI bright galaxy survey: final target selection, design, and validation, Zenodo, doi:10.5281/zenodo.4566008
Kobayashi T., Murgia R., De Simone A., Iršič V., Viel M., 2017, Phys. Rev. D, 96, 123514
Lyke B. W. et al., 2020, Astrophys. J. Suppl., 250, 8
McDonald P. et al., 2006, ApJS, 163, 80
McDonald P., 2003, Astrophys. J., 585, 34
McDonald P., Eisenstein D. J., 2007, Phys. Rev. D, 76, 063009
McQuinn M., 2016, ARA&A, 54, 313
Meiksin A. A., 2009, Rev. Mod. Phys., 81, 1405
Murgia R., Iršič V., Viel M., 2018, Phys. Rev. D, 98, 083540
Murgia R., Scelfo G., Viel M., Raccanelli A., 2019, Phys. Rev. Lett., 123, 071102
Murphy M. T., Kacprzak G. G., Savorgnan G. A. D., Carswell R. F., 2019, MNRAS, 482, 3458
Myers A. D. et al., 2015, ApJS, 221, 27
Noterdaeme P. et al., 2012, A&A, 547, L1
O'Meara J. M., Lehner N., Howk J. C., Prochaska J. X., 2021, AJ, 161, 45
Palanque-Delabrouille N. et al., 2013, A&A, 559, A85
Palanque-Delabrouille N., Yèche C., Schöneberg N., Lesgourgues J., Walther M., Chabanier S., Armengaud E., 2020, J. Cosmol. Astropart. Phys., 2020, 038
Pedersen C., Font-Ribera A., Rogers K. K., McDonald P., Peiris H. V., Pontzen A., Slosar A., 2021, J. Cosmol. Astropart. Phys., 2021, 033
Philcox O. H. E., 2021, MNRAS, 501, 4004
Philcox O. H. E., Eisenstein D. J., 2020, MNRAS, 492, 1214
Philcox O. H. E., Ivanov M. M., 2022, Phys. Rev. D, 105, 043517
Pieri M. M. et al., 2016, in Reylé C., Richard J., Cambrésy L., Deleuil M., Pécontal E., Tresse L., Vauglin I.eds, Proc. Annual meeting of the French Society of Astronomy and Astrophysics, Lyon, p. 259, preprint (arXiv:1611.09388)
Planck Collaboration VI, 2020, A&A, 641, A6
Planck Collaboration XIII, 2016, A&A, 594, A13
Puchwein E. et al., 2023, MNRAS, 519, 6162
Ramírez-Pérez C., Sanchez J., Alonso D., Font-Ribera A., 2022, J. Cosmol. Astropart. Phys., 2022, 002
Ravoux C. et al., 2023, MNRAS, 526, 5118
Rogers K. K., Bird S., Peiris H. V., Pontzen A., Font-Ribera A., Leistedt B., 2018, MNRAS, 476, 3716
Rogers K. K., Dvorkin C., Peiris H. V., 2022, Phys. Rev. Lett., 128, 171301
Seljak U. et al., 2005, Phys. Rev. D, 71, 103515
Seljak U., 2012, J. Cosmol. Astropart. Phys., 2012, 004
Slosar A. et al., 2011, J. Cosmol. Astropart. Phys., 2011, 001
Slosar A. et al., 2013, J. Cosmol. Astropart. Phys., 2013, 026
Viel M., Becker G. D., Bolton J. S., Haehnelt M. G., 2013, Phys. Rev. D, 88, 043502
Viel M., Haehnelt M. G., Springel V., 2010, J. Cosmol. Astropart. Phys., 2010, 015
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2005, Phys. Rev. D, 71, 063534
Viel M., Lesgourgues J., Haehnelt M. G., Matarrese S., Riotto A., 2006, Phys. Rev. Lett., 97, 071301
Villasenor B., Robertson B., Madau P., Schneider E., 2022, ApJ, 933, 59
Villasenor B., Robertson B., Madau P., Schneider E., 2023, Phys. Rev. D, 108, 023502
Vogt S. S. et al., 1994, in Crawford D. L., Craine E. R., eds, Proc. SPIE Conf. Ser. Vol. 2198, Instrumentation in Astronomy VIII. SPIE, Bellingham, p. 362
Walther M., Oñorbe J., Hennawi J. F., Lukić Z., 2019, ApJ, 872, 13
Wilson B., Iršič V., McQuinn M., 2022, MNRAS, 509, 2423
York D. G. et al., 2000, AJ, 120, 1579
Zaldarriaga M., 2002, ApJ, 564, 153

