

# Direct cosmological inference from three-dimensional correlations of the Lyman- $\alpha$ forest

Francesca Gerardi,<sup>1\*</sup> Andrei Cuceu,<sup>2,3,4</sup> Andreu Font-Ribera,<sup>5,1</sup> Benjamin Joachimi<sup>1</sup> and Pablo Lemos<sup>6,1</sup>

<sup>1</sup> Department of Physics & Astronomy, University College London, Gower Street, London WC1E 6BT, UK

<sup>2</sup> Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, Ohio 43210, USA

<sup>3</sup> Department of Astronomy, The Ohio State University, Columbus, Ohio 43210, USA

<sup>4</sup> Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

<sup>5</sup> Institut de Física d'Altes Energies, The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona), Spain

<sup>6</sup> Department of Physics and Astronomy, University of Sussex, Sussex House, Falmer, Brighton, BN1 9RH, UK

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

When performing cosmological inference, standard analyses of the Lyman- $\alpha$  ( $\text{Ly}\alpha$ ) three-dimensional correlation functions only consider the information carried by the distinct peak produced by baryon acoustic oscillations (BAO). In this work, we address whether this compression is sufficient to capture all the relevant cosmological information carried by these functions. We do this by performing a direct fit to the full shape, including all physical scales without compression, of synthetic  $\text{Ly}\alpha$  auto-correlation functions and cross-correlations with quasars at effective redshift  $z_{\text{eff}} = 2.3$ , assuming a DESI-like survey, and providing a comparison to the classic method applied to the same dataset. Our approach leads to a 3.5% constraint on the matter density  $\Omega_M$ , which is about three to four times better than what BAO alone can probe. The growth term  $f\sigma_8(z_{\text{eff}})$  is constrained to the 10% level, and the spectral index  $n_s$  to  $\sim 3 - 4\%$ . We demonstrate that the extra information resulting from our ‘direct fit’ approach, except for the  $n_s$  constraint, can be traced back to the Alcock-Paczynski effect and redshift space distortion information.

**Key words:** cosmological parameters – large-scale structure of universe – methods: data analysis

## 1 INTRODUCTION

Over the last couple of decades, after the discovery of the accelerated expansion of the Universe (Riess et al. 1998; Perlmutter et al. 1999), cosmology has focused on investigating the properties of dark energy. Among the multiple probes used to place a constraint on the parameters of the current  $\Lambda$ CDM model, there is the baryon acoustic oscillation (BAO) scale over different tracers<sup>1</sup> of the matter density field (Eisenstein et al. 2005; Cole et al. 2005). Measurements of this standard ruler over a range of redshifts place a constraint on the expansion history (Seo & Eisenstein 2003).

Complementary to low-redshift galaxies ( $z \lesssim 1$ ), the Lyman- $\alpha$  ( $\text{Ly}\alpha$ ) forest is a tracer of the intergalactic medium (IGM) that probes the cosmic expansion via BAO at higher redshifts, as first proposed by McDonald & Eisenstein (2007). The  $\text{Ly}\alpha$  forest is a sequence of absorption lines in high-redshift quasar (QSO) spectra, caused by the neutral hydrogen distributed along the line of sight, between the quasar and the observer. The first BAO detection from the  $\text{Ly}\alpha$  auto-correlation function and from its cross-correlation with QSOs was in the Baryon Oscillation Spectroscopic Survey (BOSS) DR9 data (Busca et al. 2013; Slosar et al. 2013; Kirkby et al. 2013) and DR11 data (Font-Ribera et al. 2014), respectively.

BAO produce a distinct feature in the correlation functions, which

we wish to measure and use to probe cosmology, in a robust and model-independent way. When performing cosmological inference, a standard method, as applied in BOSS and eBOSS (du Mas des Bourboux et al. 2020) analyses of the  $\text{Ly}\alpha$  three-dimensional correlation functions, relies on splitting them into a *peak* and a *smooth* component and only considers the information carried by the BAO feature.

Recently, Cuceu et al. (2021) (C21 hereafter) demonstrated that further cosmological information can be obtained from the broadband component using the Alcock-Paczynski (AP) effect (Alcock & Paczynski 1979). When computing the 3D correlation functions from observations, as a standard approach, we change from angular and redshift separations to comoving coordinates, based on the assumption of a fiducial cosmological model. In particular, if the latter differs from the true underlying cosmology, then the AP effect will appear as an apparent anisotropy in the correlation functions. Another source of anisotropy is redshift space distortions (RSD), which are induced by peculiar velocities and hence carry extra information. However, measuring redshift space distortions for the  $\text{Ly}\alpha$  auto-correlation alone is not informative about the growth rate of structure because of its degeneracy with an unknown velocity divergence bias (Seljak 2012). For this reason, C21 jointly employed the  $\text{Ly}\alpha$  auto- and cross-correlation with quasars to explore the potential of measuring the linear growth of structure.

All physical scales of the 3D  $\text{Ly}\alpha$  correlation functions, beyond the BAO peak, carry information about the underlying cosmology. Throughout this work we will refer to the sum of all of these scales,

\* E-mail: francesca.gerardi.19@ucl.ac.uk

<sup>1</sup> For a BAO distance ladder plot <https://www.sdss.org/science/cosmology-results-from-eboss/>.

with no compression, as the full shape of these functions. Both the  $\text{Ly}\alpha \times \text{Ly}\alpha$  auto- and  $\text{Ly}\alpha \times \text{QSO}$  cross-correlation functions can be directly used to perform cosmological inference. The work of C21 motivates a further investigation, assessing whether or not the compressed analysis based on BAO, AP and RSD successfully captures all cosmological information from the correlation functions of interest.

The same point is relevant also in the field of galaxy clustering, where the compressed standard approach extracts cosmological information from BAO, AP and RSD (Alam et al. 2017, 2021). Over the past few years, advancements in perturbation theory computations boosted the interest in fitting the observed two-point statistics and directly inferring cosmological parameters without compression (Tröster et al. 2020; Ivanov et al. 2020; d’Amico et al. 2020). In particular, Brieden et al. (2021) carried out such an investigation, identifying from where in the data vector additional information originates and extending the classic approach by introducing an extra physical parameter.

