

Lawrence Berkeley National Laboratory

LBL Publications

Title

Planting a Lyman alpha forest on AbacusSummit

Permalink

<https://escholarship.org/uc/item/8tg176f3>

Journal

Monthly Notices of the Royal Astronomical Society, 524(1)

ISSN

0035-8711

Authors

Hadzhiyska, Boryana

Font-Ribera, A

Cuceu, A

et al.

Publication Date

2023-07-04

DOI

10.1093/mnras/stad1920

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

Planting a Lyman alpha forest on ABACUSUMMIT

Boryana Hadzhiyska,^{1,2*} A. Font-Ribera,³ A. Cuceu,^{4,5,6} S. Chabanier,¹ J. Aguilar,¹ D. Brooks,⁷ A. de la Macorra,⁸ P. Doel,⁷ D. J. Eisenstein,⁹ J. E. Forero-Romero,¹⁰ S. Gontcho A Gontcho,¹ K. Honscheid,^{4,5,6} R. Kehoe,¹¹ M. Landriau,¹ R. Miquel,^{12,3} Jundan Nie,¹³ W.J. Percival,^{14,15,16} G. Rossi,¹⁷ Gregory Tarlé,¹⁸ and Zhimin Zhou¹³

¹ Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

² University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA

³ Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra Barcelona, Spain

⁴ Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

⁵ Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

⁶ The Ohio State University, Columbus, 43210 OH, USA

⁷ Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK

⁸ Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México

⁹ Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

¹⁰ Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia

¹¹ Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA

¹² Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

¹³ National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China

¹⁴ Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

¹⁵ Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

¹⁶ Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

¹⁷ Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea

¹⁸ University of Michigan, Ann Arbor, MI 48109, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The full-shape correlations of the Lyman alpha ($\text{Ly}\alpha$) forest contain a wealth of cosmological information through the Alcock-Paczynski effect. However, these measurements are challenging to model without robustly testing and verifying the theoretical framework used for analyzing them. Here, we leverage the accuracy and volume of the N -body simulation suite ABACUSUMMIT to generate high-resolution $\text{Ly}\alpha$ skewers and quasi-stellar object (QSO) catalogs. One of the main goals of our mocks is to aid in the full-shape $\text{Ly}\alpha$ analysis planned by the Dark Energy Spectroscopic Instrument (DESI) team. We provide optical depth skewers for six of the fiducial cosmology base-resolution simulations ($L_{\text{box}} = 2 h^{-1} \text{Gpc}$, $N = 6912^3$) at $z = 2.5$. We adopt a simple recipe based on the Fluctuating Gunn-Peterson Approximation (FGPA) for constructing these skewers from the matter density in an N -body simulation and calibrate it against the 1D and 3D $\text{Ly}\alpha$ power spectra extracted from the hydrodynamical simulation IllustrisTNG (TNG; $L_{\text{box}} = 205 h^{-1} \text{Mpc}$, $N = 2500^3$). As an important application, we study the non-linear broadening of the baryon acoustic oscillation (BAO) peak and show the cross-correlation between DESI-like QSOs and our $\text{Ly}\alpha$ forest skewers. We find differences on small scales between the Kaiser approximation prediction and our mock measurements of the $\text{Ly}\alpha \times \text{QSO}$ cross-correlation, which would be important to account for in upcoming analyses. The ABACUSUMMIT $\text{Ly}\alpha$ forest mocks open up the possibility for improved modelling of cross correlations between $\text{Ly}\alpha$ and cosmic microwave background (CMB) lensing and $\text{Ly}\alpha$ and QSOs, and for forecasts of the 3-point $\text{Ly}\alpha$ correlation function. Our catalogues and skewers are publicly available on Globus via the National Energy Research Scientific Computing Center (NERSC) (full link under Data Availability).

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION

Over the last few decades, we have gained an enormous amount of knowledge about the expansion history of the Universe. With the

discovery of the accelerated expansion of the Universe via distance measurements of Type Ia supernovae (Riess et al. 1998; Perlmutter et al. 1999), an additional ingredient needed to be introduced into the cosmological paradigm. This new component, dubbed “dark energy,” took on the responsibility of explaining the mysterious repulsive force these measurements were finding. A couple of decades later,

* E-mail: boryanah@berkeley.edu

the nature of dark energy is still unknown, and several ongoing and planned surveys have committed to investigating its properties as their top priority (e.g., DESI, DES, Rubin Observatory) (DESI Collaboration et al. 2016a; Levi et al. 2019; Flaugher et al. 2015; Abbott et al. 2018; Dark Energy Survey Collaboration et al. 2016; LSST Dark Energy Science Collaboration 2012).

These surveys aim to measure the baryon acoustic oscillations (BAO) (Peebles & Yu 1970), a fixed-scale imprint on large-scale structure that allows us to measure both the angular diameter distance and the Hubble parameter across cosmic time and thus map out the expansion rate of the Universe, bringing important insights into the nature of dark energy. Typically, the BAO peak is measured in the clustering of galaxies, which are used as tracers of the matter density, via the two-point correlation function or the power spectrum (see Eisenstein et al. 2005; Cole et al. 2005, for first detections in data). Of recent interest are also measurements using quasi-stellar objects (QSOs), which offer an invaluable probe of the $z \sim 1.5$ expansion history of the Universe (e.g. Ata et al. 2018). In addition, when studying the galaxy and quasar clustering, additional information can be extracted from the amplitude of the redshift-space distortions (RSD), which encodes cosmological information in the form of $f\sigma_8$, a quantity sensitive to the growth of structure. The joint analysis of the growth of structure and the expansion rate has the potential to stress-test general relativity and constrain the various components of our cosmic inventory (see e.g., DESI Collaboration et al. 2016a).

The Lyman- α forest (Ly α forest) provides a powerful alternative probe for glimpsing at our Universe’s past. Comprised of a series of absorption features in the spectra of high-redshift quasars, these spectral features trace the density of neutral hydrogen, and thus the dark matter distribution, on scales larger than the Jeans length (Bi et al. 1992).

Apart from capturing the BAO feature, quasar spectra speckled with Ly α absorption features also contain valuable information on small scales, i.e. several megaparsecs, accessible via the one-dimensional flux power spectrum, P_{1D} (Croft et al. 1998a, 1999; McDonald et al. 2000; Zaldarriaga et al. 2001; Gnedin & Hamilton 2002; Croft et al. 2002; Viel et al. 2004a; McDonald et al. 2005, 2006; Viel & Haehnelt 2006; Yèche et al. 2017; Iršič et al. 2017b; Chabanier et al. 2019). Measurements of P_{1D} , in combination with cosmic microwave background (CMB) probes, have the potential to yield tight constraints on fundamental unknowns such as the sum of the neutrino masses, the shape of the primordial power spectrum, and some exotic dark matter models (see e.g., Phillips et al. 2001; Verde et al. 2003; Spergel et al. 2003; Viel et al. 2004b; Seljak et al. 2005, 2006; Bird et al. 2011; Iršič et al. 2017a; Baur et al. 2017; Murgia et al. 2018, 2019; Nori et al. 2019; Rogers & Peiris 2021b,a).

The ongoing Dark Energy Spectroscopic Instrument (DESI) survey will achieve an unprecedented precision in the Ly α forest measurements across all scales, amassing approximately a million quasar spectra at $z > 2$ over its five years of operation (for various specifications on the experiment, see Levi et al. 2013; DESI Collaboration et al. 2016b, 2022; Silber et al. 2022; Chaussidon et al. 2023). Ahead of such immense improvements in our statistics, a factor of four larger than current surveys, it is crucial that we diligently stress test our analysis pipelines and quantify the impact of secondary astrophysical effects. The most viable path forward is through the development of synthetic mock datasets (e.g. Le Goff et al. 2011; Font-Ribera et al. 2012; Bautista et al. 2015; Peirani et al. 2014a, 2022a; Sorini et al. 2016a; Farr et al. 2020; Sinigaglia et al. 2022), which must strike the careful balance of computational efficiency and survey realism.

In this work, we provide a new mock dataset, which aims to build upon previous such efforts in several key ways. Other large-scale

mocks adopted in the literature tend to compromise on the accuracy of their Ly α forest model, for example, by utilizing lognormal realizations instead of dark matter simulations, placing a greater emphasis on volume. Our mocks, on the other hand, are generated using the N -body simulation suite, ABACUSSUMMIT, and therefore provide greater realism in the non-linear regime than the lognormal mocks while also covering a sufficient volume of $\sim 100 \text{ Gpc}^3$ to satisfy the requirements of the DESI survey. In addition, the model used to create them is calibrated on the state-of-the-art hydrodynamical simulation IllustrisTNG and thus has an advantage over standard approaches for modeling the large-scale Ly α forest signal. At the same time, it is simple enough that it can be applied to an arbitrarily large number of simulations, without this exercise becoming prohibitively expensive, as in the case of the hydrodynamical simulations used in P_{1D} analysis.

A second major goal of this work is to integrate the 1D and 3D correlation function analyses. Typically, the BAO and P_{1D} analyses are carried out as independent probes, with the BAO measurements being modeled via linear perturbation theory, while the P_{1D} ones via hydrodynamical simulations that capture the physics of the intergalactic medium (IGM). The joint analysis of these measurements would not only improve the statistical uncertainty on cosmological parameters, but also make them more robust to systematic errors (Font-Ribera et al. 2018). In order to accomplish this, however, we need a theoretical framework that can be trusted on all scales. While the mocks presented in this work lack the gas and IGM physics needed to reliably model the smallest scales targeted by P_{1D} analyses, $k \sim 10 h \text{ Mpc}^{-1}$, they still support cosmological scales spanning several orders of magnitude, $0.001 \lesssim k \lesssim 1 h \text{ Mpc}^{-1}$. They thus allow an excellent opportunity to develop novel pipelines and statistics, beyond the standard BAO analysis, for extracting cosmological information from the full shape of the 3D correlations (see e.g., Cuceu et al. 2021). Such work is planned by the DESI collaboration in the near term, and our mocks provide an important first step towards reaching these goals. As an example, these mocks provide realistic connection between the QSOs and the Ly α forest, allowing for accurate modeling of their cross-correlation down to intermediate and small scales, which typically elude more simplistic mocks. Given the high resolution and large volume of the ABACUSSUMMIT simulations, the mocks presented in this work can be used to develop high-fidelity models for analyzing upcoming measurements of the Ly α forest.

This paper is organized as follows. In Section 2, we introduce the simulations and summary statistics employed in this study. In Section 3, we detail our procedure for generating the Ly α forest mocks and present a comparison with the high-velocity Ly α skewers extracted from the hydrodynamical simulation IllustrisTNG. In Section 4, we show the outcome of applying our algorithm to six of the N -body simulation suite boxes of ABACUSSUMMIT. In particular, we examine the 1D and 3D power spectra as well as the auto- and cross-correlation of the Ly α forest and QSOs, demonstrating the impact of non-linear clustering on these observables. We summarize our findings and discuss relevant implications about future work in Section 5.

2 METHODS

2.1 Simulations

In this Section, we introduce the two simulation suites relevant to this work: IllustrisTNG and ABACUSSUMMIT.

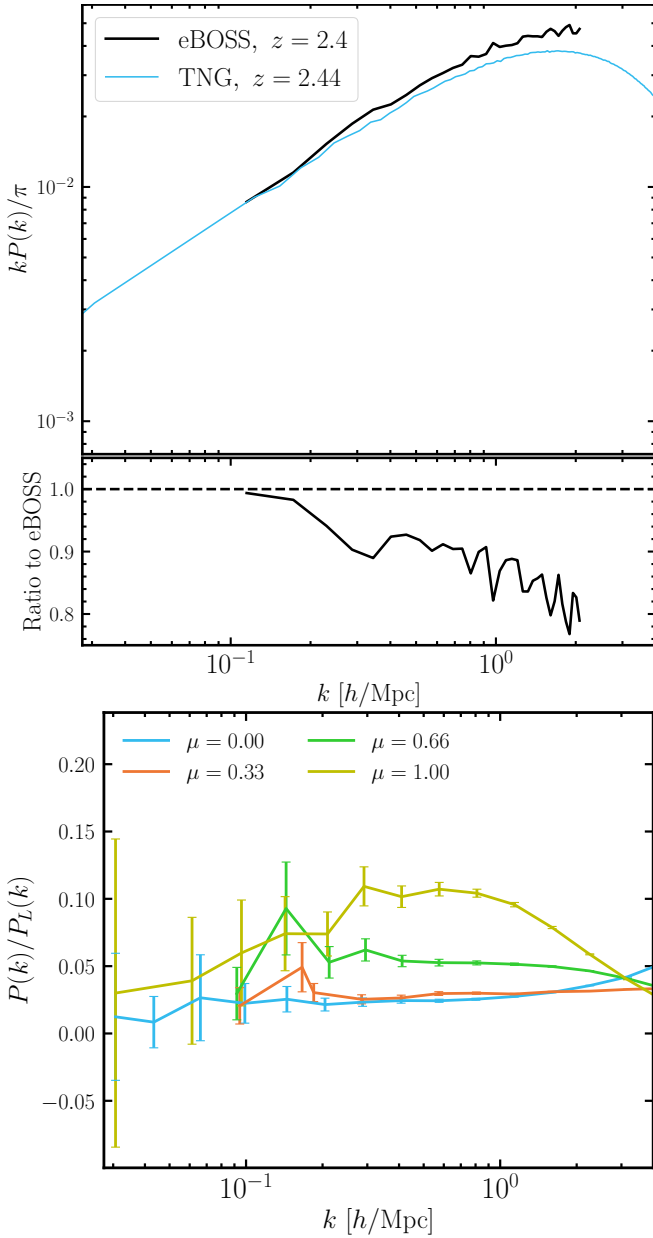


Figure 1. *Top panel:* Comparison of the 1D power spectrum of the Ly α forest skewers extracted from TNG300-1 (in blue) and eBOSS DR14 (in black) (Chabanier et al. 2019). We note that the eBOSS power spectrum is already noise and background subtracted, and the Ly α -Si correlations have been removed, facilitating the comparison with the simulation. The lower segment of the plot shows the ratio with respect to eBOSS. The agreement between simulations and observations is reasonably good. In particular, the discrepancy is $\lesssim 10\%$ for $k \lesssim 1$ h Mpc $^{-1}$, and then increases up to 20% on smaller scales, more prone to resolution effects. *Bottom panel:* 3D power spectrum of the TNG300-1 Ly α forest skewers, with values of the bias and the redshift distortion parameter given by $b_{\text{Ly}\alpha, \text{TNG}} = -0.1379$ and $\beta_{\text{Ly}\alpha, \text{TNG}} = 1.432$, respectively. These are slightly different from the eBOSS constraints of $b_{\text{Ly}\alpha, \text{eBOSS}} = -0.117$ and $\beta_{\text{Ly}\alpha, \text{eBOSS}} = 1.669$ at $z_{\text{eff}} = 2.334$ (du Mas des Bourboux et al. 2020), but still comparable and sufficiently close to the observations for the purposes of this study.