APPENDIX A: SURVEY WINDOW FUNCTION FOR GRF

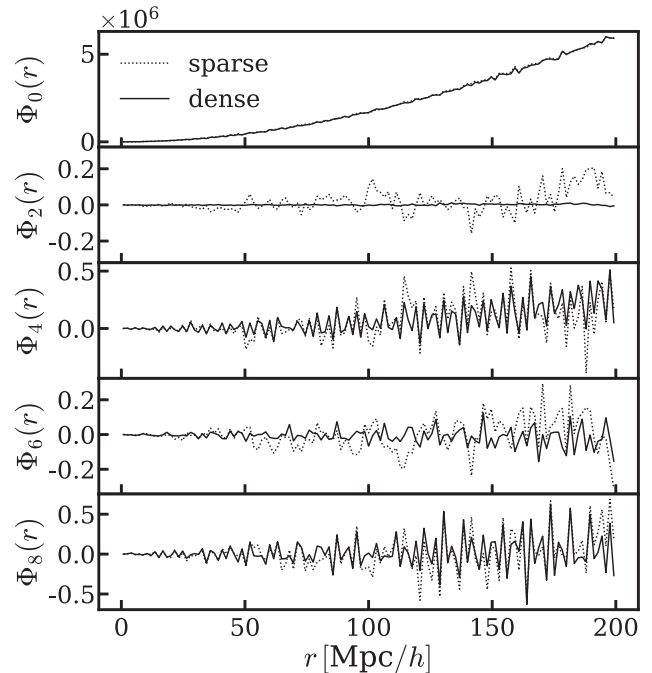In Fig. A1, we show the multipoles of the window function,

Figure A1. Window function of the data, $\Phi_\ell(r)$, for two different configurations discussed in Section 5.1: (i) sparse with $\sim 2\,\mathrm{qso\,deg^{-2}}$; and (ii) dense with an eBOSS DR16-like sampling of $30\,\mathrm{qso\,deg^{-2}}$. The dense window function has been normalized by its number density for ease of comparison.
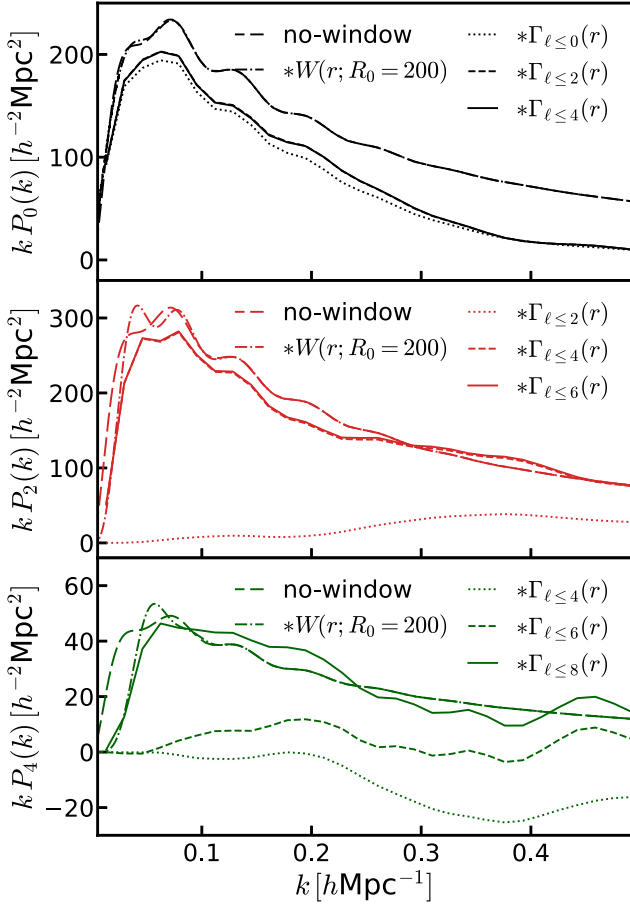
bottom panel, we show the theory convolved hexadecapole with window matrix contributions up to $\ell_{max} = 8$. Here, it is important to note that all multipoles of the window matrix contribute visibly to the final window convolved power spectrum. This stems from the integrand of the Hankel transform from the window convolved correlation function to the power spectrum multipoles. Note that the integrands of the quadrupole and hexadecapole mix a wide range of scales; this is discussed in the context of the eBOSS DR16 data in Appendix B.

