PAPER

The clustering of Lyman Alpha Emitting galaxies at =2–3

Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Received: June 5, 2024
Revised: July 21, 2024
Accepted: July 25, 2024
Published: August 13, 2024

The clustering of Lyman Alpha Emitting galaxies at z = 2–3

Martin White,[a] A. Raichoor,[b] Arjun Dey,[c] Lehman H. Garrison,[d,e] Eric Gawiser,[f] D. Lang,[g] Kyoung-soo Lee,[h] A.D. Myers,[i] D. Schlegel,[b] F. Valdes,[c] J. Aguilar,[b] S. Ahlen,[j] D. Brooks,[k] E. Chaussidon,[b] T. Claybaugh,[b] K. Dawson,[l] A. de la Macorra,[m] Biprateep Dey,[n] P. Doel,[k] K. Fanning,[o,p] A. Font-Ribera,[k,q] J.E. Forero-Romero,[r,s] S. Gontcho A Gontcho,[b] G. Gutierrez,[t] J. Guy,[b] K. Honscheid,[u,v,w] D. Kirkby,[x] A. Kremin,[b] M. Landriau,[b] L. Le Guillou,[y] M.E. Levi,[b] C. Magneville,[z] M. Manera,[aa,q] P. Martini,[u,ab,w] A. Meisner,[c] R. Miquel,[ac,q] B. Moon,[ad] J.A. Newman,[n] G. Niz,[ae,af] N. Palanque-Delabrouille,[z,b] C. Park,[ad] W.J. Percival,[ag,g,ah] F. Prada,[ai] G. Rossi,[aj] V. Ruhlmann-Kleider,[z] E. Sanchez,[ak] E.F. Schlafly,[al] M. Schubnell,[am,an] H. Seo,[ao] D. Sprayberry,[c] G. Tarlé,[an] B.A. Weaver,[c] Y. Yang,[ad] C. Yèche[z] and H. Zou[ap]

[a] *Department of Physics, University of California, Berkeley, 366 LeConte Hall MC 7300, Berkeley, CA 94720-7300, U.S.A.*
[b] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.*
[c] *NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, U.S.A.*
[d] *Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, U.S.A.*
[e] *Scientific Computing Core, Flatiron Institute, 162 5th Avenue, New York, NY 10010, U.S.A.*
[f] *Department of Physics and Astronomy, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, U.S.A.*
[g] *Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*
[h] *Department of Physics and Astronomy, Purdue University, 525 Northwestern Ave., West Lafayette, IN 47906, U.S.A.*
[i] *Department of Physics and Astronomy, University of Wyoming, 1000 E. University, Department 3905, Laramie, WY 82071, U.S.A.*
[j] *Physics Department, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, U.S.A.*
[k] *Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, U.K.*

https://doi.org/10.1088/1475-7516/2024/08/020

JCAP08(2024)020

[l]Department of Physics and Astronomy, The University of Utah,
115 South 1400 East, Salt Lake City, UT 84112, U.S.A.
[m]Instituto de Física, Universidad Nacional Autónoma de México,
Ciudad de México C.P. 04510, México
[n]Department of Physics and Astronomy and
Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC),
University of Pittsburgh,
3941 O'Hara Street, Pittsburgh, PA 15260, U.S.A.
[o]Kavli Institute for Particle Astrophysics and Cosmology, Stanford University,
Menlo Park, CA 94305, U.S.A.
[p]SLAC National Accelerator Laboratory, Menlo Park, CA 94305, U.S.A.
[q]Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology,
Campus UAB, 08193 Bellaterra Barcelona, Spain
[r]Departamento de Física, Universidad de los Andes,
Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia
[s]Observatorio Astronómico, Universidad de los Andes,
Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia
[t]Fermi National Accelerator Laboratory,
P.O. Box 500, Batavia, IL 60510, U.S.A.
[u]Center for Cosmology and AstroParticle Physics, The Ohio State University,
191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.
[v]Department of Physics, The Ohio State University,
191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.
[w]The Ohio State University, Columbus, 43210 OH, U.S.A.
[x]Department of Physics and Astronomy, University of California,
Irvine, CA 92697, U.S.A.
[y]Sorbonne Université, CNRS/IN2P3,
Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE),
Paris FR-75005, France
[z]IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette F-91191, France
[aa]Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona,
Bellaterra 08193, Barcelona, Spain
[ab]Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory,
140 W 18th Avenue, Columbus, OH 43210, U.S.A.
[ac]Institució Catalana de Recerca i Estudis Avançats,
Passeig de Lluís Companys 23, Barcelona 08010, Spain
[ad]Korea Astronomy and Space Science Institute,
776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea
[ae]Departamento de Física, Universidad de Guanajuato, DCI,
Leon C.P. 37150, Guanajuato, México
[af]Instituto Avanzado de Cosmología A. C.,
San Marcos 11, Atenas 202, Magdalena Contreras 10720, Ciudad de México, México
[ag]Department of Physics and Astronomy, University of Waterloo,
200 University Ave W, Waterloo, ON N2L 3G1, Canada
[ah]Waterloo Centre for Astrophysics, University of Waterloo,
200 University Ave W, Waterloo, ON N2L 3G1, Canada

JCAP08(2024)020

[ai] *Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía s/n, Granada E-18008, Spain*
[aj] *Department of Physics and Astronomy, Sejong University, Seoul 143-747, Korea*
[ak] *CIEMAT, Avenida Complutense 40, Madrid E-28040, Spain*
[al] *Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, U.S.A.*
[am] *Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A.*
[an] *University of Michigan, Ann Arbor, MI 48109, U.S.A.*
[ao] *Department of Physics and Astronomy, Ohio University, Athens, OH 45701, U.S.A.*
[ap] *National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing 100012, P.R. China*

E-mail: mwhite@berkeley.edu

Abstract: We measure the clustering of Lyman Alpha Emitting galaxies (LAEs) selected from the One-hundred-square-degree DECam Imaging in Narrowbands (ODIN) survey, with spectroscopic follow-up from Dark Energy Spectroscopic Instrument (DESI). We use DESI spectroscopy to optimize our selection and to constrain the interloper fraction and redshift distribution of our narrow-band selected sources. We select samples of 4000 LAEs at $z = 2.45$ and 3.1 in 9 sq.deg. centered on the COSMOS field with median Lyα fluxes of $\approx 10^{-16}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$. Covariances and cosmological inferences are obtained from a series of mock catalogs built upon high-resolution N-body simulations that match the footprint, number density, redshift distribution and observed clustering of the sample. We find that both samples have a correlation length of $r_0 = 3.0 \pm 0.2\,h^{-1}\,\mathrm{Mpc}$. Within our fiducial cosmology these correspond to 3D number densities of $\approx 10^{-3}\,h^3\,\mathrm{Mpc}^{-3}$ and, from our mock catalogs, biases of 1.7 and 2.0 at $z = 2.45$ and 3.1, respectively. We discuss the implications of these measurements for the use of LAEs as large-scale structure tracers for high-redshift cosmology.

Keywords: galaxy clustering, galaxy surveys, high redshift galaxies, redshift surveys

ArXiv ePrint: 2406.01803

JCAP08(2024)020

Contents

JCAP08(2024)020

1 Introduction

The inhomogeneous Universe, as probed by fluctuations in the cosmic microwave background (CMB) radiation or surveys of the large-scale structure of the Universe, provides one of our best windows on fundamental physics at ultra-high energies [1]. The tightest constraints on dark energy, mass limits on light dark matter particles, models of inflation, neutrino masses and light relic particles all come from one or both of these measurements. There are compelling theoretical motivations [2–4] to push the study of large-scale structure to redshifts $2 < z < 6$ using both relativistic and non-relativistic tracers. This will allow us to probe the metric, particle content and *both* epochs of accelerated expansion (Inflation and Dark Energy domination) with high precision in a regime that is not theory limited.

A high redshift survey that aims to measure the large-scale structure in three dimensions needs to be able to efficiently obtain redshifts for faint galaxies over wide areas. While we are witnessing tremendous advances in instrumental capability, this is still challenging for faint galaxies and at high redshift unless the galaxies that are targeted have bright emission lines. One such population of galaxies are Lyman Alpha Emitters (LAEs; [5]), which — as their names suggests — have prominent Lyα emission lines in their spectra. Existing surveys are

somewhat limited — cosmic variance is a major concern and many questions remain unsettled — but studies suggest that LAEs have relatively low stellar masses, low star formation rates, young ages and low dust content [5]. If LAEs populate low halo masses (compared to other galaxies that might be selected for spectroscopic follow-up at that redshift) then we might expect that they have a low and roughly scale-independent bias. Under such conditions, LAEs would make an excellent visible tracer of the underlying matter distribution.

Motivated by the possibility of using large-scale structure at early times as a cosmological probe [2–4], the Dark Energy Spectroscopic Instrument (DESI; [6, 7]) collaboration took spectra of LAE candidate galaxies selected from various surveys as part of ancillary[1] programs. In this paper, we present our measurements of the clustering of LAEs selected from the One-hundred-deg^2 DECam Imaging in Narrowbands (ODIN; [9–11]) survey. A subset of these objects were followed up spectroscopically by DESI, with the resulting data used to optimize the narrow-band selection and constrain the interloper/outlier fraction and the redshift distribution, dN/dz. After a brief introduction to the data employed (section 2; described further in our companion paper, ref. [12]) we describe the clustering analysis (section 3). Our pipeline is tested, covariance matrices computed and inferences are obtained using a series of mock catalogs built upon high resolution N-body simulations. These are described in section 4. Our main results are given in section 5 and the implications for surveys aiming to use LAEs for high-z cosmology are presented in section 6. Finally we conclude in section 7, with some technical details relegated to an appendix.

2 Data

In this section we give a brief overview of the datasets underlying our analysis. These data are drawn from two surveys: the One-hundred-square-degree DECam Imaging in Narrowbands (ODIN; [9–11]) and the Dark Energy Spectroscopic Instrument (DESI; [6, 7]). We briefly describe these data below, referring the reader to the preceeding references and companion papers for more details.

2.1 The ODIN survey

ODIN (NOIRLab Survey Program 2020B-0201) is a wide-field ($\sim 100\,\mathrm{deg}^2$), deep, imaging survey targeting seven fields in three narrow-band filters tuned to select LAEs at redshifts of $z \approx 2.4$, 3.1, and 4.5. Covering a (comoving) volume of $\approx 0.24\,\mathrm{Gpc}^3$, comparable to the Sloan Digital Sky Survey [13], the ODIN sample is designed to measure the large-scale clustering of LAEs across cosmological time. The ODIN survey is described in detail in ref. [9]; here, we just summarize the relevant details.

The survey narrow-band observations were obtained between the 2021A-2023B semesters using the Dark Energy Camera [DECam; 14] on the Victor M. Blanco 4-meter telescope at the Cerro Tololo Inter-American Observatory in Chile. The narrow-band imaging depths of 25.5, 25.7, and 25.9 AB mag ($5\,\sigma$ in a $2''$ diameter aperture) in the N419, N501 and N673 filters, respectively, correspond to Lyα line flux limits of 3.1×10^{-17}, 1.8×10^{-17}, $1.1 \times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$ at redshifts of $z \simeq 2.45$, 3.1 and 4.5 respectively. The narrow-band

[1] For DESI collaboration studies of the broader class of Lyman Break Galaxies see ref. [8].

JCAP08(2024)020

ODIN observations are complemented with existing public broad-band imaging data from the DESI Legacy Imaging Survey (LS; [15]) and the Subaru Hyper-Suprime Cam (HSC) Strategic Survey Program (SSP; [16]).

Approximately half of the ODIN fields overlap the survey footprint of the DESI spectroscopic survey, and two (COSMOS and XMM) have been targeted by DESI for spectroscopic observations of candidate LAEs. The selection and DESI spectrosopic observations of the candidate LAEs are complex and are discussed fully in a companion paper [12], but here we provide a brief summary. At the time the DESI spectroscopic observations were initiated, the ODIN observations were still in progress. Due to scheduling and resource constraints, we were only able to properly spectroscopically characterize LAE candidates in the COSMOS field (N419 and N501 candidates) and the XMM field (N419 candidates). In this paper, we focus only on the N419 and N501 LAE candidates in the COSMOS field.

We constructed photometric catalogs of the imaging data using Tractor [17], with the narrow-band imaging data as the detection image, performing forced photometry on the broad-band data. LAE candidate sources were selected for spectroscopic targeting according to a variety of photometric cuts which were designed to be "liberal", in the sense of allowing a larger fraction of interloper galaxies in order to determine how best to (re)optimize the sample selection (post redshift determination) for high purity and completeness. We refer the interested reader to ref. [12] for the (quite involved) details of the candidate selection for the spectroscopic observations.