2.1.1 IllustrisTNG

The Next Generation Illustris simulation (IllustrisTNG, TNG), which is run with the AREPO code (Springel 2010; Weinberger et al. 2020), consists of 9 simulations: 3 box sizes (300, 100 and 50 Mpc on a side), each available at 3 different resolutions, 1–3, with 1 being the highest and 3 the lowest resolution (see Springel et al. 2018; Naiman et al. 2018; Marinacci et al. 2018; Nelson et al. 2019; Pillepich et al. 2019, for details). Compared with its predecessor, Illustris (Vogelsberger et al. 2014a,b; Genel et al. 2014), TNG provides improved agreement with observations by modifying its treatment of active galactic nuclei (AGN) feedback, galactic winds and magnetic fields (Pillepich et al. 2018; Weinberger et al. 2017). In addition, various improvements of the hydrodynamical convergence have been introduced in the code.

In this work, we employ the highest-resolution hydro run of the largest TNG box, TNG300-1, as well as the lowest-resolution dark-matter-only run, TNG300-3-DM, at $z = 2.44$, which we use to calibrate our Ly α forest generation procedure (see Section 3). Having a phase-matched dark-matter-only simulation allows for a fast and direct comparison to the full hydro results. In particular, since the sample variance of the two boxes is the same, any differences observed in the power spectra can be attributed to model choices. TNG300-3-DM also has the benefit of having very similar particle resolution to the base ABACUSSUMMIT boxes: $M_{\text{part, TNG-300-3-DM}} = 3.1 \times 10^9 h^{-1} M_{\odot}$ and $M_{\text{part, Abacus}} = 2.1 \times 10^9 h^{-1} M_{\odot}$.

We make use of the noiseless mock Ly α forest spectra created and made publicly available by Qezlou et al. (2022). The spectra are obtained via the fake_spectra package (Bird et al. 2015; Bird 2017), which calculates the absorption spectra for every ion in the simulation along a chosen set of lines-of-sight. Each particle contributes to the overall absorption in the spectrum according to a Voigt profile. The cells are smoothed by an appropriate top-hat kernel. The skewers used in this study are solely due to Ly α transmission, and we leave further exploration of the effect of metal absorption lines on the Ly α observables for future work.

Throughout this work, we assume that the mean flux evolution is given by the following empirical relation, corrected for metal absorption (Faucher-Giguère et al. 2008):

$$\langle F \rangle = \exp[-1.330 \times 10^{-3} \times (1+z)^{4.094}]. \quad (1)$$

In our Ly α forest skewers, the optical depth is rescaled to match the expected observed measurement, which at $z = 2.44$ corresponds to $\langle F \rangle = 0.8101$. The high-resolution spectra assume the line-of-sight direction to be along the z axis, with a pixel width of 6.4 km/s to resolve well features along the line-of-sight. We note that ideally one would use all three axes as lines-of-sight to reduce the variance of the measurements. However, those were not provided as part of the TNG Ly α forest data release. Once the skewers are extracted, the transmission fraction is averaged over adjacent pixels to a final pixel size of 26 km/s, corresponding to $0.25 h^{-1} \text{Mpc}$. When employing its dark-matter-only counterpart, TNG300-3-DM, to generate the mock skewers, we adopt a pixel size of $0.33 h^{-1} \text{Mpc}$ (625^3 cells), corresponding also to the interparticle spacing of the simulation and approximately matching the resolution of the ABACUSSUMMIT boxes ($0.29 h^{-1} \text{Mpc}$).

2.1.2 Comparing TNG300-1 with eBOSS

We next compare the one-dimensional (1D) power spectrum measured from the TNG300-1 simulation with observational data. The observational results presented here are based on data collected by

the Sloan Digital Sky Survey (SDSS) (York et al. 2000). In particular, the sample of Ly α forest observations is selected from the quasar spectra of the DR14 catalog, which were observed either by the SDSS-III Collaboration between 2009 and 2014 (as part of BOSS) or by the SDSS-IV Collaboration in 2014-2015 (as part of eBOSS). Throughout the paper, we will be referring to this data set as ‘eBOSS.’

In Fig. 1, we illustrate the 1D power spectrum of the Ly α forest skewers measured in TNG300-1 and weigh it up against eBOSS DR14 (Chabanier et al. 2019). We note that the eBOSS power spectrum is already noise and background subtracted, and the Ly α -Si correlations have been removed, allowing for a direct comparison with the simulation. It is important to acknowledge that we do not expect a perfect agreement between observations and simulations, since uncertainties in the thermal and ionisation history of the intergalactic medium impact the correlations of the Ly α forest. We see that the agreement between simulations and observations is reasonably good, noting that we do not calibrate the mean flux of TNG to the observed mean in eBOSS data, which would change the overall normalization. We observe a discrepancy of $\lesssim 10\%$ for $k \lesssim 1 h \text{Mpc}^{-1}$, and up to 20% on smaller scales. We have checked that a finer gridding of the TNG gas cells along the line-of-sight (50 ckpc/h) does not mitigate the small-scale deviation. For the purposes of this work, this is a satisfactory result. However, we note that TNG has not been exhaustively tested against observations in the IGM regime (but rather mostly for galaxy observations) unlike other hydro simulations tailored towards mimicking the Ly α forest (see e.g., the Nyx vs. Illustris code comparison in Sorini et al. 2018). It is also worth commenting on the fact that our result exhibits greater tension with eBOSS than can be seen in the analogous figure in Qezlou et al. (2022). The reason for this difference is that the eBOSS data vector of Qezlou et al. (2022) uses a different technique for subtracting the noise contributions than the official eBOSS analysis¹ (Chabanier et al. 2019).

In the lower panel of Fig. 1, we show the 3D power spectrum defined as follows:

$$\langle \delta_{\tilde{F}}(\mathbf{k}) \delta_{\tilde{F}}^*(\mathbf{k}') \rangle = (2\pi)^3 P(k, \mu) \delta_D(\mathbf{k} - \mathbf{k}'), \quad (2)$$

where $\delta_D(\mathbf{k})$ is the three-dimensional Dirac delta function. In particular, we bin the power spectrum $P(k, \mu)$ into 20 logarithmic k bins ranging from $k \in \{(2\pi)/L_{\text{box}}, 15 h \text{Mpc}^{-1}\}$, where L_{box} is the box size of the simulation, and 16 μ bins ranging from 0 to 1, and show the estimated Gaussian error bars (see further discussions in Section 3.5 and Section 4.2). It is evident that on large scales ($k \lesssim 0.1 h \text{Mpc}^{-1}$), the error on the measurements is large due to the small size of the box. We compare the fitted values of the Ly α bias and redshift distortion parameter from TNG, $b_{\text{Ly}\alpha, \text{TNG}} = -0.1379$, $\beta_{\text{Ly}\alpha, \text{TNG}} = 1.432$, to the values measured in eBOSS data, $b_{\text{Ly}\alpha, \text{eBOSS}} = -0.117$, $\beta_{\text{Ly}\alpha, \text{eBOSS}} = 1.669$ at $z_{\text{eff}} = 2.334$ (see last column of Table 6 in du Mas des Bourboux et al. 2020). While the redshift distortion parameter is slightly lower in TNG, it is interesting to see that the values of $b(1 + \beta)$ that set the amplitude of the 3D power along the line of sight are within 10%. We note that other state-of-the-art hydro simulations report higher values for β (e.g., Chabanier et al. in prep. find $\beta = 1.8$), while Givans et al. (2022) finds $\beta = 1.35$ at $z = 2.8$ in the Sherwood suite of simulations and Arinyo-i-Prats et al. (2015) find $\beta \sim 1.3 - 1.5$ in the relevant redshift range. However, for the purposes of this study, we consider these matches good enough and calibrate our mocks to match the TNG measurements, referring to it as the ‘truth’ from hereon.

¹ Established through private communication.

2.1.3 ABACUSUMMIT

ABACUSUMMIT is a suite of high-performance cosmological N -body simulations, which was designed to meet and exceed the Cosmological Simulation Requirements of the DESI survey (Maksimova et al. 2021). The simulations were run with ABACUS (Garrison et al. 2019, 2021), a high-accuracy cosmological N -body simulation code, optimized for GPU architectures and for large-volume simulations, on the Summit supercomputer at the Oak Ridge Leadership Computing Facility.

The majority of the ABACUSUMMIT simulations are made up of the base resolution boxes, which house 6912^3 particles in a $2 h^{-1} \text{Gpc}$ box, each with a mass of $M_{\text{part}} = 2.1 \cdot 10^9 h^{-1} M_{\odot}$. While the ABACUSUMMIT suite spans a wide range of cosmologies, here we focus on the fiducial outputs (*Planck* 2018: $\Omega_b h^2 = 0.02237$, $\Omega_c h^2 = 0.12$, $h = 0.6736$, $10^9 A_s = 2.0830$, $n_s = 0.9649$, $w_0 = -1$, $w_a = 0$). In particular, we employ the 6 base boxes `AbacusSummit_base_c000_ph{000-005}`. The reason for our choice is that full particle outputs are provided for these simulations at $z = 2.5$, which is the redshift of interest for our Ly α forest study. For full details on all data products, see Maksimova et al. (2021). In future work, we plan to extend our mocks to cosmologies beyond *Planck* 2018 and adapt our method so that it utilizes only 10% of the particles (available for all ABACUSUMMIT simulations at $z = 2.5$).

2.2 Quasar catalogue

The cross-correlation function of the Ly α forest with quasars will be measured by current and next-generation experiments such as DESI. However, to ensure that our theoretical models can adequately fit the signal, we need to test our pipelines on synthetic catalogues. To this end, we generate mock quasar catalogues via ABACUSHOD, a sophisticated routine that builds upon the baseline halo occupation distribution (HOD) model by incorporating various extensions affecting both the one- and two-halo terms, and in Section 4.3, we show the cross-correlations of our mock quasar catalogue with the Ly α forest spectra. ABACUSHOD allows the user to specify different tracer types: emission-line galaxies (ELGs), luminous red galaxies (LRGs), and quasistellar objects (QSOs). The full model is described in detail in Yuan et al. (2022).

In this study, we adopt a simple HOD model for the QSO without any decorations:

$$\bar{n}_{\text{cent}}^{\text{QSO}}(M) = \frac{\text{ic}}{2} \text{erfc} \left[\frac{\log_{10}(M_{\text{cut}}/M)}{\sqrt{2}\sigma} \right], \quad (3)$$

$$\bar{n}_{\text{sat}}^{\text{QSO}}(M) = \left[\frac{M - \kappa M_{\text{cut}}}{M_1} \right]^{\alpha} \bar{n}_{\text{cent}}^{\text{QSO}}(M), \quad (4)$$

where M_{cut} characterizes the minimum halo mass to host a central galaxy, M_1 the typical halo mass that hosts one satellite galaxy, σ the steepness of the transition from 0 to 1 in the number of central galaxies, α the power law index on the number of satellite galaxies, ic the incompleteness parameter, and κM_{cut} gives the minimum halo mass to host a satellite galaxy. The parameters we choose for our QSO catalogs are in units of $h^{-1} M_{\odot}$

$$\log_{10}(M_{\text{cut}}) = 13.2, \quad \kappa = 1.11, \quad \sigma = 0.65, \quad (5)$$

$$\log_{10}(M_1) = 13.8, \quad \alpha = 0.8, \quad \text{ic} = 1.0,$$

which have been selected so as to yield a linear bias of about $b^{\text{QSO}} \approx 3.3$, roughly matching the quasar bias in du Mas des Bourboux et al. (2020), and have a number density of $1.75 \times 10^{-4} [h^{-1} \text{Mpc}]^{-3}$ (i.e.,

1.4 million quasars per box). These numbers are taken from rough fits to preliminary DESI data.

3 CREATION OF THE MOCKS

Previous large-volume Ly α forest mocks have been generated using simple, fast and computationally cheap methods such as lognormal density maps (e.g., [Farr et al. 2020](#)) augmented with approximate prescriptions to reach the volumes required by the new generation of surveys. However, models based on Gaussian random fields do not capture non-linear evolution, as they are generated solely through the initial power spectrum [Coles & Jones \(1991\)](#); [Bi & Davidsen \(1997\)](#). Slightly more complex are formalisms involving Lagrangian perturbation theory (see e.g., [Bernardeau et al. 2002](#), for a review) and COLA [Tassev et al. \(2013\)](#), which extend the modeling capabilities to the mildly non-linear regime. In pure BAO analyses, the presence of non-linear structure does not substantially affect the measurement, especially at the high-redshift regime ($z \gtrsim 2$). However, any full-shape and small-scale analysis of Ly α forest observables (including cross-correlations) will be substantially impacted by non-linear gravitational and astrophysics effects ([Cuceu et al. 2022a,b](#)).