APPENDIX B: FORWARD MODELLING DISTORTIONS FROM CONTINUUM FITTING

In Fig. B1, we compare measured power spectra of 'raw' (solid) and continuum-fitted 'eboss-0.0' (dashed) mocks to a best-fitting linear theory power spectrum. We apply the distortion matrix, denoted by DM, to the dashed–dotted linear power spectrum to test the range of validity of our distortion matrix treatment. The resulting distorted linear theory power is shown for two different maximum separations: (i) $R_{DM} = 200\,h^{-1}\,$Mpc (dotted grey line); and (ii) $R_{DM} = 400\,h^{-1}\,$Mpc (dashed). The continuum fitting approach introduced in Section 4.2 suppresses modes along the line of sight, effectively mixing δ fields resulting in distortions in the measured power spectra (or correlation functions). The resulting distortion matrix is computed from the 'eboss-0.0' mocks using equation (27).

When comparing the measured power spectra of the 'raw' to the 'eboss-0.0' mocks we find a suppression of the monopole on all scales which is most pronounced at $k \lesssim 0.1\,h\,$Mpc^{-1} by up to ~ 20 per cent. The quadrupole (red) is enhanced by ~ 20 per cent bosson all scales. The hexadecapole (green) switches sign at the largest scales and is strongly damped. We compare the measured spectra to the linear-theory power spectra (grey lines) which are only valid on large scales (since the true power of these mocks is not known). Applying the DM to the multipoles, we find that the forward modelling of the distortion is accurate at the ~ 5 per cent
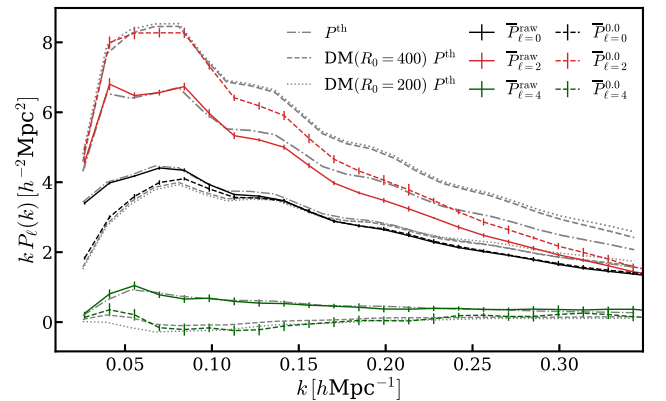

Figure A2. Effect of the convolution of the theory power spectrum with the pair-count truncation window function $W(r; R_0)$ and the survey geometry $\Gamma^2_\ell = \Phi_\ell(r, \mu)W(r; R_0)$ up to ℓ_{max} evaluated at the same wavenumbers as for the data in Section 6.1. The power spectrum monopole contributions up to $\ell_{max} = 4$, quadrupole convolution up to $\ell_{max} = 6$, and hexadecapole convolution up to $\ell_{max} = 8$ are shown from top to bottom, respectively.