2.2 DESI spectroscopy

The DESI spectroscopic observations were obtained as part of an ancillary program to target candidate high-redshift galaxy populations with the goal of investigating their suitability for high-redshift clustering studies. DESI [6, 7], a prime-focus fiber-fed multi-object spectrograph mounted on the Nicholas U. Mayall 4-meter Telescope of the Kitt Peak National Observatory, has the ability to obtain simultaneous spectra of $\approx$ 5000 targets within a 3 deg diameter field [18–21]. DESI covers a wavelength range of $\lambda\lambda$3600–9800 Å with a resolution ($R \equiv \lambda/\Delta\lambda$) varying from 2200 to 5000, and the high system throughput and high efficiency make it an ideal instrument for spectroscopic redshift surveys of faint emission line sources over large contiguous fields.

The ODIN candidates in COSMOS were observed with DESI in two campaigns [12]. During the first campaign, in March 2022, N501 targets were observed on a dedicated tile (82636) for an effective time[2] of 2.75 hrs. In the second campaign, in April 2023, which was part of a large pilot program, N419 and N501 sources were observed for either 1.1 hrs or 2.2 hrs. In total, we targeted approximately 3500 N419 and 3000 N501 sources in the COSMOS field, from which we selected our 'refined' targets.

The spectroscopic observations were reduced by the DESI pipeline and the spectrum of each LAE candidate source was visually inspected by a team of volunteers in order to measure a redshift and assign a quality flag. These "visual inspection" (VI) redshifts were combined in order to create the final redshift lists for the targeted sources [12].

[2]DESI uses the concept of "effective time" to quantify the measured squared signal-to-noise ratio in the spectra — we refer the reader to refs. [12, 22] for further details.

JCAP08(2024)020

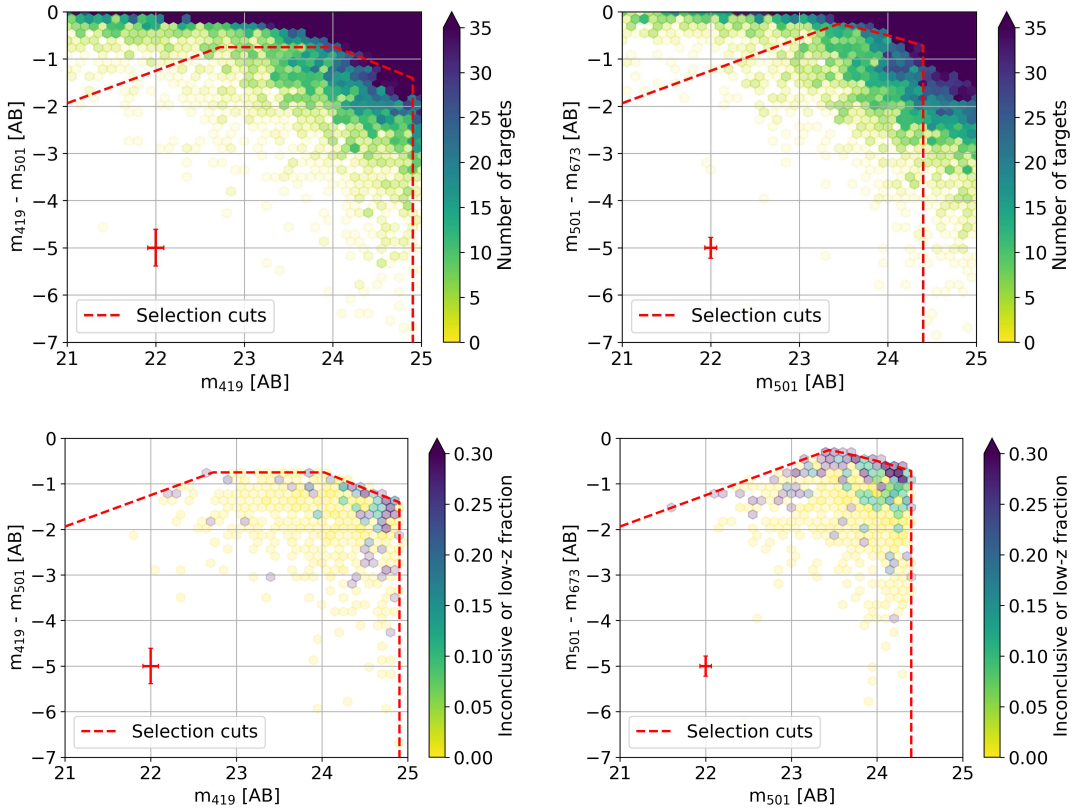

Figure 1. Color-magnitude diagrams for the refined N419 ($z \simeq 2.45$; left) and N501 ($z \simeq 3.1$; right) samples showing the number of targets (upper panels; on a linear color scale) or the 'contamination fraction' (lower panels). The red error bars show the typical observational uncertainty, computed for the median values of the selections: $(24.3, -1.6)$ for N419 and $(24.00, -1.3)$ for N501. See text for further details and the exact cuts adopted on the narrow bands.

2.3 LAE target selections

Using the final lists of high-quality redshifts, we optimized the selection criteria to result in high purity samples with a well-measured interloper fraction. We shall refer to this as the "refined" sample below, and caution that it has a complex relationship to a sample selected only on rest-frame equivalent width or one which has been selected primarily for galaxy-evolution studies. This reflects the optimization for a different purpose — to enable cosmological studies we desire targets with a high probability of being at high redshift and that are likely to give good redshifts with modest cost. To interpret the clustering measurements we desire a well-controlled interloper fraction. The chosen selections are reported below, and the ODIN color-magnitude diagrams are displayed in figure 1. Further details are provided in [12].

The adopted selection criteria for the $z \simeq 2.45$ sample are:

$$19 < m_{419} < 24.906 \tag{2.1a}$$

$$m_{419} - m_{501} < -0.75 \tag{2.1b}$$

$$m_{419} - m_{501} < -16.375 + 0.6875\, m_{419} \tag{2.1c}$$

$$m_{419} - m_{501} < 17.27 - 0.75\, m_{419} \tag{2.1d}$$

JCAP08(2024)020

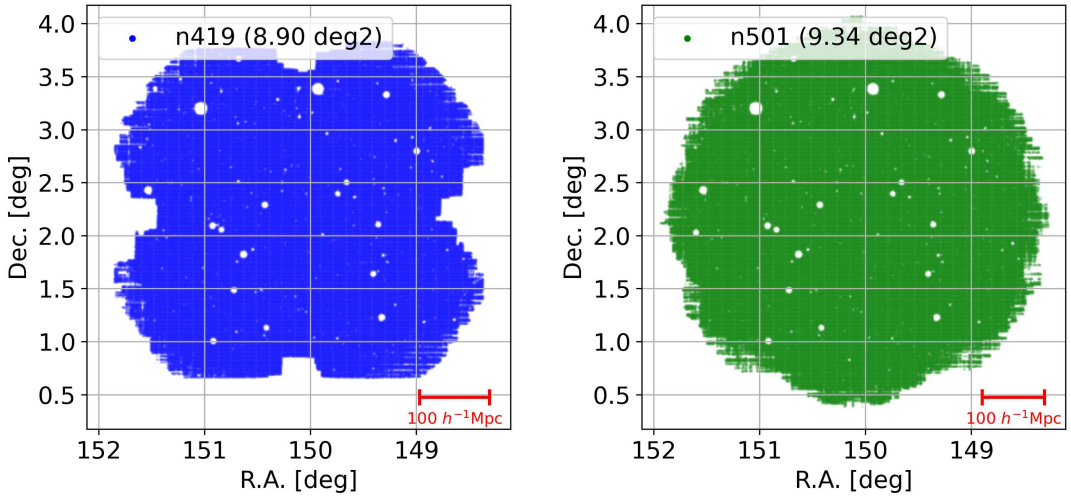

Figure 2. Survey geometry for our analysis for each ODIN band. The blue and green regions indicate the footprint, as determined from our random catalog, with the small (white) holes due to bright stars or other masksed regions. Note the footprint for N419 differs from that for N501 due to a further restriction from the HSC imaging coverage (see text). The inset bar shows $100\,h^{-1}$ Mpc, comoving at the mean redshifts of the sample ($z \simeq 2.45$ and 3.1 for the left and right panels, respecitvely).

Filter	Area [deg^2]	$N_{\rm targ}$	$N_{\rm obs}$	$N_{\rm VI,ok}$	$N_{\rm VI,zok}$	Interloper fraction
N419	8.90	2382	897	838	822	0.02–0.08
N501	9.34	1956	1240	1145	1099	0.04–0.11

Table 1. Selected LAE candidates in COSMOS for our 'refined' sample (see text). We report for each band the number of selected candidates ($N_{\rm targ}$), the number of observed candidates ($N_{\rm obs}$), the number of those with a VI-validated redshift ($N_{\rm VI,ok}$) and with a VI-validated redshift in the relevant redshift range ($N_{\rm VI,zok}$ and $z_{\rm min} < z_{\rm VI,ok} < z_{\rm max}$). We lastly report our estimated interloper fraction range (see text for details).

The adopted selection criteria for the $z \simeq 3.1$ sample are:

$$18 < m_{501} < 24.40 \tag{2.2a}$$

$$m_{501} - m_{673} < -60.5 + 5.5\,m_{501} - 0.125\,m_{501}^2 \tag{2.2b}$$

$$m_{501} - m_{673} < -16.375 + 0.6875\,m_{501} \tag{2.2c}$$

We present in table 1 the number of candidates ($N_{\rm targ}$) and the number of observed candidates with more than 1.5 hrs of effective time ($N_{\rm obs}$). Among those, $N_{\rm VI,ok}$ is the number having a VI-validated redshift and $N_{\rm VI,zok}$ is the number of those with $2.35 < z_{\rm VI,ok} < 2.50$ for N419 and $3.075 < z_{\rm VI,ok} < 3.175$ for N501.

By definition, we do not know the redshift of the spectra with a non-conclusive VI. We thus provide two values of the interloper fraction, which should bracket the truth. An optimistic estimation is to assume that the spectra with a non-conclusive VI have a similar

JCAP08(2024)020

redshift distribution to the ones with a conclusive VI: the interloper fraction would then be $f_{\rm int} = 1 - N_{\rm VI,zok}/N_{\rm VI,ok}$. A pessimistic estimate assumes that all the spectra with a non-conclusive VI are interlopers: the interloper fraction would then be $1 - N_{\rm VI,zok}/N_{\rm VI}$. From the VI and the pattern of failures, with the interloper and high-z fractions being steady with exposure time but the redshift failures depending strongly on exposure time [12], it appears the optimistic scenario is more likely for the N419 sample for which we thus take a fiducial $f_{\rm int} = 0.02$. The N501 sample is more ambiguous and we take $f_{\rm int} = 0.08$ as our fiducial value. Variations over the full range $f_{\rm int} = 0.04$–0.11 lead to changes in the amplitude of the inferred clustering of -8% or $+7\%$ of the fiducial. This translates into a 3–4% systematic uncertainty in the clustering length (r_0), which is smaller than our statistical errors. We shall neglect this henceforth, pending further spectroscopic follow-up (though we list this as a further systematic uncertainty in the measured clustering amplitude in table 2).

2.4 Survey geometry

The survey geometry is defined by the coverage of the narrow- and broad-band imaging, along with the angular masking done in the target selection. The considered geometry for each band is displayed in figure 2; we describe below how we construct it.

First, we generate 10^6 randoms per deg^2 uniformly within a 1.9 deg radius from our COSMOS field center (R.A., Dec.) = (150.11°, 2.173°). We use the approach described in ref. [24], adapted to our custom Tractor products. In particular, in the `MASKBITS` content, the three ODIN bands (N419, N501, N673) replace the g, r, and z broad bands of the Legacy Surveys. This allows us to propagate the angular quantities used in the target selection, and apply those cuts on the randoms: (1) `NOBS` $\geq$ `10` requiring at least ten images at the central pixel location for each of the two ODIN bands used to build the ODIN color in the target selection; (2) `MASKBITS` = 0, which discards unreliable (e.g. saturated) pixels of the imaging, along with regions masked by the `BRIGHT` and `MEDIUM` Legacy Surveys stellar masks; (3) an additional stellar mask rejecting objects within a radius

$$R = 0.07^\circ \left(\frac{6.3}{\texttt{GAIA\_PHOT\_G\_MEAN\_MAG}} \right)^2 \tag{2.3}$$

of a star brighter than 9 mag. For the N419 selection, we add a mask mapping the HSC imaging coverage. Since the N501 selection is based on a "LS or HSC" set of cuts and the union of the LS and HSC images fully covers our 1.9°-radius region, there is no need for an extra mask. Finally, based on those cut randoms, we build a high-resolution HealPix pixel map [25] at Nside = 8192 which can apply to mocks in the clustering analysis.

