Prepared for submission to JCAP

The effects of the non-linear evolution on the BAO measurement from the Lyman-α forest at $z = 2.5$

Oleg Burgueño,[a,1] **and Others** [b]

[1]Corresponding author.
[2]Also at Some University.

Contents

abstract

1 Introduction

The discovery that the Universe is expanding [1] and the posterior confirmation that it is also accelerating [2] has brought up the idea of Dark Energy (DE), a component that accounts for two thirds of the energy content of the Universe. Several efforts to understand the physical properties of DE are now in the mainstream cosmology research [3], [4], [5], with surveys whose goal is to measure the Baryon Acoustic Oscillations (BAO) at the forefront [6], [7], [8], [9].

The Baryon Acoustic Oscillations arise from sound waves that propagate in the hot baryon-photon plasma in the early Universe. After recombination, the Universe becomes cold enough for the photons to decouple from the baryons and causing the sound waves to freeze in place, leaving an imprint in the Cosmic Microwave Background (CMB) anisotropies [10], [11]. At late times, the anisotropies can be observed as wiggles in the matter power spectrum, or as a well defined peak in the two point correlation function (2PCF) of large scale clustering. The BAO signal is often referred as a standard ruler, as measuring the length of the feature along the line of sight (LoS) allow us to calculate the Hubble parameter $H(z) = c\Delta z/s_{\parallel}(z)$ and a with a measurement of the transverse length we can calculate the angular diameter distance $d_A(z) = s_{\perp}/\Delta\theta(z)$, where $(s_{\parallel}, s_{\perp})$ are the position of the peak along and transverse to the line of sight, respectively; $\Delta\theta$ is the angular size of the feature, Δz is redshift range it spans and c is the velocity of light. Also, by measuring $H(z)$ and $d_A(z)$ one can put constrains on the equation of state of dark energy, $w(z)$.

Many surveys have been (or soon will be) deployed with the primary goal of measuring BAO using different tracers of the underlying dark matter density field in a large redshift range. The first measurement of BAO in the power spectrum was reported by the 2dF Galaxy

Redshift Survey (2dFGRS) [12], while the first detection in the correlation function was done by the Sloan Digital Sky Survey (SDSS) [13] collaboration; both using tracer galaxies. Later, better measurements have been made by improving the catalogs [14] and adding different tracers such as high redshift QSO's, yielding better constrains on the cosmological parameters.

Aside of discrete tracers such as galaxies and quasars, another kind of tracer is available in the form of the Lyman-α forest [15], [16]; a continuous absorption feature that appears in the spectra of high redshift QSO's. This signal is the result of QSO radiation traveling along the line of sight, being redshifted and then absorbed by diffuse neutral hydrogen present in the Inter Galactic Medium (IGM) producing a Lyman-α transition. Similar to galaxies tracers, analytical modeling of the forest have shown that there exist a relation between the Lyman-α forest flux and the matter density field, thus making the forest an appropriate BAO tracer. The Baryon Acoustic Spectroscopic Survey (BOSS) program of SDSS reported the first measurement of the Lyman-α forest flux auto-correlation function [17] and subsequently they also reported the first detection of BAO [18].

A great amount of work is invested in the characterization of systematic effects and noise that could alter the data and the data analysis. As surveys start to increase the density of observed targets to obtain a higher level of accuracy of the measurements, previous less study effects in the data are becoming an important subject of study. The broadening of the BAO peak is a well known non-linear effect present in the measurement of the 2PCF, that when not taken into consideration can bias the parameter inference. Previous work has been done to understand how much this systematic effect modifies the measurment correlation function of tracer galaxies [19], showing a reduction of correlation $\sim 4\%$ around the BAO peak when considering non-linearity. In the case the the Lyman-α forest, some work has been done to add the broadening of the signal at the level of the model [20], but there is little to no effort directed to measure it directly on the the flux auto-correlation function.

The aim of this work is to study the effects of the BAO broadening on the Lyman-α flux auto-correlation function and compare if the theoretical modeling for the broadening holds up (as the description from [20] is directly based in the formalism of [19]), as well as to test if it is possible to measure such effect in modern survey such as the Dark Energy Spectroscopic Survey (DESI) [8]. For that goal we present a methodology to generate Lyman-α mocks using the Comoving Lagrangian Acceleration (COLA) method. This algorithm combines second order Lagrangian perturbation theory solution (2LPT) and a exact N-body integrator to solve for the evolution of the dark matter density field. Compared to a full N-body simulation the COLA method can reproduce nearly the same clustering at a fraction of time steps three times faster [21]. This is of great advantage to us, as it allow us to compute the large amount of realizations needed to calculate the expected covariance matrix of a modern survey without being limited by today's available computational resources. Gaussian random fields (GRF) with a lognormal transformation has been used to generate galaxy and Lyman-α mock catalogs [22],[23], [24], [25], as an alternative of running expensive N-body simulations. The advantage of this mocks is that running realizations with high densities and large volumes is very cheap as they do not solve for the evolution density field. The downside is that they have to be tuned to reproduce a particular clustering distribution, hence GRF's are not useful to measure the broadening effect.Andrei mocks

