UC Irvine
UC Irvine Previously Published Works

Title
DESI DR2 baryon acoustic oscillations from the Lyman alpha forest multipoles

Permalink
https://escholarship.org/uc/item/40x7r216

Journal
Physical Review D, 114(4)

ISSN
2470-0010

Authors
Karaçayl?, Naim Göksel
Cuceu, Andrei
Aguilar, J
et al.

Publication Date
2026-08-15

DOI
10.1103/yl3v-cyck

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

DESI DR2 Baryon Acoustic Oscillations from the Lyman Alpha Forest Multipoles

Naim Göksel Karaçaylı,[1,2,3,*] Andrei Cuceu,[4] J. Aguilar,[4] S. Ahlen,[5] S. Bailey,[4] S. BenZvi,[6] D. Bianchi,[7,8] A. Brodzeller,[4] D. Brooks,[9] T. Claybaugh,[4] A. de la Macorra,[10] B. Dey,[11,12] P. Doel,[9] J. Estrada,[13] S. Ferraro,[4,14] A. Font-Ribera,[15,16] J. E. Forero-Romero,[17,18] E. Gaztañaga,[19,20,21] S. Gontcho A Gontcho,[22] A. X. Gonzalez-Morales,[23] G. Gutierrez,[13] C. Hahn,[24] H. K. Herrera-Alcantar,[25,26] K. Honscheid,[1,3] C. Howlett,[27] M. Ishak,[28] R. Joyce,[29] R. Kehoe,[30] D. Kirkby,[31] T. Kisner,[4] A. Kremin,[4] O. Lahav,[9] M. Landriau,[4] J.M. Le Goff,[26] L. Le Guillou,[32] M. E. Levi,[4] M. Manera,[33,16] P. Martini,[1,2] A. Meisner,[29] R. Miquel,[15,16] J. Moustakas,[34] A. Muñoz-Gutiérrez,[10] S. Nadathur,[20] N. Palanque-Delabrouille,[26,4] W. J. Percival,[35,36,37] F. Prada,[38] I. Pérez-Ràfols,[39] G. Rossi,[40] E. Sanchez,[41] D. Schlegel,[4] M. Schubnell,[42,43] H. Seo,[44] J. Silber,[4] D. Sprayberry,[29] G. Tarlé,[43] B. A. Weaver,[29] R. Zhou,[4] and H. Zou[45]

[1] *Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[2] *Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, USA*
[3] *Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*
[4] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*
[5] *Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02215 USA*
[6] *Department of Physics & Astronomy, University of Rochester, 206 Bausch and Lomb Hall, P.O. Box 270171, Rochester, NY 14627-0171, USA*
[7] *Dipartimento di Fisica "Aldo Pontremoli", Università degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy*
[8] *INAF-Osservatorio Astronomico di Brera, Via Brera 28, 20122 Milano, Italy*
[9] *Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*
[10] *Instituto de Física, Universidad Nacional Autónoma de México, Circuito de la Investigación Científica, Ciudad Universitaria, Cd. de México C. P. 04510, México*
[11] *Department of Astronomy & Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada*
[12] *Department of Physics & Astronomy and Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA*
[13] *Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA*
[14] *University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA*
[15] *Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain*
[16] *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Edifici Cn, Campus UAB, 08193, Bellaterra (Barcelona), Spain*
[17] *Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia*
[18] *Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia*
[19] *Institut d'Estudis Espacials de Catalunya (IEEC), c/ Esteve Terradas 1, Edifici RDIT, Campus PMT-UPC, 08860 Castelldefels, Spain*
[20] *Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK*
[21] *Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain*
[22] *University of Virginia, Department of Astronomy, Charlottesville, VA 22904, USA*
[23] *Departamento de Física, DCI-Campus León, Universidad de Guanajuato, Loma del Bosque 103, León, Guanajuato C. P. 37150, México*
[24] *Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA*
[25] *Institut d'Astrophysique de Paris. 98 bis boulevard Arago. 75014 Paris, France*
[26] *IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*
[27] *School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*
[28] *Department of Physics, The University of Texas at Dallas, 800 W. Campbell Rd., Richardson, TX 75080, USA*
[29] *NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA*
[30] *Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA*
[31] *Department of Physics and Astronomy, University of California, Irvine, 92697, USA*
[32] *Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), FR-75005 Paris, France*

arXiv:2603.04281v2 [astro-ph.CO] 30 Jun 2026

[33] *Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain*
[34] *Department of Physics and Astronomy, Siena University, 515 Loudon Road, Loudonville, NY 12211, USA*
[35] *Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[36] *Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
[37] *Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*
[38] *Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain*
[39] *Departament de Física, EEBE, Universitat Politècnica de Catalunya, c/Eduard Maristany 10, 08930 Barcelona, Spain*
[40] *Department of Physics and Astronomy, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul 05006, Republic of Korea*
[41] *CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain*
[42] *Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, MI 48109, USA*
[43] *University of Michigan, 500 S. State Street, Ann Arbor, MI 48109, USA*
[44] *Department of Physics & Astronomy, Ohio University, 139 University Terrace, Athens, OH 45701, USA*
[45] *National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing, 100101, P. R. China*

(Dated: July 1, 2026)

We present an alternative measurement of the Baryon Acoustic Oscillation (BAO) using the Legendre multipole representation of the Lyα forest correlation functions from the second data release (DR2) of the Dark Energy Spectroscopic Instrument survey. Compressing the auto- and cross-correlation functions into Legendre multipoles yields a positive-definite covariance matrix without any smoothing—unlike the baseline DR2 analysis—thanks to a significantly reduced data vector size. We introduce the statistical corrections required to debias the finite-sample covariance matrix estimate and demonstrate that monopole and quadrupole terms for both auto- and cross-correlations can be used even when the correlation functions are distorted by continuum errors and contaminated by metals. This formalism has slightly diminished the constraining power of the BAO scale, while considerably weakening constraints on nuisance parameters. We measure the isotropic BAO scale with 0.93% precision at $z_{\rm eff} = 2.35$, the Hubble parameter $H(z_{\rm eff}) = (239.5 \pm 3.4)$ (147.09 Mpc/r_d) km s^{-1} Mpc^{-1}, and the transverse comoving distance $D_M(z_{\rm eff}) = (5.80 \pm 0.10)$ (r_d/147.09 Mpc) Gpc for a given value of the sound horizon (r_d). Our BAO results are entirely consistent with the baseline DR2 analysis.

I. INTRODUCTION

The Dark Energy Spectroscopic Instrument (DESI; [1–4]) is currently delivering measurements of the baryon acoustic oscillation (BAO) scale with sub-percent precision [5–8]. In combination with the cosmic microwave background (CMB) and type Ia supernovae (SNe), the Data Release 2 (DR2) results of DESI, with over 14 million galaxies and 1.2 million quasars, favor "dynamical" dark energy models instead of the cosmological constant [8–10]. As DESI observations continue to improve, we revisit the choices for the Lyman-α (Lyα) forest analysis, specifically the dimensionality of the correlation function covariance matrix, to ensure the robustness of current and future DESI results.

The Lyα forest is the collection of the absorption lines of neutral hydrogen gas at wavelengths shorter than the Lyα emission line. It provides a unique window into the matter distribution and the tightest constraints on the cosmic expansion rate at the redshift range $2 < z < 4$ [6, 7]. The BAO scale is extracted from the three-dimensional auto-correlation of Lyα forest (Lyα×Lyα) and the cross-correlation between Lyα forest and quasars (Lyα×QSO) [6, 7, 11–14]. Beyond the BAO feature, the full shape of the correlation function carries additional cosmological information in the form of the Alcock-Paczynski (AP) effect [15] and redshift space distortions [16]. Including these leads to tighter constraints on the cosmic expansion rate and a measurement of the growth rate of cosmic structure, $f\sigma_8$ [17–20].

The standard approach measures the Lyα forest correlation functions in bins of transverse and line-of-sight separations, $(r_\perp, r_\parallel)$. This basis is better suited to describe distortions due to quasar continuum-fitting errors and metal-line systematics, which predominantly affect small $r_\perp$ scales. However, it yields a larger data vector (2,500 elements for autocorrelations and 5,000 elements for cross-correlations) than the independent realizations possible with subsampling of real or mock data. This poses a significant challenge for covariance matrix estimation, since at least as many independent samples as the data vector size are needed to suppress random noise in the estimated covariance. In practice, an ad-hoc smoothing is applied to the noisy covariance esti-

* karacayli.1@osu.edu

mates in $(r_\perp, r_\parallel)$, which is shown to be adequate for the current precision level of the data [21, 22]. However, this smoothing procedure is not guaranteed to work in the high signal-to-noise regime. For example, Fig. 11 of Cuceu *et al.* [21] visually indicates that the off-diagonal structure of the Ly$\alpha\times$QSO covariance matrix changes with smoothing. Additionally, alternative continuum fitting methods [e.g., 23]) in which the large-scale information is preserved rather than projected out may manifest a different structure than the one assumed in smoothing [24].

To address these limitations, we explore an alternative measurement scheme that dramatically reduces the dimensionality of the correlation function data vector by decomposing it into Legendre multipoles. Multipole compression is a well-established technique in galaxy and quasar clustering analyses [5, 8, 25] and has been previously considered in the context of Lyα forest correlations [11]. Nevertheless, it has been set aside due to its limitations in describing distortions arising from continuum fitting errors and metal contamination. For example, ionized silicon systems cause offsets in the correlation function most strongly along the line of sight (small $r_\perp$). In principle, a multipole representation that uses all components would work equally well. However, when the multipoles are truncated to $\ell \leq 4$, which is our main focus for reducing the data size, this effect is diluted by angular integration. Despite these challenges, we argue that the reduced data dimensionality improves the covariance matrix estimation and makes the multipole-based analysis a valuable alternative. By connecting Lyα forest analyses to the well-developed galaxy clustering literature, this approach has the potential to enhance the robustness of future DESI Lyα forest results.

This paper is organized as follows. We provide an overview of the DR2 Lyα forest data, including catalogs of astrophysical contaminants such as damped Lyα absorption systems and broad absorption line quasars, in Sec. II. Sec. III describes our method for measuring the correlation function multipoles and the necessary changes to the existing pipeline. This section also includes a summary of the modeling of correlations. We validate our method, investigate the optimal number of multipoles, and the goodness of fit using ten mock realizations in Sec. IV. We present our BAO measurements using this technique and compare them with the baseline DR2 results in Sec. V. We discuss advantages, disadvantages, and possible extensions in Sec. VI.