In this work, we aim to address whether or not compression is a suitable approach in the field of  $\text{Ly}\alpha$  forest cosmology, by performing a direct fit to the full shape of synthetic  $\text{Ly}\alpha \times \text{Ly}\alpha$  auto- and  $\text{Ly}\alpha \times \text{QSO}$  cross-correlation functions, and by subsequently comparing the constraints with those obtained using the standard approach. In real data,  $\text{Ly}\alpha$  correlations have contaminants that affect the amount of cosmological information extractable from a given scale. We discuss here an optimistic scenario without contaminants, and leave for future work a detailed study of their impact on the cosmological constraints.

The paper is structured as follows. We start in Sect. 2 by outlining the methodology and explaining the inference framework we use. We proceed to apply this method to noiseless synthetic correlation functions and present the forecasts in Sect. 3. We finally compare our main results to the compressed analyses in Sect. 4 and draw our conclusions in Sect. 5.

## 2 METHOD

We use the full shape of the 3D  $\text{Ly}\alpha$  correlation functions to directly infer cosmological parameters, without the usual compression methods. Based on the modelling of C21 and with the aid of CAMB (Lewis et al. 2000; Howlett et al. 2012), we construct a likelihood for the Cobaya framework (Torrado & Lewis 2021; Torrado & Lewis 2019), and we perform Markov chain Monte Carlo (MCMC) sampling (Lewis & Bridle 2002; Neal 2005; Lewis 2013) using the  $\text{Ly}\alpha \times \text{Ly}\alpha$  auto- and  $\text{Ly}\alpha \times \text{QSO}$  cross-correlations.

In this section we outline the key features of the method. In Sect. 2.1 we describe how the synthetic correlation data was generated, and focus on the modelling in Sect. 2.2. In particular, throughout the analysis, we use the code VEGA<sup>2</sup> (Cuceu et al. 2020), which is based on the code used in eBOSS DR16 (du Mas des Bourboux et al. 2020) for fitting and modelling the  $\text{Ly}\alpha$  correlation functions. We finally motivate the choice of the sampled parameter space and describe the likelihood in Sect. 2.3.

### 2.1 Synthetic data vector and covariance

In this work, we focus on idealised 3D  $\text{Ly}\alpha$  synthetic correlations in flat  $\Lambda$ CDM, without contaminants. However, we do include the

distortion due to quasar continuum fitting. The latter filters out information and ‘distorts’ the true correlation function (Bautista et al. 2017; du Mas des Bourboux et al. 2017). Our synthetic data was generated using the framework of C21, and is given by an uncontaminated model based on the best fit of eBOSS DR16 (see Tab. 1). We did not add noise to the data vector, as we are only interested in forecasting. We used covariance matrices based on DESI mocks similar to those used in Youles et al. (2022). These mocks were created with the CoLoRe (Ramírez-Pérez et al. 2022) and LyaCoLoRe (Farr et al. 2020) packages, covering 14000 sq. degrees with a target density of  $\sim 50$  QSOs / sq. degree (DESI Collaboration et al. 2016). The covariance was computed using the community package `picca` (du Mas des Bourboux et al. 2020). In this analysis, we limit ourselves to linear scales, assuming  $r_{\min} = 30h^{-1}\text{Mpc}$ , up to  $r_{\max} = 180h^{-1}\text{Mpc}$ . The effective redshift of the correlation functions is  $z_{\text{eff}} = 2.3$ .

### 2.2 Modelling

To infer cosmology from these synthetic correlations, we first need a theory to model the data given any cosmology  $\mathbf{p}_{\mathcal{C}}$ . The theoretical 3D  $\text{Ly}\alpha$  correlation functions are computed from the isotropic matter power spectrum  $P(k)$  and then compared against data to evaluate the likelihood.

When modelling and fitting  $\text{Ly}\alpha$  correlations, we must match the coordinate grid for the theoretical correlation  $\xi$  with the grid of the data. When measuring the 3D  $\text{Ly}\alpha$  correlation functions from observations, we change from angular  $\Delta\theta$  and redshift  $\Delta z$  separations to a set of comoving coordinates  $(r_{\parallel}, r_{\perp})$ , respectively defined along and across the line of sight. This is motivated by the fact that both the radial comoving distance  $D_{\mathcal{C}}(z) = c \int_0^z dz/H(z)$  and the comoving angular diameter distance  $D_{\mathcal{M}}(z)$  are redshift dependent, where  $c$  is the speed of light and  $H(z)$  the Hubble parameter. For this reason, we wish to refer instead to a set of comoving coordinates. Given two locations at redshift  $z_i$  and  $z_j$  separated by an angle  $\Delta\theta$ , these are defined as

$$r_{\parallel} = [D_{\mathcal{C},\text{fid}}(z_i) - D_{\mathcal{C},\text{fid}}(z_j)] \cos \frac{\Delta\theta}{2}; \quad (1)$$

$$r_{\perp} = [D_{\mathcal{M},\text{fid}}(z_i) + D_{\mathcal{M},\text{fid}}(z_j)] \sin \frac{\Delta\theta}{2}, \quad (2)$$

where both  $D_{\mathcal{C}}$  and  $D_{\mathcal{M}}$  are computed using an assumed fiducial cosmology (fid subscript). In our case, the fiducial cosmology coincides with the cosmological model that was used to generate the data vector. Given that the sampled cosmology that generated the theoretical  $\xi$  can be different from the fiducial one, we need to match coordinate grids of data and  $\xi$  by rescaling at each sampling step the coordinates of the correlation via

$$q_{\parallel} = D_{\mathcal{H}}(z_{\text{eff}})/D_{\mathcal{H}}^{\text{fid}}(z_{\text{eff}}); \quad (3)$$

$$q_{\perp} = D_{\mathcal{M}}(z_{\text{eff}})/D_{\mathcal{M}}^{\text{fid}}(z_{\text{eff}}), \quad (4)$$

where  $D_{\mathcal{H}}(z) = c/H(z)$ , such that  $r'_{\parallel,\perp} = q_{\parallel,\perp} r_{\parallel,\perp}$ .