$\Phi_\ell(r)$, for a sparse and a dense survey configuration with $\sim$2–3 and $30\,$qso$\,$deg^{-2}, respectively. The window function for the dense configuration is normalized by its number density to illustrate that reducing the number of skewers only increases the noise in each r bin.

In Fig. A2, we consider the impact of window function convolution on the linear theory power spectrum. We show five cases in the Figure: (i) the raw linear theory power spectrum; (ii) the convolution with the pair truncation window function with $R_0 = 200\,h^{-1}\,$Mpc; (iii) the window-convolved spectra, defined in equations (20)–(22), including only ξ_0 contributions; (iv) as (iii), but adding ξ_2 contributions; and (v) as (iv) but adding ξ_4 contributions. Here, Γ^2 denotes the window matrix obtained from the survey geometry through pair-counting without the δ_F weights. It is interesting to note that the window pair separation function only affects very large scales below $k \lesssim 0.08\,h\,$Mpc^{-1}; this is as expected since the $W(r; R_0)$ function given in equation (7) was chosen to reduce aliasing (Philcox & Eisenstein 2020). For the monopole spectra (top panel), we find large contributions from the monopole and quadrupole windows *at all scales*, though negligible effects from the hexadecapole. In the middle panel, we show the power spectrum quadrupole with contributions up to $\ell_{max} = 6$ where the hexadecapole contributions are strongest and higher order terms have vanishing power. In the

Figure B1. Comparison of power spectra measured from 'raw' (solid coloured lines) and continuum-fitted 'eboss-0.0' (dashed coloured lines) LyaCoLoRe mocks shown for the monopole (black), quadrupole (red), and hexadecapole (green lines), respectively. The linear-theory Kaiser power spectrum, P^{th}, shown as a dashed–dotted grey line fits well the large scales of the 'raw' mocks. To illustrate the fidelity of the forward modelling of the distortion matrix, we show two variants of the analysis with distortion matrices computed for two different maximum separations: (i) $R_{DM} = 200\,h^{-1}\,$Mpc shown as dotted line in grey; (ii) $R_{DM} = 400\,h^{-1}\,$Mpc as dashed grey line.
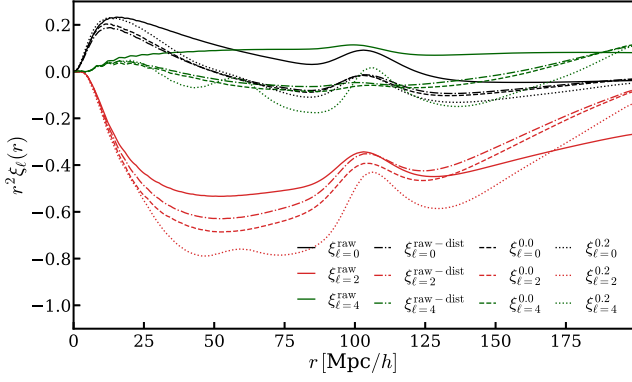
Figure B2. Effect of the distortion matrix and contaminants on the theory correlation function multipoles: monopole (black), quadrupole (red), and hexadecapole (green lines), respectively. The solid lines are the best-fitting correlation functions obtained from the mean correlation function measurements (dMdB20) which agree well with the mean measured power spectrum for each mock configuration in Fig. 3: solid lines correspond to the ones from 'raw' mocks, dashed–dotted lines from 'raw-dist' mocks including the mean subtraction displaying the effect of the distortion matrix, dashed lines from the 'eboss-0.0' mocks which include the effect of noise and continuum fitting through the distortion matrix and the dotted lines are the correlation function multipoles of the 'eboss-0.2' mocks which include contaminants and effects from continuum fitting.