This work aims to enable the full-shape analysis of the Ly α forest power spectrum, planned to be conducted as part of the DESI Y3 Ly α science program. While ideally one would strive to generate as realistic mocks as possible, which would mean employing state-of-the-art hydrodynamical simulations, this is unfortunately not a viable path forward, as the computational expense associated with generating Ly α forest skewers in a volume sufficiently large for modern surveys is tremendous. In this work, we therefore seek a middle path of using fully evolved N -body simulations and adopting an approximate technique calibrated to a hydro simulation.

In our Ly α forest mocks on ABACUSSUMMIT, we opt for a resolution of 6912^3 cells per box, corresponding to an average of one particle per cell and a mean interparticle distance of $0.29 h^{-1}$ Mpc. The density and velocity field grids are obtained as described in Section 3.1. The resolution is chosen to be comparable to (though still larger than) the Jeans length at that redshift ($100 \text{ kpc}/h$) while avoiding the creation of too many empty cells, as that would contribute substantial noise to the density field and the derived optical depth, subsequently. Since our resolution is limited by the simulation resolution, we are unable to obtain an accurate estimate of the field at scales lower than $\sim 0.3 h^{-1}$ Mpc. Thus, the power spectrum of the skewers $P_{1D}(k_{\parallel})$ measured from modes lying along the line of sight is suppressed, which also affects the 3D flux power spectrum ([Farr et al. 2020](#)). For this reason, we boost the power spectrum by adding small-scale fluctuations to the density field, as discussed in Section 3.2.

Next, to convert from dark matter density to optical depth, we adopt the simple fluctuating Gunn-Peterson approximation (FGPA) [Croft et al. \(1998b\)](#). While this method is simplistic, it offers a fast and transparent way of connecting the matter density to that of neutral hydrogen. More complex techniques do exist, including the Ly α Mass Association Scheme (LyMAS; [Peirani et al. \(2014b, 2022b\)](#)), the Iteratively Matched Statistics (IMS; [Sorini et al. \(2016b\)](#)) method, and Hydro-BAM ([Sinigaglia et al. 2022](#)). These use a variety of approaches tuned using smaller hydro simulations that range from matching the Ly α forest probability distribution function and/or power spectrum to using a supervised machine learning method. However, these methods have yet to be applied to simulations with the purpose of making large scale DESI mocks. Therefore, in this first work we focus on using the simpler FGPA approach and leave the

application of these more complex recipes to future work. We adopt two slight variations of the FGPA approach discussed in Section 3.3.

Finally, we add RSDs to our skewers and convert them to transmission flux spectra in Section 3.4. Those are the result of peculiar velocities in the inter-galactic medium (IGM) projected along the line of sight, and manifest themselves as an anisotropy in the power spectrum and correlation function measurements.

3.1 Calculating the density and velocity fields

The first step in applying the FGPA method to an N -body simulation (in our case, ABACUSSUMMIT and TNG300-3-DM) is the deposition of particles onto a grid. In this study, we adopt triangular shape cloud (TSC) interpolation, to obtain both the density, $\rho_{\text{dm}}(\mathbf{x})$, and peculiar velocity $v_r(\mathbf{x})$ fields. Some studies (e.g., [Sorini et al. 2016a](#)) apply smoothing to the fields to mimic the effect of baryonic pressure on small scales. However, similarly to [Qezlou et al. \(2022\)](#); [Stark et al. \(2015\)](#); [Newman et al. \(2020\)](#), we find that when simulating the neutral hydrogen absorption on scales of ~ 1 Mpc, smoothing the fields has negligible effects. In future work, we plan to revisit our choice of a particle-to-grid deposition method. While TSC has clear advantages (especially in the low-density regime) over the lower-order kernels, i.e. nearest grid point (NGP) and cloud-in-cell (CIC), tessellation-based methods are even better suited for obtaining a near exact estimate of the low-density dark matter field (e.g., phase-space sheet tessellation as in [Abel et al. \(2012\)](#)), which is the most relevant for Ly α forest analysis (see e.g., [Chabanier et al. 2023](#), for an evaluation of these effects).

3.2 Adding small-scale noise

Similarly to [Farr et al. \(2020\)](#), we add small-scale noise to the initial density field so as to make up for the deficit in the 1D power spectrum. This deficit is the result of the effective smoothing on small scales imposed by the relatively large size of the gas blobs ($\sim 0.3 h^{-1}$ Mpc), which suppresses the power on small scales. We start by generating independent Gaussian skewers δ_{ϵ} for each line-of-sight preserving the resolution of the original density field. We impose that the 1D power spectrum of the skewers obeys the following equation [McDonald et al. \(2006\)](#):

$$P_{1D}(k) \propto [1 + (k/k_1)^n]^{-1}, \quad (6)$$

normalized so that $\text{Var}[\delta_{\epsilon}] = 1$. These Gaussian skewers are then all scaled by a factor σ_{ϵ} to control the variance in the extra power added. This factor, together with n and k_1 is a free parameter in our model and takes a single value determined by the procedure described in Section 3.5.

The new density skewers are then obtained by multiplying the true density field skewers by the lognormal field:

$$\rho(\mathbf{x}) = \rho_{\text{dm}}(\mathbf{x})(1 + \delta_{\text{ln}}(\mathbf{x})), \quad (7)$$

where the lognormal field is given by the lognormal transformation

$$\delta_{\text{ln}}(\mathbf{x}) = \exp \left[\delta_{\epsilon}(\mathbf{x})\sigma_{\epsilon} - \frac{\sigma_{\epsilon}^2}{2} \right] - 1 \quad (8)$$

to ensure zero mean.

Note that the lognormal skewers are generated independently of each other, so there is no correlation across different lines-of-sight. We note that the same effect could have been achieved by adding small-scale fluctuations to the velocity field. However, since only one of our methods directly uses the velocity field, we opt to be consistent

and add small-scale noise to only the density field. We also note that one of the models for generating synthetic Ly α skewers (see Model 1 in Table 3.5) does not include lognormal noise. The effect of turning off the extra noise is visible in the power spectrum measurements shown in Fig. 4.

3.3 Deriving the observed optical depth

To convert the fluctuations in the density field into optical depth, we adopt two different, but closely related approaches. The first one of them follows the standard FGPA prescription, while the second introduces a small modification to it. We detail the two methods below.

- **Method I:** Two key assumptions go into the FGPA approach (Gunn & Peterson 1965): adiabatic expansion of the gas and photoionization equilibrium in the IGM. The first one implies that the relationship between density and temperature is well approximated by (Hui & Gnedin 1997)

$$T(\mathbf{x}) \propto \rho(\mathbf{x})^{\gamma-1}, \quad (9)$$

where γ is the slope of the temperature-density relation, while the second dictates the connection between the temperature of the gas and the number of neutral hydrogen atoms:

$$n_{\text{HI}}(\mathbf{x}) \propto \rho(\mathbf{x})^2 T(\mathbf{x})^{-0.7}. \quad (10)$$

Here, ρ is the baryonic matter density (Hui et al. 1997). However, we note that in a collisionless dark-matter simulation, we can only access the total matter field, as defined in Section 3.1, which we assume traces the baryonic field reasonably well. Combining these two equations and noting that the optical depth, τ , is proportional to the neutral hydrogen column density, n_{HI} , we arrive at the final form of the FGPA method (Bi & Davidsen 1997; Croft et al. 1998b)

$$\tau(\mathbf{x}) = \tau_0 \rho(\mathbf{x})^\alpha, \quad (11)$$

where τ_0 is the overall normalization resulting from the temperature-density and photoionization rate assumptions and $\alpha \equiv 2 - 0.7(\gamma - 1)$. In our analysis, τ_0 and γ are free parameters chosen according to the description in Section 3.5. As a helpful reference, we note that the value of the temperature-density slope in TNG can be measured to be $\gamma \approx 1.5$ (Gouin et al. 2022).

The final step in converting the matter field, $\rho(x)$, into the “observed” optical depth $\tau(s)$ involves converting the real-space $\tau(x)$ into its redshift-space equivalent, $\tau(s)$. In addition to the redshifting of the Ly α absorption features due to cosmic expansion, $\lambda_{\text{obs}} = \lambda_\alpha(1+z)$, with λ_α being the Ly α wavelength and z the absorption redshift, there is an additional effect of RSDs caused by the peculiar velocities of neutral hydrogen clouds. We introduce RSDs into our mocks by treating each cell in our three-dimensional grid as a gas blob with a mean velocity along the line of sight as calculated in Section 3.1. The conversion to redshift-space of each skewer can be expressed as an integral over velocity space of the real-space optical depth multiplied by some kernel, K :

$$\tau(s) = \int \tau(x) K\left(s - x - v_r(x)(1+z)/H(z)\right) dx, \quad (12)$$

where v_r is the peculiar velocity along the line of sight, while x and s are the spatial coordinates in real- and redshift-space, respectively. A typical choice for the kernel in Ly α mock generation is the Voigt profile, a Gaussian kernel with a Lorentzian term, or the Doppler profile, just a Gaussian kernel, both of which aim to account for the effects of thermal broadening due to the random thermal velocities

of the gas atoms. We implement convolution with the Doppler profile as an option in our ABACUS-SUMMIT mocks, but find that it has little effect on our observables (e.g., the 1D power spectrum), since the width of the kernel is comparable or smaller than the size of the cells. We show this in Appendix A. Thus, to simplify our process, we set $K(x) = \delta^D(x)$, where δ^D is the Dirac delta function, which amounts to shifting the optical depth of each blob according to its peculiar velocity. In practice, we need to adopt some particle deposition technique due to the discreteness of the cells. A standard choice is to employ a nearest-grid point scheme; however, we opt to use TSC, as it is higher-order than CIC and NGP.

- **Method II:** Similarly to the first method, here we also assume that the optical depth is related to the density field as $\tau(\mathbf{x}) \propto \rho(\mathbf{x})^\alpha$. However, the main difference is that in this version, we go directly from the particle positions and their velocities to the final “observed” optical depth in redshift space. In particular, we compute a weight for each particle given by $\rho_{\text{dm}}(\mathbf{x})^{\alpha-1} \times [1 + \delta_{\text{ln}}(\mathbf{x})]^\alpha$, where $\rho_{\text{dm}}(\mathbf{x})$ is the dark matter density field in real space (see Section 3.1) and $\delta_{\text{ln}}(\mathbf{x})$ is the lognormal noise field in real space introduced in Eq. 6. We then displace the line of sight coordinate component of each particle according to its peculiar velocity as follows:

$$s = x + v_r(1+z)/H(z), \quad (13)$$

where $H(z)$ is the Hubble parameter at redshift z . Adopting TSC interpolation, we deposit the weighted and displaced particles on the three-dimensional grid to obtain the observed optical depth $\tau(s)$. Thus, this method yields the optical depth directly in redshift-space and as such is less computationally intensive. We note that the reason that this approach leads to the correct form is that the usual particle deposition results in a density field $\rho_{\text{dm}} \propto (1 + \delta_{\text{dm}})$. Therefore, upon weighting and displacing the particles, we arrive at a field behaving as $[\rho_{\text{dm}}(\mathbf{x})^\alpha \times [1 + \delta_{\text{ln}}(\mathbf{x})]^\alpha](s)$.

The main difference between the two methods is that **Method I** treats the individual grid cells as Ly α absorption clouds with a velocity determined by the mean in the cell, whereas in **Method II**, each particle is approximated as an individual absorber of Ly α photons. A downside of the first method is that at low densities, averaging the velocities of sparsely distributed particles results in a suppression of the peculiar velocities of dark matter substructures, which translates as a deficiency in the RSD signal. On the other hand, the second method can potentially lead to a stronger RSD signal than the true Ly α forest, as the thermal velocities of individual particles will be larger compared with the gas clouds due to the lack of baryonic pressure in the N -body simulation. In an idealized scenario, one could consider identifying substructures via some halo-finding algorithm and deriving the absorption cloud velocities from that, but even this method would not be able to capture correctly the underlying physics, as it would lack important gas and baryonic physics.

3.4 Obtaining the flux skewers

Finally, we need to transform the optical depth, $\tau(s)$, into the transmitted flux fraction, $F(s)$, following:

$$F(s) = \exp[-\tau(s)]. \quad (14)$$

When computing power spectra of the Ly α forest, we typically work with the transmitted flux contrast:

$$\delta_F(s) = \frac{F}{\langle F \rangle} - 1, \quad (15)$$

which is characterized by having a zero mean. As can be seen in Eq. 11 and Eq. 14, the optical depth is saturated in over-dense regions

yielding zero flux and hence no information. On the other hand, more information can be gleaned from low- and intermediate-density regions, where there is some absorption but not enough to cause the signal to be saturated.

3.5 Parameter tuning

Our model consists of a number of free parameters defined in Section 3.2 and Section 3.3, namely, τ_0 , σ_ϵ , γ , n , k_1 (see Eq. 11 and Eq. 6 for descriptions). To decide on the values of these parameters, we aim to match several key properties of the hydro simulation Ly α forest skewers: the mean transmitted flux fraction $\langle F \rangle$, the variance of the low-pass-filtered flux with a cutoff at $k_{1D} = 1 \text{ h Mpc}^{-1}$, $\langle F^2 \rangle$, the 1D power spectrum, P_{1D} (or equivalently, $P(k)$), up to $k < 0.8 \text{ h Mpc}^{-1}$, and the 3D power spectrum, P_{3D} (or equivalently, $P(k, \mu)$), up to $k < 0.8 \text{ h Mpc}^{-1}$ for selected values of μ . We note that we tune our model by comparing the FGPA-derived skewers on TNG300-3-DM with the full-physics TNG300-1 skewers, as the two simulations share many properties such as initial seed and cosmic variance, which enables a direct comparison. We prefer matching the full shape of the power spectra rather than a compressed statistics such as the bias. The reason for this choice is that due to the limited sizes of the box, low-wavemode quantities are noisy to measure. Only once we are satisfied with the match between TNG300-1 and TNG300-3-DM, do we apply our method to the large boxes of ABACUSUMMIT to obtain the final products. In addition, we find that the shape of the FGPA-derived power spectra also differs across the different models, which is an additional advantage of matching to the 1D and 3D power spectra. We describe our process in more detail below.