In modelling the  $\text{Ly}\alpha$  correlation functions of interest we follow Eq. (27) of du Mas des Bourboux et al. (2020), adopting the prescriptions of C21. For any cosmology  $\mathbf{p}_{\mathcal{C}}$ , the power spectra of the tracers are computed from the isotropic linear matter power spectrum  $P(k, z)$  as

<sup>2</sup> <https://github.com/andreicuceu/vega>

Parameter	Fiducial	Prior	68% limits
$H_0$ [km/(s $\times$ Mpc)]	67.31	$\mathcal{U}(40, 100)$	$67.69^{+5.5}_{-3.16}$
$\Omega_M$	0.3144	$\mathcal{U}(0.01, 0.99)$	$0.318 \pm 0.011$
$\Omega_B h^2$	0.02222	$\mathcal{U}(0.01, 0.05)$	$0.0229^{+0.0064}_{-0.0038}$
$A_s$	$2.196 \cdot 10^{-9}$	$\mathcal{U}_{\log(10^{10} A_s)}(0.5, 6)$	$(2.06^{+0.42}_{-0.46}) \cdot 10^{-9}$
$n_s$	0.9655	$\mathcal{U}(0.8, 1.2)$	$0.958^{+0.025}_{-0.035}$
$b_{Ly\alpha}$	-0.117	$\mathcal{U}_{\log(-b_{Ly\alpha})}(-2, 0)$	$-0.111^{+0.011}_{-0.012}$
$\beta_{Ly\alpha}$	1.67	$\mathcal{U}(0, 5)$	$1.67 \pm 0.03$
$b_{QSO}$	3.8	$\mathcal{U}_{\log(b_{QSO})}(-2, 1.3)$	$3.61^{+0.47}_{-0.32}$
$\sigma_v$ (Mpc/h)	6.86	$\mathcal{U}(0, 15)$	$6.75^{+0.64}_{-0.55}$

**Table 1.** Full set of sampled parameters, alongside with the fiducial values used to compute the synthetic correlations and the uniform ( $\mathcal{U}$ ) priors adopted for the sampling procedure. When sampling in logarithmic space we add a ‘log’ subscript to  $\mathcal{U}$ . In the last column, we provide the one-dimensional marginals (68% c.l.) for all the parameters sampled, where for any asymmetric posteriors we report the posterior maximum with lower and upper 68% limits.

$$P_{Ly\alpha}(k, \mu_k, z) = b_{Ly\alpha}^2 \left(1 + \beta_{Ly\alpha} \mu_k^2\right)^2 F_{nl, Ly\alpha}^2(k, \mu_k) P(k, z); \quad (5)$$

$$P_{\times}(k, \mu_k, z) = b_{Ly\alpha} \left(1 + \beta_{Ly\alpha} \mu_k^2\right) \times (b_{QSO} + f(z) \mu_k^2) F_{nl, QSO}(k_{\parallel}) P(k, z), \quad (6)$$

with  $f(z) = \frac{d \ln D}{d \ln a}$  being the logarithmic growth rate and  $\mu_k = k_{\parallel}/k$ , where  $k$  and  $k_{\parallel}$  are the modulus of the wave vector and its projection along the line of sight respectively. Focusing first on the Ly $\alpha$  $\times$ Ly $\alpha$  power spectrum in Eq. (5), we identify the Ly $\alpha$  forest linear bias,  $b_{Ly\alpha}$ , and its RSD term,  $\beta_{Ly\alpha} = \frac{b_{\eta, Ly\alpha} f(z)}{b_{Ly\alpha}}$ , where  $b_{\eta, Ly\alpha}$  is an extra unknown bias, the velocity divergence bias. The choice of using  $\beta_{Ly\alpha}$  in the RSD term comes from the fact that  $b_{\eta, Ly\alpha}$  is fully degenerate with the growth rate  $f(z)$ . We treat both  $b_{Ly\alpha}$  and  $\beta_{Ly\alpha}$  as nuisance parameters and marginalize over them. The  $F_{nl, Ly\alpha}$  term encodes the non-linear corrections according to the model of Arinyo-i Prats et al. (2015). The parameters involved in this model are kept constant for simplicity, but in principle they should also be varied. However, this should not have a major impact in our analysis, as we are restricting the analysis to linear scales ( $r_{\min} = 30h^{-1}$  Mpc). Moving on to the Ly $\alpha$  $\times$ QSO power spectrum  $P_{\times}$  in Eq. (6),  $b_{QSO}$  is the quasar linear bias, another nuisance parameter. In contrast to the Ly $\alpha$  RSD term, the QSO RSD term is instead simply  $f(z)$  as by definition  $b_{\eta, QSO} = 1$ . Following du Mas des Bourboux et al. (2020), we model the impact of redshift errors and non-linear peculiar velocities of quasars with a damping term  $F_{nl, QSO}(k_{\parallel})$ . We use a Lorentzian function

$$F_{nl, QSO}(k_{\parallel}) = \sqrt{\left[1 + (k_{\parallel} \sigma_v)^2\right]^{-1}}, \quad (7)$$

with a free parameter  $\sigma_v$ , which represents the velocity dispersion and is an extra nuisance parameter.

At each sampling step then, theoretical correlation functions  $\xi$  are computed in VEGA. At its core, VEGA decomposes the power spectrum into multipoles, transforms them into correlation function multipoles using the FFTLog algorithm (Hamilton 2000) and finally reconstructs the two-dimensional correlation function.