A brief description of the COLA method is presented in section 2, and in section 2.1 we shown how to transform the output of the simulations to a Lyman-α flux field. In section 3 we show how to obtain the correlation from the mocks; in section 4 we present the model we used for the correlation function and finally in section 5 we describe the fitting procedure and

the results of our analysis.

Trough-out this work we used a standard flat ΛCDM cosmology with parameters: $h = 0.67$, $\Omega_\Lambda = 0.69$, $\Omega_b = 0.048$, $\sigma_8 = 0.87$, $n_s = 0.96$.

2 Lyman-α COLA mocks.

In this section we briefly describe the COLA method its main differences with respect to to a full N-body simulation. We also present the steps made to go from a the dark matter density field of a simulation box to the Lyman-α flux density field.

The main goal of a N-body code is to solve the equation of motion of the dark matter particles:

$$\nabla \left(\frac{d^2\Psi_{i,j}}{d\tau^2} + H(\tau)\frac{d\Psi_{i,j}}{d\tau} \right) = -\frac{3}{2}\Omega_{0,m}H(\tau)\delta(q), \tag{2.1}$$

where τ is the conformal time; H is the Hubble parameter; $\Omega_{0,m}$ is the matter density parameter; $\delta(q)$ is the density contrast in the Lagrangian frame with q representing the Lagrangian spatial coordinate; Ψ is the Lagrangian displacement field and $\Psi_{i,j} = \frac{\partial \Psi_i}{\partial x_j}$. The Lagrangian coordinates relate to the Eulerian frame as:

$$x(\tau) = q + \Psi(q, \tau). \tag{2.2}$$

Using perturbation theory up to second order, the displacement field can be expanded as $\Psi = \Psi^{(1)} + \Psi^{(2)}$. Then equation 2.1 can be rewritten as:

$$\nabla \cdot \Psi^{(1)} = -D_1(\tau)\delta_1(q), \tag{2.3}$$

$$\nabla \cdot \Psi^{(2)} = \frac{1}{2}D_2(\tau)\sum_{i\neq j}(\Psi^{(1)}_{i,i}\Psi^{(1)}_{j,j} - \Psi^{(1)}_{i,j}\Psi^{(1)}_{j,i}), \tag{2.4}$$

where $D_1(\tau)$ is the linear growth factor and $D_2(\tau) \approx -3D_1^2(\tau)\Omega_{0,m}^{-1/143}/7$, is the second order growth factor for a flat ΛCDM Universe. It is important to notice that the dynamic variable of equation 2.1 is no longer coupled to the displacement field in equations 2.3 and 2.4; all the information is now contained in the growth factor. For a more detailed description of how to derive the solutions for Ψ refer to Appendix C.

The COLA method solve the equation of motion of DM particles as:

$$\partial_\tau^2\Psi_{res} + \partial_\tau^2 D_1(\tau)\Psi^{(1)} + \partial_\tau^2 D_2(\tau)\Psi^{(2)} + \nabla\phi = 0, \tag{2.5}$$

where $\Psi^{(1)}$ and $\Psi^{(2)}$ are the 2LPT displacement fields used to solve for scales larger than 100 Mpch^{-1}; Ψ_{res} is the resultant of subtracting the perturbative solutions from the complete displacement field, $\Psi_{res} = \Psi - \Psi^{(1)} - \Psi^{(2)}$, and

$$\partial_\tau^2 X = \frac{d^2X}{d\tau^2} + H(\tau)\frac{dX}{d\tau}. \tag{2.6}$$

To solve equation 2.5, an initial matter power spectrum at redshift $z = 0$ is used to obtain the perturbative displacement fields, which then are shifted back in time using the growth factors. Ψ_{res} is solved using a full N-body algorithm.

To generate the simulation set used throughout this work, we choose the publicly available code L-PICOLA [26]. This software is a parallel implementation of the COLA method with an option to perform lightcone simulations on the fly. In the future this feature in could allow us to include the effects of redshift evolution in our analysis and to generate more realistic mocks with an specific survey geometry, but this is beyond the scope of this work. Another similar code to L-PICOLA in the literature is ICE-COLA. It has already been used to produce galaxy [27] and weak lensing [28] mock catalogs, and it could be a good alternative to L-PICOLA, but for the moment ICE-COLA is not publicly available.