- We first measure the 1D and 3D power spectra from TNG300-1 and quantify their error bars. In the case of the 1D power spectrum, the process is straightforward: we Fourier transform the flux contrast δ_F along each skewer and compute the power spectrum, averaging over all lines-of-sight. We bin the 1D power spectrum into 400 linear bins ranging from $k \in \{0, 12.26 \text{ h Mpc}^{-1}\}$, i.e. spaced by $(2\pi)/L_{\text{box}}$. To obtain the error bars on the 1D measurement, we apply jackknives on the available skewers. In the case of the 3D power spectrum, we work with the quantity $P(k, \mu)$ defined in Eq. 2. As before, we bin the power spectrum into 20 k bins ranging from $k \in \{(2\pi)/L_{\text{box}}, 15 \text{ h Mpc}^{-1}\}$ and 16 μ bins ranging from 0 to 1. We assume that the error bars on this measurement are well approximated by the Gaussian error:

$$\Delta P(k, \mu) = \sqrt{\frac{2}{N_k}} P(k, \mu), \quad (16)$$

where we calculate the number of k modes in each k and μ bin as $N_k = k^2 dk d\mu / (2\pi/L_{\text{box}})^3$ with L_{box} being the box size.

We next split the tuning process into a slow and a fast step, with the fast step varying τ_0 and σ_ϵ to match the mean and variance of the flux, and slow step varying γ , n , and k_1 to additionally match the 1D and 3D power spectra.

- **Fast parameters:** For a given choice of slow parameters, γ , n , and k_1 , and a Method (I or II as defined in Section 3.3), we vary the parameters τ_0 and σ_ϵ , so as to minimize the function:

$$\chi_{\text{mean, std}}^2 = [\langle F_{\text{TNG Hydro}} \rangle - \langle F_{\text{TNG FGPA}} \rangle]^2 + [\sqrt{\text{Var}[F_{\text{TNG Hydro}}]} - \sqrt{\text{Var}[F_{\text{TNG FGPA}}]}]^2, \quad (17)$$

where the mean flux $\langle F_{\text{TNG Hydro}} \rangle = 0.8101$ is taken from the empirical relation, i.e. Eq. 1, while the variance,

$(\text{Var}[F_{\text{TNG Hydro}}])^{1/2} = 0.1878$, is computed from the low-pass filtered flux skewers described above. We note that the fast parameters are optimized separately from the slow ones, which is the reason that we do not worry about normalizing the χ^2 . Furthermore, we implicitly assume that the error on the mean and standard deviation measurements is comparable. We test this assumption for our default resolution and find that the two differ only by a small $O(1)$ factor.

- **Slow parameters:** To decide on the values of the slow parameters, we generate a three-dimensional uniform grid with possible values they can take: we allow γ to vary between 1.4 and 1.7 and test 6 values in that range (typically, one assumes that $\alpha \approx 1.6$, which corresponds to $\gamma \approx 1.56$); n is allowed to vary between -3 and 3, and we test 200 values in that range; finally, k_1 varies between 0.001 to 1, and we test 200 values in that range. We do not sample γ as densely as the other two parameters, as we find that our observables are weakly affected by this choice.

For each of the two methods (see Section 3.3) and each point in the three-dimensional grid, we first fit for the mean and variance of the flux so as to calibrate τ_0 and σ_ϵ and then record the mean and the flux alongside the contribution to the χ^2 of the 1D and 3D power spectra, computed as follows:

$$\chi_{1D}^2 = \sum_k \left[\frac{(P_{1D, \text{TNG Hydro}}(k) - P_{1D, \text{TNG FGPA}}(k))}{\Delta P_{1D, \text{TNG Hydro}}(k)} \right]^2 \quad (18)$$

$$\chi_{3D}^2 = \sum_{k, \{\mu\}} \left[\frac{(P_{\text{TNG Hydro}}(k, \mu) - P_{\text{TNG FGPA}}(k, \mu))}{\Delta P_{\text{TNG Hydro}}(k, \mu)} \right]^2$$

for four selected values of μ , namely $\{0.03, 0.33, 0.66, 0.97\}$ with a bin width of $\Delta\mu \approx 0.06$. We have checked that the parameter selection is negligibly affected by whether we use only a handful of μ values or the full $P(k, \mu)$ vector.

Note: The reason we classify n and k_1 as slow parameters is that the step of generating the Gaussian noise skewers is relatively slow. Similarly, changing γ in **Method II** requires a rerunning of the TSC particle deposition step, which is computationally expensive.

- In the final step of this process, we select the values of the three slow parameters, which will be used in the ABACUSUMMIT Ly α forest mocks. To do so, we combine the χ^2 values from the 1D and 3D power spectra. In particular, for each of the two methods (see Section 3.3), we choose two sets of slow parameters: the first set corresponds to the best-fit parameters we obtain when minimizing $\chi_{\text{Model 1}}^2 = \chi_{3D}^2$, while the second set comes from minimizing $\chi_{\text{Model 2}}^2 = \chi_{3D}^2 + \chi_{1D}^2$. We quote the best-fit values for all four models (two per method) in Table 3.5. We note that we do not include the mean and variance χ^2 contributions, as those are already calibrated individually for each model. Additionally, since the errors on the mean and the variance are much smaller, they would dominate the selection, and our final objective is to match the power spectra.

We note that while the values of the fast parameters τ_0 and σ_ϵ in the large boxes are quite similar to the TNG-DM boxes, we opt to recalibrate them to make sure we match the mean and variance of the flux as best as we can.

To summarize the tuning process, we start by generating a three-dimensional regular grid, for which each point corresponds to a set of predetermined values for the three slow parameters. For each set of three slow parameters, we minimize the absolute difference with TNG300-1 of the mean and the variance of the flux, adopting the Nelder-Mead scheme, to find the values of the fast parameters and record the resulting 1D and 3D power spectrum difference (in terms of χ^2). For each of the 4 models considered in this work, we then simply report the set of slow and fast parameters that correspond

Model #	Method	Fit	$\langle F \rangle$	$\sqrt{\text{Var}[F]}$	τ_0	σ_ϵ	γ	n	k_1	χ^2_{1D}	χ^2_{3D}	$b_{\text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$
1	Method I	P3D	0.801	0.168	0.387	0.000	1.650	—	—	242.276	61.138	-0.146	0.920
2	Method I	P1D+P3D	0.811	0.187	0.391	0.772	1.450	1.000	0.063	27.536	104.887	-0.129	0.949
3	Method II	P3D	0.825	0.212	0.385	1.696	1.500	1.500	1.000	770.975	25.186	-0.130	2.022
4	Method II	P1D+P3D	0.810	0.187	0.654	2.116	1.550	0.000	—	131.403	100.800	-0.126	2.330

Table 1. Specifications of the four models used in the creation of the Ly α forest synthetic catalogs. In particular, we indicate the values of the slow and fast parameters, γ , n , k_1 , τ_0 and σ_ϵ , defined in Section 3.2 and Section 3.3. Descriptions of the two FGPA-based methods (I and II) can be found in Section 3.3, while the fitting procedure is detailed in Section 3.5. The target values of the mean and variance for these mocks are derived from the hydro simulation TNG300-1 and are $\langle F \rangle = 0.8101$ ($\text{Var}[F]^{1/2} = 0.1878$). Model 1 has effectively no small-scale noise added (hence the blanks), while in the case of Model 3, we effectively add “white” noise, i.e. with no scale-dependence. We also share the measurements of the bias and the redshift distortion parameter, $b_{\text{Ly}\alpha}$ and $\beta_{\text{Ly}\alpha}$, for each model and compare them with the TNG values of -0.1379 and 1.432, respectively. The number of data points fitted in the 1D power spectrum is 27, while that for the 3D power is 20, suggesting that Model 2 provides a good fit for the 1D power, while Model 3 does well with matching the 3D power. We note that some state-of-the-art hydro simulations report higher $\beta_{\text{Ly}\alpha}$ values (e.g., Chabanier et al. in prep. find $\beta_{\text{Ly}\alpha} = 1.8$), which agree better with our Models 3 and 4.

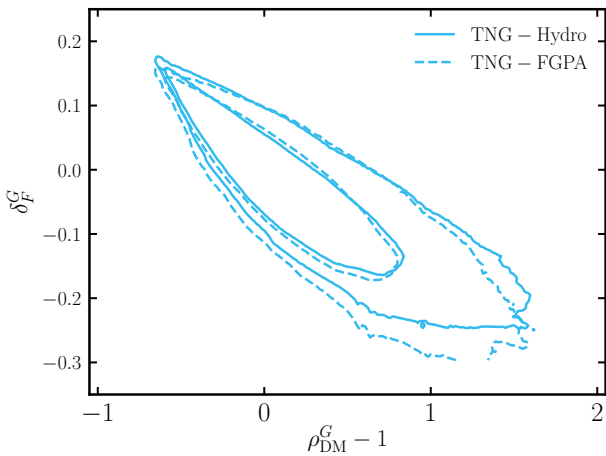


Figure 2. Two-dimensional PDF contours comparing the dark-matter-field-Ly α -flux relation for the TNG300-1 hydro simulation (solid) and one of our FGPA-based synthetic catalogs (dashed) applied to the low-resolution dark-matter-only counterpart TNG300-3-DM (Model 3; see Table 3.5). The levels shown correspond to 2% and 68%. The voxel resolution of the maps is $0.33 h^{-1} \text{Mpc}$, and both maps are smoothed with a Gaussian kernel of size $\sigma_G = 3 h^{-1} \text{Mpc}$ for visualization purposes. The similarity between the two curves confirms that the gas physics has small effect on Ly α observables on megaparsec scales.

to the smallest χ^2 across all grid points. In the future, we plan to adopt a more flexible iterative process rather than the preset three-dimensional grid. However, that would require that we substantially speed up the power spectrum computation, for example, by adopting analytical approximations (e.g., Farr et al. 2020). We defer these ideas for later work, where we explore a more complex model and utilize larger boxes for calibration.

In Fig. 2, we show the two-dimensional PDF contours comparing the dark-matter-field-Ly α -flux relation for the TNG300-1 hydro simulation and one of our FGPA-based synthetic catalogs applied to the low-resolution dark-matter-only counterpart TNG300-3-DM (Model 3; see Table 3.5). The levels shown correspond to 2% and 68%. The voxel resolution of the maps is $0.33 h^{-1} \text{Mpc}$ and both are smoothed with a Gaussian kernel of size $\sigma_G = 3 h^{-1} \text{Mpc}$. Note that the smoothing is applied for visualization purposes and is not used for

any other figure in this paper. The similarity between the two curves confirms that the gas physics has small effect on Ly α observables on megaparsec scales.

In Fig. 3, we show a couple of skewers passing through the entire TNG300 box for the “true” Ly α spectra extracted from TNG300-1 and the synthetic ones obtained using our Model 3 (see Section 3 and Table 3.5) on the low-resolution counterpart TNG300-3-DM. For visualization purposes, we plot $\delta F(r) + 1 \equiv F(r)/\bar{F}(r)$ for the skewer on top and δF for the skewer on the bottom. Reassuringly, the simplified FGPA model does an adequate job of matching the majority of the features present in the full hydrodynamical spectra. Visible in the comparison of the two is that the true skewers appear smoother than the FGPA ones due to the extra noise added to the latter. As we will see in Fig. 4, the Model 3 FGPA 1D power spectrum overshoots the true 1D power spectrum partly due to the addition of small-scale power. Reassuringly, we have inspected (not shown) the skewers for Model 1, which has no small-scale power added, and found the opposite: the FGPA skewers lack small-scale features compared with the true skewers (and their 1D power spectrum, shown in Fig. 4, is lower, as expected).