3 Clustering analysis

The DESI spectroscopy provides us with constraints on the redshift distribution and interloper fraction for the ODIN galaxies. We see from figure 3 that the ODIN galaxies lie in thin shells of widths of $\Delta z = 0.06$ at $z = 2.45$ and 3.1. These translate into comoving distance 'depths' of $\Delta \simeq 50$ and $38\,h^{-1}$ Mpc (table 2). For such narrow slices in redshift the majority of the large-scale clustering information lies in the angular clustering. While for such small sky areas the corrections for beyond-plane-parallel are small, the narrow depth of the ODIN

JCAP08(2024)020

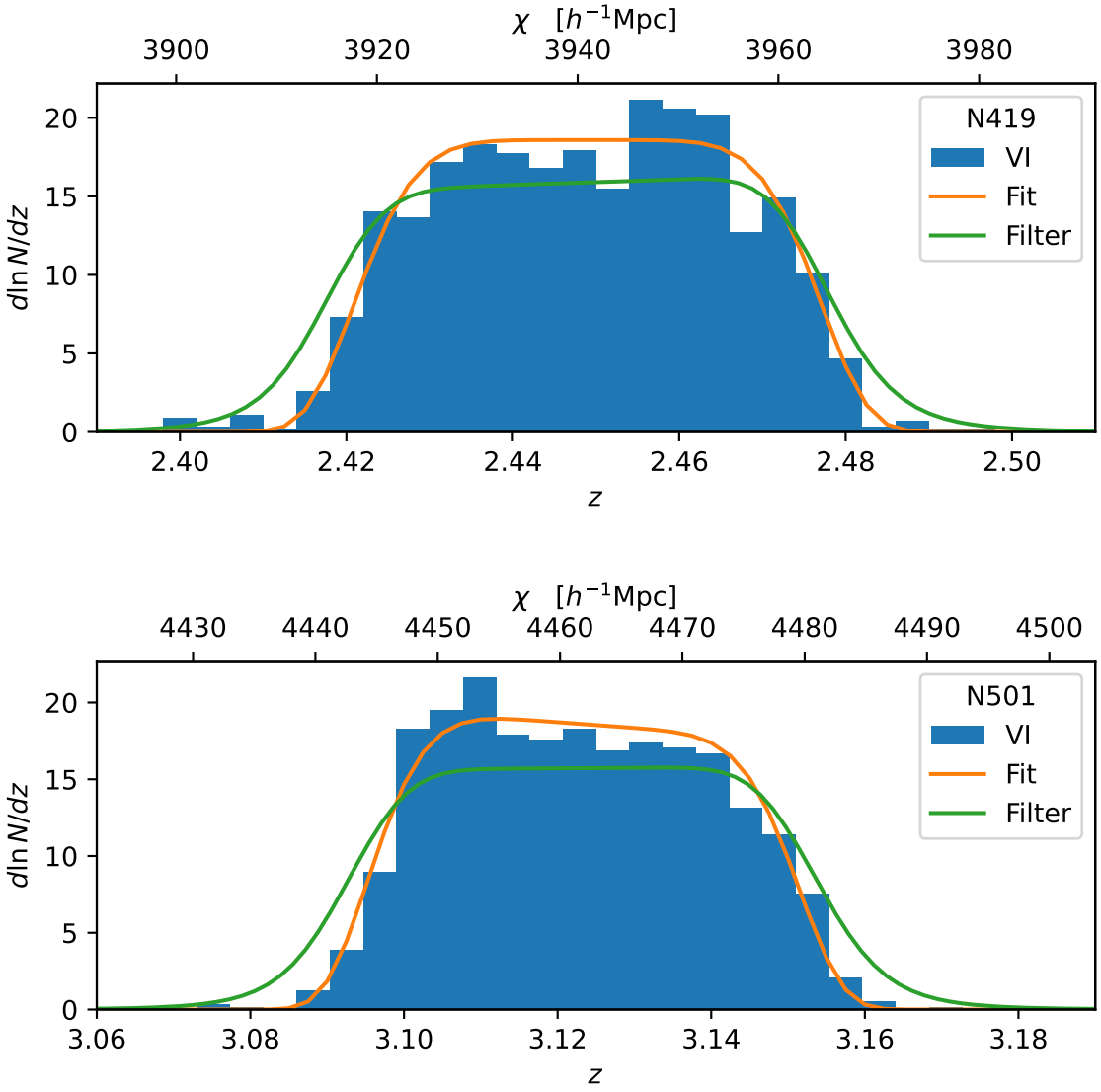

Figure 3. The (normalized) redshift distributions for the N419 ($z \simeq 2.45$) and N501 ($z \simeq 3.1$) samples described in the text. The blue histograms show the 'secure' redshifts of the visually inspected sources (VI), while the orange line is a simple analytic fit used in the theoretical model and the green line shows the filter transmission curve converted to redshift at the Lyα wavelength. On each panel the lower x-axis shows the redshift while the upper x-axis shows the corresponding comoving, radial distance (in h^{-1} Mpc) assuming our fiducial cosmology [23].

survey means that we cannot neglect redshift-space distortions[3] (RSD). Further, the Limber approximation [26], which assumes a depth much larger than the transverse scale of interest, does not hold. However because the clustering is not diluted by projection over a large depth, the magnification bias is very small.

[3]For a top-hat selection of width $\Delta\chi$ the average line-of-sight separation is $\Delta\chi/3 \sim 10\,h^{-1}$ Mpc. Thus on scales $\mathcal{O}(10\,\mathrm{Mpc})$ redshift-space distortions give a correction to the clustering of tens of percent.

JCAP08(2024)020

Filter	Redshift	$\lg F^{\rm med}_{{\rm Ly}\alpha}$	$f_{\rm int}$	χ_0	Δ	$d\chi/dz$	r_0	$10^3\,\bar{n}$	b
N419	2.4 ± 0.03	-16.0	0.02	3941	50	829	3.0 ± 0.2	1.2	1.7 ± 0.2
N501	3.1 ± 0.03	-16.1	0.08	4448	38	633	3.0 ± 0.2	1.0	2.0 ± 0.2

Table 2. Inferred properties of the refined LAE samples in each redshift range. The median Lyα flux, $F^{\rm med}_{{\rm Ly}\alpha}$, is in erg s^{-1} cm^{-2} and lg indicates $\log_{10}$. The interloper fraction, $f_{\rm int}$, is inferred from DESI spectroscopy. The comoving distance to the center of the shell (χ_0), the FWHM of the shell (Δ), the distance-redshift slope, the correlation length, r_0, and the 3D number density are in h^{-1} Mpc units to match the conventions typically employed in large-scale structure. The large-scale bias, b, is inferred from the mock catalogs (see text). For the N501 sample, the full range of uncertainty in $f_{\rm int}$ would lead to an additional $\approx$ 3–4% systematic error on r_0 or b.

Assuming $\phi(\chi) \propto H(z)\,dN/dz$ is the radial distribution as a function of comoving distance, normalized to $\int \phi\, d\chi = 1$, the (angular) power spectrum of LAEs observed within the shell in the flat-sky approximation is

$$C_\ell = \int_0^\infty \frac{dk_\parallel}{\pi\chi^2}\, P\left(k_\perp = \ell\chi^{-1}, k_\parallel\right) \left|\widetilde{\phi}(k_\parallel)\right|^2 \tag{3.1}$$

with $\widetilde{\phi}$ the Fourier transform of ϕ, χ the comoving distance to the shell center and $P(k_\perp, k_\parallel)$ is the 3D power spectrum. Variation of dN/dz over the field is negligible for our study [27]. Note for shells of full-width Δ the line-of-sight window, $\widetilde{\phi}^2(k_\parallel)$, falls off for $k_\parallel\,\Delta \gg 1$ and so we expect the integral to depend on $k_\parallel < \mathcal{O}(\Delta^{-1}) \sim 0.1\,h\,{\rm Mpc}^{-1}$. If the shell is further divided into redshift slices with distributions ϕ_a then $|\widetilde{\phi}|^2$ is replaced by a product. In the specific case of disjoint, top-hat distributions of width Δ whose centers are separated by χ_{ab} the line-of-sight window function becomes $\cos[k_\parallel\chi_{ab}]\, j_0^2(k_\parallel\Delta/2)$ with $j_0 = \sin x/x$ the spherical Bessel function of order 0.

The angular correlation function can be written as

$$\begin{aligned} w(\theta) &= \sum_\ell \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos\theta) \simeq \int \frac{\ell\, d\ell}{2\pi} C_\ell J_0(\ell\tilde{\omega}) \\ &= \int \frac{d^3k}{(2\pi)^3}\, P\left(k_\perp, k_\parallel\right)\, \widetilde{\phi}^2(k_\parallel) J_0(k_\perp\chi\,\tilde{\omega}) \end{aligned} \tag{3.2}$$

with J_0 the cylindrical Bessel function of order 0 and $\tilde{\omega} = 2\sin(\theta/2)$. Note this has the expected form of an integral of the power spectrum multiplied by a line-of-sight window function ($\widetilde{\phi}^2$) and a transverse window function (J_0). This can also be rewritten directly in terms of the 3D correlation function as

$$w_\theta(R) = 2\int_0^\infty dy\; W(y)\,\xi(\sqrt{R^2+y^2}, \mu_y)\,, \quad W(y) = \int_0^\infty d\bar{\chi}\; \phi(\bar{\chi} - y/2)\phi(\bar{\chi} + y/2) \tag{3.3}$$

where $\mu_y = y/\sqrt{R^2+y^2}$ and we have converted angles to (transverse) distances with $R = \chi\theta$, denoting this statistic $w_\theta(R)$. In the limit of a top-hat redshift distribution of full-width Δ this further simplifies to

$$w_\theta(R) = 2\int_0^1 dY\; (1-Y)\,\xi(\sqrt{R^2 + Y^2\Delta^2}, \mu) \qquad (y = Y\Delta) \tag{3.4}$$

JCAP08(2024)020

The integral can be done with simple quadrature given a model for $\xi(s,\mu)$, and the case of power-law power spectra or correlation functions is treated in appendix A. The generalization for the cross-spectrum of two shells is straightforward.

From the data and random catalogs we compute the angular power spectrum (using the method of ref. [28]) and correlation functions, $w_\theta(R)$ (using the Landy-Szalay estimator [29]). The field is large enough that the integral correction in $w_\theta(R)$ can be neglected on the scales of interest, which we have verified with our N-body-based mock catalogs. Angles are converted to distances assuming the median redshift (χ_0; listed in table 2). Since the non-LAE targets are physically well separated (in redshift) from the LAEs the measured clustering can be well approximated as $w_\theta \simeq (1-f_{\rm int})^2 w_{\rm LAE} + f_{\rm int}^2 w_{\rm int}$, where $f_{\rm int}$ is the interloper fraction. Neglecting the clustering of the interlopers, since $f_{\rm int}$ is small and the interlopers appear to be spread over redshift, we can compute the LAE clustering by dividing our measured $w_\theta(R)$ by $(1-f_{\rm int})^2$. The plotted correlation functions have had this correction applied, assuming the fiducial value of $f_{\rm int}$ quoted in table 2.

4 Mock catalogs

During the planning of the observations and for the analysis we have used a set of mock catalogs. These mock catalogs are built upon halo catalogs from the AbacusSummit cosmological N-body simulation suite [30], produced with the Abacus N-body code [31]. Our fiducial simulation was of the ΛCDM family employing 6300^3 dark matter particles in a box of side length $1\,h^{-1}\,$Gpc ($m_{\rm part} = 3.5\times10^8\,h^{-1}\,M_\odot$; see ref. [30] for more details). As well as the simulations released in ref. [30], we use two additional simulations that were run on the Perlmutter supercomputer at the National Energy Research Scientific Computing center. The first simulation utilized 6000^3 particles in a $1.25\,h^{-1}\,$Gpc box ($m_{\rm part} = 7.9\times10^8\,h^{-1}\,M_\odot$) with outputs from $z=5$ to $z=2$. The force softening was held fixed in proper coordinates, reaching 0.025 of the mean interparticle spacing at $z=2$. The second utilized 6912^3 particles in a $750\,h^{-1}\,$Mpc box ($m_{\rm part} = 1.1\times10^8\,h^{-1}\,M_\odot$), also with outputs down to $z=2$. We used these additional simulations to confirm that our results are converged with respect to halo mass resolution and finite volume effects. In all cases, the halos are populated with mock galaxies using the AbacusUtils[4] software.

The halo occupation distribution of LAEs is quite uncertain, so our modeling is at best approximate. Recent simulations [32–39] and observations [5, 40–42] suggest that LAEs occupy a fraction of the low-mass halo population with the number of galaxies growing more slowly than halo mass towards large masses. We build such catalogs from the AbacusUtils software, which implements [43] a standard halo occupation distribution modeling [44] with occupancy following the form introduced by ref. [45]. Each halo, of mass M_h, has a central galaxy selected from a binomial distribution and a Poisson-distributed number of satellites with respective means

$$\bar{n}_{\rm cen}(M_h) = \frac{1}{2}{\rm erfc}\left[\frac{\ln M_{\rm cut}/M_h}{\sqrt{2}\,\sigma}\right], \quad \bar{n}_{\rm sat}(M_h) = \left[\frac{M_h - \kappa M_{\rm cut}}{M_1}\right]^\alpha \bar{n}_{\rm cen}(M_h) \tag{4.1}$$

[4] https://abacusutils.readthedocs.io/en/latest/.