II. DATA

Our sample is the processed data from the first three years of the DESI survey, which will be publicly available in the future as DR2. DESI DR2 covers a larger area than DR1 [26] and has nearly 1.3 million quasar spectra with $z > 1.77$ and 825,000 with $z > 2.09$, typically observed 2–3 times. This dataset was used in the key science paper measuring the BAO scale from the Lyα forest [7, DESI-DR2-Lyα from now on]. The instrumentation and connection to the science requirements are overviewed in ref. [4], with additional papers detailing the robotic focal plane assembly [27], the wide-field corrector system [28], the highly-efficient fiber system [29], and the survey operations [30]. A sophisticated data processing pipeline helps to minimize instrumental systematic errors [31]. Multiple observations of a given quasar are coadded and organized by their `HEALPix` pixel [32]. This pixelization scheme is also used to split correlation function measurements into subsamples when estimating the covariance matrix.

The Lyα forest measurements are contaminated by several astrophysical effects. Among them, damped Lyα absorption (DLA) systems and broad absorption lines (BALs) are identified and masked out from the analysis. DLAs have neutral hydrogen column densities $N_{\rm HI} > 2 \times 10^{20}$ cm^{-2} and are denser objects that are more strongly clustered than the Lyα forest [33]. They are also saturated absorption lines with damping wings extending to large separations, which impact the continuum level. DLAs are identified, and their column densities and redshifts are measured through a combination of a template-based approach [34], a convolutional neural network [35], and a Gaussian process [36] method.

BAL quasars have absorption features that arise from quasar outflows and are unrelated to the matter distribution in the intergalactic medium. These are identified in the same manner as the DESI DR1 analysis [6]. Ref. [37] describes the method in detail and demonstrates that incompleteness in the BAL catalog has minimal impact on the BAO measurements.

We use the same forest measurements of DESI-DR2-Lyα that are in the 3600–5772 Å wavelength range. The astrophysical contaminants, such as DLAs and BALs, cosmic rays, and major sky emission lines, are masked before the quasar continuum is estimated. Forests that are shorter than 120 Å are discarded to ensure a sufficient length in continuum fitting. The transmitted flux fluctuations $\delta_q(\lambda)$ for each quasar q are given by

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

where $\overline{F}$ is the mean intergalactic medium (IGM) transmission and C_q is the quasar continuum. The fitting procedure is described in detail by ref. [38] and performed in the forest region. This removes the underlying large-scale density field from each forest, which distorts the measured Lyα clustering statistics through mixing of short- and long-range correlations.

We analyze only Region A, which is defined as the 1040–1205 Å wavelength range of the quasar rest-frame, and exclude Region B (920–1020 Å) from our analysis. This is to ensure the size of our data vector remains small and to ascertain that the covariance matrix can be reliably estimated from the sub-samples. Excluding the B region has a minor impact, which will be demonstrated

in Sec. IV, since the measurements in this region are lower in SNR and contaminated by higher-order Lyman transitions. Nevertheless, Region B can be reintroduced through more advanced techniques of covariance matrix estimation in future work.

III. METHOD

In order to measure the multipoles of the correlation function, we first make minor adjustments to our previous approach. In DESI-DR2-Lyα (and earlier work [6, 14, 39]), we employed a rectangular grid in directions parallel ($r_\parallel$) and transverse ($r_\perp$) to the line-of-sight in clustering measurements of the Lyα forest. For a pixel pair (i, j) with angular separation θ_{ij}, these are defined as

$$r_\parallel = (D_C(z_i) - D_C(z_j))\cos(\theta_{ij}/2) \quad (2)$$

$$r_\perp = (D_M(z_i) + D_M(z_j))\sin(\theta_{ij}/2), \quad (3)$$

where D_C is the comoving distance and D_M is the transverse comoving distance, which are identical for a flat ΛCDM cosmology. We adopt the same fiducial cosmology as the DESI-DR2-Lyα, which is based on ref. [40].

We first modify the grid to measure the anisotropic two-point correlation function $\xi(r, \mu)$ at a separation distance $r = \sqrt{r_\parallel^2 + r_\perp^2}$ and $\mu = r_\parallel/r$. The estimator is the same weighted average for each (r, μ) bin, where the effective $r_{\rm eff}$, $\mu_{\rm eff}$, and $z_{\rm eff}$ are calculated using the same weights (a smoothed model is evaluated at these effective points as discussed later).

We use 50 r-bins of 4 h^{-1} Mpc up to 200 h^{-1} Mpc and 100 μ bins in $[0, 1]$ for the auto-correlations and 200 μ bins in $[-1, 1]$ for cross-correlations. The distortion matrices are calculated using the same bin sizes up to 300 h^{-1} Mpc to capture the contributions from larger separations.[1,2] Then, we project $\xi(r, \mu)$ onto the Legendre polynomial basis:

$$\xi_\ell(r) = \frac{2\ell + 1}{2} \int \mathrm{d}\mu\ \xi(r, \mu)\mathcal{L}_\ell(\mu), \quad (4)$$

where $\mathcal{L}_\ell(\mu)$ is the Legendre polynomial. Following ref. [5], this integral is computed as a finite sum over μ-bins, weighted by the integral $\frac{2\ell+1}{2}\int \mathrm{d}\mu\mathcal{L}_\ell(\mu)$ within each bin. This projection is performed in the `Vega` package[3], which models the correlations and infers parameters. It follows that the distortion and covariance (without smoothing in our case) matrices, as well as the model, are calculated as before in the full (r, μ) grid, and then projected onto the Legendre basis in a coherent manner. These minimal changes seamlessly integrate multipoles into our existing pipeline.

Our approach differs from ref. [11] in which the multipoles are calculated through a fit. Although this approach may enhance the signal by incorporating weights into the multipole compression, it also compromises the orthonormality of the multipoles, complicates the window function, and hinders a precise connection to the model, since the parameters differ across fitting setups. These least-squares solutions may be preferred in the low-signal regime, but are not needed at the statistical power of DR2.

The multipole representation of the correlation function dramatically reduces the dimension of our data vector, which forms the basis of our covariance matrix study. We estimate the covariance matrix using the sample covariance estimator over `HEALPix` pixels as before:

$$C_{AB} = \frac{\sum_h w_A^h w_B^h (\xi_A^h - \bar{\xi}_A)(\xi_B^h - \bar{\xi}_B)}{\sum_h w_A^h \sum_h w_B^h}, \quad (5)$$

where the indices A and B represent (r, μ) bins. This estimator is formally biased and needs a sample bias correction. However, as we show in Appendix A, this correction is insignificant ($\sim 0.1\%$) and can be ignored for DR2. This covariance matrix is then compressed in both axes to obtain the covariance of multipoles: $\mathbf{C} \to \mathbf{P}_\ell \mathbf{C} \mathbf{P}_\ell^{\rm T}$, where $\mathbf{P}_\ell$ is the matrix representation of the finite-sum Legendre decomposition. We choose `nside` $= 16$ such that the average size of a pixel (250 h^{-1} Mpc at $z = 2.3$) is larger than the largest separation used in the correlation function (180 h^{-1} Mpc), resulting in $N_h = 1147$ "independent" samples for DR2.

The $(r_\perp, r_\parallel)$ binning yields a correlation function with over 7000 elements, the covariance of which becomes noisy since it cannot be reliably estimated using only about 1000 samples. In the baseline DESI analyses, we smooth this noisy covariance estimate by averaging the non-diagonal elements of the correlation matrix $(C_{AB}/\sqrt{C_{AA}C_{BB}})$ of the same $|r_\parallel^A - r_\parallel^B|$ and $|r_\perp^A - r_\perp^B|$ differences. In parallel, our formalism in (r, μ) binning still yields a noisy covariance matrix. In contrast, however, the Legendre compression of this matrix ($\mathbf{P}_\ell \mathbf{C} \mathbf{P}_\ell^{\rm T}$) is a well-motivated "averaging" operation and not an assertion about the correlation matrix's structure. This allows us to focus on the corrections to the covariance matrix below.

Using the sample mean and covariance estimates of a Gaussian-distributed dataset for parameter inference leads to a non-Gaussian likelihood function with heavier tails, similar to the Student's t-distribution [42]. To approximate this widening in the distribution, we multiply our covariance matrix by the inverse Hartlap factor [43] and keep our likelihood Gaussian following ref. [10]:

$$C_{\rm Hartlap} = \frac{N_h - 1}{N_h - N_b - 2}, \quad (6)$$

[1] Including separations $r_\parallel > 300$ h^{-1} Mpc may improve the cross-correlation accuracy (see ref. [41]). We prefer following the DESI-DR2-Lyα as closely as possible in this work.

[2] The window function can be more accurately propagated by estimating the correlation function in narrower r bins as done in DESI's galaxy and quasar clustering analysis [5, 8, 26]. However, the distortion matrix presents additional complications and increases the computational time.

[3] `https://github.com/andreicuceu/vega`

where N_h is the number of `HEALPix` samples, N_b is the number of bins, and $\mathbf{C} \to C_{\rm Hartlap}\mathbf{C}$. Furthermore, we propagate the uncertainty in the covariance matrix to the inferred parameters by multiplying the *parameter* covariance matrix by the following correction factor:

$$C_{\rm Percival} = \frac{1 + C_1(N_b - N_{\rm params})}{1 + C_0 + C_1(N_{\rm params} - 1)}, \tag{7}$$

where $N_{\rm params}$ is the number of inferred parameters and

$$C_0 = \frac{2}{(N_h - N_b - 1)(N_h - N_b - 4)} \tag{8}$$

$$C_1 = \frac{N_h - N_b - 2}{2} C_0 \tag{9}$$

as introduced by ref. [44]. $C_{\rm Hartlap}$ and $C_{\rm Percival}$ factors range between $1.1 - 1.25$ in our work.

When $N_h < N_b$ as is the case in DESI-DR2-Lyα, these corrections become negative, hence inapplicable. This is another manifestation of insufficient samples for covariance matrix estimation, alongside non-positive-definite outcomes. However, these corrections themselves are approximations of the underlying true likelihood, not exact quantifications of biases in covariance matrix estimates. Similarly, the smoothing procedure in DESI-DR2-Lyα serves as an approximation for the true likelihood and covariance matrix, replacing these corrections. It has been shown that this procedure yields unbiased errors on the fitted parameters using hundreds of survey realizations [21, 22]. We are concerned that it is unclear when this approximation will break down and what to replace it with, and we find that the truncated multipole representation offers a promising avenue.

A. Modeling the correlations

In this section, we provide an overview of our model, which was applied to DESI DR1 and DR2 Lyα forest analyses to extract the BAO scale [6, 7] as well as the full-shape information [20]. The Lyα forest correlation model is fundamentally based on the linear theory, where non-linearities are modeled with an empirical model built from simulations. The major modeling difficulties arise from distortions due to continuum fitting and contaminations from metal and high-column-density systems. As to our multipole representation, the full model is first computed on an (r, μ) grid, as in previous DESI analyses, and then projected onto the Legendre multipole basis, as mentioned in the previous section. This ensures that the model and data are consistently handled and the window function is accurately propagated. This model is publicly accessible in the `Vega` software package.