### 2.3 Parameter space and likelihood

The BAO feature is able to constrain  $\alpha_{\parallel} = D_H(z) r_d^{\text{fid}} / D_H^{\text{fid}}(z) r_d$  along the line of sight and  $\alpha_{\perp} = D_M(z) r_d^{\text{fid}} / D_M^{\text{fid}} r_d$  in the transverse

direction, where  $r_d$  is the sound horizon at the drag epoch. Since in flat  $\Lambda$ CDM, at low redshifts, the Hubble parameter  $H(z)$  can be expressed as a function of  $H_0$  and  $\Omega_M$ ,  $\alpha_{\parallel}$  and  $\alpha_{\perp}$  will ultimately place a constraint on  $\{H_0 r_d, \Omega_M\}$ . Additionally,  $r_d$  can be numerically approximated as a function of  $\Omega_{\nu} h^2$ ,  $\Omega_M h^2$  and  $\Omega_B h^2$  (Aubourg et al. 2015), which are the neutrino, matter and baryon densities respectively, all evaluated at redshift  $z = 0$  by definition. This further motivates the choice of sampling  $\{H_0, \Omega_M, \Omega_B h^2\}$ , where  $\Omega_{\nu} h^2$  is constant for a given choice of the neutrino mass. Extra information on  $\Omega_M$  also comes from the AP effect (Alcock & Paczynski 1979)

$$F_{AP} = \frac{\alpha_{\perp}}{\alpha_{\parallel}} = \frac{D_M(z_{\text{eff}}) / D_M^{\text{fid}}(z_{\text{eff}})}{D_H(z_{\text{eff}}) / D_H^{\text{fid}}(z_{\text{eff}})} = \frac{[D_M(z_{\text{eff}}) H(z_{\text{eff}})]}{[D_M(z_{\text{eff}}) H(z_{\text{eff}})]_{\text{fid}}}, \quad (8)$$

which is an apparent anisotropy present if the sampled cosmology differs from the fiducial one. For the same assumptions as before, the AP parameter will only be a function of  $\Omega_M$ . As we fit the full shape of the correlation functions directly and the amplitude of primordial fluctuations  $A_s$  and their spectral index  $n_s$  affect the functional form of  $\xi$ , we will sample the full set of parameters  $\mathbf{p}_C = \{H_0, \Omega_M, \Omega_B h^2, A_s, n_s\}$ . On the other hand,  $\{b_{Ly\alpha}, b_{QSO}, \beta_{Ly\alpha}, \sigma_v\}$  are treated as nuisance parameters  $\mathbf{p}_A$  to marginalize over.

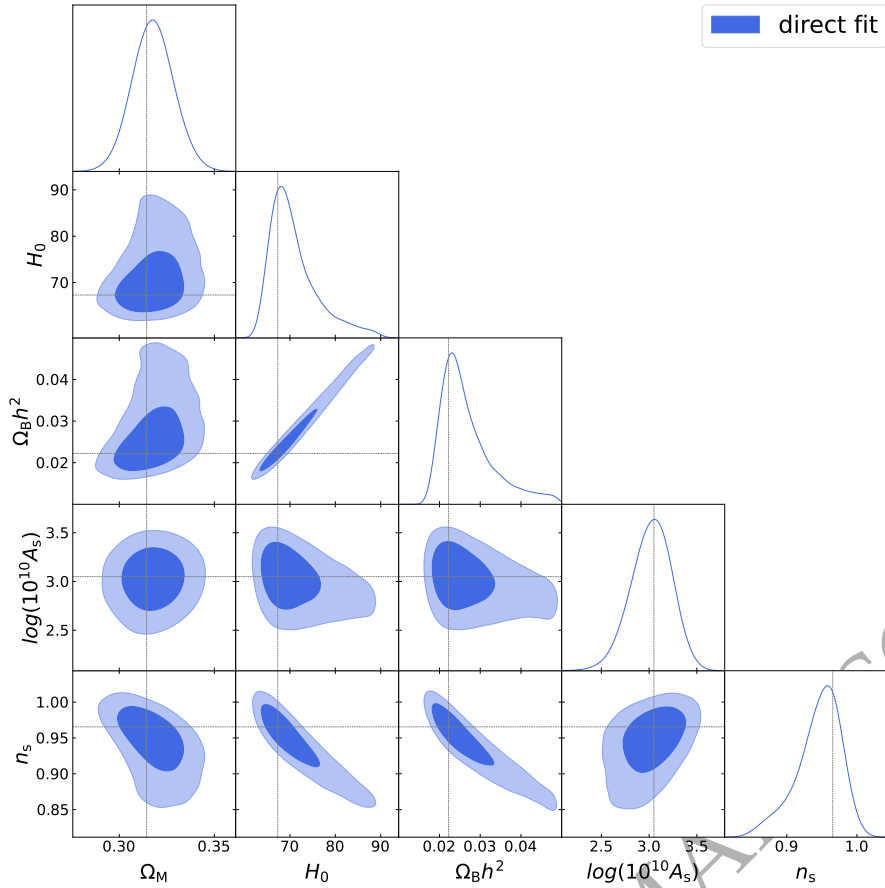
For all these parameters we choose uniform priors, which are listed in Tab. 1. As is common,  $A_s$  is sampled in logarithmic space, and we made the choice of doing the same with the two linear biases because they are degenerate with  $A_s$  and span over several orders of magnitude.

In this work, we assume a Gaussian likelihood, which is also computed using VEGA. A likelihood evaluation, via CAMB, first computes the comoving distances, to calculate  $q_{\parallel}$  and  $q_{\perp}$ , and then the isotropic linear matter power spectrum, along with  $r_d$ , the growth rate at  $z_{\text{eff}}$  and  $f \sigma_8(z_{\text{eff}})$ . Then, VEGA computes the correlation functions, based on the modelling description in Sect. 2.2, and the  $\chi^2$  value.