2.1 Simulations and flux transformation

To obtain the dark matter density field needed to generate a Lyman-α flux field, we used L-PICOLA (described in a previous section) to simulate boxes of 1 Gpch^{-1} cubed with 1024^3 particles and a mesh partition of the same size. This configuration gives a mean interparticle distance of $\sim$ 1Mpch^{-1}, LOOK UP THE PIXEL RESOLUTION FOR DESI. which should allow us to have enough small scale resolution to study the effects of non-linear evolution on the BAO feature of the Lyman-α flux auto-correlation function.

The initial conditions were generated internally by the code using a modified version of 2LPTIC, which takes as input a linear matter power spectrum at z=0 that was obtained from CLASS [29]. The simulations were started from z=63 with outputs at z=3 and z=2.5 taking 30 times steps between each snapshot. We tested that the number of steps taken between snapshots were not to bias our statistics by computing the power spectrum and correlation function of a set of simulations with different number of steps between outputs, ranging from 10 to 60 steps. We found a difference of less than 0.01% between the presented configuration from ones with more steps at the relevant scales for our analysis.

To transition from the dark matter only simulations to the Lyman-α flux, we used the fluctuating Gunn-Peterson approximation (FGPA) and the relation between the temperature of the gas and the density, $T = \Delta^\gamma$, to relate the optical depth of the Lyman-α transition to the dark matter density field as

$$\tau_{\mathrm{Ly}\alpha}(z) = A(z)\Delta^\beta, \tag{2.7}$$

where Δ is the mean normalized mesh density of the dark matter field; $\beta = 2-\gamma$ is a parameter that arise from the relation between temperature and density and the normalization factor is:

$$A(z) = 1.41\frac{(1+z)^6(\Omega_{0,b}h^2)^2}{T^{0.7}hE(z)\Gamma_{-12}(z)}, \tag{2.8}$$

where h is the Hubble parameter; $\Omega_{0,b}$ is the baryon density parameter; $E(z) = \sqrt{\Omega_{0,m}(1+z)^3 + \Omega_{DE}}$, with Ω_{DE} as the dark energy density; Γ_{-12} is the photo-ionization rate of neutral hydrogen and $T^{0.7}$ is a reference temperature for the gas.

The transmitted flux fraction overdensity is then calculated as:

$$\delta_F(\mathbf{x}) = \frac{F(\mathbf{x})}{\bar{F}(\mathbf{x})} - 1, \tag{2.9}$$

where

$$F(\mathbf{x}) = e^{-\tau_{\mathrm{Ly}\alpha}}, \tag{2.10}$$

is the transmitted flux fraction and $\bar{F}(\mathbf{x})$ the mean flux. For this work $\gamma = 1.6$ and the photo-ionization rate and the reference temperature were fixed by imposing that equation 2.7 matched a mean flux of $\bar{F} = 0.8$ [30], [22].

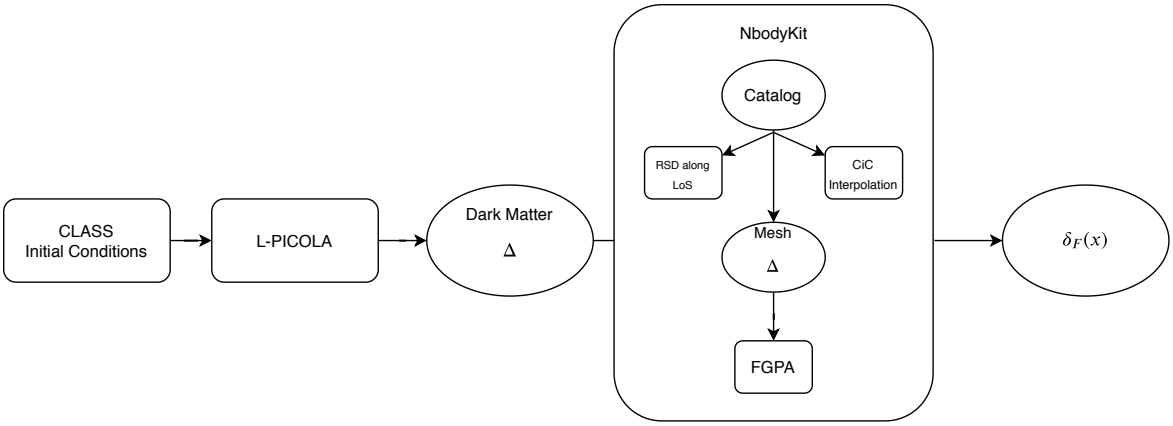

Figure 1. Flowchart of the mock generation and analysis pipeline. The rounded squares represents codes or scripts and the ovals represent outputs and/or inputs. The big square that have substructure represents