We calculate a template isotropic linear matter power spectrum using `CAMB` [45] at the fiducial cosmology [40]. The BAO information is isolated by decomposing this into peak and smooth components, while applying a Gaussian smoothing to the peak component to model the non-linear broadening of the BAO feature as described by ref. [46]:

$$P_{\rm QL}(k_\perp, k_\parallel) = P_{\rm peak}(k) e^{-(\Sigma_\perp^2 k_\perp^2 + \Sigma_\parallel^2 k_\parallel^2)/2} + P_{\rm sm}(k). \tag{10}$$

We introduce two BAO scaling parameters $(\alpha_\perp, \alpha_\parallel)$ that multiply the coordinates in transverse and parallel directions only for the peak component of the correlation function. As a deviation from previous DESI analyses, we directly measure the isotropic scaling parameter $\alpha_{\rm iso} = (\alpha_\parallel \alpha_\perp^2)^{1/3}$ and the Alcock-Paczynski parameter $\alpha_{\rm AP} = \alpha_\parallel / \alpha_\perp$, since $\alpha_{\rm iso}$ is the natural parameter constrained by the monopole. The $\alpha_\perp$ and $\alpha_\parallel$ parameters are derived from $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$.

For two tracers X and Y, the anisotropic power spectrum is built on the Kaiser kernel at linear order $K_i(\mu) = b_i(1 + \beta_i \mu^2)$ [16], where $\mu \equiv k_\parallel / k$:

$$P_{XY} = K_X(\mu) K_Y(\mu) P_{\rm QL}(k, \mu) F_{\rm NL}^{XY}(k, \mu) G(k, \mu), \tag{11}$$

where $F_{\rm NL}^{XY}$ is the aforementioned non-linear fitting fuction. There are two possibilities for XY combinations in our work. For Ly$\alpha\times$Lyα auto-correlations, we use the fitting function in ref. [47] for $F_{\rm NL}^{\alpha\alpha}$. For Ly$\alpha\times$QSO cross correlations, we model the impact of quasar redshift errors and fingers-of-god effect due to peculiar velocities using a Lorentzian ($F_{\rm NL}^{Q\alpha} = \sqrt{1 + (k_\parallel \sigma_v)^2}$) for real data following DESI-DR2-Lyα, but using a Gaussian ($F_{\rm NL}^{Q\alpha} = \exp\left[-(k_\parallel \sigma_v)^2/4\right]$) for mock data following ref. [22]. When the tracer refers to quasars, the RSD parameter is given by $\beta_Q = f/b_Q$, where f is the linear growth rate. Following previous Lyα forest BAO analyses [7], we fix f to its fiducial value and marginalize over b_Q.[4]. On the other hand, the Lyα RSD parameter has an extra unknown velocity divergence bias, b_η, and is given by $\beta_\alpha = b_\eta f / b_\alpha$ [see e.g., 48, 49]. Therefore, in this case, we marginalize over both b_α and β_α, following [7].

The smoothing due to binning the correlation function is incorporated through the $G(k)$ term. We include the smoothing of radial bins by multiplying the model spectrum by $G(k) = {\rm sinc}(k\Delta r/2)$, which is a sufficient approximation for our investigation, as validation tests prove. The smoothing in μ bins is already included in the finite sum of the Legendre decomposition. We obtain the correlation function by transforming the first four even multipoles of the anisotropic power spectrum in Eq. (11) using the FFTLog algorithm [50]. The correlation function on the (r, μ) grid is then reconstructed from these multipoles, which will be reprojected onto the Legendre multipole basis after accounting for systematics and distortions.

Major model complications associated with the multipole representation arise from (1) residual HCDs, (2)

[4] Previous analyses have found this does not have a significant impact on BAO measurements [6], and measuring this parameter is generally only done in full-shape analyses [e.g., 20]

metal contamination, and (3) distortions resulting from biased continuum estimates. HCDs are tracers of overdense regions of the matter field with a different bias, $b_{\rm HCD}$, than the forest. They exhibit saturated absorption lines with long damping wings, leading to a suppression in power along the line of sight. The linear order model follows a simple rescaling of b_α and β_α:

$$b_\alpha \to b_\alpha + b_{\rm HCD} F_{\rm HCD}(k_\parallel), \tag{12}$$

$$b_\alpha\beta_\alpha \to b_\alpha\beta_\alpha + b_{\rm HCD}\beta_{\rm HCD} F_{\rm HCD}(k_\parallel), \tag{13}$$

where we approximate the suppression in the line of sight with $F_{\rm HCD}(k_\parallel) = \exp(-L_{\rm HCD}k_\parallel)$, which is the Fourier transform of the Lorentzian broadening in the Voight profile, and $L_{\rm HCD} \sim 5\ h^{-1}$ Mpc is the typical length of the HCD absorption [14, 51–53]. We note that the BAO broadening in eq. (10) is also applied to HCDs. In configuration space, the $F_{\rm HCD}(k_\parallel)$ term smooths the 2D correlation function, both the peak and smooth components, along the line-of-sight direction with a scale set by $L_{\rm HCD}$. In order to detect this change in shape, one needs high enough precision and/or sufficiently dense sampling of the correlation function. Consequently, our typical $\Delta r = 4\ h^{-1}$ Mpc bin size effectively sets a lower threshold for detecting HCDs even for full 2D analyses. Furthermore, we expect our multipole representation to be significantly less effective at constraining these HCD parameters. In fact, we fail to detect $b_{\rm HCD}$ in the real data application as discussed in Sec. V.

In terms of the metal contamination, we follow the recipe in refs. [6, 7]. Metal systems are misplaced in the radial direction because the Lyα transition wavelength is used as opposed to their true transition wavelength. This results in an offset in the $r_\parallel$ direction of the correlation function. Our recipe loops over all possible pairs and computes the relative weight of pairs with true separations in bin B among those with misestimated separations in bin A. These weights construct a metal matrix, M_{AB}, in the $(r_\perp, r_\parallel)$ grid that maps "true" correlation function to a measured correlation, $\xi_A^{\rm meas} = \sum_B M_{AB}\xi_B^{\rm true}$, for each pair of transitions. We then interpolate $\xi^{\rm meas}$ from the $(r_\perp, r_\parallel)$ grid onto the (r, μ) grid.

The continuum fitting procedure introduces distortions to measured correlation functions [39, 41, 54]. These are modeled with a distortion matrix, $\mathbf{D}^{XY}$, that multiplies the correlation function: $\boldsymbol{\xi}^{XY}_{\rm dist} = \mathbf{D}^{XY}\cdot\boldsymbol{\xi}^{XY}_{\rm undist}$. This distortion matrix varies for each XY combination and is computed from the dataset on the (r, μ) grid, reflecting the weights and redshift distribution of the sample. We use the same resolution as the correlation function while computing $\mathbf{D}^{XY}$, i.e., $\Delta r = 4\ h^{-1}$ Mpc spacing. However, we compute $\mathbf{D}^{XY}$ up to $r_{\rm max} = 300\ h^{-1}$ Mpc as noted in the beginning of Sec. III. Lastly, the projection onto the Legendre multipoles is represented by another matrix operation, $\mathbf{P}_\ell^{XY}$, with $\Delta\mu = 0.01$. This operation is combined with the distortion matrix, so that our final model vector is given by $(\mathbf{P}_\ell\mathbf{D})\cdot\boldsymbol{\xi}_{\rm undist}$.

The undistorted model can also be represented by Legendre multipoles with the addition of an interpolation matrix in future work. The highly anisotropic nature of metal contamination makes it difficult to determine how many multipoles would be needed in this case. There are other systematics (e.g., correlated noise from data processing [55] and quasar redshift errors [56]) that we have not detailed in this section, which require similar studies.

IV. VALIDATION WITH DR2 MOCKS

The DR2 Lyα forest BAO measurement validation was built on the already extensive work for DR1 [21]. In DR2, we generated 400 synthetic realizations of the dataset using two different methods: 100 realizations with the `Saclay` method [57] and 300 with the `CoLoRe-QL` method. The `CoLoRe-QL` method has improved quasar clustering on small scales and captures the BAO broadening due to non-linear structure growth, and is an improvement of the `LyaCoLoRe` method employed in DR1 [58, 59]. These thorough validation efforts are presented in the DR2 companion paper [22]. Building on this foundation, we use only the first ten of the `CoLoRe-QL` mocks to demonstrate that the multipole representation recovers most of the information without bias.

As mentioned before, we operate only in the A region and measure the Lyα forest auto-correlation (Ly$\alpha\times$Lyα) and the Lyα forest-quasar cross-correlation (Ly$\alpha\times$QSO). For our multipole compression, we consider two cases for both correlations: 1) The $\ell_{\rm max} = 2$ case refers to using the monopole and quadrupole, and 2) the $\ell_{\rm max} = 4$ case refers to including the hexadecapole. We find that the dipole term for Ly$\alpha\times$QSO is featureless and lacks the BAO peak; therefore, we do not include it. We will discuss what dipole has to offer in Sec. VI, including which parameters are improved in a variation on real data. In short, the dipole does not improve the BAO constraints as expected, but provides some improvement on metal biases. Throughout this section (and this work), we analyze the correlation function in the range $30\ h^{-1}$ Mpc $< r < 180\ h^{-1}$ Mpc following DESI-DR2-Lyα.

A. Covariance matrix

The average of ten `CoLoRe-QL` covariances for the $\ell_{\rm max} = 2$ case is shown in the top panel of Fig. 1 as a correlation matrix. This and the $\ell_{\rm max} = 4$ case produce a positive definite covariance matrix without any smoothing, even for one realization, as we sought. The bottom panel of this figure shows the correlation matrix obtained from actual data. We find that the median absolute difference between the two correlation matrices is 3%, indicating that the mocks reasonably capture the expected correlations in the data.