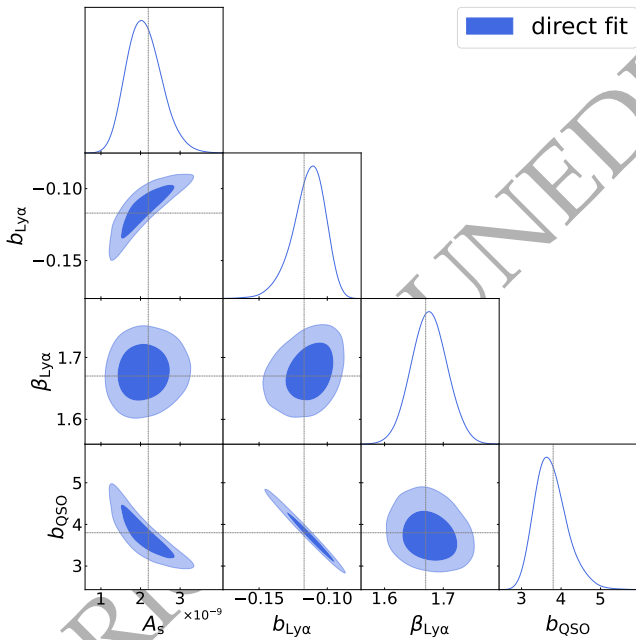
## 3 RESULTS

In this section we present the forecasts produced using the method outlined in Sect. 2 on 3D Ly $\alpha$  $\times$ Ly $\alpha$  and Ly $\alpha$  $\times$ QSO simplified synthetic correlation functions. We sample over the cosmological parameters  $\mathbf{p}_C = \{H_0, \Omega_M, \Omega_B h^2, A_s, n_s\}$ , marginalizing over the astrophysical model parameters  $\mathbf{p}_A = \{b_{Ly\alpha}, b_{QSO}, \beta_{Ly\alpha}, \sigma_v\}$ . The fiducial values of these parameters, along with the priors, are listed in Tab. 1.

In Fig. 1 we show the results for  $\{H_0, \Omega_M, \Omega_B h^2, A_s, n_s\}$  using the noiseless mock data vector. In the last column of Tab. 1 we list



**Figure 1.** Triangle plot of the cosmological parameters of interest  $\{H_0, \Omega_M, \Omega_B h^2, A_s, n_s\}$ , marginalizing over the nuisance parameters  $\mathbf{p}_A$ . The blue contours refer to the results obtained performing the inference using the method outlined in Sect. 2, which we denote as ‘direct fit’. The grey dashed lines mark the fiducial values used to generate data.



**Figure 2.** Correlations among the nuisance parameters  $\{b_{\text{Ly}\alpha}, \beta_{\text{Ly}\alpha}, b_{\text{QSO}}\}$  and  $A_s$ , where ‘direct fit’ refers to the inference method described in Sect. 2. The grey dashed lines mark the fiducial values used to generate the data.

the one-dimensional marginal constraints for the full set of sampled parameters.

From Fig. 1, it can be seen that we do recover the true values (shown in grey dashed lines) of cosmological parameters  $\mathbf{p}_C$  well within  $1\sigma$  (Tab. 1). The analysis provides a 3.5% constraint on the matter density  $\Omega_M$ . Clear degeneracies are present between  $H_0$  and  $\Omega_B h^2$ . As previously mentioned, the baryon acoustic oscillation peak measures the product  $H_0 r_d$ , which can be expressed as a function of  $H_0$ ,  $\Omega_B h^2$  and  $\Omega_M$ . Given that we have a good measurement of  $\Omega_M$  from the AP information, the remaining degeneracy is between  $H_0$  and  $\Omega_B h^2$ : if there were no other information on either one of them, these two parameters would be fully degenerate. The fact that both  $H_0$  and  $\Omega_B h^2$  are strongly correlated with the spectral index  $n_s$  could hint that the turnover of the power spectrum is the feature partially breaking the degeneracy. Despite this correlation, we are able to place a constraint on the spectral index, namely  $n_s = 0.958^{+0.025}_{-0.035}$ . We obtain a 21% constraint on the amplitude of fluctuations  $A_s$ , with a corresponding constraint on the amplitude of linear matter fluctuations in spheres of  $8 h^{-1} \text{Mpc}$  of  $\sigma_8(z_{\text{eff}}) = 0.317 \pm 0.032$ . We will further analyze the constraining power of our analysis against the state-of-the-art results later in Sect. 4.1.

In Fig. 2 we show the strong correlation among the linear biases,  $b_{\text{Ly}\alpha}$  and  $b_{\text{QSO}}$ , and  $A_s$ , which is expected given the functional form of Eqs. (5-6). The  $\text{Ly}\alpha$  auto-correlation alone would not be able to place a constraint on  $A_s$  since, for  $\mu = 0$ , we would measure the combination  $A_s b_{\text{Ly}\alpha}^2$  only, whereas its anisotropy would provide a



Parameter	BAO	BAO+AP+RSD	direct fit
$\Omega_M$	12%	4%	3.5 %
$H_0 r_d/c$	4.5%	1.65%	1.43 %
$f\sigma_8(z_{\text{eff}})$	–	12.5%	10.4%

**Table 2.** Constraining power of our method (‘direct fit’) on the listed parameters, against those from the standard analysis (BAO) and the one of C21 (BAO+AP+RSD).

measurement of  $\beta_{\text{Ly}\alpha}$ . On the other hand, the transverse mode of the cross-power spectrum (Eq. 6), combined with the auto-correlation, measures the combination of  $A_s b_{\text{Ly}\alpha}^2$  and  $b_{\text{QSO}}/b_{\text{Ly}\alpha}$ , while through the RSD term we are able to constrain  $A_s$  (or  $f\sigma_8$  in compressed analyses).

## 4 DISCUSSION

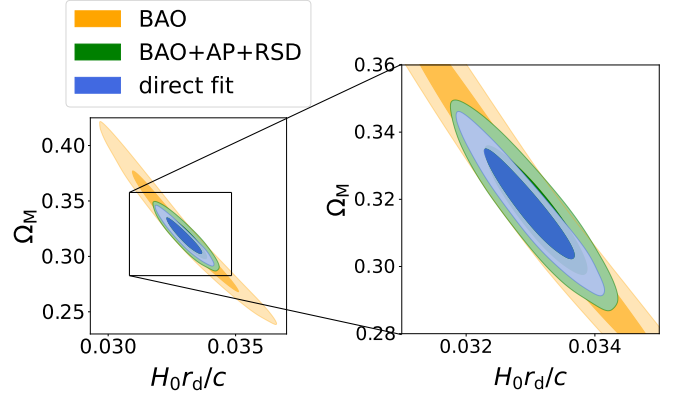
In Sect. 4.1 we provide a direct comparison of the forecasts presented in Sect. 3 and the literature. In particular, we will focus on a comparison with the results on the same synthetic data obtained using the standard BOSS and eBOSS analysis first, as well as the C21 approach. Our goal is to understand whether the compressed analyses successfully capture the cosmological information carried by the 3D Ly $\alpha$  correlation functions and discuss which components of the data are the most informative. In Sect. 4.2 we discuss how results change when marginalizing over the growth of structure. This is instructive to further understand from where extra information originates.