3.5.1 Comparing the power spectrum of TNG300-1 and TNG-300-3-DM

We produce ABACUSSUMMIT Ly α forest mocks for four different models: two FGPA-based methods (see Section 3.3) calibrated to match the 3D power spectrum individually and the 1D and 3D power spectra jointly of the TNG300-1 skewers. In Fig. 4, we illustrate the level of agreement between the hydro run TNG300-1 and the FGPA mocks run on the dark-matter-only simulation TNG300-3-DM. We note that the sample variance is the same in both boxes, which facilitates the comparison of the models to the “truth.” In addition, on small scales, the measurement is affected by an interlacing effect due to the size of the cells and an aliasing effect due to the sparseness of the lines-of-sight, which scales as $P_{\text{1D}}/n_{\text{2D}}$ (McDonald & Eisenstein 2007), where n_{2D} is the two dimensional density of skewers. Subtracting the aliasing effect theoretically is not trivial, as the skewers in our mocks are placed in a regular grid rather than randomly, as would be the case in observations, and thus the formula in McDonald & Eisenstein (2007) does not hold. In this work, we opt not to do this, as the affected k -modes are beyond our scales of interest. Similarly, interlacing affects our measurements beyond $k \sim 4 - 5 h \text{Mpc}^{-1}$ and can be numerically corrected by offsetting the grid by half a cell size

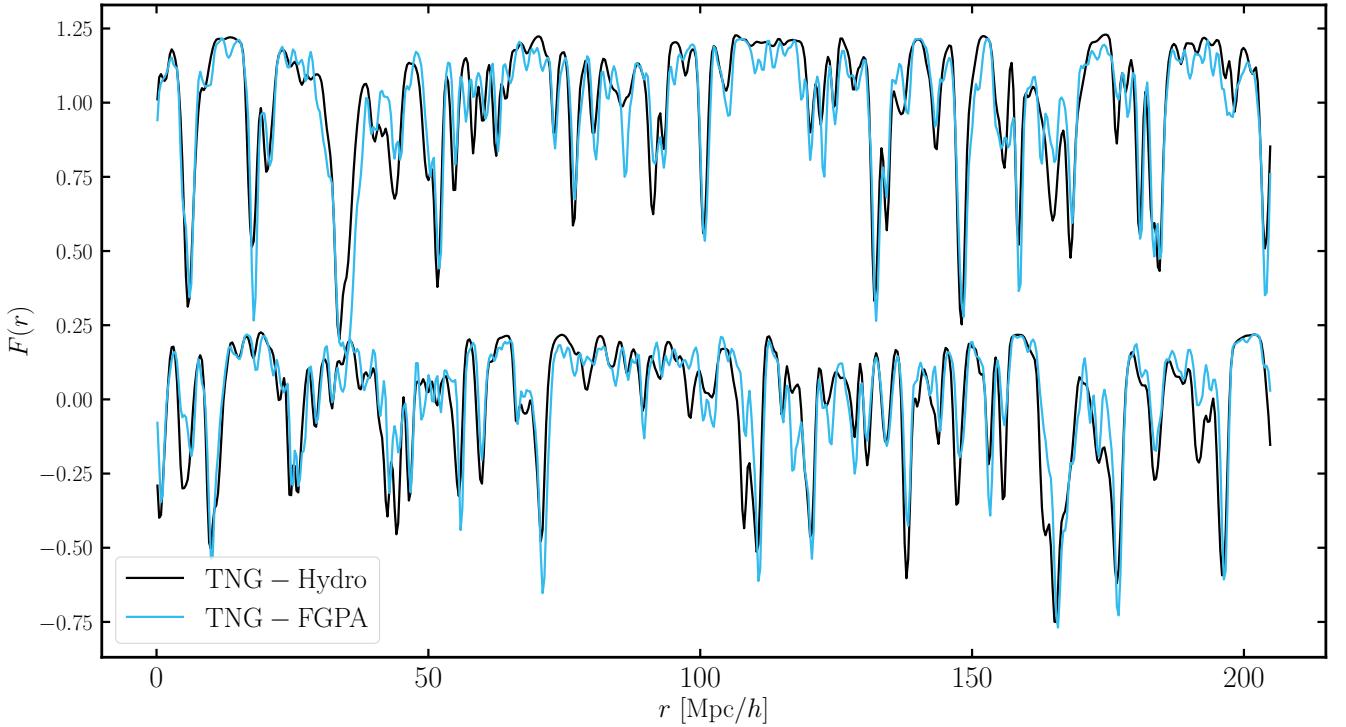


Figure 3. Skewers of the “true” Ly α forest flux (solid) and the FGPA prediction (dashed) for our Model 3 (see Table 3.5). It is reassuring to see that the simplified FGPA model does an adequate job of matching the majority of the features present in the full hydrodynamical spectra. For easier visualization, we are plotting $\delta F(r) + 1 \equiv F(r)/\bar{F}(r)$ and δF for the skewer on top and bottom, respectively.

and recomputing the Ly α observables. As this is prohibitively expensive in the case of the AbacusSummit boxes, which are generated using a lightweight, single-node script, we choose not to apply it to the TNG case either, as we try to make the comparison as consistent as possible (for example, by choosing similar resolution and grid size).

As expected, Models 2 and 4, which are aiming to fit both the 1D and the 3D power spectrum, exhibit closest agreement to the “true” (TNG300-1) 1D power spectrum for $k < 1 h \text{Mpc}^{-1}$. Model 4 overpredicts the 1D power on the smallest scales, not included in the fits, probably because of the large amount of extra power added (large value of σ_ϵ).

In terms of the 3D power, as expected, Models 1 and 3 show better agreement with TNG300-1, respectively, than Models 2 and 4. Models 1 and 2, moreover, have noticeably smaller redshift-space distortions (lower value of the parameter β). This was also the case in the FGPA-lognormal mocks presented in Farr et al. (2020), where the authors addressed this issue by artificially boosting the velocities by 30%. Overall, the four models exhibit a reasonable agreement with the hydro “truth,” providing a wide selection of synthetic catalogs for the users of these mocks to have at their disposal.

4 VALIDATION OF THE MOCKS

In this Section, we study observable summary statistics of our Ly α forest mocks relevant for current and future surveys, namely, the 1D and 3D power spectrum, and the correlation function, in order to validate our mocks. In particular, we first introduce the available large-volume synthetic products on ABACUSUMMIT. We then show measurements of these statistics from our mock skewers and discuss their shortfalls and successes in recovering the “true” statistics com-

ing from the hydro simulation, IllustrisTNG. We then compute the real-space clustering of our ABACUSUMMIT Ly α forest skewers and study the effect of non-linear broadening on the BAO peak. We also compare our measurements against observations from eBOSS (du Mas des Bourboux et al. 2020).

4.1 Available ABACUSUMMIT products

All ABACUSUMMIT Ly α forest mocks are generated on a single node of the National Energy Research Scientific Computing (NERSC) Centre’s cori machine using specially developed python scripts with no external dependencies apart from `scipy`, `numba`, and the specialized package for reading ABACUSUMMIT products, `abacusutils`². The maximum RAM consumption is capped at 70 GB for any of the scripts and the total size of all products (6 simulations, 4 models, 2 lines-of-sight) after applying ASDF ‘blsc’ compression is 50 TB³. As discussed in Section 3.5, we generate mocks for four separate models (adopting **Method I** and **II** to fit the 1D power spectrum and the 1D+3D power spectrum, subsequently). Our products are available for each of the six fiducial cosmology simulations `AbacusSummit_base_c000_ph000-005` ($2 h^{-1} \text{Gpc}$, 6912^3) at $z = 2.5$ and each of the four models, and can be downloaded via Globus (see Data Availability). Each Ly α forest mock has a resolution of $0.29 h^{-1} \text{Mpc}$ per cell, corresponding to a total of 6912^3 grid cells. Below, we list the specifications of our Ly α forest and QSO catalogs for each simulation:

² The package and instructions for installing it can be found here: <https://github.com/abacusorg/abacusutils>.

³ Available via the package `abacusutils`.

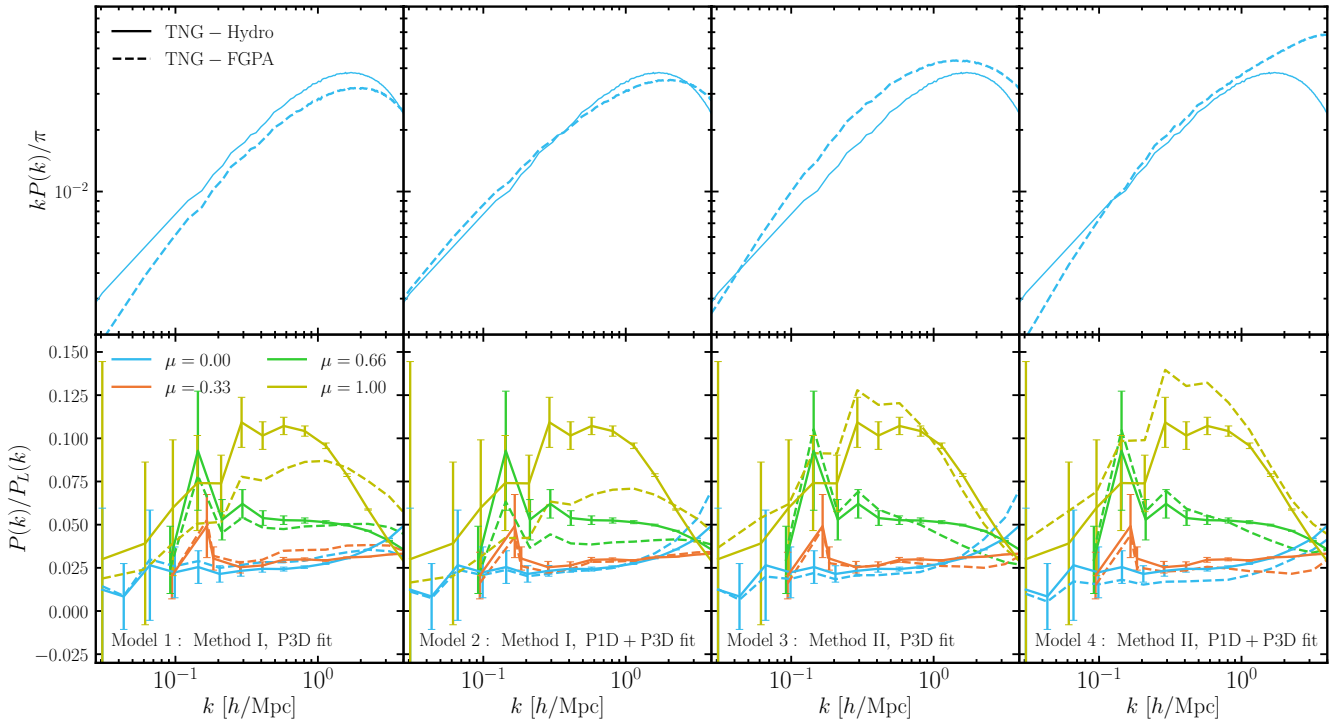


Figure 4. 1D and 3D power spectrum of the Ly α forest generated for the four models presented in this work (see Table 3.5 and Section 3) when applied to the N -body simulation TNG300-3-DM (dashed curves) and the “true” measurements from the hydro simulation (solid curves). In the top panels, we show the 1D power spectrum, $P_{1D}(k)$, with the shaded regions indicating jackknife errors, whereas in the bottom ones, we show the ratio between the linear power spectrum and $P(k, \mu)$ for four different values of $\mu = \{0, 0.33, 0.66, 1\}$, with the error bars coming from the Gaussian prediction. Note that the sample variance is the same in both boxes, which facilitates the comparison, and that on small scales, the measurement is affected by the effects of interlacing and aliasing due to the sparseness of the lines-of-sight (McDonald & Eisenstein 2007) and the cell size. Overall, the four models exhibit a good agreement with the “truth” and provide a wide variety of synthetic catalogs for the users of these mocks.

- Two full sets of redshift-space optical depth skewers (6912², with 6912 line-of-sight pixels), one placing the observer along the y axis and one along the z axis⁴. These skewers can be easily converted into flux transmission skewers according to Eq. 14. Each map takes up 1 TB of disk space and is split into 144 pieces each containing 48×6912 lines-of-sight.

- Two sets of complex $\tilde{\delta}_F(\mathbf{k})$ maps (as before, provided for line-of-sight along y and z directions) generated by Fourier transforming the flux contrast field $\delta_F(\mathbf{x})$ (6912³ cells) and then low-pass filtering the result, i.e. removing the small-scale modes, $k_{\max, \text{los}} > 4 h \text{ Mpc}^{-1}$ and $k_{\max, \text{perp}} > 2 h \text{ Mpc}^{-1}$ along and perpendicular to the line-of-sight, respectively⁵. We filter out small scales, which we know are dominated by baryonic effects missing in our simulations, to save disk space (each of the Fourier maps is 13 GB). We note that DESI will measure the Ly α forest power spectrum down to 2-3 $h \text{ Mpc}^{-1}$, so these complex maps provide sufficient small-scale information for modeling the DESI measurements.

- QSO catalogs containing the positions, velocities and host halo masses of each quasar with RSD effects applied along the y and z axis. The sample is generated via ABACUSHOD as described in

Section 2.2 with a number density of $1.75 \times 10^{-4} [h^{-1} \text{ Mpc}]^{-3}$ (i.e., 1.4 million quasars per box) and a bias and redshift distortion parameter of $b^{\text{QSO}} \approx 3.3$ (i.e., $\beta^{\text{QSO}} = f/b_q \approx 0.294$), chosen to be close to the eBOSS measurement (du Mas des Bourboux et al. 2020).

- Similarly to the complex maps we generate for the Ly α forest quantities, $\tilde{\delta}_F(\mathbf{k})$, we also generate complex maps of the quasar overdensity field, $\tilde{\delta}_g^{\text{QSO}}(\mathbf{k})$, calculated by Fourier transforming the quasar overdensity field, $\delta_g^{\text{QSO}}(\mathbf{x})$, obtained through the TSC interpolation of the redshift-distorted quasar positions on the three-dimensional grid (6912³ cells), and applying a low-pass filter of $k_{\max, \text{los}} < 4 h \text{ Mpc}^{-1}$ and $k_{\max, \text{perp}} < 2 h \text{ Mpc}^{-1}$.

4.2 Power spectrum

As described in Section 3.5, when deciding on the values of the free parameters introduced in our model, we try to maximize the similarity between the power spectrum measurements from the “true” Ly α forest skewers extracted from TNG300-1 and TNG300-3-DM equipped with our two FGPA-based methods (see Section 3.3 for descriptions). As a reminder, the 1D power spectrum of the Ly α forest is measured by Fourier transforming each skewer along the line-of-sight and averaging over all lines-of-sight to arrive at the final quantity. Thus, each skewer is treated independently and this statistic does not take into account any cross-correlation between different lines-of-sight. On the other hand, the second statistic, $P(k, \mu)$ (see Eq. 2), incorporates the correlation between skewers: $P(k, \mu = 0)$ measures the power in the transverse direction, whereas $P(k, \mu = 1)$ measures it in the direction parallel to the line-of-sight.