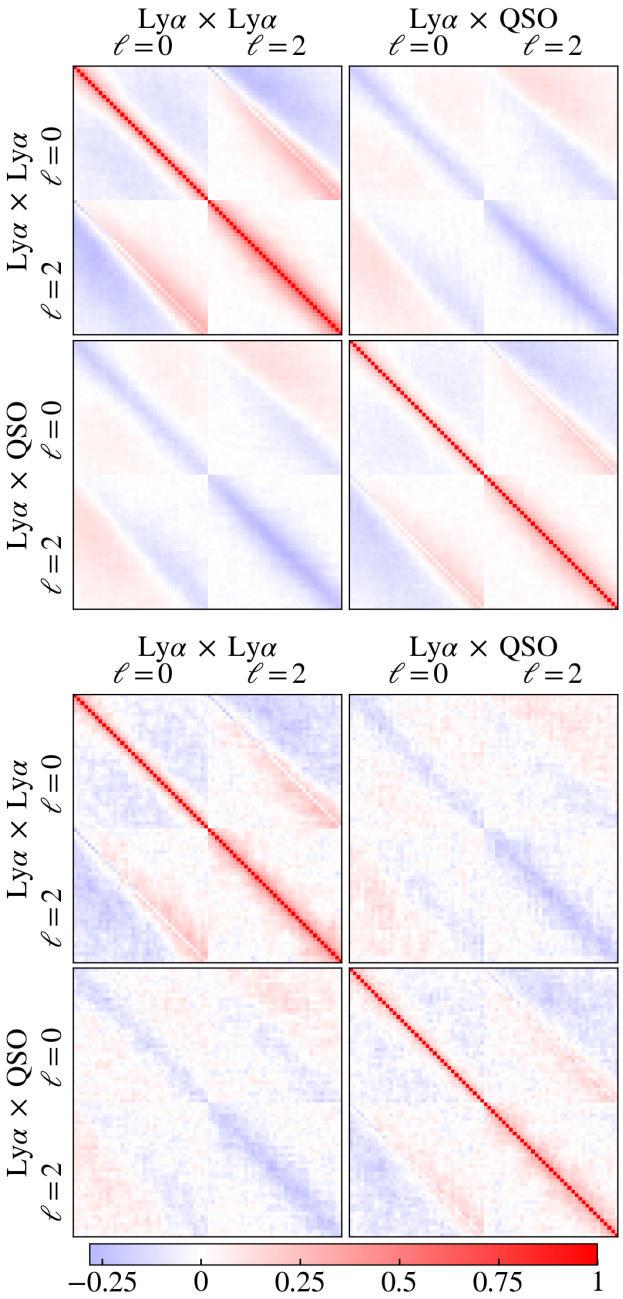


FIG. 1. The average correlation matrix of ten `CoLoRe-QL` mocks (*top*) and the correlation matrix obtained from data (*bottom*). This demonstrates that mocks reasonably capture the correlations in the data covariance matrix. However, the variance in the data is considerably larger than in the mocks. Both matrices are positive definite without any ad hoc smoothing.

However, the amplitude of the diagonal is off by a large factor — on average, the Ly$\alpha\times$Lyα variance is 75% larger, and the Ly$\alpha\times$QSO variance is 30% larger in data estimates before the 15% Hartlap correction (which is negligible at the 1% level for the average of ten mocks). The amplitude offset is even larger for the $(r_\perp, r_\parallel)$ binning, by 95% and 40%, respectively. These offsets suggest that, if mock realizations are to be used to improve covariance matrix estimation in the future, the amplitude of the variance needs to be scaled to match the data variance. Lastly, for this section, the data covariance matrix remains positive definite, although it is noisier.

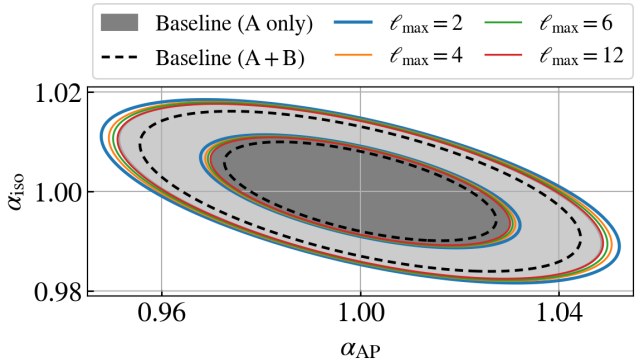


FIG. 2. BAO constraining power for various analysis choices when modeling errors and noise in measured correlations are eliminated using a data vector at the fiducial model. The constraints weaken slightly when the B region is excluded (*gray shaded*) from the DR2 baseline analysis (*black dashed lines*), and weaken slightly more when using Legendre multipoles $\ell_{\rm max} = 2$ (*blue solid lines*). The baseline and Legendre multipole contours converge at $\ell_{\rm max} = 12$.

B. BAO constraining power

Let us begin by assessing the potential information loss due to the projection from the full grid $(r_\perp, r_\parallel)$ to Legendre multipoles. We first eliminate modeling error complications and stochastic noise in measured correlations by generating a data vector at the fiducial model. In order to judge both methods on equal footing, we use covariance matrices derived from mock data, since larger statistics mitigate the non-Gaussianity of the likelihood function, and we ignore the Hartlap and Percival corrections.

Fig. 2 demonstrates that the information loss in the BAO parameters is small in the Legendre multipole space, and the monopole and quadrupole contain most of the BAO constraining power of the data. The baseline DR2 analysis, which includes the B-region, constrains $\alpha_{\rm iso}$ with 0.68% precision and $\alpha_{\rm AP}$ with 1.8% precision. These deteriorate by 10% for an A-region-only analysis. A monopole and quadrupole-only analysis (using just the A-region) does minimally worse compared to this, where $\alpha_{\rm iso}$ precision worsens by an additional 4% and $\alpha_{\rm AP}$ by 7%. Including the hexadecapole improves both BAO constraints by only 2% and 4%, respectively, while equivalence between the baseline and truncated multipole analyses is reached at a very high $\ell_{\rm max} = 12$. However, the resulting data vector is still approximately seven times smaller than the baseline data vector. For completeness,

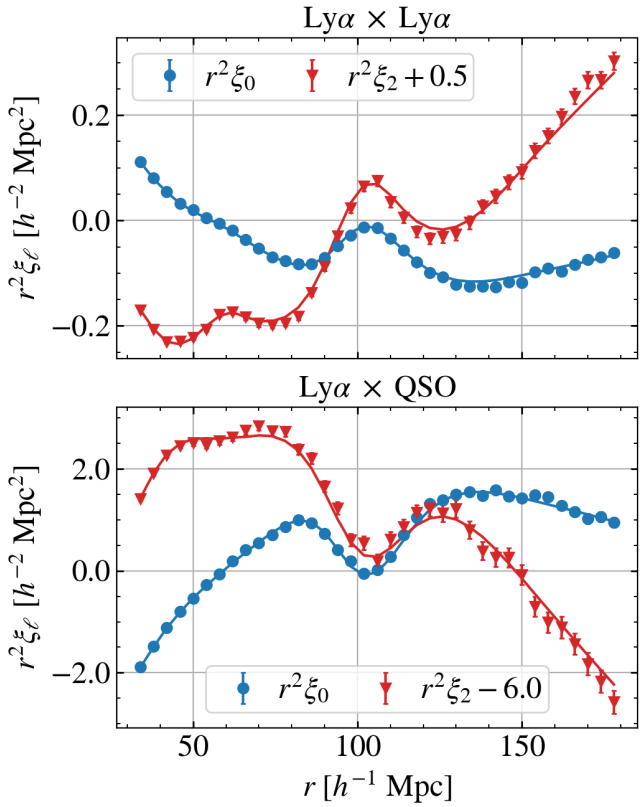


FIG. 3. Average of ten Lyα auto (*top*) and Lyα×QSO cross (*bottom*) correlation functions measured from the mocks. The quadrupoles are shifted to match the monopole scale. The $\ell_{\rm max} = 2$ best-fit model (*solid lines*) provides a good fit with PTE of 22%, whereas the $\ell_{\rm max} = 4$ has a near-zero PTE.

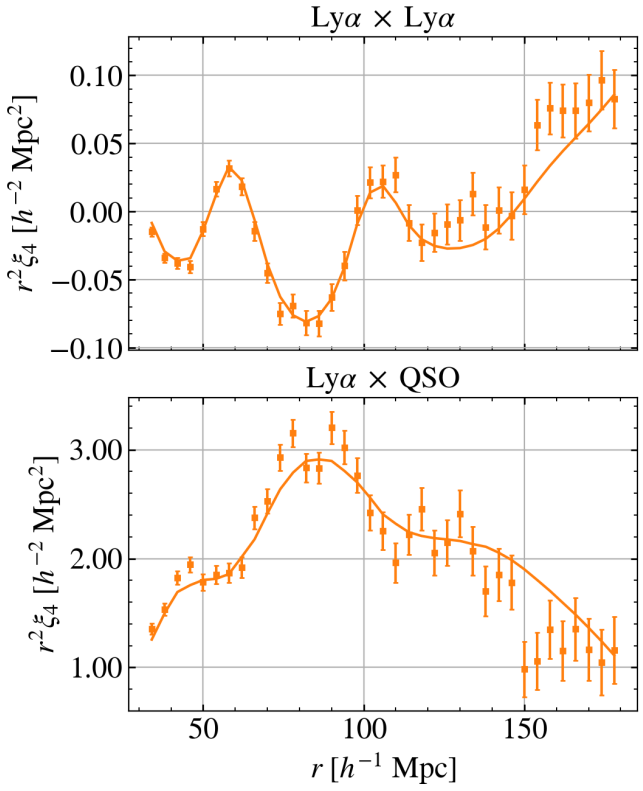


FIG. 4. Average of ten Lyα auto (*top*) and Lyα×QSO cross (*bottom*) hexadecapoles measured from the mocks. The best-fit model (*solid lines*) yields a reasonable $\chi^2_\nu = 36.4/37$ for the auto, while a poor $\chi^2_\nu = 72.0/37$ for the cross correlation function. Including both in the fit yields a PTE of zero. Including only the auto-correlation function still performs considerably worse than the $\ell_{\rm max} = 2$ case with a PTE of 4%.

the $\ell_{\rm max} = 2$ case constrains $\alpha_{\rm iso}$ with 0.75% precision and $\alpha_{\rm AP}$ with 2.1% precision, and is expected to perform worse by 14.5% and 17.6% than the baseline DR2 choices.

However, the constraints on all nuisance parameters become less stringent in the multipole representation with $\ell_{\rm max} = 2$, which, in particular, underperforms at constraining the bias parameters for metal absorbers. We illustrate the correlation function with and without metal contamination in Appendix B, which shows that the BAO peak vanishes for $\ell \geq 4$ from the correlation function, but metal peaks remain prominent. This explains how the metal contamination can be effectively isolated and quantified using higher-order multipoles. Including the hexadecapole improves these constraints by 40% for the Si III (1207 Å) bias parameter and by 30% for the remaining metal biases, but they are not as good as the baseline DR2 case. Furthermore, the parameter covariance matrix indicates slightly stronger correlations between BAO (most notably $\alpha_{\rm AP}$) and nuisance parameters. The correlation coefficient between $\alpha_{\rm AP}$ and the bias parameter for the Si II (1260 Å) line rise in absolute value from −7.4% for the baseline analysis to −15% for the $\ell_{\rm max} = 2$ case, and to −19% for the $\ell_{\rm max} = 4$ case. This line creates a spurious peak in the correlation function at 104 h^{-1} Mpc, particularly near the BAO peak, and its correlation with $\alpha_{\rm AP}$ may pose a challenge for future analyses that use multipole decomposition.