### 4.1 Cosmological information

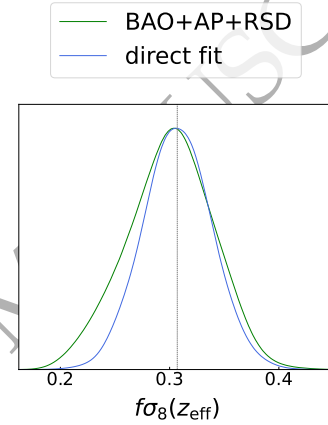
As mentioned above, in flat  $\Lambda$ CDM the BAO scale, along and across the line of sight, identifies a banana-shaped degeneracy in the  $[H_0 r_d/c, \Omega_M]$  plane. This justifies that any comparison among methods which have the BAO as a primary feature should necessarily happen in this plane. On the other hand, the Ly $\alpha$ -QSO cross-correlation can in principle measure  $f\sigma_8(z_{\text{eff}})$  because of the functional form of Eq. (6). However, because of the degeneracy with the linear biases, the combination with the Ly $\alpha$  auto-correlation is needed. In what follows we will focus on a comparison based on the derived parameters  $\mathbf{p}_d = \{H_0 r_d/c, f\sigma_8(z_{\text{eff}})\}$  and  $\Omega_M$ . In particular, in Fig. 3 we plot the two-dimensional contours of  $\{H_0 r_d/c, \Omega_M\}$  and in Fig. 4 the one-dimensional marginal of  $f\sigma_8(z_{\text{eff}})$  for the methods we want to compare.

As discussed above, standard BOSS and eBOSS analyses focus on the peak component of the 3D Ly $\alpha$  correlation functions only. For this reason, we will refer to this approach as ‘BAO’ for simplicity. We run this analysis using our noiseless mock data, and the most important result is shown in Fig. 3. This approach provides  $\Omega_M = 0.32 \pm 0.04$ ,  $H_0 r_d/c = 0.0329 \pm 0.0015$ , with a constraining power on  $H_0 r_d/c$  of 4.5%, a factor of three worse compared to our direct fit (summary in Tab. 2).

Such an improvement was already found by C21, who demonstrated that considering the AP effect from the *smooth* component in addition to the *peak* provides significantly tighter constraints. We present results using their method with the additional RSD information, and we will refer to it as ‘BAO+AP+RSD’. By running their analysis over our noiseless data, we find that ‘BAO+AP+RSD’ is able to place a constraint on  $H_0 r_d/c$  of 1.65%, which is of the same order as for our analysis (Tab. 2). As it can be seen in Fig. 3,  $H_0 r_d/c$  and  $\Omega_M$  are strongly correlated, with correlation coefficient 0.95,



**Figure 3.** Two-dimensional contour plots of  $\{H_0 r_d/c, \Omega_M\}$ , comparing our method (‘direct fit’) in blue against standard BOSS and eBOSS analysis (‘BAO’) in orange and C21 (‘BAO+AP+RSD’) in green. On the right, there is a zoom-in to further highlight the differences among the ‘direct fit’ and ‘BAO+AP+RSD’ methods.



**Figure 4.** Posterior plot for  $f\sigma_8(z_{\text{eff}})$ , comparing our method (‘direct fit’) in blue against C21 (‘BAO+AP+RSD’) in green.

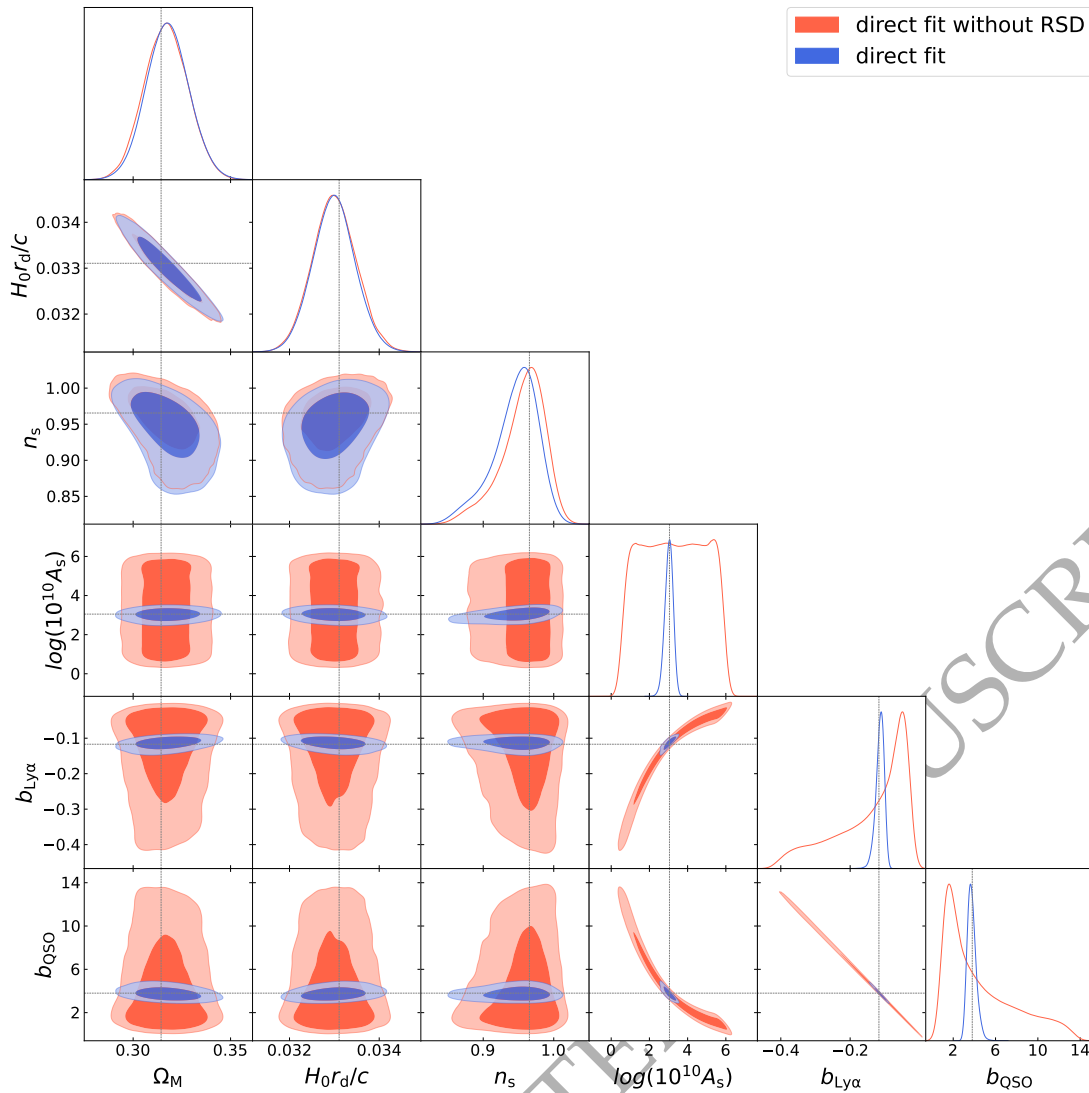
along the same direction as observed by C21. Our method constrains the growth of structure (Eq. 5-6) as well. For this reason, we also compare the RSD information of C21 with ours. The  $f\sigma_8(z_{\text{eff}})$  one-dimensional marginals for both methods are plotted in Fig. 4 and the constraining power is given for completion in Tab. 2. Overall, our method provides tighter constraints, on  $H_0 r_d/c$ ,  $\Omega_M$  and  $f\sigma_8(z_{\text{eff}})$ , with respect to the ‘BAO+AP+RSD’ analysis, by about 16%, 17% and 18%, respectively.