⁴ We do not generate maps with the line-of-sight direction being along the x axis, as the ABACUSUMMIT particle outputs are split into slabs along the x -axis that we analyze independently for the sake of efficiency.

⁵ In order to perform the Fourier transform of a 6912³ map on a single node, we consecutively load each slab in x , apply Fourier transformation in y and z and then low-pass filter the resulting array, until we are finished with all slabs and can apply one final low-pass filter along x .

The end goal of this project is to generate Ly α forest mocks in volumes sufficiently large to aid the analysis of large-scale surveys targeting quasars such as DESI. For this reason, it is of utmost importance that we can scale up our algorithm and run it successfully on the $2 h^{-1} \text{Gpc}$ ABACUSSUMMIT boxes. We note that as the cell grid of the data increases substantially between TNG300-1-DM and ABACUS-SUMMIT, it is necessary to refactor and rewrite our scripts altering the straightforward implementation described in Section 3. Therefore, verifying that our method yields results comparable to TNG300-3-DM is an essential step before our mocks are declared satisfactory. An additional complexity is that the resolution of TNG300-3 (mean particle distance of $0.33 h^{-1} \text{Mpc}$) and the ABACUSSUMMIT base boxes differs slightly (mean particle distance of $0.29 h^{-1} \text{Mpc}$). Ideally, one would want to recalibrate the slow parameters (see Section 3.5 for a definition of “slow” vs. “fast”) for each distinct simulation, but that would constitute a substantial computational burden. Here, we demonstrate that the behavior of the ABACUSSUMMIT mocks is sufficiently similar given our targeted precision, so we defer a more complex treatment to future work.

In Fig. 5, we study the 1D and 3D power spectrum of the “true” Ly α forest skewers from TNG300-1 and the skewers obtained for each of our four models from Section 3.5 applied to ABACUSSUMMIT. We find that the agreement of our mocks with TNG300-1 is similar to the agreement between TNG300-1 and TNG300-3-DM (see Fig. 4). We cut the smallest scales shown to $k < 4 h \text{Mpc}^{-1}$, as for these measurements, we employ the complex δ_F maps, which are available up to $k_{\text{max,los}} = 4 h \text{Mpc}^{-1}$ and easier to handle than the real-space τ skewers. It is reassuring that the agreement with TNG300-1 is comparable to our findings in Fig. 4, suggesting that the implementation of the mocks in the larger ABACUSSUMMIT boxes has been successful. Remaining differences in the intermediate regime, shared by both TNG300-3-DM and ABACUSSUMMIT can be attributed to differences in the resolution and the cosmological parameters.

We perform an additional test of dividing the power spectrum by the linear theory prediction with matching best-fit bias and β . We find that the mock power spectra agree within 10% with the linear theory result up to $k \lesssim 0.4 h \text{Mpc}^{-1}$, after which they begin to diverge noticeably from linear theory. The agreement within the Method I Models (i.e., 1 and 2) and within the Method II Models (i.e., 3 and 4) is excellent until $k \lesssim 2 h \text{Mpc}^{-1}$, but the two methods show larger deviations between each other beyond $k \sim 0.7 h \text{Mpc}^{-1}$, especially for high μ values.

4.3 Correlation function

Modern surveys will be capable of measuring the spectra of millions of distant objects and make subpercent measurements of the flux decrement correlation function over a wide range of scales and redshifts. This provides a handle of crucial cosmological measurements, such as the angular and redshift scale of the BAO, cosmic expansion, and the effect of neutrinos on the power spectrum. When measuring the small-scale clustering of galaxies, one can directly relate the redshift distortion parameter β to the growth of structure; however, in the case of the Ly α forest, β depends on a second bias factor that must be determined independently, which comes from a more general linear theory calculation of RSDs in which the distorted field, in this case τ , undergoes a non-linear transformation, in this case $F = \exp(-\tau)$ (McDonald et al. 2000). One viable way of doing so is by jointly analyzing the two-point correlation function of Ly α -Ly α , Ly α -QSO, and QSO-QSO in a “3 \times 2-pt” fashion (Cuceu et al. 2021). However, to do so reliably, we need to extensively test our analysis tools on realistic mocks. Hence, this is one of the main

objectives of our data products. In addition, it is well known that non-linear evolution causes a broadening of the BAO peak in the correlation function of galaxies. Therefore, it is interesting to ask whether the BAO peak in the flux correlation function is similarly broadened. In this work, we explore the auto- and cross-correlation function of Ly α and QSO and demonstrate the non-linear broadening of the Ly α -measured BAO peak for the first time in simulations. This is crucial to incorporate in and test through our theoretical models, as we expect that real Ly α observations will also be affected.

We summarize the flow of the section here to make it easier for the reader to follow. In Section 4.3.1, we sketch out the calculation connecting the theoretical power spectrum $P(k, \mu)$ to the correlation function multipoles, ξ_ℓ . The utility of this calculation is two-fold: to convert our simulated power spectrum into the simulated correlation function via the Hankel transform, we need a smooth function on very large scales, which we supply via linear theory by fitting the bias and β parameters to the simulated $P(k, \mu)$. On the other hand, we want to compare the simulated correlation function near the BAO scale with some theoretical model, so use these equations to calculate the linear theory prediction and also two models of the BAO peak broadening, defined in Section 4.3.2.

4.3.1 Measuring the correlation function from the power spectrum

To obtain the two-point correlation function measurement from our mocks, we start by calculating the power spectrum, $P(k, \mu)$, as before (see Eq. 2). We adopt maximum $k_{\text{max}} = 1.6 h \text{Mpc}^{-1}$ to speed up the calculations and because for this part of the validation, we are mostly interested in the BAO scales. We also calculate the multipoles of the redshift-distorted power spectrum, $P(k, \mu)$, via:

$$P_\ell(k) = \frac{2\ell + 1}{2} \int_{-1}^{+1} P(k, \mu) L_\ell(\mu) d\mu, \quad (19)$$

where L_ℓ is the Legendre polynomial and $P_\ell(k)$ are the multipoles of the redshift-distorted power spectrum $P(k, \mu)$.

Approximating the error on the measurement as Gaussian, we fit the b and β parameters to linear theory with the Kaiser approximation Kaiser (1987):

$$\tilde{P}(k, \mu) = b^2 (1 + \beta\mu^2)^2 \tilde{P}(k), \quad (20)$$

where the \tilde{P} signifies this is a theory prediction. Similarly, we fit the cross-power spectrum between Ly α and quasars via

$$\tilde{P}_q(k, \mu) = b b_q (1 + \beta\mu^2)(1 + \beta_q\mu^2) \tilde{P}(k), \quad (21)$$

to obtain the parameters b_q and β_q . For a given choice of b , b_q , β and β_q , where b_q and β_q are related through $\beta_q = f/b_q$, we can calculate the theory-predicted multipoles of the auto-power spectrum, truncating at the hexadecapole, $\ell = 4$,

$$\tilde{P}_\ell(k) = b^2 C_\ell(\beta) \tilde{P}(k) \quad (22)$$

where

$$C_\ell(\beta) \equiv \frac{2\ell + 1}{2} \int_{-1}^{+1} (1 + \beta\mu^2)^2 L_\ell d\mu = \begin{cases} 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 & \ell = 0 \\ \frac{4}{3}\beta + \frac{4}{7}\beta^2 & \ell = 2 \\ \frac{8}{35}\beta^2 & \ell = 4 \end{cases}. \quad (23)$$

Similarly, we can express the cross-power spectrum with QSOs as:

$$\tilde{P}_{q,\ell}(k) = b b_q C_{q,\ell}(\beta, \beta_q) \tilde{P}(k) \quad (24)$$

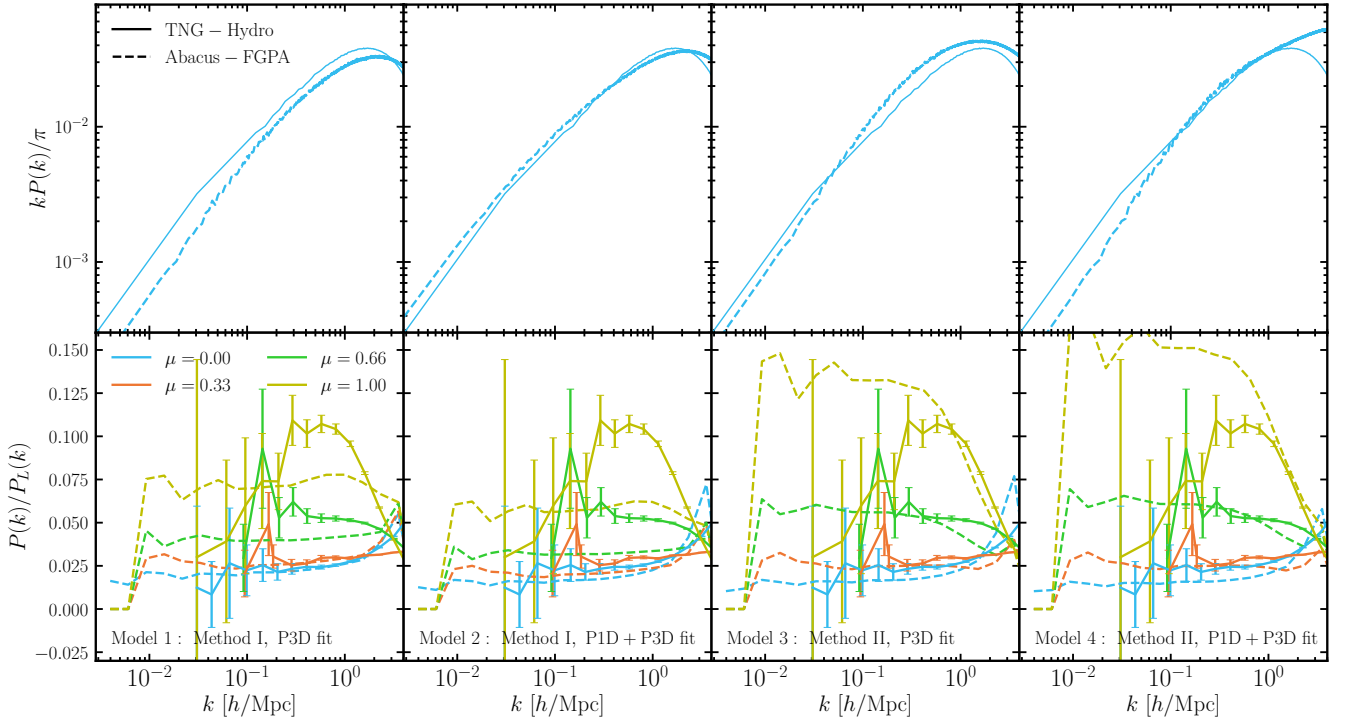


Figure 5. Similarly to Fig. 4, 1D and 3D power spectrum of the Ly α forest skewers extracted from our ABACUSUMMIT mocks using four different models (see Table 3.5 and Section 3). In this case, the N -body simulation (ph000) is much larger than the hydro one ($2 h^{-1}\text{Gpc}$ vs. $205 h^{-1}\text{Mpc}$), which allows us to extend the power spectrum measurements by about an order of magnitude to larger scales. We cut the smallest scales shown to $k < 4 h \text{Mpc}^{-1}$, as for this measurement, we use the complex δ_F maps, which are available up to $k_{\text{max,los}} = 4 h \text{Mpc}^{-1}$. Reassuringly, the agreement with TNG300-1 is comparable to our findings in Fig. 4 and provides us with confidence as to the validity of our Ly α forest mocks.

with

$$C_{q,\ell}(\beta) \equiv \begin{cases} 1 + \frac{1}{3}\beta + \frac{1}{3}\beta_q + \frac{1}{5}\beta \beta_q & \ell = 0 \\ \frac{2}{3}\beta + \frac{2}{3}\beta_q + \frac{4}{7}\beta \beta_q & \ell = 2 \\ \frac{8}{35}\beta \beta_q & \ell = 4 \end{cases} . \quad (25)$$

Having obtained both the measured and the theoretical multipoles of the power spectrum, we can combine them into a single data vector, so as to supplement the poorly measured large scales ($k \lesssim 0.01 h \text{Mpc}^{-1}$) with the linear theory fit,

$$P_\ell^{\text{comb}}(k) = (1 - w(k))P_\ell(k) + w(k)\tilde{P}_\ell(k), \quad (26)$$

where P_ℓ and \tilde{P}_ℓ are the predictions from ABACUSUMMIT and linear theory, respectively, and weighting function

$$w(k) \equiv \frac{1}{2} \left[1 - \tanh \left(\frac{k - k_{\text{pivot}}}{\Delta k_w} \right) \right], \quad (27)$$

which ensures smooth interpolation between the two limits. We use $\Delta k_w = 0.01 h \text{Mpc}^{-1}$, but manually fine-tune values of k_{pivot} for the different multipoles, based on their noisiness:

$$k_{\text{pivot},\ell=0} = 0.03 h \text{Mpc}^{-1}, \quad (28)$$

$$k_{\text{pivot},\ell=2} = 0.06 h \text{Mpc}^{-1}, \quad k_{\text{pivot},\ell=4} = 0.09 h \text{Mpc}^{-1}. \quad (29)$$

We have checked that the choice of a pivot scale for the monopole and quadrupole $\ell = 0, 2$ has negligible effect on the BAO feature. On the other hand, the hexadecapole, $\ell = 4$ is trickier to measure and hence a rather conservative scale cut is needed to ensure that the Hankel transform does not misbehave.

Finally, we Hankel transform the power spectrum multipoles into

correlation function multipoles, according to

$$\xi_\ell(r) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) P_\ell^{\text{comb}}(k) dk, \quad (30)$$

where j_ℓ is the spherical Bessel function, and

$$\xi_{q,\ell}(r) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) P_{q,\ell}^{\text{comb}}(k) dk, \quad (31)$$

for the quasar cross-correlation. The interpolation between theory and simulations ensures smooth integration and does not affect the clustering near the BAO scale and for smaller separations, r .