JCAP08(2024)020

and we take $\alpha < 1$ and apply a random downsampling of the galaxies to match the observed number density [46, 47]. This allows fractional occupancy even at low halo mass and has LAEs preferentially avoiding high mass halos.[5] In the above, $M_{\rm cut}$ parameterizes the mass threshold below which halos do not host LAEs, with a threshold sharpness set by σ. The value of M_1 indicates where halos host ≈ 1 satellite in addition to a central, with κ setting the low-mass suppression of the satellite occupancy. Ref. [39] show that such a form matches star-forming galaxies selected in cosmological hydrodynamic simulations. Physically such occupancy arises due to selections on narrow and broad band color for which only a fraction of the parent population passes (due to a combination of the fraction of time spent in an appropriately low-dust burst of star formation and any orientation effects that influence the "opening angle" over which Lya photons escape after radiative transfer through the ISM) and on Lyα flux and on equivalent width [48], for which our procedure amounts to a random selection near threshold. The central galaxy is then assigned to the center of mass of the halo, with the velocity vector also set to that of the center of mass of the halo. Satellite galaxies are assigned to particles of the halo with equal weights. Following the preliminary investigations reported in [39], and in the absence of observational evidence to the contrary, we do not include secondary/assembly bias or velocity bias in the model [43].

Since the number density provides only an upper limit on $M_{\rm cut}$, the information on the characteristic mass of LAE-hosting halos is mostly provided by the large-scale bias. The scale-dependence of the clustering could in principle provide constraints on the detailed halo occupation, but with current error bars the problem is underconstrained. Rather than attempt a detailed exploration of the parameter space, we have selected from a grid of HOD models a subset that provide good fits to our data ($\chi^2 < 18$ for 10 d.o.f.). These provide us nearly noise-free 'observations' and a qualitative sense of the model uncertainty and have been made publicly available (see section 8).

The mock catalogs are generated from single snapshots of the N-body simulation at fixed redshift, using the closest redshift to the observed galaxies (i.e., 2.5 and 3.0 respectively). Real- and redshift-space correlation functions and power spectra are computed directly from the (periodic box) outputs for each model, using the technique of Zeldovich control variates [49–51] to reduce the sampling fluctuations at large scales. In addition, mock observations are performed by projecting the galaxies onto the sky. By applying random offsets parallel to each of the box axes we generate multiple independent volumes. For each volume we apply the same masks as for the real data and randomly downsample the objects to match the observed number density and redshift distribution. This ensures the clustering amplitude, shotnoise and projection of distances to angles matches that in the data. We compute the covariance of the clustering, with the same binning and statistics as used in the real data, for 1024 mocks in order to generate a Monte-Carlo covariance matrix. Given the small number of bins that we measure this number is more than sufficient to give a converged covariance matrix.

While we have set the mean number of objects per field to match the observed angular number density, the actual number of mock LAEs per field fluctuates by $\approx 10\%$ from

[5]The exact behavior of the HOD for high masses is not too important for our purposes. Our LAEs have $b > 1$ and so populate halos on the exponentially declining part of the halo mass function. Halos significantly more massive than $M_{\rm cut}$ are very rare, and thus our results are quite insensitive to the behavior of $\bar{n}_{\rm cen}$ or $\bar{n}_{\rm sat}$ for these masses.

JCAP08(2024)020

realization to realization. This is due to sample variance from the small observational volume of a single field. We make no attempt to correct for this effect. The left panel of figure 4 shows the errors on the clustering inferred for a single field (i.e., not the error on the mean). We see that our pipeline returns an unbiased measure of the clustering from the mocks and a high-precision measurement of the clustering is anticipated from the data. At small scales the bins are largely uncorrelated (right panel of figure 4) as expected for shot-noise. As we move to larger scales the points become increasingly correlated as we enter the sample-variance dominated regime and the finite survey volume becomes important. We use the full covariance matrix for all of our inferences, thus including these correlations.

Since the main purpose of our simulations is to provide mock catalogs that match the clustering of the samples under consideration, we regard them as adequate. However future work in this direction is clearly desirable. A preliminary investigation of the occupancy of LAEs within the MTNG simulation [39] suggests that our modeling is reasonable under the approximation that Lyα line strength closely follows the star-formation rate (which is what the simulation determines best), but additional investigation of the LAE population in hydrodynamic simulations including modeling of the observed Lyα line flux would obviously be valuable [32–38].

5 Results

Figure 5 shows the measured configuration-space clustering, $w_\theta(R)$, with angles converted to tranverse distances using the χ_0 listed in table 2, compared to the predictions of our mock catalogs. The error bars are determined from the mock catalog Monte-Carlo observations and represent the square root of the diagonal of the covariance matrix (see figure 4). Figure 6 shows the angular power spectra, and the window function, for the samples in the COSMOS field. Again the error bars represent the square root of the diagonal of the covariance matrix. The lines indicating the mock catalog angular power spectra are for the model fit to $w_\theta(R)$, i.e., the same model shown in figure 5, and not a separate fit to the angular power spectrum data itself. For completeness, figure 7 shows the mean halo occupation, $\langle N_{\rm gal}(M_h)\rangle$, of the HOD models that provide acceptable fits to the $w_\theta(R)$ data — the parameters and $w_\theta(R)$ predictions for the entire grid as well as the best-fit model are publicly available (see section 8).

The clustering is relatively well described by a power-law for $1 < R < 10\,h^{-1}\,$Mpc, and it has been common in the literature to fit a power-law to the data (see also appendix A). Several further approximations need to be made in order to fit a power-law model, most notably how to handle redshift-space distortions. The size of the distortion depends upon scale and bias, with the largest effect arising when $R \sim \Delta$ and the bias is small. To handle this we infer r_0 from the projected clustering, $w_p(R)$, measured in our mock catalogs. Specifically for each HOD model we derive the goodness-of-fit to the data (χ^2) and also the projected clustering, $w_p(R)$. From $w_p(R_{\rm fid} = 5\,h^{-1}\,{\rm Mpc})$ we infer r_0 assuming $\xi_{\rm real}(r) = (r_0/r)^\gamma$ with $\gamma = 1.8$, matching the assumptions most commonly employed in the literature:

$$r_0 = \left[\frac{\Gamma(\gamma/2)\; w_p(R_{\rm fid})}{R_{\rm fid}^{1-\gamma}\sqrt{\pi}\,\Gamma([\gamma-1]/2)}\right]^{1/\gamma} \tag{5.1}$$

JCAP08(2024)020

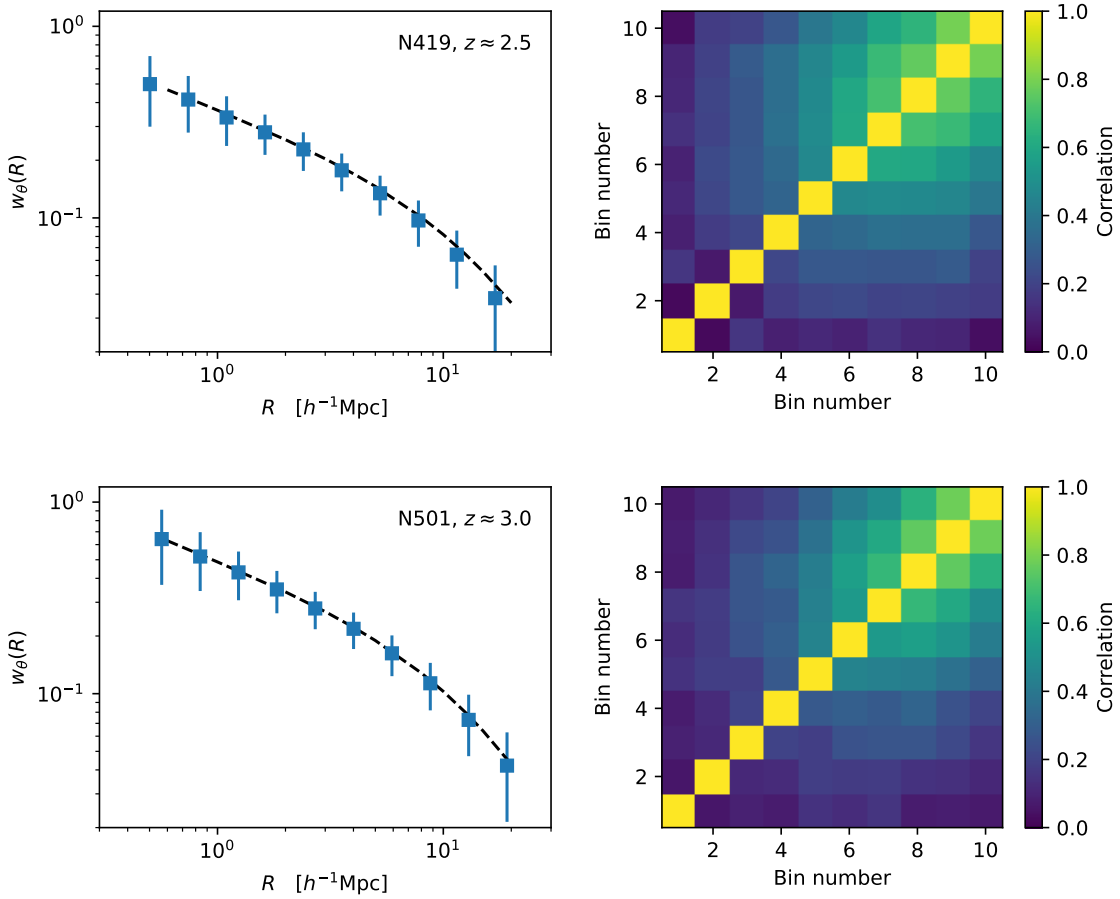

Figure 4. (Left) The average angular clustering measured from the 1024 mock catalogs for the N419 and N501 samples in the COSMOS field using the same pipeline as the observations, compared to the prediction from the correlation function multipoles measured directly in the N-body simulation using eq. (3.3). The agreement indicates the pipeline returns unbiased answers. The errors on the points indicate the standard deviation measured from the mocks, as appropriate for a single field, i.e. not the error on the mean. (Right) The correlation matrix. Note that at small scales the bins are close to independent, as expected for shot-noise, while at larger scales neighboring bins become more correlated as expected in the sample variance dominated regime.

We find a good correlation between r_0 and χ^2, suggesting that this clustering amplitude is a good summary statistic for the HOD models we investigate (see also ref. [52]). We then compute a best fit and standard deviation from the χ^2 minimum and $\Delta\chi^2$ assuming Gaussian statistics. These values are reported in table 2 and compared to earlier estimates of the same quantity in figure 8. We find that our measured clustering is lower than the predictions for similar samples in the simulation of ref. [39], though within the envelope of previous observational determinations even though our selection is slightly different.

A key quantity of interest for cosmological applications of LAEs is the large-scale bias, which impacts the degree of anisotropy from redshift-space distortions and in combination with the 3D number density the 'signal to noise' ratio for the power spectrum. The uncertainties from the ODIN observations become large on quasi-linear scales where we expect the bias

JCAP08(2024)020

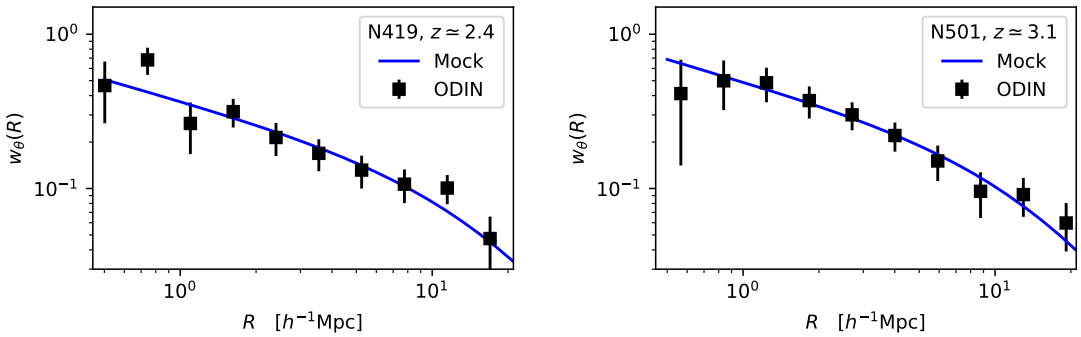

Figure 5. Clustering predicted in the fiducial mock catalogs (lines) compared to the ODIN data (black points with error bars) for the N419 and N501 samples. The error bars indicate the diagonal entries of the covariance matrix derived from the same mock catalogs and described further in the text.