C. Modeling accuracy

We now turn to testing whether our modeling of the Lyα forest multipoles is sufficiently accurate to yield an unbiased BAO measurement. We average ten correlation functions measured from the mocks to obtain a data vector that is ten times as powerful as DR2 (which would translate to $1/\sqrt{10}$ scaling in standard deviation). Similar to ref. [22], we set our threshold for $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$ to be equal to one within one-third of the actual DR2 standard deviation, which we measure using the $\ell_{\rm max} = 2$ setting. One-ninth of the covariance matrix derived from the minimizer establishes the contour for our validation test.

The $\ell_{\rm max} = 2$ case provides a reasonable fit to this mock data vector with $\chi^2_\nu = 148.5/(148 - 12) = 1.092$ and a probability of having a larger value (PTE) of 22%.

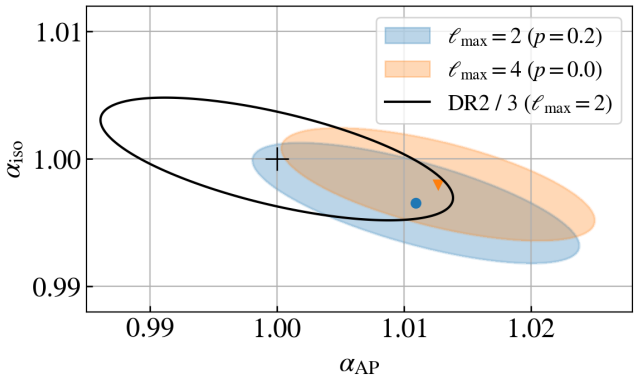


FIG. 5. The best-fit BAO parameters to the average of ten correlation function measurements from the mocks. Both the $\ell_{\rm max} = 2$ case (*blue*) and $\ell_{\rm max} = 4$ case (*orange*) yield unbiased values, that is, one within one-third of the DR2 uncertainty using $\ell_{\rm max} = 2$ (*black line*).

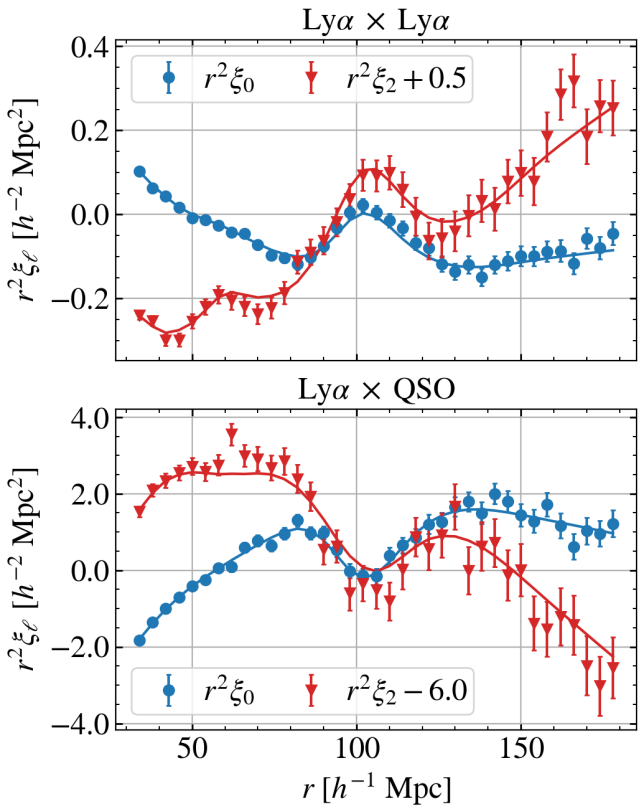


FIG. 6. DESI DR2 best-fitting model vs data for Lyα auto (*top*) and Lyα×QSO cross (*bottom*) correlation functions. The $\ell_{\rm max} = 2$ model provides an excellent fit with a PTE of 53%.

Fig. 3 shows the best-fitting model against the mock data points. Lyα×Lyα multipoles are on the top panel, and Lyα×QSO multipoles are on the bottom panel. Including the hexadecapole significantly degrades the goodness of fit, with $\chi^2_\nu = 287.2/(222 - 12) = 1.37$, which has nearly zero probability (0.03%) of being larger than this value. Fig. 4 displays the hexadecapole for the auto and cross correlation functions with respective best-fitting model curves in this case. This extremely poor fit is largely caused by the Lyα×QSO hexadecapole with $\chi^2_\nu = 72/37 = 1.95$. Removing only this component improves the probability to 4%, which is still not as good as using only the monopole and quadrupole.

Finally, Fig. 5 demonstrates that we recover unbiased BAO parameters for both cases. The $\alpha = 1$ point remains within the one-sigma region for the $\ell_{\rm max} = 2$ case, whereas it is marginally outside for the $\ell_{\rm max} = 4$ case. The best-fitting BAO values are equal to one within one-third of the DR2 uncertainty for both.

We conclude this section by selecting $\ell_{\rm max} = 2$ as our "baseline" for the multipole analysis, since it provides a better χ^2 value and an unbiased BAO measurement within one-third of the DR2 uncertainty.

V. RESULTS

As our validation tests on mocks indicated, we employ the $\ell_{\rm max} = 2$ case in our data application. In this setting, the Hartlap correction in Eq. (6) becomes 1.15 on our estimated covariance matrix, and the Percival correction in Eq. (7) is 1.12 on the best-fit parameter covariance. All of our reported uncertainties include this term.

Following DESI-DR2-Lyα, we adopt Gaussian priors on the effective bias of C IV systems ($b_{\rm C\ IV(eff)}$), the quasar bias (b_Q), and the systematic shift in quasar redshifts ($\Delta r_\parallel$), as noted in Table I. We confirmed that removing these priors does not affect the BAO measurement, as discussed in Sec. VI. In addition, two nuisance parameters related to HCDs cannot be constrained in our work, unlike in DESI-DR2-Lyα. Since the multipole decomposition largely weakens the constraints on the nuisance parameters, this loss of constraining power is not unexpected. These parameters are $\beta_{\rm HCD}$, the redshift-space distortions parameter, and $L_{\rm HCD}$, the width of a filter that is used to model the wings of the HCDs. As we will see shortly below, our best-fitting value for the HCD bias is zero, indicating an inability to detect the HCD contribution to the correlation function in a multipole configuration. Consequently, $\beta_{\rm HCD}$ and $L_{\rm HCD}$ become unconstrained. These parameters were constrained with the help of strong priors in previous DESI publications [6, 7] and were fixed in the eBOSS analysis [14]. Similarly, we fix these to the mean of the prior of DESI-DR2-Lyα: $\beta_{\rm HCD} = 0.5$ and $L_{\rm HCD} = 5\ h^{-1}$ Mpc. However, these values do not matter as their contribution is multiplied by $b_{\rm HCD} = 0$.

This leaves us with 148 data points and 15 free parameters. Our best-fitting BAO parameters are

$$\alpha_{\rm iso} = 1.0005 \pm 0.0093, \quad (14)$$

$$\alpha_{\rm AP} = 0.9929 \pm 0.0269, \quad (15)$$

with a correlation coefficient $\rho = -0.623$ at an effective

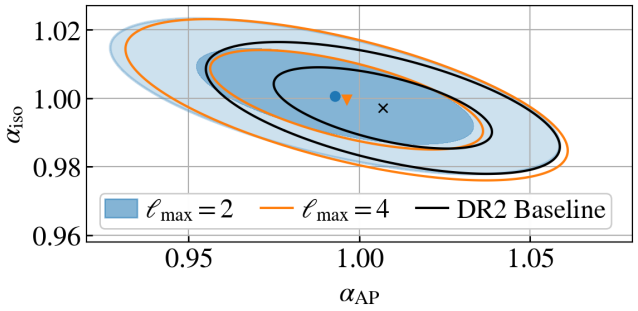


FIG. 7. Our DR2 $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$ measurements using $\ell_{\rm max}=2$ (*blue*). Confidence intervals of 68% and 95% are plotted on equal axis scale to highlight the correlation coefficient $\rho=-0.633$. Our findings are consistent with the DESI DR2 baseline analysis (*black*). We perform the $\ell_{\rm max}=4$ case as a variation (*orange*), which slightly shifts our main findings towards the DR2 baseline.

redshift $z_{\rm eff}=2.348$. We obtain an excellent fit with $\chi^2=127.7$ and a PTE of 61%. Fig. 6 shows the best-fitting model against the data points for Ly$\alpha\times$Lyα on the top and Ly$\alpha\times$QSO on the bottom panel. The fitting performance is exquisite even in the quadrupole for both correlation functions. Fig. 7 shows the correlation between $\alpha_{\rm iso}$ and $\alpha_{\rm AP}$, and demonstrates the agreement between our results and the DESI-DR2-Lyα results. We test the $\ell_{\rm max}=4$ case as a variation and find that it moves BAO parameters slightly towards DESI-DR2-Lyα, which can be seen in Fig. 7. The fit in this case is worse than the $\ell_{\rm max}=2$ case, as expected, but unlike mock results, it has a reasonable PTE of 34% with $\chi^2/\nu=214.8/(222-15)$.

Table I lists the best-fitting values for all the nuisance parameters and compares them to the DESI-DR2-Lyα results. As mentioned in previous sections, our multipole configuration provides less constraining power across the board. We have performed worse than our forecasts in Sec. IV B in terms of BAO precision: $\alpha_{\rm iso}$ precision has degraded by 18% and $\alpha_{\rm AP}$ has degraded by 27% as opposed to 14% and 18% respectively. Our best-fitting HCD bias ($b_{\rm HCD}$) is zero with 0.02 uncertainty, which is within the one-sigma region of the DESI-DR2-Lyα value. We continue discussing the implications of this in Sec. VI. In addition to the HCD bias, the best-fitting value for the amplitude of the proximity effect ($\xi_0^{\rm TP}$) is also zero with a large uncertainty. Although these differences are noteworthy, they ultimately do not impact the BAO parameters, as we demonstrated in the previous section.

TABLE I. All best-fit values of our work compared to those of DESI-DR2-Lyα. We quote only the statistical errors for cosmological distance ratios D_H/r_d and D_M/r_d.