### 4.2 Direct fit analysis without RSD

In order to further investigate where some of the information in the ‘direct fit’ alone is coming from, we repeat the same analysis, marginalizing over the growth of structure. We refer to this case as ‘direct fit without RSD’. If we use the quasar RSD parameter defined as  $\beta_{\text{QSO}} = f(z)/b_{\text{QSO}}$ , Eq. (6) can be rewritten as

$$P_{\times}(k, \mu_k, z) = b_{\text{Ly}\alpha} \left( 1 + \beta_{\text{Ly}\alpha} \mu_k^2 \right) \times b_{\text{QSO}} \left( 1 + \beta_{\text{QSO}} \mu_k^2 \right) F_{\text{nl, QSO}} P(k, z), \quad (9)$$

where  $\beta_{\text{QSO}}$  is now an extra nuisance parameter.



**Figure 5.** Triangle plot comparing the constraints on  $\{\Omega_M, H_0 r_d/c, A_s, n_s\}$  and the two linear biases  $\{b_{\text{Ly}\alpha}, b_{\text{QSO}}\}$  using our ‘direct fit’ approach in blue against the same fitting method but marginalizing over the growth of structure (‘direct fit without RSD’) in red. A further explanation of the second approach can be found in Sect. 4.2.

The implications of this choice can be directly seen in Fig. 5. In the no-RSD case (red in Figure), the degeneracy among  $A_s$  and the two linear biases extends up to the prior limits, while the information on the other cosmological parameters is preserved. Marginalizing over the growth of structure washes out any constraining power on the linear biases and  $A_s$ , while RSD are not constraining any other parameter.

## 5 CONCLUSIONS

Given that baryon acoustic oscillations (BAO) produce a distinct feature in 3D  $\text{Ly}\alpha$  correlation functions, and its properties are well understood, BOSS and eBOSS analyses so far considered only the peak for cosmological inference (‘BAO’ analysis). A previous analysis conducted by Cuceu et al. (2021) (C21 throughout the paper) highlighted the importance of also considering the broadband component, as it significantly contributes to the overall cosmological information via the Alcock-Paczynski (AP) effect (‘BAO+AP’ analysis

– ‘BAO+AP+RSD’ if redshift space distortions (RSD) are included). Given these premises, in this paper we addressed the question about whether or not the compressed analyses based on BAO, AP and RSD parameters are able to capture all the cosmological information brought by the  $\text{Ly}\alpha$  correlation functions. We performed a full shape analysis without any of the above parameters and instead directly inferred cosmology (‘direct fit’ analysis).

The inference framework we used is Cobaya, for which we implemented an *ad-hoc* gaussian likelihood based on the VEGA package, as extensively described in Sect. 2.

We performed ‘BAO’, ‘BAO+AP+RSD’ and ‘direct fit’ analyses on the same set of synthetic  $\text{Ly}\alpha \times \text{Ly}\alpha$  auto- and  $\text{Ly}\alpha \times \text{QSO}$  cross-correlations, which include distortion effects due to continuum fitting, but no other contaminants, and ran the inference over  $\mathcal{PC} = \{H_0, \Omega_M, \Omega_B h^2, A_s, n_s\}$ . We also marginalized over the nuisance astrophysical parameters  $\mathcal{p}_A = \{b_{\text{Ly}\alpha}, b_{\text{QSO}}, \beta_{\text{Ly}\alpha}, \sigma_v\}$ , which are the  $\text{Ly}\alpha$  and quasars linear biases, the  $\text{Ly}\alpha$  RSD term and the velocity dispersion of quasars respectively.

We were able to measure the matter density parameter  $\Omega_M$  and the

amplitude of primordial fluctuations  $A_s$  with a precision of 3.5% and 21%, respectively, and  $f\sigma_8(z_{\text{eff}})$  at 10.4%, which is a noteworthy result given the few measurements of this parameter at  $z > 2$ . For these parameters we do not obtain a significant improvement in constraining power with respect to the ‘BAO+AP+RSD’ approach. However, we obtain a constraint of the spectral index  $n_s$  at the  $\sim 3 - 4\%$  level. Similarly to the findings and solutions put forward in Brieden et al. (2021), we could account for this extra information by adding a slope parameter to the compressed analysis.