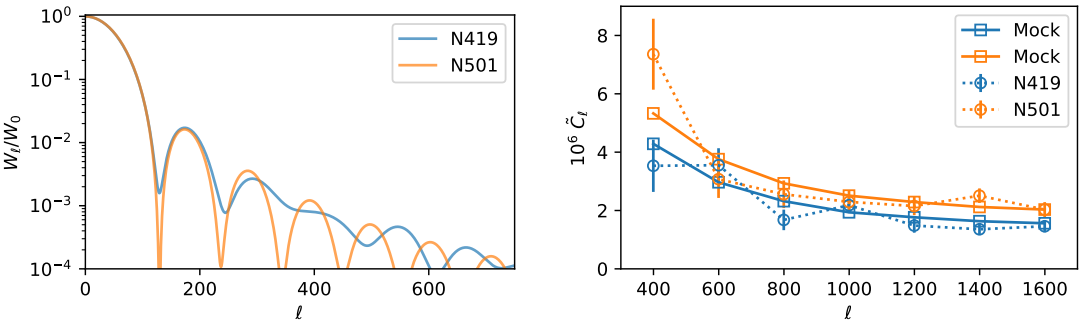

Figure 6. Window function (left) and angular power spectra (right) for the N419 and N501 samples in the COSMOS field. The small angular extent of the survey means that multipoles are coupled, with an extent determined by W_ℓ. For this reason we bin the power spectrum in bins of $\Delta\ell = 200$. The circles connected by dashed lines indicate the pseudo-C_ℓ measured on the data while the squares with error bars indicate the mean and standard deviation from the mocks described in the text (which were fit to the $w_\theta(R)$ measurements of figure 5). The errors are correlated at the $\approx 20\%$ level, and the spectra are shot-noise dominated beyond $\ell = 10^3$.

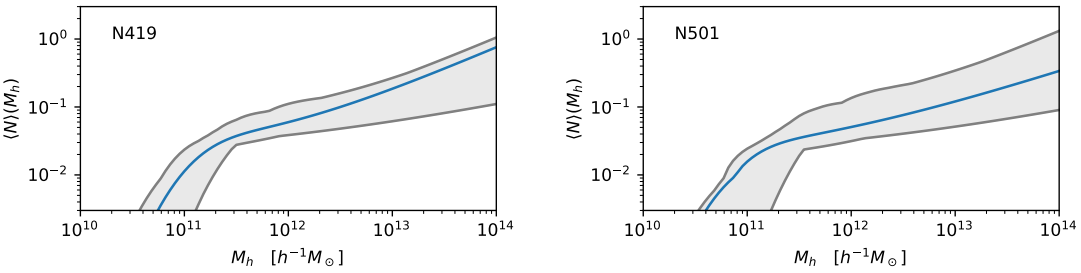

Figure 7. Mean LAE halo occupation in the mock catalogs. The blue lines show the best-fitting model in our grid, the grey bands show the range of $\langle N_{\rm gal} \rangle$ in the HOD models that provide acceptable fits to the ODIN data ($\chi^2 < 18$ for 10 d.o.f.) for the N419 (left) and N501 (right) samples. As discussed in the text, the high-M_h behavior is not well constrained by the data but this presents little difficult for cosmological interpretation.

JCAP08(2024)020

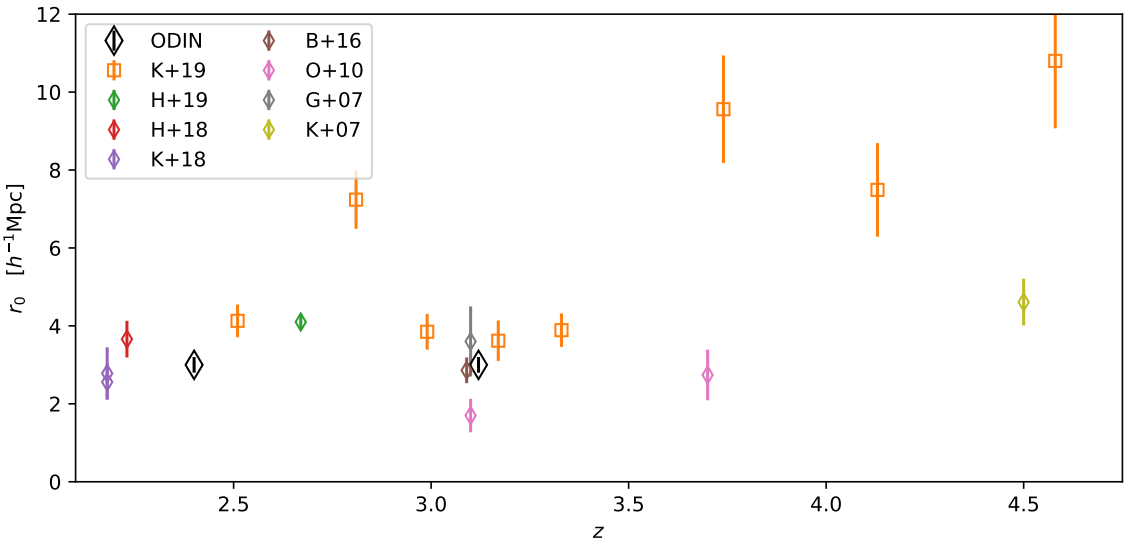

Figure 8. Comparison of our correlation length estimates to a selection of earlier work [40–42, 46, 47, 53–55]. Our results are generally consistent with earlier estimates, despite differences in selection. At 9 sq.deg. our results represent the largest contiguous area amongst these surveys.

to approach its scale-independent, large-scale value. However we can use the mock catalogs to measure the bias of the models that provide good fits to the ODIN data over a wider range of scales by directly comparing the mock galaxy clustering to that of the underlying matter field in the simulations — to the extent that these models describe the halo occupancy of LAEs in the real Universe this can give a guide to the degree of scale-dependence we expect. We discuss this further below.

6 Forecasts

Since they are numerous, have a bias similar to that of massive galaxies at $z < 1$ and prominent emission lines for ease of redshifting, LAEs make appealing targets for future redshift surveys aimed at high-z large-scale structure and primordial physics [2–4, 52]. In this section we look at the implications of our observations, interpreted within the context of the mock catalogs we have produced. While they are only our first attempt at modeling the LAE population, since in those catalogs we have access to the underlying dark matter, have no observational artifacts and almost no sample variance (due to our use of control variates, discussed above) they allow us to make inferences about how well future surveys could perform. The surveys, and analytic theory, also provide a plausible extension of the observed clustering to larger scales where the cosmological signals largely lie. For brevity we focus on the $z \simeq 3.1$ population here, though the results for $z \simeq 2.45$ are qualitatively similar.

6.1 Real-space clustering

Figure 9 shows the real-space auto-spectrum of our mock LAEs, as well as the cross-spectrum with the (non-linear) dark matter field, in both Fourier (left) and configuration (right) space.

JCAP08(2024)020

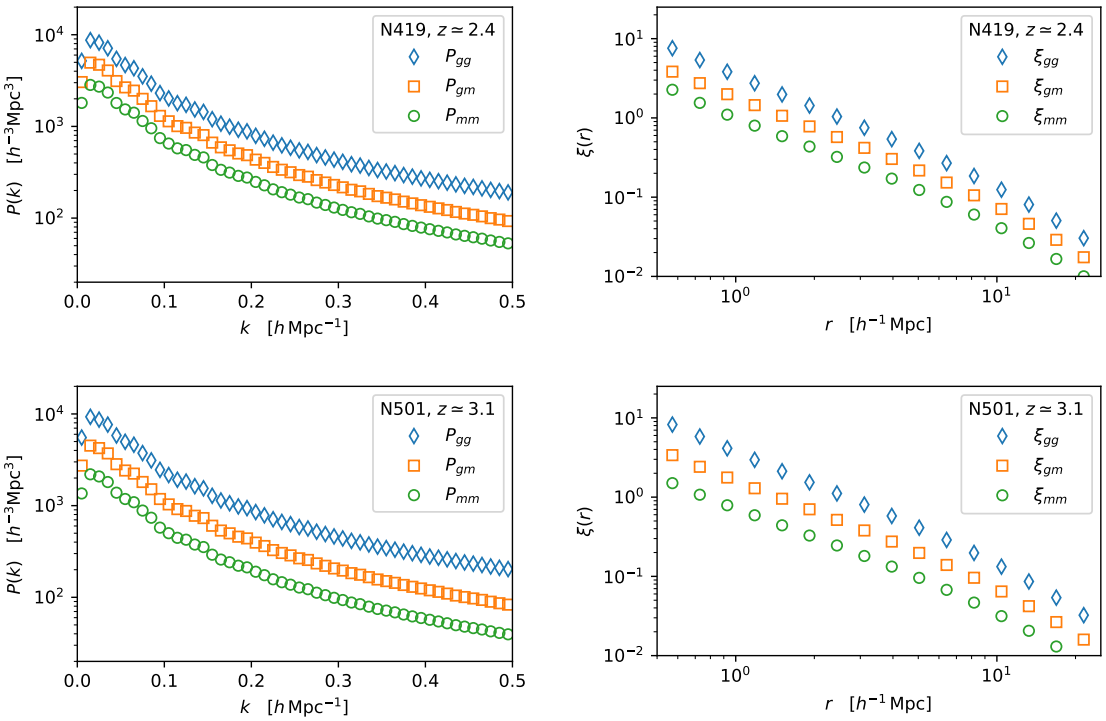

Figure 9. (Left) The real-space auto and cross power spectra of mock galaxies (with shot-noise suppressed; see text) and the dark matter for the N419 ($z \simeq 2.45$) and N501 ($z \simeq 3.1$) samples (upper and lower). (Right) The real-space auto and cross correlation functions.

In Fourier space, to better show the clustered component, we have suppressed shot-noise in the auto-spectrum by not randomly downsampling the mock LAEs in the calculation. The left panel of figure 10 shows the implied biases, $b_a = \sqrt{P_{gg}/P_{mm}}$ and $b_\times = P_{gm}/P_{mm}$, as a function of k. At large scales the two estimates agree and are scale-independent with a value consistent with that in our observations. At small scales (larger k) the two display scale dependence and begin to differ as the galaxy field decorrelates from the non-linear matter field ($r = P_{gm}/\sqrt{P_{gg}P_{mm}} = b_\times/b_a < 1$). The galaxy field decorrelates even more rapidly with the linear matter field, or with the initial conditions, and the presence of shot-noise leads to additional decorrelation over that shown in figure 10. The right panel of figure 10 shows the scale-dependent bias inferred from the configuration space statistics, where we see a similar level of scale-dependent bias but better agreement between b_a and $b_\times$ (though the interpretation of the ratio $b_\times/b_a$ as a correlation coefficient is not valid in configuration space). Both the scale-dependence of the bias and the decorrelation of the galaxy and matter fields is less pronounced than for a more biased population, such as we might expect of luminous LBGs (e.g. figures 8 and 9 of ref. [2]). This makes LAEs a good candidate for probing primordial physics [3, 52], though the samples most useful for cosmology would likely be selected using broader filters than the ODIN filters in order to increase the line-of-sight depth and hence survey volume.

Given the highly complex spectroscopic selection we have not attempted to measure the small-scale, line-of-sight clustering to provide constraints on the virial velocities in the sample

JCAP08(2024)020

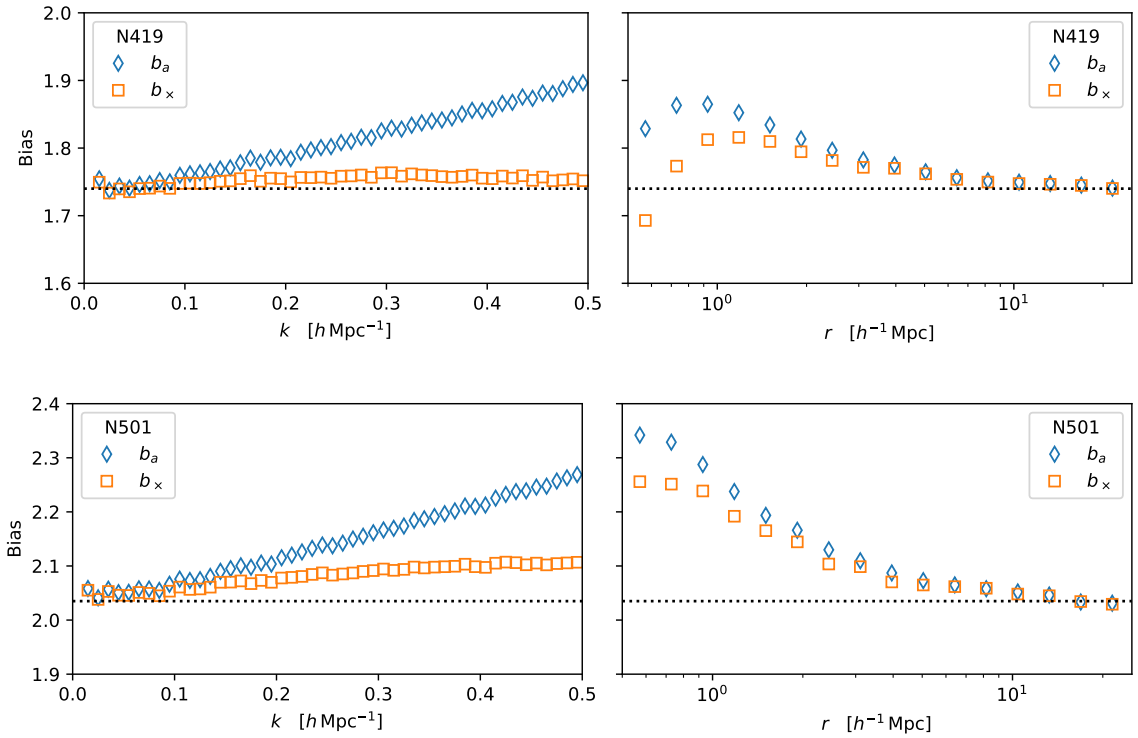

Figure 10. Scale dependent bias(es) for the N419 ($z \simeq 2.45$) mock sample (upper) and N501 ($z \simeq 3.1$) mock sample (lower). (Left) the Fourier-space biases, $b_a \equiv \sqrt{P_{gg}/P_{mm}}$ and $b_\times \equiv P_{gm}/P_{mm}$. (Right) The configuration-space biases, $b_a \equiv \sqrt{\xi_{gg}/\xi_{mm}}$ and $b_\times \equiv \xi_{gm}/\xi_{mm}$. The non-linear matter 2-point function is used in all cases. The biases are noticeably scale-dependent on small scales, and $b_a(k) \geq b_\times(k)$ which indicates that $r_{gm}(k) < 1$ on small scales.