Parameter	This work	DESI-DR2-Lyα
$\alpha_{\rm iso}$	1.0005 ± 0.0093	0.9971 ± 0.0079
$\alpha_{\rm AP}$	0.9929 ± 0.0269	1.0069 ± 0.0212
ρ	-0.623	-0.591
b_α	-0.180 ± 0.008	-0.135 ± 0.007
β_α	0.926 ± 0.081	1.445 ± 0.064
$10^3 b_{\rm Si\ II(1190)}$	-5.44 ± 1.40	-3.70 ± 0.39
$10^3 b_{\rm Si\ II(1193)}$	-0.89 ± 1.26	-3.18 ± 0.38
$10^3 b_{\rm Si\ III(1207)}$	-5.92 ± 2.82	-7.30 ± 1.48
$10^3 b_{\rm Si\ II(1260)}$	-5.33 ± 1.70	-3.67 ± 0.40
$10^3 b_{\rm C\ IV(eff)}$ [a]	-18.7 ± 5.2	-18.6 ± 4.8
$b_{\rm HCD}$	-0.00 ± 0.02	-0.02 ± 0.01
b_Q [b]	3.43 ± 0.18	3.545 ± 0.054
σ_v [h^{-1} Mpc]	2.12 ± 2.47	3.18 ± 0.63
$\Delta r_\parallel$ [h^{-1} Mpc] [c]	0.0 ± 1.0	0.53 ± 0.18
$\xi_0^{\rm TP}$	0.0 ± 2.1	0.453 ± 0.046
$10^4 a_{\rm noise}$	3.15 ± 2.50	2.21 ± 0.15
$z_{\rm eff}$	2.35	2.33
$D_H(z_{\rm eff})/r_d$	8.51 ± 0.12	8.63 ± 0.10
$D_M(z_{\rm eff})/r_d$	39.46 ± 0.65	38.99 ± 0.52
ρ_{HM}	-0.55	-0.46

[a] Prior: $\mathcal{N}(-19.0, 5.0)$
[b] Prior: $\mathcal{N}(3.4, 0.2)$
[c] Prior: $\mathcal{N}(0.0, 1.0)$

VI. DISCUSSION

In this work, we compressed the Lyα forest correlation functions into Legendre multipoles and achieved our primary goal of obtaining a positive-definite covariance matrix without ad hoc smoothing. We then investigated how well these multipoles can be modeled, given the highly anisotropic nature of metal contamination and the distortions due to continuum fitting errors. We found that the monopole and quadrupole terms can be modeled well and efficiently capture BAO information, whereas including the hexadecapole yields diminishing returns, improving BAO information by only a few percent. However, nuisance parameters degrade noticeably in the multipole representation and exhibit stronger correlations between BAO parameters, particularly $\alpha_{\rm AP}$.

The Legendre compression reduces the number of data points and, therefore, the size of the covariance matrix substantially. This approach introduces an important advantage: it is now feasible to supplement covariance matrix estimation with that from mocks, as is done in galaxy clustering statistics. However, we also demonstrated that the variance of our mock data is lower than that of the real data. Consequently, the variance amplitude needs to be scaled to match the data while the mock fidelity is improved from first principles.

Our BAO parameter measurements are in complete agreement with DESI-DR2-Lyα, deviating by $\delta\alpha_{\rm iso}=0.0034$ and $\delta\alpha_{\rm AP}=-0.014$, both of which are less than $\sigma/2$ differences. Our α measurements are slightly more correlated and have worse precision: by 18% for $\alpha_{\rm iso}$ and 27% for $\alpha_{\rm AP}$. The two main reasons for this are (1) not using the B region and (2) truncating the data to two multipoles for each correlation function. This is not a trivial amount of precision loss. However, the compression efficiency is noteworthy, considering we reduced the

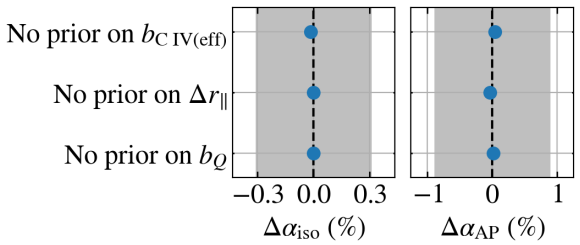


FIG. 8. Shifts in the BAO parameters when a prior is removed from the analysis. The grey area represents one-third of the uncertainty in our baseline analysis. This confirms the robustness of the multipoles framework to prior choices.

data vector by 98% (from 9,306 elements to 148).

In terms of the sound horizon at the drag epoch, r_d, our α measurements correspond to the following transverse comoving distance (D_M), Hubble distance ($D_H = c/H$), and isotropic BAO distance ($D_V \equiv (zD_M^2 D_H)^{1/3}$) measurements:

$$D_M(z_{\mathrm{eff}})/r_d = 39.46 \pm 0.65, \quad (16)$$
$$D_H(z_{\mathrm{eff}})/r_d = 8.51 \pm 0.12, \quad (17)$$
$$D_V(z_{\mathrm{eff}})/r_d = 31.46 \pm 0.29, \quad (18)$$

where $z_{\mathrm{eff}} = 2.35$. The correlation coefficient between D_H and D_M is $\rho_{HM} = -0.55$. For $r_d = 147.09$ Mpc, it follows that the Hubble parameter at this redshift is $H(z_{\mathrm{eff}}) = 239.5 \pm 3.4$ km s^{-1} Mpc^{-1} and the transverse comoving distance is $D_M(z_{\mathrm{eff}}) = 5.80 \pm 0.10$ Gpc.

We confirmed that our framework is robust to prior choices by removing each prior and refitting the model. The BAO measurement hardly moves, as shown in Fig. 8, where the grey shaded region corresponds to one third of the error from fits using $\ell_{\mathrm{max}} = 2$. Our findings parallel those of DESI-DR2-Lyα.

The absence of HCD bias detection in our real data analysis is accompanied by lower β_α and larger absolute b_α values. As we mentioned in Sec. III, the distinguishing feature between the HCD contribution and the Lyα forest signal is the extra smoothing along the line of sight, which becomes indistinguishable in the multipoles representation of the real data. This results in the b_α and β_α parameters compensating for the HCD contribution, lowering β_α's, which then raises the absolute value of b_α.

The weaker constraints on nuisance parameters imply that analyses that exploit the full shape of the correlation functions to extract more cosmological information will require higher-order multipoles. The hexadecapole will be necessary for the full-shape analysis as it provides tighter constraints on nuisance parameters. Our forecasted improvement of 30% on the metal biases from mocks remains approximately correct for the data as well. The hexacadepole improves the precision of metal biases by about 25% on average. Furthermore, the precision of the quasar proximity effect improves the most, becoming $\xi_0^{\mathrm{TP}} = 0.42 \pm 0.22$, which is an improvement of 85% in precision.

When analyzing the stack of ten mocks, we found that modeling the Ly$\alpha\times$Lyα hexadecapole performs similarly well to the monopole and quadrupole, but modeling the Ly$\alpha\times$QSO hexadecapole performs significantly worse. However, we also achieved a good fit to the data while including both hexadecapoles in the model. This can indicate two things: (1) Modeling errors are not important for the current dataset, but will gradually become important as we collect more data. (2) The model is a good description for the Ly$\alpha\times$QSO hexadecapole for real data, but not a good description for simulated data. We would need to improve our simulations, focusing on quasar and metal clustering, to reliably test our model.

Relatedly, we ignored the dipole of Ly$\alpha\times$QSO in this work. This term could potentially carry interesting asymmetric "relativistic" effects such as gravitational redshift and gravitational potential terms [60–62]. However, in practice, it is caused by more trivial effects, such as omnipresent distortions, quasar proximity effects, systematic shift in quasar redshifts, and differences in the time evolution of quasar, forest, and metal biases. Indeed, we confirm that the Ly$\alpha\times$QSO dipole is clearly detected from real data, though it is mostly featureless and cannot constrain the $\Delta r_\parallel$ parameter. When the dipole term is included in the analysis, metal bias constraints improve by about 10% on average, while a maximum of 30% improvement is achieved for $b_{\mathrm{Si\,III(1207)}}$. However, the BAO parameters and their uncertainties remain completely unchanged. Relativistic effects can be detected using the dipole only if these trivial effects are accurately modeled.

Lastly, one might be interested in "weighting" the multipoles in Eq. (4), for example, using inverse variances of each $\xi(r,\mu)$ bin, in the hopes that it will improve the precision of best-fitting parameters by suppressing noisy μ bins. This weighting is easy to implement and propagate from model to data. However, it also breaks the orthonormality of the Legendre polynomial basis. Nevertheless, we test this formalism using $\ell_{\mathrm{max}} = 2$ with mocks and find that it does not improve the BAO constraints over regular Legendre decomposition. In fact, performs worse for the α_{AP} parameter. The non-orthonormality of this weighted basis also leads to increased correlations between the monopole and quadrupole.

DATA AVAILABILITY

The data used in this analysis will be made public, along with DR2 (details available at `https://data.desi.lbl.gov/doc/releases/`). The data points shown in the figures will be available in a Zenodo repository after publication at `https://doi.org/10.5281/zenodo.18989744`.

ACKNOWLEDGMENTS

This material is based upon work supported by the U.S. 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 Humanities, 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`.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgments can be found at `https://www.legacysurvey.org/`.

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

Appendix A: Sample bias correction

We outline the steps to derive the sample bias correction for covariance matrix estimators belonging to the class of Eq. (5). That is, for a type of data that is a vector (index i), organized by subsamples (s) and associated weights ($\boldsymbol{w}_s$). The mean is given by $\hat{\bar{x}}_i = \sum_s w_{s,i} x_{s,i} / \sum_s w_{s,i}$, and the (biased) covariance estimator is given by:

$$\hat{C}^*_{ij} = \frac{\sum_s w_{s,i} w_{s,j} (x_{s,i} - \bar{x}_i)(x_{s,j} - \bar{x}_j)}{\sum_s w_{s,i} \sum_s w_{s,j}}. \tag{A1}$$

Without loss of generality, the weights can be normalized to one such that $\sum_s w_{s,i} = 1$ for all i. The simplified expressions are as follows:

$$\hat{\bar{x}}_i = \sum_s w_{s,i} x_{s,i}, \tag{A2}$$

$$\hat{C}^*_{ij} = \sum_s w_{s,i} w_{s,j} (x_{s,i} - \bar{x}_i)(x_{s,j} - \bar{x}_j) \tag{A3}$$

$$= \sum_s w_{s,i} w_{s,j} (x_{s,i} x_{s,j} - \bar{x}_i x_{s,j} - x_{s,i} \bar{x}_j + \bar{x}_i \bar{x}_j). \tag{A4}$$

We need to calculate the expectation value of this estimator $\langle \hat{C}^*_{ij} \rangle$ to derive its bias, assuming subsamples are statistically independent, which means the following:

$$\langle x_{s,i} x_{s',j} \rangle = \begin{cases} \overline{x_i x_j} & \text{if } s = s', \\ \bar{x}_i \bar{x}_j & \text{if } s \neq s', \end{cases} \tag{A5}$$

such that the **true sample** covariance matrix is given by $C^s_{ij} = \overline{x_i x_j} - \bar{x}_i \bar{x}_j$. The covariance of the **mean** is multiplied by $\sum_s w_{s,i} w_{s,j}$ for independent and identically distributed subsamples, analgous to the variance propagation rule: $\bar{x} = \sum w_s x_s \Rightarrow \mathrm{Var}(\bar{x}) = \sum w_s^2 \mathrm{Var}(x_s)$.