The robustness of the ‘direct fit’ method against systematics must be tested. In forthcoming work, it would be interesting to investigate whether the compressed analysis is more robust, given it measures specific physical effects that are well understood. A further natural next step should be including contaminants to the analysis. The constraining power in the spectral index that we achieve could be affected in particular by Damped Ly $\alpha$  System (DLA) contamination (McQuinn & White 2011; Font-Ribera et al. 2012) and fluctuations in the UV background (Pontzen & Governato 2014; Gontcho et al. 2014), which would change the correlation function in a way that could mimic  $n_s$ . Further improvements to the analysis could come from varying  $r_{\text{min}}$  and  $r_{\text{max}}$ , and also checking how effects of continuum distortion consequently behave.

We conclude by recalling that soon DESI will provide even better measurements of Ly $\alpha$  correlations. Therefore this kind of study is key to finding the optimal approach to infer cosmology from the data.

## ACKNOWLEDGEMENTS

This work was partially enabled by funding from the UCL Cosmoparticle Initiative. AC acknowledges support from the United States Department of Energy, Office of High Energy Physics under Award Number DE-SC-0011726. AFR acknowledges support by the program Ramon y Cajal (RYC-2018-025210) of the Spanish Ministry of Science and Innovation 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. BJ acknowledges support by STFC Consolidated Grant ST/V000780/1. PL acknowledges STFC Consolidated Grants ST/R000476/1 and ST/T000473/1. For the purpose of open access, the authors have applied a creative commons attribution (CC BY) licence to any author-accepted manuscript version arising.

## DATA AVAILABILITY

The code is publicly available at [https://github.com/frgerardi/LyA\\_directfit](https://github.com/frgerardi/LyA_directfit). The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

Alam S., et al., 2017, *MNRAS*, 470, 2617  
 Alam S., et al., 2021, *Phys. Rev. D*, 103, 083533  
 Alcock C., Paczynski B., 1979, *Nature*, 281, 358  
 Arinyo-i Prats A., Miralda-Escudé J., Viel M., Cen R., 2015, *JCAP*, 12, 017  
 Aubourg E., et al., 2015, *Phys. Rev. D*, 92, 123516  
 Bautista J. E., et al., 2017, *Astron. Astrophys.*, 603, A12  
 Brieden S., Gil-Marín H., Verde L., 2021, *Journal of Cosmology and Astroparticle Physics*, 2021, 054  
 Busca N. G., et al., 2013, *Astronomy & Astrophysics*, 552, A96

Cole S., et al., 2005, *Mon. Not. Roy. Astron. Soc.*, 362, 505  
 Cuceu A., Font-Ribera A., Joachimi B., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 035  
 Cuceu A., Font-Ribera A., Joachimi B., Nadathur S., 2021, *Monthly Notices of the Royal Astronomical Society*, 506, 5439–5450  
 DESI Collaboration et al., 2016, The DESI Experiment Part I: Science, Targeting, and Survey Design, doi:10.48550/ARXIV.1611.00036, <https://arxiv.org/abs/1611.00036>  
 Eisenstein D. J., et al., 2005, *Astrophys. J.*, 633, 560  
 Farr J., et al., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 068  
 Font-Ribera A., et al., 2012, *Journal of Cosmology and Astroparticle Physics*, 2012, 059  
 Font-Ribera A., et al., 2014, *JCAP*, 05, 027  
 Gontcho S. G. A., Miralda-Escudé J., Busca N. G., 2014, *Monthly Notices of the Royal Astronomical Society*, 442, 187  
 Hamilton A. J. S., 2000, *Monthly Notices of the Royal Astronomical Society*, 312, 257  
 Howlett C., Lewis A., Hall A., Challinor A., 2012, *Journal of Cosmology and Astroparticle Physics*, 2012, 027  
 Ivanov M. M., Simonović M., Zaldarriaga M., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 042  
 Kirkby D., et al., 2013, *JCAP*, 03, 024  
 Lewis A., 2013, *Phys. Rev. D*, 87, 103529  
 Lewis A., Bridle S., 2002, *Phys. Rev. D*, 66, 103511  
 Lewis A., Challinor A., Lasenby A., 2000, *The Astrophysical Journal*, 538, 473  
 McDonald P., Eisenstein D. J., 2007, *Phys. Rev. D*, 76, 063009  
 McQuinn M., White M., 2011, *Monthly Notices of the Royal Astronomical Society*, 415, 2257  
 Neal R. M., 2005, Taking Bigger Metropolis Steps by Dragging Fast Variables, doi:10.48550/ARXIV.MATH/0502099, <https://arxiv.org/abs/math/0502099>  
 Perlmutter S., et al., 1999, *The Astrophysical Journal*, 517, 565  
 Pontzen A., Governato F., 2014, *Nature*, 506, 171  
 Ramírez-Pérez C., Sanchez J., Alonso D., Font-Ribera A., 2022, *Journal of Cosmology and Astroparticle Physics*, 2022, 002  
 Riess A. G., et al., 1998, *The Astronomical Journal*, 116, 1009  
 Seljak U., 2012, *Journal of Cosmology and Astroparticle Physics*, 2012, 004  
 Seo H.-J., Eisenstein D. J., 2003, *ApJ*, 598, 720  
 Slosar A., et al., 2013, *JCAP*, 04, 026  
 Torrado J., Lewis A., 2019, Cobaya: Bayesian analysis in cosmology, Astrophysics Source Code Library, record ascl:1910.019 (ascl:1910.019)  
 Torrado J., Lewis A., 2021, *Journal of Cosmology and Astroparticle Physics*, 2021, 057  
 Tröster T., et al., 2020, *A&A*, 633, L10  
 Youles S., et al., 2022, The effect of quasar redshift errors on Lyman- $\alpha$  forest correlation functions, doi:10.48550/ARXIV.2205.06648, <https://arxiv.org/abs/2205.06648>  
 d’Amico G., Gleyzes J., Kokron N., Markovic K., Senatore L., Zhang P., Beutler F., Gil-Marín H., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 005  
 du Mas des Bourboux H., et al., 2017, *A&A*, 608, A130  
 du Mas des Bourboux H., et al., 2020, *The Astrophysical Journal*, 901, 153