(also known as "fingers of god" or "stochastic velocities"). We anticipate that these will be small, because the large-scale bias implies the halos hosting our galaxies lie primarily on the steeply falling, high-mass tail of the halo mass function. If satellite LAEs live in even more massive halos than central LAEs their contribution would be highly suppressed (indeed in our best-fitting HOD models the satellite fraction is only a few percent). If LAE activity is underrepresented in the satellite population this would provide further suppression of the satellite fraction. To the extent that central galaxies move with their host halo, most of the impact of virial velocities comes from central-satellite pairs and these pairs would be correspondingly suppressed. This bodes well for using LAEs as cosmological tracers, since virial velocities imply an unavoidable loss of cosmological constraining power on small scales. It will be very valuable to investigate this issue further with more spectroscopic data.

6.2 Cross-correlation with CMB lensing

These LAEs trace large-scale structure in a regime where the kernel of the CMB lensing is still large. This allows us to measure the cross-correlation between the galaxy and (projected) matter density, which helps to isolate the lensing contributions to a narrow redshift slice. Comparison of the amplitude of clustering inferred by the relativistic tracers (photons) with

JCAP08(2024)020

that inferred from the non-relativistic tracers (LAEs) provides constraints on the theory of gravity. The real-space power spectra allow us to predict the angular power spectra that would be measured for such galaxies. For example, for a thin shell of galaxies at distance χ_0 of width $\Delta\chi \ll \chi_0$ we have

$$C_\ell^{gg} \approx \mathcal{V}^{-1} P_{gg}(k = \ell/\chi_0)\,, \quad C_\ell^{g\kappa} \approx W^\kappa(\chi_0)\chi_0^{-2} P_{gm}(k = \ell/\chi_0) \tag{6.1}$$

where $\mathcal{V} = \chi_0^2\,\Delta\chi$ is the volume per steradian and

$$W^\kappa(\chi) = \frac{3}{2}\Omega_{m0}H_0^2(1+z)\frac{\chi(\chi_* - \chi)}{\chi_*} \tag{6.2}$$

is the CMB lensing kernel with $\chi_* = \chi(z_*{\approx}1100) \approx 9400\,h^{-1}\,\mathrm{Mpc}$ the distance to the surface of last scattering (we have neglected the magnification terms for simplicity). Note the galaxy auto-spectrum increases as we decrease the width of the shell (and there is less "washing out" of the clustering signal) while the cross-spectrum amplitude is independent of $\Delta\chi$. A very similar expression holds for the cross-spectrum with cosmic shear, i.e. galaxy-galaxy lensing, though we don't anticipate having competitive shear measurements at high redshift.

The error on $C_\ell^{g\kappa}$ depends upon C_ℓ^{gg}, $C_\ell^{g\kappa}$ and $C_\ell^{\kappa\kappa}$ and thus indirectly upon $\Delta\chi$. A very narrow slice in redshift, such as the ODIN samples analyzed above, has little cross-correlation with CMB lensing since it represents such a tiny fraction of the path traversed by the photons. However if similar galaxies could be selected over a broader redshift range, e.g. by using a series of narrow- or medium-band filters, then very high significance detections could be obtained. Figure 11 shows $C_\ell^{g\kappa}$ and its error (per ℓ-mode) for the best-fitting model at $z = 3$ and CMB noise levels appropriate to the Simons Observatory [56] or CMB-S4 [57] assuming they overlap on 30% of the sky and the galaxies cover $\Delta z = 0.5$. In either scenario the cross-spectrum would be detected at very high significance. For smaller fractions of the sky surveyed the errors scale as $f_{\rm sky}^{-1/2}$, so even smaller surveys would provide highly significant detections.

6.3 Redshift-space clustering

Moving into 3D and redshift space, figure 12 (left) shows the predicted monopole and quadrupole moments of the redshift-space power spectrum. The ratio of the two on large scales allows us to measure the growth rate, $f\sigma_8$, with the SNR in this regime being larger the lower the bias and the larger the volume. The comparison of the two moments at larger k allows constraints on the satellite fraction, virial motions within the halo and scale-dependent bias [39, 44] though with little constraining power on $f\sigma_8$ [58]. As discussed previously, since such virial motions represent a key limit to our modeling it would be good to revisit the observational constraints with more data. In the meantime, figure 12 also shows a perturbative model [59–61] fit to the N-body data over the range $0.02 < k < 0.45\,h\,\mathrm{Mpc}^{-1}$. The fit is excellent, and shows that the shift of the non-linear scale to high k at high z means that perturbative models can fit the N-body data over a broad range of scales, providing further support for the forecasting framework used in ref. [52]. The right hand panel of figure 12 shows the redshift-space clustering in configuration space, with the points being the N-body results and the lines showing linear theory and 1-loop perturbation theory. The BAO peak visible at $s \sim 100\,h^{-1}\,\mathrm{Mpc}$ in ξ_0 is only modestly broadened by non-linear evolution as the non-linear scale is smaller than the Silk damping scale.

JCAP08(2024)020

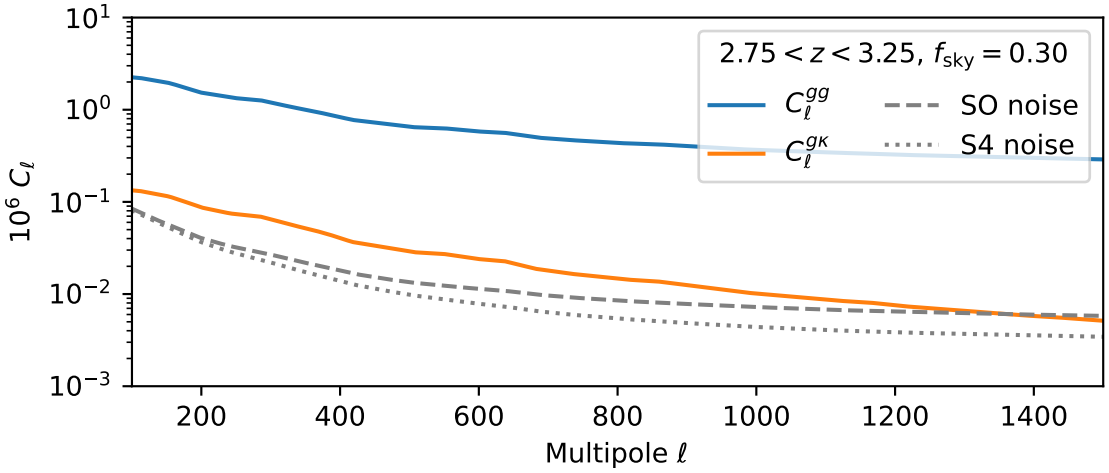

Figure 11. The projected auto spectrum (C_ℓ^{gg}) and galaxy-kappa cross spectrum ($C_\ell^{g\kappa}$) assuming LAEs in a slice $2.75 < z < 3.25$ clustering in the same way as our best-fit LAE sample at $z = 3$. The grey dashed and dotted lines show the error-per-ℓ-mode on $C_\ell^{g\kappa}$ for noise levels appropriate to Simons Observatory (SO) and CMB-S4 assuming the surveys overlap on 30% of the sky. This is to be compared to the signal plotted as the orange line. Such a cross-correlation would be detected at very high significance by either experiment.

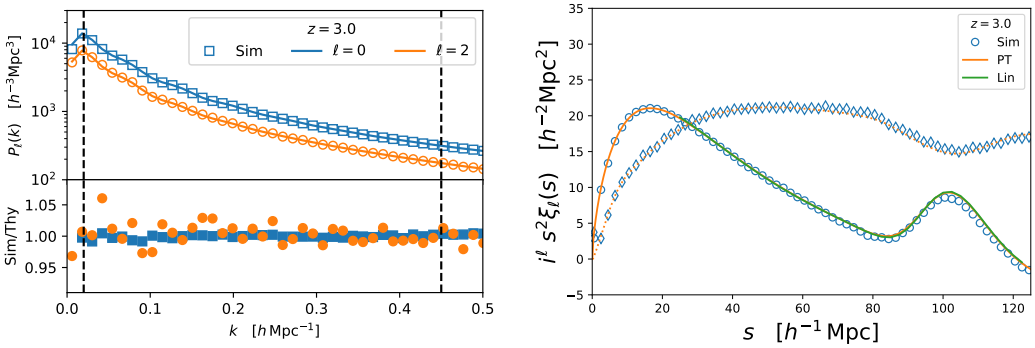

Figure 12. (Left) The monopole and quadrupole moments of the redshift-space power spectrum measured in the simulation (points; using control variates to reduce noise at large scales), compared to the best-fitting perturbation theory model (lines). The lower panel shows the ratio of simulation to theory while the vertical dashed lines denote the range of scales fit. (Right) The monopole and quadrupole of the correlation function measured in the simulation (points) compared to the theoretical prediction from the model of the left panel and linear theory with scale-independent bias. The broadening of the BAO peak by non-linear evolution is very weak at $z = 3$.

The large-scale bias measurements above also help determine what the simulation requirements would be in order to support a mock catalog effort for future LAE surveys. Since the central halo occupancy of LAEs is < 1 even in massive halos, the number density alone cannot be used to infer the characteristic mass of LAE-hosting halos. That information must come from the large-scale clustering. Our results (figure 5), and the HOD models

JCAP08(2024)020

that fit them, suggest that the majority of the LAEs we have selected from the ODIN survey reside in halos with $M_h \geq 10^{10.5}\,h^{-1}\,M_\odot$, though with a tail to lower masses. This in turn provides a target for the resolution of N-body simulations aimed at modeling this population. We note that the clustering measured here is lower than the predictions of ref. [39], suggesting that LAEs in those simulations live in more massive halos than LAEs selected from the ODIN survey.

One potential cause for concern in the use of LAEs as tracers of the matter field is the impact of radiative transfer (RT) on the population of galaxies that are selected. Early work [62] suggests that RT effects can modulate the number of galaxies in the large-scale structure catalog in a manner that depends upon the local density and line-of-sight velocity divergence. This latter effect is degenerate with the signal used in redshift-space distortion studies to measure $f\sigma_8$ (as the velocities are 'protected' by the equivalence principle and thus independent of galaxy bias). Follow-up work [63] argues that such a large effect was due to the limited resolution of those early simulations. In higher resolution simulations the regions where Lyα photons are produced tend to have higher density, such that the photons escaping from the galaxy have diffused further from the resonance and are less susceptible to the surrounding environment. Direct observations of these RT effects are currently not available, though the impact of HI absorption on LAE clustering has been detected [64]. Further theoretical and observational work on this is urgently needed, since the impact on forecasts for the science reach of future surveys employing LAEs as tracers is large [52].

7 Conclusions

Narrow or medium band[6] surveys provide an efficient route to selecting star-forming galaxies at high redshift [5]. Such galaxies can be valuable tracers of large-scale structure and enable constraints on cosmology and fundamental physics [52]. In this paper we have measured the clustering of galaxies selected from the ODIN survey, using follow-up spectroscopy from DESI to refine our selection, to infer the rate of interlopers or contaminants and calibrate the redshift distribution (section 2).

Our galaxies reside in thin shells of redshift at $z \simeq 2.45$ and 3.1. We measure the angular correlation function of the sample (section 3), converting from angles to transverse, comoving distance using the mean distance to the shell. This allows us to side-step the complex spectroscopic selection while paying little penalty in signal-to-noise ratio — for such shells the clustering on the large scales of most interest for cosmology is dominated by the transverse modes. We test our pipeline, measure the covariance of our clustering measurement and perform inference using a suite of high-resolution N-body simulations (section 4) that match the footprint, number density, redshift distribution and clustering of the observed sample.