After calculating each term in Eq. (A1), we are left with the following expression:

$$\langle \hat{C}^*_{ij} \rangle = C^s_{ij} \left[\sum_s w_{s,i} w_{s,j} - \sum_s w^2_{s,i} w_{s,j} \right. \tag{A6}$$

$$\left. - \sum_s w^2_{s,j} w_{s,i} + \left(\sum_s w_{s,i} w_{s,j} \right)^2 \right]. \tag{A7}$$

Note that C^s_{ij} here is the covariance of subsamples, and the covariance of the mean is $C^{\mathrm{mean}}_{ij} = C^s_{ij} \sum_s w_{s,i} w_{s,j}$. Therefore, the bias factor of Eq. (A1) is

$$F_{ij} = 1 + \sum_s w_{s,i} w_{s,j} - \frac{\sum_s w_{s,i} w_{s,j} (w_{s,i} + w_{s,j})}{\sum_s w_{s,i} w_{s,j}}, \tag{A8}$$

such that unbiased covariance estimate is given by $\hat{C}_{ij}^{\rm unbiased} = \hat{C}_{ij}^{*}/F_{ij}$. The equal weight case ($w_s = 1/N$) recovers the sample vs population variance bias factor:

$$F_{ij} = 1 + \frac{1}{N} - \frac{2/N^2}{1/N} = 1 - \frac{1}{N}. \tag{A9}$$

Finally, we calculate this factor and find it to be insignificant for DR2, $1 - F \sim 0.0014$.

Appendix B: The effect of metals in the correction function multipoles

To illustrate the effect of metal contamination in the Lyα forest auto-correlation function, we plot a fiducial model with and without metals. This is shown in Fig. 9, where error bars are reflective of one mock DR2 realization. The BAO peak vanishes for $\ell \geq 4$ from the correlation function, but metal peaks remain. This is because metal contaminations are sharply localized at $\mu \approx 1$, which requires higher multipoles to represent. This figure also demonstrates that since the BAO peak disappears at $\ell \geq 4$, metal contamination can be effectively isolated and quantified using higher-order multipoles.

[1] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, D. Schlegel, and representing the DESI collaboration, The DESI Experiment, a whitepaper for Snowmass 2013, arXiv e-prints , arXiv:1308.0847 (2013), arXiv:1308.0847 [astro-ph.CO].

[2] DESI Collaboration, The DESI Experiment Part I: Science,Targeting, and Survey Design, arXiv e-prints , arXiv:1611.00036 (2016), arXiv:1611.00036 [astro-ph.IM].

[3] DESI Collaboration, The DESI Experiment Part II: Instrument Design, arXiv e-prints , arXiv:1611.00037 (2016), arXiv:1611.00037 [astro-ph.IM].

[4] DESI Collaboration, Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument, Astron. J. **164**, 207 (2022), arXiv:2205.10939 [astro-ph.IM].

[5] DESI Collaboration, DESI 2024 III: baryon acoustic oscillations from galaxies and quasars, J. Cosmol. Astropart. Phys. **2025**, 012 (2025), arXiv:2404.03000 [astro-ph.CO].

[6] DESI Collaboration, DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman alpha forest, J. Cosmol. Astropart. Phys. **2025**, 124 (2025), arXiv:2404.03001 [astro-ph.CO].

[7] DESI Collaboration, DESI DR2 results. I. Baryon acoustic oscillations from the Lyman alpha forest, Phys. Rev. D **112**, 083514 (2025), arXiv:2503.14739 [astro-ph.CO].

[8] DESI Collaboration, DESI DR2 results. II. Measurements of baryon acoustic oscillations and cosmological constraints, Phys. Rev. D **112**, 083515 (2025), arXiv:2503.14738 [astro-ph.CO].

[9] DESI Collaboration, DESI 2024 VI: cosmological constraints from the measurements of baryon acoustic oscillations, J. Cosmol. Astropart. Phys. **2025**, 021 (2025), arXiv:2404.03002 [astro-ph.CO].

[10] DESI Collaboration, DESI 2024 V: Full-Shape galaxy clustering from galaxies and quasars, J. Cosmol. Astropart. Phys. **2025**, 008 (2025), arXiv:2411.12021 [astro-ph.CO].

[11] N. G. Busca, T. Delubac, J. Rich, S. Bailey, A. Font-Ribera, D. Kirkby, J.-M. Le Goff, M. M. Pieri, A. Slosar, *et al.*, Baryon acoustic oscillations in the Lyα forest of BOSS quasars, Astron. Astrophys. **552**, A96 (2013).

[12] A. Font-Ribera, D. Kirkby, N. Busca, J. Miralda-Escudé, N. P. Ross, A. Slosar, J. Rich, *et al.*, Quasar-Lyman α forest cross-correlation from BOSS DR11: Baryon Acoustic Oscillations, J. Cosmol. Astropart. Phys. **2014**, 027 (2014), arXiv:1311.1767 [astro-ph.CO].

[13] A. Slosar, V. Iršič, D. Kirkby, *et al.*, Measurement of baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9, J. Cosmol. Astropart. Phys. **2013**, 026 (2013).

[14] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau, J.-M. Le Goff, A. Cuceu, *et al.*, The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Lyα Forests, Astrophys. J. **901**, 153 (2020), arXiv:2007.08995 [astro-ph.CO].

[15] C. Alcock and B. Paczynski, An evolution free test for non-zero cosmological constant, Nature (London) **281**, 358 (1979).

[16] N. Kaiser, Clustering in real space and in redshift space, Mon. Not. R. Astron. Soc. **227**, 1 (1987).

[17] A. Cuceu, A. Font-Ribera, B. Joachimi, and S. Nadathur, Cosmology beyond BAO from the 3D distribution of the Lyman-α forest, Mon. Not. R. Astron. Soc. **506**, 5439 (2021), arXiv:2103.14075 [astro-ph.CO].

[18] A. Cuceu, A. Font-Ribera, S. Nadathur, B. Joachimi, and P. Martini, Constraints on the Cosmic Expansion Rate at Redshift 2.3 from the Lyman-α Forest, Phys. Rev. Lett. **130**, 191003 (2023), arXiv:2209.13942 [astro-ph.CO].

[19] A. Cuceu, A. Font-Ribera, P. Martini, B. Joachimi, S. Nadathur, J. Rich, A. X. González-Morales, H. du Mas des Bourboux, and J. Farr, The Alcock-Paczyński effect from Lyman-α forest correlations: analysis validation with synthetic data, Mon. Not. R. Astron. Soc. **523**, 3773 (2023), arXiv:2209.12931 [astro-ph.CO].

[20] A. Cuceu, H. K. Herrera-Alcantar, C. Gordon, C. Ramírez-Pérez, E. Armengaud, A. Font-Ribera, J. Guy, B. Joachimi, P. Martini, S. Nadathur, I. Pérez-Ràfols, J. Rich, *et al.*, DESI DR1 Lyα forest: 3D full-shape analysis and cosmological constraints, arXiv e-prints , arXiv:2509.15308 (2025), arXiv:2509.15308 [astro-ph.CO].

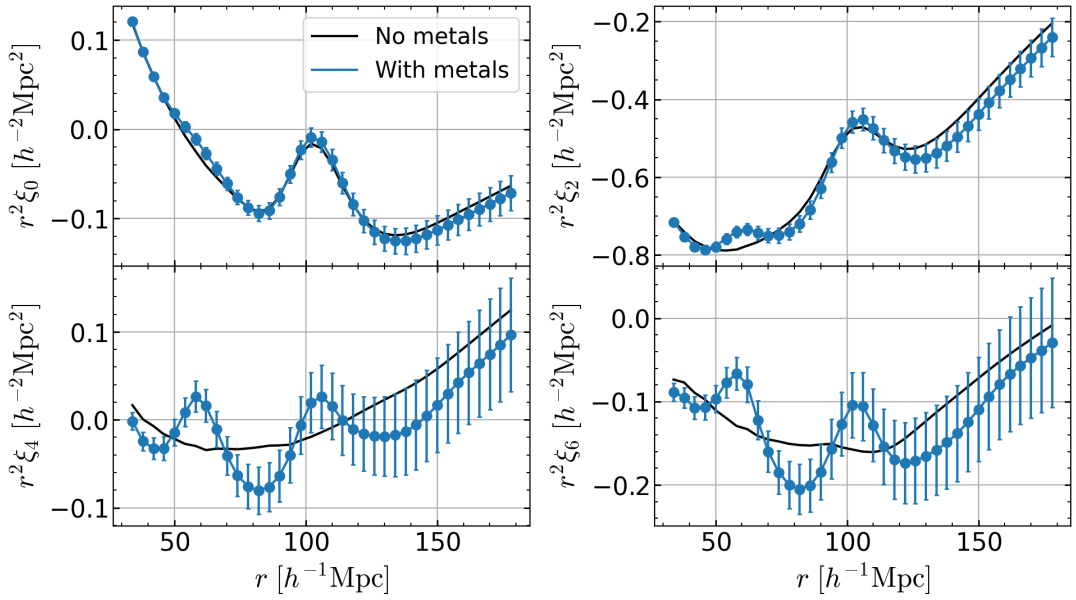


FIG. 9. Multipoles of the Lyα forest auto-correlation function for $\ell = 0, 2, 4, 6$, comparing a fiducial model without metals (*black curve*) and with metals (*blue points with error bars*) where error bars correspond to one DR2 mock realization. The BAO peak, visible at $r \approx 100\,h^{-1}$Mpc in the monopole and quadrupole, disappears for $\ell \geq 4$, while metal contamination features persist.

[21] A. Cuceu, H. K. Herrera-Alcantar, C. Gordon, P. Martini, J. Guy, A. Font-Ribera, A. X. Gonzalez-Morales, *et al.*, Validation of the DESI 2024 Lyα forest BAO analysis using synthetic datasets, J. Cosmol. Astropart. Phys. **2025**, 148 (2025), arXiv:2404.03004 [astro-ph.CO].

[22] L. Casas, H. K. Herrera-Alcantar, J. Chaves-Montero, A. Cuceu, A. Font-Ribera, M. Lokken, M. Abdul Karim, C. Ramírez-Pérez, *et al.*, Validation of the DESI DR2 Lyα BAO analysis using synthetic datasets, arXiv e-prints , arXiv:2503.14741 (2025), arXiv:2503.14741 [astro-ph.IM].

[23] W. Turner, P. Martini, N. G. Karaçaylı, *et al.*, New Measurements of the Lyα Forest Continuum and Effective Optical Depth with LyCAN and DESI Y1 Data, Astrophys. J. **976**, 143 (2024), arXiv:2405.06743 [astro-ph.CO].

[24] W. Turner, A. Cuceu, P. Martini, *et al.*, Probing the limits of cosmological information from the Lyman-α forest 2-point correlation functions, arXiv e-prints , arXiv:2509.14322 (2025), arXiv:2509.14322 [astro-ph.CO].

[25] S. Alam *et al.*, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample, Mon. Not. R. Astron. Soc. **470**, 2617 (2017), arXiv:1607.03155 [astro-ph.CO].