level. For the quadrupole and the hexadecapole, we observe that the discrepancy between the distorted theory and the measurement for the 'eboss-0.0' mocks is reduced at the largest scales by increasing the computed maximum separation of the distortion matrix from $200\,h^{-1}\,\mathrm{Mpc}$ to $400\,h^{-1}\,\mathrm{Mpc}$ in $4\,h^{-1}\,\mathrm{Mpc}$ bins in $\{r_\parallel, \mathbf{r}_\perp\}$. We additionally tested that including higher order contributions of the theory correlation function, $\xi_\ell(r)$ up to $\ell_{max} = 6$, and including all the window matrix multipoles in equations (20)–(22), did have a negligible effect on the theory prediction of the multipoles. This additional level of complexity can be removed from the analysis pipeline by using a continuum fitting method that does not introduce distortions, e.g. a PCA-based continuum fitting at the expense of obtaining noisier continuum estimates.

For comparison, we show the effect of the distortion matrix and contaminants on the multipoles of the theory correlation function in Fig. B2 for the four LyaCoLoRe mock configurations. It is interesting to note that the application of the DM introduces a correlation between modes that are much further apart than a skewer
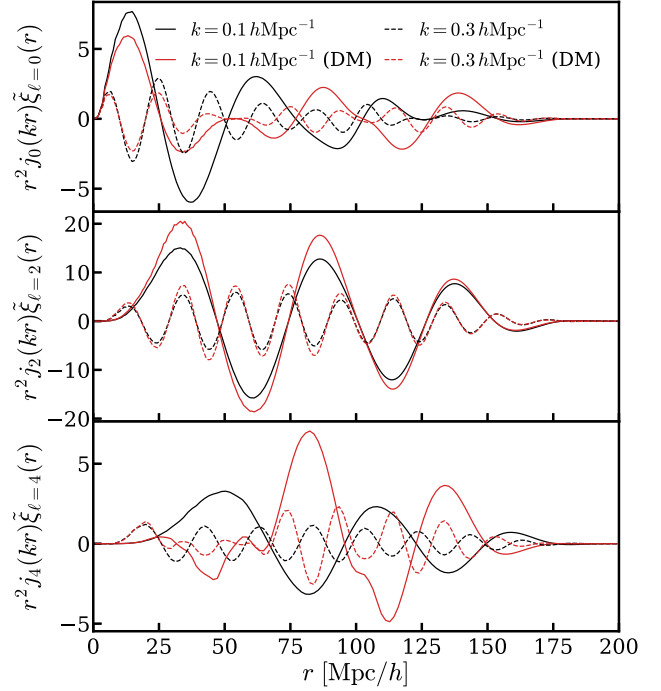

Figure B3. Integrand of the Hankel transform given in equation (23) from the multipole of the correlation function to the corresponding multipole of the power spectrum (monopole top, quadrupole centre, hexadecapole bottom). The pair truncation window $W(r; R_0 = 200\,h^{-1}\,\mathrm{Mpc})$ is applied to the theory correlation functions. The solid (dashed) lines are for $k = 0.1\,h\,\mathrm{Mpc}^{-1}$ ($k = 0.3\,h\,\mathrm{Mpc}^{-1}$); red (black) lines are with (with out) the distortion matrix.

length of a few hundred $h^{-1}\,\mathrm{Mpc}$'s, i.e. in Fig. B1 quadrupole modes at $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ and $k \sim 0.35\,h\,\mathrm{Mpc}^{-1}$ are both affected by the DM. We conjecture that for a given mode k the power spectrum quadrupole is sensitive to larger scales in the correlation function quadrupole than for the monopole. Therefore, we show in Fig. B3 the integrand of the Hankel transform of the window-convolved correlation function to the power spectrum multipoles for both wavenumbers $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ as a black solid and $k \sim 0.35\,h\,\mathrm{Mpc}^{-1}$ as a dashed line, respectively. The application of the distortion matrix (black to red) mixes a wide range of scales for quadrupole and hexadecapole.

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