We find that ODIN-selected LAEs have low clustering amplitudes, indicative of galaxies living in low-mass halos. Assuming a power-law correlation function with slope -1.8, we infer $r_0 = (3.0 \pm 0.2)h^{-1}\,\mathrm{Mpc}$ at both $z \simeq 2.45$ and 3.1. In the "Planck" [65] cosmology used in the N-body simulations these correspond to large-scale biases of 1.7 ± 0.2 and 2.0 ± 0.2. We

[6]The Rubin Observatory's Legacy Survey of Space and Time (LSST) will reach faint-enough magnitudes to detect a fraction of the LAE population as Lyman break galaxies (LBGs), but optimizing the selection of faint broad-band selected sources which exhibit Lyα emission remains to be done.

JCAP08(2024)020

compare our clustering measurements to earlier values from the literature (figure 8), with our results being in the same range as — though slightly lower than — previous measurements at comparable redshifts even though our selection differs slightly.

LAEs similar to those studied here make appealing targets for future spectroscopic surveys [52]. We use a combination of analytic theory and mock catalogs, constrained by the ODIN-DESI data, to look at the cosmological constraining power of samples with similar properties to our data (section 6). We anticipate that LAEs similar to those studied here could potentially form excellent targets for future, high-z, spectroscopic surveys aimed at constraining cosmology and fundamental physics if they could be efficiently selected over a wider redshift range and the impacts of radiative transfer on the selection function were small.

8 Data availability

Data from the plots in this paper are available as part of DESI's Data Management Plan. The grid of models, predictions and scripts for the analysis and making of the plots can be found at https://github.com/martinjameswhite/AnalyzeLAE and the remaining data are available at 10.5281/zenodo.11043784.

Acknowledgments

M.W. is supported by the DOE. AD's research activities are supported by the NSF NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. AD's research is also supported in part by Fellowships from the John Simon Guggenheim Memorial Foundation. EG and KSL acknowledge financial support from the National Science Foundation under Grant Nos. AST-2206222 and 2206705. EG acknowledges support from the U.S. Department of Energy, Office of Science, Office of High Energy Physics Cosmic Frontier Research program under award No. DE- SC0010008. ADM was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0019022. This research has made use of NASA's Astrophysics Data System and the arXiv preprint server. This research is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy 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.

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 (CONAHCYT); the Ministry of Science and Innovation of Spain (MICINN), and

JCAP08(2024)020

by the DESI Member Institutions: https://www.desi.lbl.gov/collaborating-institutions. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation, the U.S. Department of Energy, or any of the listed funding agencies.

The authors are 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.

This project is based in part on observations using the Dark Energy Camera (DECam) at the Victor M. Blanco telescope at the NSF Cerro Tololo Inter-American Observatory, NSF NOIRLab (NOIRLab Prop. ID 2020B-0201; PI: K.S. Lee), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the U.S. National Science Foundation. DECam was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute for Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Enérgeticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF NOIRLab, the University of Nottingham, the Ohio State University, the OzDES Membership Consortium, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

Facilities: Blanco (DECam), Mayall (DESI)

A Power-law model

It is common in the literature to fit power-law models to LAE clustering data. To make contact with that practice, here we develop the expressions in the main text for the power-law case. Specifically we assume $P(k) \propto k^n$, with $-3 < n < -1$, and $\Omega_m(z) \approx 1$, appropriate for high redshift observations. We need to further make an assumption about bias and redshift-space distortions. We shall take the simplest case with scale-independent bias and linear redshift-

JCAP08(2024)020

space distortions though the validity of these assumptions on small scales is questionable. With these approximations we can derive expressions for all of our observables analytically.

In terms of the variance per $\log k$, $\Delta^2_{\rm real}(k) = k^3 P_{\rm real}(k)/(2\pi^2) \equiv (k/k_\star)^{3+n}$, the multipoles of the redshift-space correlation function are related to the power spectrum multipoles via a Hankel transform

$$\xi_\ell(r) = i^\ell \int \frac{dk}{k} \Delta^2_\ell(k)\, j_\ell(kr) = k_\ell \mathcal{I}_\ell(n)\, (k_\star r)^{-(3+n)} \tag{A.1}$$

with j_ℓ is the spherical Bessel function of order ℓ and the k_ℓ

$$k_{\rm real} = b^2\,, \quad k_0 = \left(b^2 + \frac{2}{3}bf + \frac{f^2}{5}\right)\,, \quad k_2 = \left(\frac{4}{3}bf + \frac{4}{7}f^2\right)\,, \quad k_4 = \frac{8f^2}{35} \tag{A.2}$$

with $f \simeq \Omega_m^{0.55}(z) \approx 1$ and we have defined

$$\mathcal{I}_\ell(n) \equiv i^\ell \int_0^\infty x^{2+n}\, j_\ell(x)\, dx \qquad \text{for} \quad -3 < n < -1 \tag{A.3}$$

$$= i^\ell\, 2^{1+n}\, \sqrt{\pi}\, \frac{\Gamma([\ell+n+3]/2)}{\Gamma([\ell-n]/2)} \tag{A.4}$$

If we choose $n = -1.2$ then $\xi \propto r^{-1.8}$, which is a frequently used slope as it closely matches the observed correlation functions of several galaxy samples. We have $\mathcal{I}_0(-1.2) \simeq 1.10725$, $\mathcal{I}_2(-1.2) \simeq -1.66087$ and $\mathcal{I}_4(-1.2) \simeq 1.97229$. Thus

$$\xi_{\rm real}(r) = b^2 \left(\frac{r_\star}{r}\right)^{1.8} \quad \Rightarrow \quad \frac{w_p(R)}{R} \simeq 3.679\, b^2 \left(\frac{r_\star}{R}\right)^{1.8} \tag{A.5}$$

$$\xi_0(s) = \left(b^2 + \frac{2}{3}bf + \frac{f^2}{5}\right) \left(\frac{r_\star}{s}\right)^{1.8} \tag{A.6}$$

$$\xi_2(s) = -\frac{3}{2}\left(\frac{4}{3}bf + \frac{4}{7}f^2\right) \left(\frac{r_\star}{s}\right)^{1.8} \tag{A.7}$$

$$\xi_4(s) = \frac{57}{32}\,\frac{8f^2}{35} \left(\frac{r_\star}{s}\right)^{1.8} \tag{A.8}$$

where $r_\star \simeq (0.945\, k_\star)^{-1}$ is the matter correlation length and we have used s for the redshift-space separation vector to distinguish it from the real-space correlation.

For a top-hat redshift distribution of full-width Δ the angular correlation function is then

$$w_\theta(R) = 2\int_0^1 dy\ (1-y) \sum_\ell \xi_\ell(\sqrt{R^2 + y^2\Delta^2}) \mathcal{L}_\ell \left(\frac{y\Delta}{\sqrt{R^2+y^2\Delta^2}}\right) \tag{A.9}$$

where $\mathcal{L}_\ell$ is the Legendre polynomial of order ℓ. The integral can be written in closed form using Gaussian hypergeometric functions though the expressions are cumbersome. If we

JCAP08(2024)020

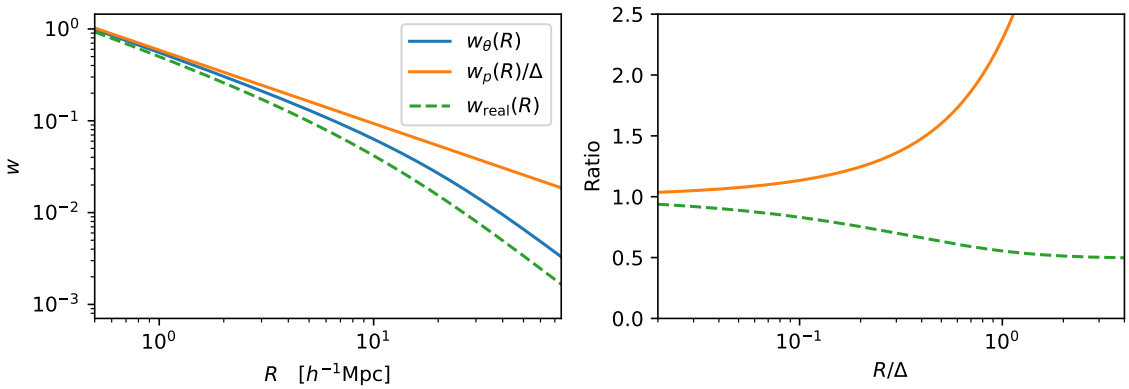

Figure 13. Power-law example with $\xi_{\rm real} = (1\,h^{-1}\,{\rm Mpc}/r)^{1.8}$, $b = 2$ and $\Delta = 25\,h^{-1}\,{\rm Mpc}$. (Left) The full calculation (eq. (A.9); blue) compared to two approximations. The solid orange line shows the approximation $w_\theta \approx w_p/\Delta$ that holds on small scales (eq. (A.11)) while the dashed green line shows w_θ computed using the real-space correlation function rather than the redshift-space correlation function. (Right) The ratio of the two approximations to the full calculation as a function of R/Δ. Note that both approximations agree with the full result in the limit $R \ll \Delta$, i.e. a "thick" shell. Neglecting RSD leads to an underestimate of the clustering on scales approaching the shell thickness while the w_p approximation overestimates power in the same limit. For lower bias the impact of RSD is larger.

expand for large R we find

$$\begin{aligned} w_\theta(R) \to k_0 \left(\frac{r_\star}{R}\right)^{1.8} \left(1 - 0.15\frac{\Delta^2}{R^2} + \cdots\right) \\ + \frac{3\,k_2}{4} \left(\frac{r_\star}{R}\right)^{1.8} \left(1 - 0.65\frac{\Delta^2}{R^2} + \cdots\right) \\ + \frac{171\,k_4}{256} \left(\frac{r_\star}{R}\right)^{1.8} \left(1 - 1.82\frac{\Delta^2}{R^2} + \cdots\right) + \cdots \quad \text{as } R \to \infty \end{aligned} \tag{A.10}$$

Note that the contribution from the quadrupole and hexadecapole are similar to that of the monopole, depending upon the value of the bias (and hence the relative sizes of k_0 and k_2, k_4). For example, for $b = 2$ we have $(k_0 + 3k_2/4 + 171k_4/256) \simeq 2\,k_{\rm real}$. Expressions in the limit $\Delta \gg R$ can also be derived simply

$$w_\theta(R) \to \frac{3.679\,R}{\Delta} \left(\frac{r_\star}{R}\right)^{1.8} \left[k_0 - \frac{k_2}{2} + \frac{3k_4}{8}\right] = \frac{w_p(R)}{\Delta} \quad \text{as } R \to 0 \tag{A.11}$$

On scales much smaller than the shell width, Δ, the effect of RSD is reduced and $w_\theta(R) \propto w_p(R)$. We caution that our simple linear model for RSD is likely not appropriate on small scales, and a more reasonable model would include scale-dependent bias, non-linear RSD and the finger-of-god effect (i.e. stochastic terms).

To aid with visualizing these results, figure 13 shows these approximations for a power-law model with $\xi_{\rm real} = (1\,h^{-1}\,{\rm Mpc}/r)^{1.8}$ and $b = 2$. The solid blue line shows the full calculation (eq. (A.9)), including (linear) redshift-space distortions and the finite shell width for a shell of

JCAP08(2024)020

width $\Delta = 25\,h^{-1}\,\mathrm{Mpc}$. The solid orange line shows the approximation $w_\theta \approx w_p/\Delta$ that holds on small scales, as derived above (eq. (A.11)). The dashed green line shows w_θ computed using the real-space correlation function rather than the redshift-space correlation function in eq. (A.9). Note that both approximations agree with the full result in the limit $R \ll \Delta$, i.e. a "thick" shell. Neglecting redshift-space distortions leads to an underestimate of the clustering on scales approaching the shell thickness while the w_p approximation overestimates power in the same limit. For lower bias the impact of redshift-space distortions is larger, while for higher bias it is reduced.

References

[1] A.S. Chou, M. Soares-Santos, T.M.P. Tait et al., *Snowmass Cosmic Frontier Report*, in the proceedings of the *Snowmass 2021*, Seattle, WA, U.S.A., 17–26 July 2022, arXiv:2211.09978 [INSPIRE].

[2] M.J. Wilson and M. White, *Cosmology with dropout selection: straw-man surveys & CMB lensing*, *JCAP* **10** (2019) 015 [arXiv:1904.13378] [INSPIRE].

[3] N. Sailer, E. Castorina, S. Ferraro and M. White, *Cosmology at high redshift — a probe of fundamental physics*, *JCAP* **12** (2021) 049 [arXiv:2106.09713] [INSPIRE].

[4] S. Ferraro, N. Sailer, A. Slosar and M. White, *Snowmass2021 Cosmic Frontier White Paper: Cosmology and Fundamental Physics from the three-dimensional Large Scale Structure*, arXiv:2203.07506 [INSPIRE].

[5] M. Ouchi, Y. Ono and T. Shibuya, *Observations of the Lyman-α Universe*, *Ann. Rev. Astron. Astrophys.* **58** (2020) 617 [arXiv:2012.07960] [INSPIRE].

[6] M. Levi, C. Bebek, T. Beers et al., *The DESI Experiment, a whitepaper for Snowmass 2013*, arXiv:1308.0847 [INSPIRE].

[7] DESI collaboration, *The DESI Experiment. Part I. Science, Targeting, and Survey Design*, arXiv:1611.00036 [INSPIRE].