[26] DESI Collaboration, Data Release 1 of the Dark Energy Spectroscopic Instrument, arXiv e-prints , arXiv:2503.14745 (2025), arXiv:2503.14745 [astro-ph.CO].

[27] J. H. Silber, P. Fagrelius, K. Fanning, M. Schubnell, *et al.*, The Robotic Multiobject Focal Plane System of the Dark Energy Spectroscopic Instrument (DESI), Astron. J. **165**, 9 (2023), arXiv:2205.09014 [astro-ph.IM].

[28] T. N. Miller *et al.*, The Optical Corrector for the Dark Energy Spectroscopic Instrument, Astron. J. **168**, 95 (2024), arXiv:2306.06310 [astro-ph.IM].

[29] C. Poppett, L. Tyas, *et al.*, Overview of the Fiber System for the Dark Energy Spectroscopic Instrument, Astron. J. **168**, 245 (2024).

[30] E. F. Schlafly, D. Kirkby, D. J. Schlegel, A. D. Myers, A. Raichoor, K. Dawson, *et al.*, Survey Operations for the Dark Energy Spectroscopic Instrument, Astron. J. **166**, 259 (2023), arXiv:2306.06309 [astro-ph.CO].

[31] J. Guy, S. Bailey, A. Kremin, *et al.*, The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument, Astron. J. **165**, 144 (2023), arXiv:2209.14482 [astro-ph.IM].

[32] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, Astrophys. J. **622**, 759 (2005), arXiv:astro-ph/0409513 [astro-ph].

[33] A. Font-Ribera, J. Miralda-Escudé, E. Arnau, B. Carithers, K.-G. Lee, P. Noterdaeme, I. Pâris, P. Petitjean, J. Rich, E. Rollinde, N. P. Ross, D. P. Schneider, M. White, and D. G. York, The large-scale cross-correlation of Damped Lyman alpha systems with the Lyman alpha forest: first measurements from BOSS, J. Cosmol. Astropart. Phys. **2012**, 059 (2012), arXiv:1209.4596 [astro-ph.CO].

[34] A. Brodzeller, M. Wolfson, D. M. Santos, M. Ho, T. Tan, M. M. Pieri, A. Cuceu, *et al.*, Construction of the Damped Lyα Absorber Catalog for DESI DR2 Lyα BAO, arXiv e-prints , arXiv:2503.14740 (2025), arXiv:2503.14740 [astro-ph.CO].

[35] B. Wang, J. Zou, Z. Cai, J. X. Prochaska, Z. Sun, J. Ding, A. Font-Ribera, A. Gonzalez, H. K. Herrera-Alcantar, V. Irsic, X. Lin, D. Brooks, S. Chabanier, R. d. Belsunce, N. Palanque-Delabrouille, G. Tarle, and Z. Zhou, Deep Learning of Dark Energy Spectroscopic Instrument Mock Spectra to Find Damped Lyα Systems, Astrophys. J. Suppl. Ser. **259**, 28 (2022).

[36] M.-F. Ho, S. Bird, and R. Garnett, Damped Lyman-α absorbers from Sloan digital sky survey DR16Q with Gaussian processes, Mon. Not. R. Astron. Soc. **507**, 704 (2021), arXiv:2103.10964 [astro-ph.GA].

[37] P. Martini, A. Cuceu, L. Ennesser, A. Brodzeller, *et al.*, Validation of the DESI 2024 Lyman alpha forest BAL masking strategy, J. Cosmol. Astropart. Phys. **2025**, 137 (2025), arXiv:2405.09737 [astro-ph.CO].

[38] C. Ramírez-Pérez, I. Pérez-Ràfols, A. Font-Ribera, *et al.*, The Lyman-α forest catalogue from the Dark Energy Spectroscopic Instrument Early Data Release, Mon. Not. R. Astron. Soc. **528**, 6666 (2024), arXiv:2306.06312 [astro-ph.CO].

[39] J. E. Bautista, N. G. Busca, J. Guy, J. Rich, M. Blomqvist, H. du Mas des Bourboux, M. M. Pieri, A. Font-Ribera, *et al.*, Measurement of baryon acoustic oscillation correlations at z = 2.3 with SDSS DR12 Lyα-Forests, Astron. Astrophys. **603**, A12 (2017), arXiv:1702.00176 [astro-ph.CO].

[40] Planck Collaboration, Planck 2018 results. VI. Cosmological parameters, Astron. Astrophys. **641**, A6 (2020), arXiv:1807.06209 [astro-ph.CO].

[41] N. Busca, J. Rich, J. Bautista, A. Cuceu, A. Font-Ribera, J. Guy, H. K. Herrera-Alcantar, J. Stermer, C. Balland, *et al.*, The effects of continuum fitting on Lyman-α forest correlations, J. Cosmol. Astropart. Phys. **2025**, 020 (2025), arXiv:2506.15262 [astro-ph.CO].

[42] E. Sellentin and A. F. Heavens, Parameter inference with estimated covariance matrices, Mon. Not. R. Astron. Soc. **456**, L132 (2016), arXiv:1511.05969 [astro-ph.CO].

[43] J. Hartlap, P. Simon, and P. Schneider, Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix, Astron. Astrophys. **464**, 399 (2007), arXiv:astro-ph/0608064 [astro-ph].

[44] W. J. Percival, A. J. Ross, A. G. Sánchez, L. Samushia, *et al.*, The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: including covariance matrix errors, Mon. Not. R. Astron. Soc. **439**, 2531 (2014), arXiv:1312.4841 [astro-ph.CO].

[45] A. Lewis, A. Challinor, and A. Lasenby, Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models, Astrophys. J. **538**, 473 (2000), arXiv:astro-ph/9911177 [astro-ph].

[46] D. Kirkby, D. Margala, A. Slosar, S. Bailey, N. G. Busca, T. Delubac, J. Rich, *et al.*, Fitting methods for baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9, J. Cosmol. Astropart. Phys. **2013**, 024 (2013), arXiv:1301.3456 [astro-ph.CO].

[47] A. Arinyo-i-Prats, J. Miralda-Escudé, M. Viel, and R. Cen, The non-linear power spectrum of the Lyman alpha forest, J. Cosmol. Astropart. Phys. **2015**, 017 (2015), arXiv:1506.04519 [astro-ph.CO].

[48] U. Seljak, Bias, redshift space distortions and primordial nongaussianity of nonlinear transformations: application to Ly-α forest, J. Cosmol. Astropart. Phys. **2012**, 004 (2012), arXiv:1201.0594 [astro-ph.CO].

[49] S.-F. Chen, Z. Vlah, and M. White, The Lyα forest flux correlation function: a perturbation theory perspective, J. Cosmol. Astropart. Phys. **2021**, 053 (2021), arXiv:2103.13498 [astro-ph.CO].

[50] A. J. S. Hamilton, Uncorrelated modes of the non-linear power spectrum, Mon. Not. R. Astron. Soc. **312**, 257 (2000).

[51] A. Font-Ribera and J. Miralda-Escudé, The effect of high column density systems on the measurement of the Lyman-α forest correlation function, J. Cosmol. Astropart. Phys. **2012**, 028 (2012).

[52] K. K. Rogers, S. Bird, H. V. Peiris, A. Pontzen, A. Font-Ribera, and B. Leistedt, Correlations in the three-dimensional Lyman-alpha forest contaminated by high column density absorbers, Mon. Not. R. Astron. Soc. **476**, 3716 (2018), arXiv:1711.06275 [astro-ph.CO].

[53] T. Tan, J. Rich, E. Chaussidon, J. M. Le Goff, C. Balland, E. Armengaud, *et al.*, Modeling of the high column density systems in the Lyman-Alpha forest, J. Cosmol. Astropart. Phys. **2025**, 074 (2025), arXiv:2506.13005 [astro-ph.CO].

[54] A. Slosar, A. Font-Ribera, M. M. Pieri, J. Rich, J.-M. Le Goff, *et al.*, The Lyman-α forest in three dimensions: measurements of large scale flux correlations from BOSS 1st-year data, J. Cosmol. Astropart. Phys. **2011**, 001 (2011), arXiv:1104.5244 [astro-ph.CO].

[55] J. Guy, S. G. A. Gontcho, E. Armengaud, A. Brodzeller, A. Cuceu, A. Font-Ribera, H. K. Herrera-Alcantar, N. G. Karaçaylı, A. Muñoz-Gutiérrez, M. M. Pieri, I. Pérez-Ràfols, C. Ramírez-Pérez, C. Ravoux, J. Rich, M. Walther, *et al.*, Characterization of contaminants in the Lyman-alpha forest auto-correlation with DESI, J. Cosmol. Astropart. Phys. **2025**, 140 (2025), arXiv:2404.03003 [astro-ph.CO].

[56] A. Bault, D. Kirkby, J. Guy, A. Brodzeller, *et al.*, Impact of systematic redshift errors on the cross-correlation of the Lyman-α forest with quasars at small scales using DESI Early Data, J. Cosmol. Astropart. Phys. **2025**, 130 (2025), arXiv:2402.18009 [astro-ph.CO].

[57] T. Etourneau, J.-M. L. Goff, J. Rich, T. Tan, A. Cuceu, *et al.*, Mock data sets for the Eboss and DESI Lyman-α forest surveys, J. Cosmol. Astropart. Phys. **2024**, 077 (2024), arXiv:2310.18996 [astro-ph.CO].

[58] J. Farr, A. Font-Ribera, H. du Mas des Bourboux, A. Muñoz-Gutiérrez, F. J. Sánchez, A. Pontzen, A. Xochitl González-Morales, *et al.*, LyaCoLoRe: synthetic datasets for current and future Lyman-α forest BAO surveys, J. Cosmol. Astropart. Phys. **2020**, 068 (2020), arXiv:1912.02763 [astro-ph.CO].

[59] C. Ramírez-Pérez, J. Sanchez, D. Alonso, and A. Font-Ribera, CoLoRe: fast cosmological realisations over large volumes with multiple tracers, J. Cosmol. Astropart. Phys. **2022**, 002 (2022), arXiv:2111.05069 [astro-ph.CO].

[60] C. Bonvin, L. Hui, and E. Gaztañaga, Asymmetric galaxy correlation functions, Phys. Rev. D **89**, 083535 (2014), arXiv:1309.1321 [astro-ph.CO].

[61] V. Iršič, E. Di Dio, and M. Viel, Relativistic effects in Lyman-α forest, J. Cosmol. Astropart. Phys. **2016**, 051 (2016), arXiv:1510.03436 [astro-ph.CO].

[62] M. Blomqvist, H. du Mas des Bourboux, N. G. Busca, V. de Sainte Agathe, J. Rich, *et al.*, Baryon acoustic oscillations from the cross-correlation of Lyα absorption and quasars in eBOSS DR14, Astron. Astrophys. **629**, A86 (2019), arXiv:1904.03430 [astro-ph.CO].