4.3.2 Comparing with linear and perturbation theory

One can model the effects of non-linear structure growth on the BAO feature via an anisotropic Gaussian smoothing of the linear power spectrum, effectively modifying Eq. 20 (Eisenstein et al. 2007):

$$\tilde{P}_{\text{nl}}(k, \mu) = \exp[-k^2 \Sigma^2(\mu)/2] \cdot \tilde{P}(k) \quad (32)$$

where $\tilde{P}(k)$ is the linear power spectrum and

$$\Sigma^2(\mu) = \mu^2 \Sigma_{\parallel}^2 + (1 - \mu^2) \Sigma_{\perp}^2, \quad (33)$$

where at redshift $z = 2.4$, we expect $\Sigma_{\parallel} \simeq 6.41 h^{-1}\text{Mpc}$ and $\Sigma_{\perp} \simeq 3.26 h^{-1}\text{Mpc}$. In principle, this equation implies that $\tilde{P}_\ell(k)$, Eq. 22, cannot be decomposed into a β - and a k -dependent factor, but instead that the integrals should be re-evaluated for each value of β . However, since $\Sigma \simeq 5 h^{-1}\text{Mpc}$ is expected to be smaller than the peak full-width half-maximum, following Kirkby et al. (2013), we adopt the approximation:

$$P_{\text{nl},\ell}(k) \simeq \exp(-k^2 \Sigma_\ell^2(\beta)/2) \cdot \tilde{P}_\ell(k), \quad (34)$$

so that each multipole undergoes a different amount of isotropic broadening according to:

$$\Sigma_\ell^2(\beta) \equiv f_\ell(\beta) \cdot \Sigma_{\parallel}^2 + (1 - f_\ell(\beta)) \cdot \Sigma_{\perp}^2 \quad (35)$$

where

$$f_\ell(\beta) \equiv \frac{\int_{-1}^{+1} \mu^2 (1 + \beta\mu^2)^2 L_\ell(\mu) d\mu}{\int_{-1}^{+1} (1 + \beta\mu^2)^2 L_\ell(\mu) d\mu} = \begin{cases} \frac{35+42\beta+15\beta^2}{105+70\beta+21\beta^2} & \ell = 0 \\ \frac{7+12\beta+5\beta^2}{14\beta+6\beta^2} & \ell = 2 \\ \frac{15}{11} + \frac{2}{\beta} & \ell = 4 \end{cases} \quad (36)$$

Similarly, we can perform the analogous integral to arrive at the equations predicting the non-linear broadening in the Ly α -QSO correlation function:

$$f_{q,\ell}(\beta) \equiv \begin{cases} \frac{35+21\beta+21\beta_q+15\beta\beta_q}{105+35\beta+35\beta_q+21\beta\beta_q} & \ell = 0 \\ \frac{7+6\beta+6\beta_q+5\beta\beta_q}{7\beta+7\beta_q+6\beta\beta_q} & \ell = 2 \\ \frac{15}{11} + \frac{1}{\beta} + \frac{1}{\beta_q} & \ell = 4 \end{cases} \quad (37)$$

Next, we study the correlation functions computed from our N -body mocks and compare them with linear and perturbation theory, with the latter following the derived form above.

In Fig. 6, we show the correlation function of the Ly α -Ly α and Ly α -QSO tracers for ABACUSUMMIT, linear theory (Eq. 22) and two BAO models based on Lagrangian Perturbation Theory (LPT): 1) Eisenstein et al. (2007), which applies the Gaussian smoothing to the entire linear power (Eq. 32); 2) Kirkby et al. (2013), which decomposes the power spectrum into a ‘wiggle’ and ‘no-wiggle’ component, and only applies the Gaussian smoothing to the peak component. We refer to these two models as LPT-based, but stress that they do not adopt the full LPT toolkit to model small scales, but instead offer BAO scale correction to recover the non-linear broadening of the BAO peak. The ABACUSUMMIT measurements are obtained by combining all six boxes for one of the four models (in particular, Model 1; see Table 3.5), performing a Hankel transform and averaging over them, to get smoother behavior. The BAO feature is visible in all curves except for $\ell = 4$, which is both noisier and has a weaker BAO signal. We note that the perturbation theory curve for $\ell = 4$ is also lacking a visible peak. It is clear that the sharpness of the linear theory prediction is substantially suppressed in the simulation, providing strong evidence of non-linear broadening. It is further reassuring that the perturbation theory predictions are in good agreement with the simulation at the BAO scale, especially for $\ell = 0$ and $\ell = 2$.

On scales smaller than $r \lesssim 80 h^{-1}\text{Mpc}$, the simplified perturbation theory model of Eisenstein et al. (2007) is inaccurate, as it over-smooths the power spectrum on small scales. Models that only smooth the peak component (Kirkby et al. 2013) can be trusted down to smaller scales. Remaining differences between the simulations and the Kirkby et al. (2013) model on small scales, i.e. $r \lesssim 30 h^{-1}\text{Mpc}$, can be attributed to the various simplifications of the Kaiser model, which neglects non-linear effects. Nonetheless, these appear to be quite small for the Ly α auto-correlation function, indicating that the Kaiser approximation works surprisingly well in that regime. They are, however, more pronounced when studying cross-correlations with the QSOs, suggesting that the QSO field is more affected by non-linear effects, as one would expect. One of the main uses of our mocks will be to test the scales at which the Kaiser approximation breaks down, as this is a central question for the analysis pipelines being developed.

In Fig. 7, we explore how the broadening changes for two of the four

different models we have adopted in generating the Ly α forest mocks, namely, Model 1 and Model 3. We find that this choice has little to no effect on the broadened BAO feature, and thus the comparison with perturbation theory remains qualitatively unchanged. Furthermore, the $\ell = 0$ and $\ell = 2$ multipoles of Models 1 and 3 are very consistent with each other across a wide range of scales, suggesting that the Ly α painting technique hardly affects these multipoles. Larger differences are seen for the $\ell = 4$ case, which is noisier and hence more difficult to measure, so we leave a more detailed study for the future. As in Fig. 6, the Kaiser approximation of the BAO model provides a poor match below $r \lesssim 30 h^{-1}\text{Mpc}$, indicating that the cross-correlations may need to be modeled beyond the Kaiser approximation with non-linear effects properly accounted for. Note that when showing the difference curves, we have rescaled the Model 1 multipoles by the pre-factors $C_\ell(\beta)$ and $C_{q,\ell}(\beta)$ (see Eq. 23 and Eq. 25) to account to linear order for the different values of β (see Table 3.5) and make the comparison with Model 3 more straightforward to see.

5 SUMMARY

The absorption of Ly α photons by hydrogen clouds imparts a characteristic signature on the spectra of high-redshift sources, known as the Ly α forest. These features, revealing the cosmic web of filamentary structures, have become a powerful tool for the study of large-scale structure in observational cosmology through measurements of their power spectrum and clustering. Accurately measuring these requires careful accounting of the systematic errors and is essential in order to extract cosmological constraints. The only reliable way of doing this is to generate random realizations of multiple Ly α absorption spectra in a survey and decorate them with various systematic effects so as to obtain a maximum realism data set. Examples of such systematics include a thorough modeling of the quasar continuum, which is used to infer the transmitted flux fraction, a modeling of the variable spectral resolution and noise, a calibration of the flux errors, an evaluation of the impact of redshift evolution, Damped Lyman alpha systems (DLAs), Lyman limit systems (LLS), metal absorption lines, and the cosmic ionizing background. The mock surveys needed to investigate these questions must include a large number of lines-of-sight over a large volume so as to satisfy the ambitious requirements set by surveys such as DESI, while also including small-scale fluctuations which contain a lot of valuable information through redshift-space distortions and the suppression of the power spectrum. Needless to say, this is extremely computationally challenging, and cosmologists typically resolve to having two sets of simulations for modeling the large- and small-scale observables. While our mocks do not provide accuracy down to the smallest scales needed to constrain neutrino and dark matter models, they present a first step to reconciling the large scales necessary for a BAO peak study and the small scales used to extract structure growth rate information in a single suite of mock catalogs.

Our Ly α forest synthetic catalogs are generated on the largest N -body simulation suite ABACUSUMMIT and are publicly available for six of the base boxes, $L_{\text{box}} = 2 h^{-1}\text{Gpc}$, at the fiducial Planck 2018 cosmology. Mock skewers are available on a regular grid with 6912^3 cells, and we output four different versions of our recipe per each observer location (at infinity along the z axis and along the y axis). In particular, we utilize the Fluctuating Gunn-Peterson Approximation (FGPA) and a modification thereof to transform the dark matter density field into a Ly α forest catalog (see Section 3 for details on our methods) and aim to match various Ly α observables extracted from a hydrodynamical simulation. Namely, we employ the high-

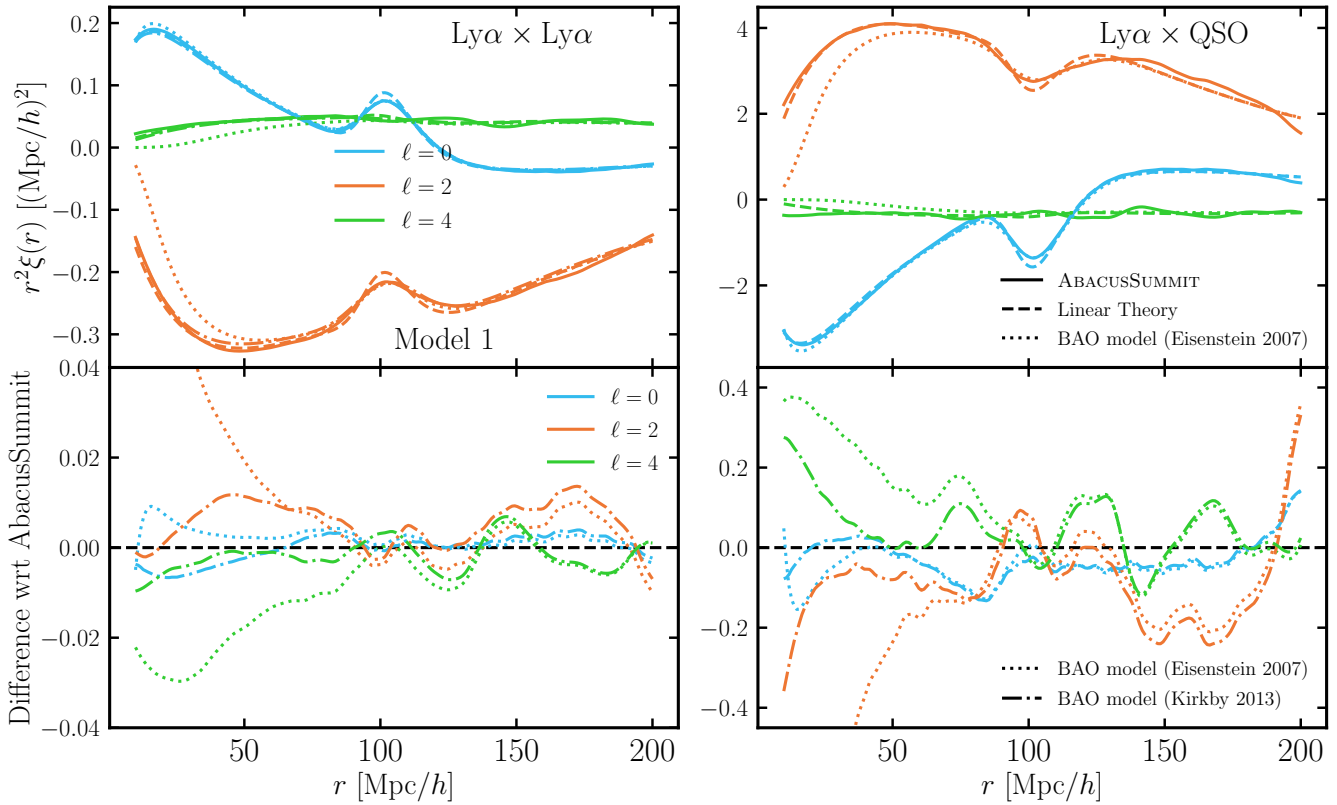


Figure 6. Multipoles of the $\text{Ly}\alpha$ - $\text{Ly}\alpha$ and $\text{Ly}\alpha$ -QSO correlation function, ξ_ℓ , comparing measurements from our $\text{Ly}\alpha$ forest mocks on ABACUSSUMMIT for Model 1 (see Table 3.5) with the theoretical prediction from linear theory and the BAO broadening (LPT-based) models of Eisenstein et al. (2007) and Kirkby et al. (2013). Difference with respect to the simulations is shown in the bottom. We find clear evidence for the broadening of the BAO peak in our simulations and excellent agreement with the LPT-based models. Note that the Eisenstein et al. (2007) is not suitable on small scales, i.e., below $r \lesssim 80 h^{-1}\text{Mpc}$, as it overly suppresses the power. On small scales below $r \lesssim 30 h^{-1}\text{Mpc}$, the Kirkby et al. (2013) BAO model prediction of the cross-correlation with quasars (right panel) deviates from the simulation, which we can attribute to the breakdown of the Kaiser approximation and small-scale physics effects (such as non-linear and Fingers-of-God effects).

realism $\text{Ly}\alpha$ forest skewers produced by Qezlou et al. (2022) for the hydro run TNG300-1 and calibrate our FGPA-based mocks against the mean and standard deviation of the transmission flux as well as the 1D and 3D $\text{Ly}\alpha$ power spectrum. We make the prime choices for our model parameters through the comparison between TNG300-1 and its low-resolution dark-matter-only counterpart TNG300-3-DM and find that our simplistic recipe yields a satisfactory agreement between the power spectra, as presented in Fig. 4. We then go on to apply this prescription to the ABACUSSUMMIT boxes, which have similar resolution to TNG300-3-DM, finding that the level of agreement is retained and the largest scales reachable largely extended by an order of magnitude (see Fig. 5). Next, we study the correlation function multipoles of $\text{Ly}\alpha$ - $\text{Ly}\alpha$ and $\text{Ly}\alpha$ -QSO in Fig. 6, which demonstrates for the first time in $\text{Ly}\alpha$ simulations the effect of non-linear clustering on the BAO peak. We find differences on small scales between the linear model (i.e., Kaiser approximation) and our mocks, especially in cross-correlations with the QSO population, which would be important to account for in the analysis of $\text{Ly}\alpha$ forest data.