[8] V. Ruhlmann-Kleider, C. Yèche, C. Magneville et al., *High redshift LBGs from deep broadband imaging for future spectroscopic surveys*, arXiv:2404.03569 [INSPIRE].

[9] K.-S. Lee, E. Gawiser, C. Park et al., *The One-hundred-*deg^2 *DECam Imaging in Narrowbands (ODIN): Survey Design and Science Goals*, *Astrophys. J.* **962** (2024) 36 [arXiv:2309.10191].

[10] V. Ramakrishnan, B. Moon, S.H. Im et al., *ODIN: Where Do Lyα Blobs Live? Contextualizing Blob Environments within the Large-Scale Structure*, *Astrophys. J.* **951** (2023) 119 [arXiv:2302.07860].

[11] N.M. Firestone, E. Gawiser, V. Ramakrishnan et al., *ODIN: Improved Narrowband Lyα Emitter Selection Techniques for* $z = 2.4$, 3.1, *and* 4.5, arXiv:2312.16075.

[12] A. Dey et al., *DESI Spectroscopy of Lyman Alpha Emitting Galaxy Candidates*, in preparation.

[13] SDSS collaboration, *The First data release of the Sloan Digital Sky Survey*, *Astron. J.* **126** (2003) 2081 [astro-ph/0305492] [INSPIRE].

[14] B. Flaugher, H.T. Diehl, K. Honscheid et al., *The Dark Energy Camera*, *Astron. J.* **150** (2015) 150 [arXiv:1504.02900] [INSPIRE].

[15] A. Dey, D.J. Schlegel, D. Lang et al., *Overview of the DESI Legacy Imaging Surveys*, *Astron. J.* **157** (2019) 168 [arXiv:1804.08657] [INSPIRE].

JCAP08(2024)020

[16] H. Aihara, N. Arimoto, R. Armstrong et al., *The Hyper Suprime-Cam SSP Survey: Overview and Survey Design*, *Publ. Astron. Soc. Jpn.* **70** (2018) S4 [arXiv:1704.05858] [INSPIRE].

[17] D. Lang, D.W. Hogg and D. Mykytyn, *The Tractor: Probabilistic astronomical source detection and measurement*, in *Astrophysics Source Code Library*, ascl:1604.008 (2016).

[18] DESI collaboration, *The DESI Experiment. Part II. Instrument Design*, arXiv:1611.00037 [INSPIRE].

[19] DESI collaboration, *Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **164** (2022) 207 [arXiv:2205.10939] [INSPIRE].

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

[21] T.N. Miller, P. Doel, G. Gutierrez et al., *The Optical Corrector for the Dark Energy Spectroscopic Instrument*, arXiv:2306.06310 [INSPIRE].

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

[23] PLANCK collaboration, *Planck 2018 results. Part I. Overview and the cosmological legacy of Planck*, *Astron. Astrophys.* **641** (2020) A1 [arXiv:1807.06205] [INSPIRE].

[24] A.D. Myers, J. Moustakas, S. Bailey et al., *The Target-selection Pipeline for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **165** (2023) 50 [arXiv:2208.08518] [INSPIRE].

[25] K.M. Górski, E. Hivon, A.J. Banday et al., *HEALPix — A Framework for high resolution discretization, and fast analysis of data distributed on the sphere*, *Astrophys. J.* **622** (2005) 759 [astro-ph/0409513] [INSPIRE].

[26] D.N. Limber, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field*, *Astrophys. J.* **117** (1953) 134 [INSPIRE].

[27] A. Baleato Lizancos and M. White, *The impact of anisotropic redshift distributions on angular clustering*, *JCAP* **07** (2023) 044 [arXiv:2305.15406] [INSPIRE].

[28] A. Baleato Lizancos and M. White, *Harmonic analysis of discrete tracers of large-scale structure*, *JCAP* **05** (2024) 010 [arXiv:2312.12285] [INSPIRE].

[29] S.D. Landy and A.S. Szalay, *Bias and variance of angular correlation functions*, *Astrophys. J.* **412** (1993) 64 [INSPIRE].

[30] N.A. Maksimova, L.H. Garrison, D.J. Eisenstein et al., *AbacusSummit: a massive set of high-accuracy, high-resolution N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4017 [arXiv:2110.11398] [INSPIRE].

[31] L.H. Garrison, D.J. Eisenstein, D. Ferrer et al., *The abacus cosmological N-body code*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 575 [arXiv:2110.11392] [INSPIRE].

[32] M. McQuinn, L. Hernquist, M. Zaldarriaga and S. Dutta, *Studying Reionization with Lyα Emitters*, *Mon. Not. Roy. Astron. Soc.* **381** (2007) 75 [arXiv:0704.2239] [INSPIRE].

[33] K. Nagamine, M. Ouchi, V. Springel and L. Hernquist, *Lyman-α Emitters and Lyman-break Galaxies at $z = 3$–6 in Cosmological SPH Simulations*, *Publ. Astron. Soc. Jpn.* **62** (2010) 1455 [arXiv:0802.0228] [INSPIRE].

[34] H. Yajima, J.-H. Choi and K. Nagamine, *Effect of radiative transfer on damped Lyα and Lyman limit systems in cosmological SPH simulations*, *Mon. Not. Roy. Astron. Soc.* **427** (2012) 2889 [arXiv:1112.5691] [INSPIRE].

[35] E. Lake, Z. Zheng, R. Cen et al., *On the Diffuse Lyα Halo Around Lyα Emitting Galaxies*, *Astrophys. J.* **806** (2015) 46 [arXiv:1502.01349] [INSPIRE].

[36] T. Garel, J. Blaizot, B. Guiderdoni et al., *The UV, Lyman α, and dark matter halo properties of high-redshift galaxies*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 1279 [arXiv:1503.06635] [INSPIRE].

[37] L.H. Weinberger, G. Kulkarni, M.G. Haehnelt et al., *Lyman-α emitters gone missing: the different evolution of the bright and faint populations*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 2564 [arXiv:1803.03789] [INSPIRE].

[38] L.H. Weinberger, M.G. Haehnelt and G. Kulkarni, *Modelling the observed luminosity function and clustering evolution of Lyman-α emitters: growing evidence for late reionization*, *Mon. Not. Roy. Astron. Soc.* **485** (2019) 1350 [arXiv:1902.05077] [INSPIRE].

[39] J. Ravi, B. Hadzhiyska, M. White et al., *Examining Lyman-alpha Emitters through MillenniumTNG in anticipation of DESI-II*, arXiv:2403.02414 [INSPIRE].

[40] R.M. Bielby, P. Tummuangpak, T. Shanks et al., *The VLT LBG redshift survey. Part V. Characterizing the $z = 3.1$ Lyman α emitter population*, *Mon. Not. Roy. Astron. Soc.* **456** (2016) 4061 [arXiv:1501.01215] [INSPIRE].

[41] H. Kusakabe, K. Shimasaku, M. Ouchi et al., *The stellar mass, star formation rate and dark matter halo properties of LAEs at $z \sim 2$*, *Publ. Astron. Soc. Jpn.* **70** (2018) 4 [arXiv:1707.09373] [INSPIRE].

[42] A.A. Khostovan, D. Sobral, B. Mobasher et al., *The clustering of typical Lyα emitters from $z \sim 2.5$–6: host halo masses depend on Lyα and UV luminosities*, *Mon. Not. Roy. Astron. Soc.* **489** (2019) 555 [arXiv:1811.00556].

[43] S. Yuan, L.H. Garrison, B. Hadzhiyska et al., *AbacusHOD: a highly efficient extended multitracer HOD framework and its application to BOSS and eBOSS data*, *Mon. Not. Roy. Astron. Soc.* **510** (2022) 3301 [arXiv:2110.11412] [INSPIRE].

[44] M. Asgari, A.J. Mead and C. Heymans, *The halo model for cosmology: a pedagogical review*, *Open J. Astrophys.* **6** (2023) 39 [arXiv:2303.08752] [INSPIRE].

[45] Z. Zheng, A.L. Coil and I. Zehavi, *Galaxy Evolution from Halo Occupation Distribution Modeling of DEEP2 and SDSS Galaxy Clustering*, *Astrophys. J.* **667** (2007) 760 [astro-ph/0703457] [INSPIRE].

[46] E. Gawiser, H. Francke, K. Lai et al., *Lyα-Emitting Galaxies at $z = 3.1$: L^* Progenitors Experiencing Rapid Star Formation*, *Astrophys. J.* **671** (2007) 278 [arXiv:0710.2697] [INSPIRE].

[47] K. Kovač, R.S. Somerville, J.E. Rhoads et al., *Clustering of Lyman alpha emitters at $z \sim 4.5$*, *Astrophys. J.* **668** (2007) 15 [arXiv:0706.0893] [INSPIRE].

[48] M. Dijkstra and E. Westra, *Star Formation Indicators and Line Equivalent Width in Lyman Alpha Galaxies*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 2343 [arXiv:0911.1357] [INSPIRE].

[49] N. Kokron, S.-F. Chen, M. White et al., *Accurate predictions from small boxes: variance suppression via the Zel'dovich approximation*, *JCAP* **09** (2022) 059 [arXiv:2205.15327] [INSPIRE].

[50] J. DeRose, S.-F. Chen, N. Kokron and M. White, *Precision redshift-space galaxy power spectra using Zel'dovich control variates*, *JCAP* **02** (2023) 008 [arXiv:2210.14239] [INSPIRE].

[51] B. Hadzhiyska, M.J. White, X. Chen et al., *Mitigating the noise of DESI mocks using analytic control variates*, *Open J. Astrophys.* **6** (2023) 38 [arXiv:2308.12343] [INSPIRE].

[52] H. Ebina and M. White, *Cosmology before noon with multiple galaxy populations*, *JCAP* **06** (2024) 052 [arXiv:2401.13166] [INSPIRE].

[53] S. Hong, A. Dey, K.-S. Lee et al., *Statistics of two-point correlation and network topology for Ly α emitters at $z \approx 2.67$*, *Mon. Not. Roy. Astron. Soc.* **483** (2019) 3950 [arXiv:1811.10631] [INSPIRE].

[54] C.-N. Hao, J.-S. Huang, X. Xia et al., *A Deep Lyα Survey in ECDF-S and COSMOS. Part I. General Properties of Lyα Emitters at $z \sim 2$*, *Astrophys. J.* **864** (2018) 145 [arXiv:1808.02704].

[55] M. Ouchi, K. Shimasaku, H. Furusawa et al., *Statistics of 207 Lya Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models*, *Astrophys. J.* **723** (2010) 869 [arXiv:1007.2961] [INSPIRE].

[56] Simons Observatory collaboration, *The Simons Observatory: Astro2020 Decadal Project Whitepaper*, *Bull. Am. Astron. Soc.* **51** (2019) 147 [arXiv:1907.08284] [INSPIRE].

[57] K. Abazajian, G. Addison, P. Adshead et al., *CMB-S4 Science Case, Reference Design, and Project Plan*, arXiv:1907.04473 [INSPIRE].

[58] S.-F. Chen, Z. Vlah and M. White, *A new analysis of galaxy 2-point functions in the BOSS survey, including full-shape information and post-reconstruction BAO*, *JCAP* **02** (2022) 008 [arXiv:2110.05530] [INSPIRE].

[59] S.-F. Chen, Z. Vlah and M. White, *Consistent Modeling of Velocity Statistics and Redshift-Space Distortions in One-Loop Perturbation Theory*, *JCAP* **07** (2020) 062 [arXiv:2005.00523] [INSPIRE].

[60] S.-F. Chen, Z. Vlah, E. Castorina and M. White, *Redshift-Space Distortions in Lagrangian Perturbation Theory*, *JCAP* **03** (2021) 100 [arXiv:2012.04636] [INSPIRE].

[61] M. Maus, S. Chen, M. White et al., *An analysis of parameter compression and full-modeling techniques with Velocileptors for DESI 2024 and beyond*, arXiv:2404.07312 [INSPIRE].

[62] Z. Zheng, R. Cen, H. Trac and J. Miralda-Escude, *Radiative Transfer Modeling of Lyα Emitters. Part II. New Effects in Galaxy Clustering*, *Astrophys. J.* **726** (2011) 38 [arXiv:1003.4990] [INSPIRE].

[63] C. Behrens, C. Byrohl, S. Saito and J.C. Niemeyer, *The impact of Lyman-α radiative transfer on large-scale clustering in the Illustris simulation*, *Astron. Astrophys.* **614** (2018) A31 [arXiv:1710.06171] [INSPIRE].

[64] R. Momose, K. Shimasaku, K. Nagamine et al., *Catch Me if You Can: Biased Distribution of Lyα-emitting Galaxies according to the Viewing Direction*, *Astrophys. J. Lett.* **912** (2021) L24 [arXiv:2104.10580].

[65] Planck collaboration, *Planck 2018 results. Part VIII. Gravitational lensing*, *Astron. Astrophys.* **641** (2020) A8 [arXiv:1807.06210] [INSPIRE].