Apart from being useful for testing systematics and the analysis pipeline, our mocks also open the doors for modeling novel statistics and joint probes with other tracers. As an example, developing and testing summary statistics that maximally use the 3D information in the $\text{Ly}\alpha$ forest such as the P_\times estimator of Font-Ribera et al. (2018) would be crucial to fully realizing the potential of the $\text{Ly}\alpha$ probe. The question of whether the $\text{Ly}\alpha$ forest measurements from current sur-

veys such as DESI can be utilized to constrain the growth rate f (e.g., as done in 3×2 -pt $\text{Ly}\alpha$ -QSO analysis), is also not yet fully resolved. By grafting the survey properties onto our mocks and performing the analysis on them as if on real data, we can tackle this problem and quote forecasts for the expected constraining power. It is also essential that we understand the scales at which linear theory breaks from our mocks and develop theoretical models that can recover the small-scale clustering correctly. Finally, a particularly exciting venue to explore is the development of combined analysis tools for $\text{Ly}\alpha$ forest and CMB lensing, which promises to break important degeneracies in our models (such as the two bias parameters characteristic of $\text{Ly}\alpha$ observables). As an initial step in near-term work, we plan to develop a model able to reproduce the joint data vector, using already available CMB and light cone products (Hadzhiyska et al. 2022, 2023), and successfully glean cosmological information from it.

Another important direction in which we could further develop our mocks is by adding realistic observational effects that are otherwise difficult to study analytically. These include (but are not limited to) the effects of Damped Lyman alpha systems (DLAs), Lyman limit systems (LLS), metal absorption lines, and the cosmic ionizing (UVB) background, which could be added to our mocks following prescriptions similar to the FGPA method employed in this work. In later versions of our mocks, we also plan to adopt density estimation techniques better suited for low-density regions such as phase-

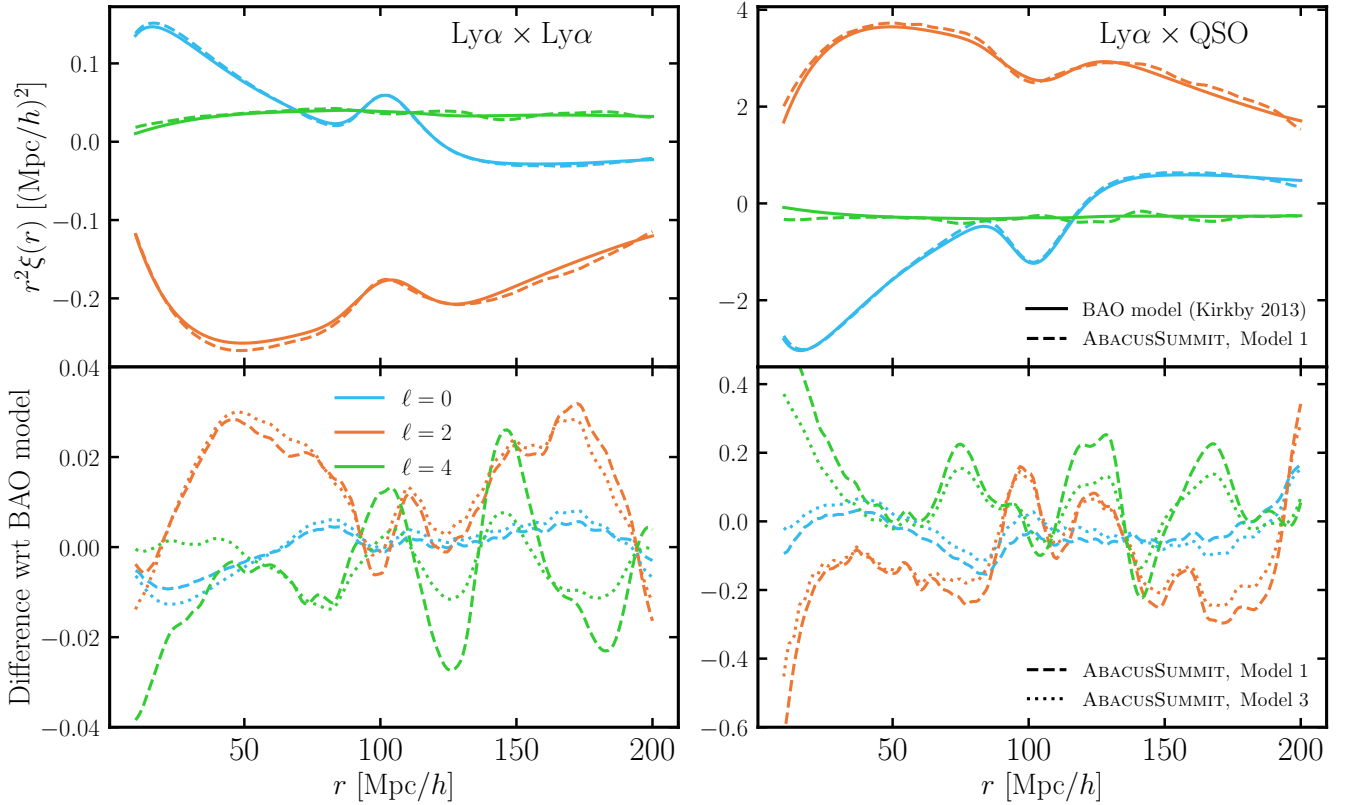


Figure 7. Multipoles of the Ly α -Ly α and Ly α -QSO correlation function, ξ_ℓ , comparing measurements from our Ly α forest mocks on ABACUSUMMIT for Model 1 and Model 3 (see Table 3.5) with the theoretical prediction from the BAO broadening (LPT-based) model of Kirkby et al. (2013). The bottom panels show the difference between the two mock measurements and the Kirkby et al. (2013) BAO model, while the top shows Model 1 and the Kirkby et al. (2013) prediction matching the bias and β values. Reassuringly, the $\ell = 0$ and $\ell = 2$ multipoles of Models 1 and 3 are very consistent with each other, suggesting that the Ly α painting technique has little effect on intermediate and large scales. Larger differences between the two are seen for the $\ell = 4$ case, which needs to be studied in more details. As in Fig. 6, the Kaiser approximation provides a poor match below $r \lesssim 30 h^{-1}\text{Mpc}$ for the cross-correlation, indicating that more careful modeling needs to be done to achieve sufficient precision. Note that we have rescaled the difference curves (bottom panels) for Model 1 by the pre-factors $C_\ell(\beta)$ and $C_{q,\ell}(\beta)$ (see Eq. 23 and Eq. 25) to account to linear order for the different values of β .

space tessellation, improved velocity field estimation schemes, and machine-learning methods for painting hydro simulation results on N -body simulations, which would improve the small-scale synthetic absorption signal. That way, we also hope to make our model more flexible to matching the observed power spectrum and correlation function across a wider range of scales and with a higher level of accuracy. In addition, we have planned to run our model on the light cone, so as to enable maximum realism including redshift evolution and curved sky effects, as well as facilitate joint studies with other tracers such as the CMB and weak lensing. The Ly α forest is still a largely unexplored resource that is brimming with astrophysical and cosmological information, waiting to be relinquished and utilized to uncover fundamental truths about our Universe.

ACKNOWLEDGEMENTS

We would like to thank Mahdi Qezlou, Vid Iršič, Patrick McDonald, John Peacock, Anže Slosar, Julien Guy, Zarija Lukić and Martin White for illuminating discussions over the course of this project. AFR acknowledges support from the Spanish Ministry of Science and Innovation through the program Ramon y Cajal (RYC-2018-025210) and from the European Union’s Horizon Europe research and innovation programme (COSMO-LYA, grant agreement 101044612).

IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. AC acknowledges support from the United States Department of Energy, Office of High Energy Physics under Award Number DE-SC-0011726.

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; additional support for DESI is provided by the U.S. National Science Foundation, Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technologies Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Science and Technology of Mexico (CONACYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: <https://www.desi.lbl.gov/collaborating-institutions>.

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.

- Vogelsberger M., et al., 2014a, *MNRAS*, 444, 1518
 Vogelsberger M., et al., 2014b, *Nature*, 509, 177
 Weinberger R., et al., 2017, *MNRAS*, 465, 3291
 Weinberger R., Springel V., Pakmor R., 2020, *ApJS*, 248, 32
 Yèche C., Palanque-Delabrouille N., Baur J., du Mas des Bourboux H., 2017, *J. Cosmology Astropart. Phys.*, 2017, 047
 York D. G., et al., 2000, *AJ*, 120, 1579
 Yuan S., Garrison L. H., Hadzhiyska B., Bose S., Eisenstein D. J., 2022, *MNRAS*, 510, 3301
 Zaldarriaga M., Hui L., Tegmark M., 2001, *ApJ*, 557, 519
 du Mas des Bourboux H., et al., 2020, *ApJ*, 901, 153

APPENDIX A: CONVOLUTION WITH THE DOPPLER PROFILE

When applying redshift space distortions in our FGPA-based models, we ignore the effects of thermal broadening due to the random thermal velocities of the gas atoms. In this Appendix, we illustrate the impact thermal broadening, as implemented through a Doppler profile convolution, has on our measured 1D power spectrum. The obvious advantage of the Voigt profile is that it incorporates a physical effect and thus adds more realism to the very small-scale behavior. On the other hand, it is computationally more expensive than the alternative and ends up yielding qualitatively similar results to what we obtain when we add small-scale noise (see Section 3.2). Below, we will show that adopting the Voigt profile has an almost negligible effect on the power spectrum given the current level of accuracy of our mocks and scales of interest for the power spectrum, $k \lesssim 3 h \text{Mpc}^{-1}$, and thus, we can justify omit it from the present mocks.

To obtain the Doppler-profile-convolved optical depth in redshift space, we perform the integral over velocity space for each skewer:

$$\tau(s) = \int dx \frac{\tau(x)}{b(x)} \exp \left[- \left(\frac{(s-x-v_r(x))}{b(x)} \right)^2 \right], \quad (\text{A1})$$

where s and x are velocity coordinates, v_r is the peculiar velocity in the line-of-sight direction, and

$$b(x) \equiv \sqrt{\frac{2k_B T(x)}{m_p}} \quad (\text{A2})$$

is the thermal velocity of the atoms, k_B the Boltzmann constant, and m_p the proton mass. We approximate the temperature of the gas as:

$$T(x) = T_0 [1 + \delta_{\text{dm}}(x)]^{\gamma-1} \quad (\text{A3})$$

with δ_{dm} being the dark-matter overdensity obtained via TSC interpolation and $T_0 = 1.94 \times 10^4$ K being the normalization factor (Qezlou et al. 2022).

From Fig. A1, we see that indeed the effect of convolving with the Doppler profile on the 1D power spectrum is negligible for our model, although it does appear to boost slightly the power near $k \sim 1 h \text{Mpc}^{-1}$, reducing the discrepancy between the two curves. We note that after applying the convolution, we refit τ_0 and σ_ϵ (see Section 3) to fit the mean and variance of the flux. Additionally, we expect thermal broadening to lead to a suppression of the power on scales smaller than $k > 1 h \text{Mpc}^{-1}$, which would add more realism to our mocks and lead to a better agreement with the hydro simulation on these scales (see the top panels of Fig. 5). We plan to incorporate this effect into future versions of our mocks, as the change is negligible given our current precision and scales of interest.

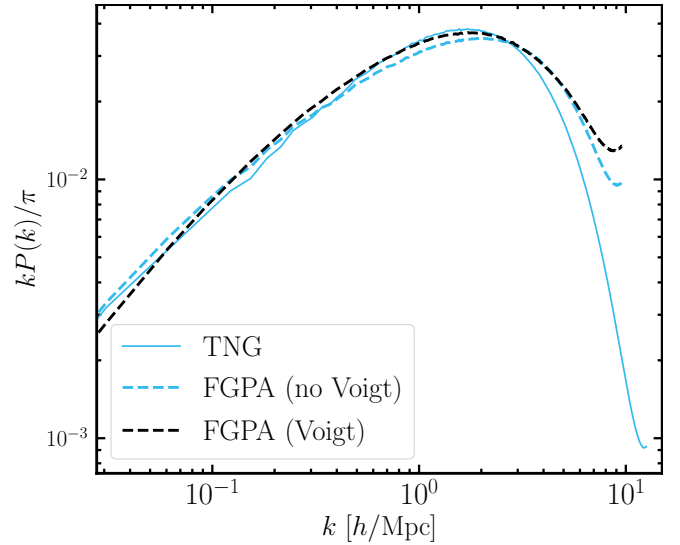


Figure A1. 1D power spectrum comparison between the “true” Ly α forest extracted from TNG300-1 and the FGPA-generated skewers for the fiducial model (blue solid; see Table 3.5) with and without a convolution with the Doppler profile (black and blue dashed, respectively; see Eq. A1). Note that we refit τ_0 and σ_ϵ (see Section 3) to fit the mean and variance of the flux.