It is important to notice that the mocks do not include Redshift Space Distortions (RSD) which are needed to study the broadening of BAO. We included RSD by shifting each particle by its peculiar velocity along the z-axis of the simulation box (the direction of our line of sight). To avoid low absorption zones, that could hinder the statistics of interest, the density field is smoothed via a Cloud-in-Cell (CiC) interpolation. By construction this mocks have 1024^2 parallel LoS, contrary to the way QSO catalogs are structured. This means that it would be very difficult to use these simulations where a realistic QSO sky distribution is required. This is not of great importance in this work, as with this set of simulations we aim to measure the global effect of non-linear evolution of the density field in the auto-correlation and not to explore if this is possible for a particular survey.

In Figure 1 we show the pipeline used to generate the Lyman-α mocks, staring from DM only simulation boxes from L-PICOLA, up to generating the transmitted flux fraction over density, $\delta_F(\mathbf{x})$. All the steps seen after obtaining the DM density were performed using the `NbodyKit` [31] Python package. The same code was used to compute the correlation function from the mocks in later sections.

3 Measuring the correlation from Mocks

To calculate the flux auto-correlation the mocks, $\xi(r,\mu) = \langle \delta_F \delta'_F \rangle$, we used the `FFTCorr` function of `Nbodykit`. This function transform an input mesh object to Fourier space to calculate the power spectrum, $P_F(k,\mu) = \langle \tilde{\delta}_F \tilde{\delta}'_F \rangle$. Once it calculates the power spectrum, an inverse Fourier transform in applied to get the corresponding correlation function. The measured correlation from our simulations consist of 22 spatial bins ranging from 50 to 160 $\mathrm{Mpc}\,h^{-1}$ with a bin size of 5 $\mathrm{Mpc}\,h^{-1}$ over four averaged angular bins. The periodic boundary conditions of the simulations were maintained when calculating the correlation. In Figure 2 we show the sample mean and standard error of the correlation function calculated from the mocks, as well as the best fit for a linear and non-linear model.

Next we calculated the sample covariance matrix of our mocks using the standard unbiased estimator:

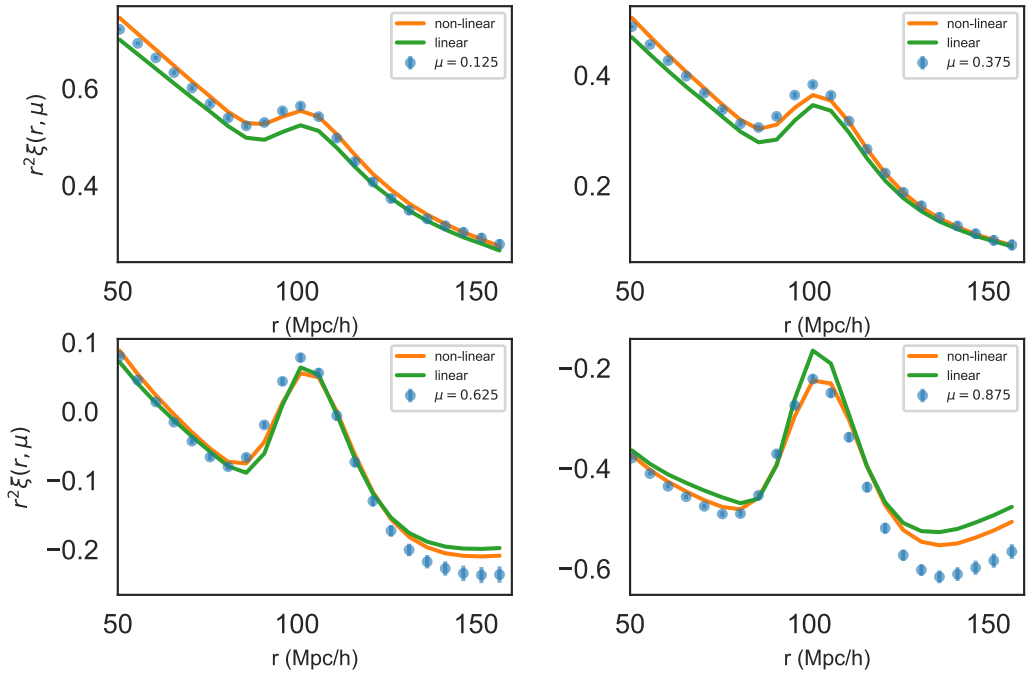

Figure 2. Flux auto-correlation function obtained form a set of 200 mocks (blue dots). The orange line represent the best fit for a non-linear model given as a combination of equation 4.1 and equation 4.3, whilst the green line represents the best fit for a linear model given only by equation 4.1.

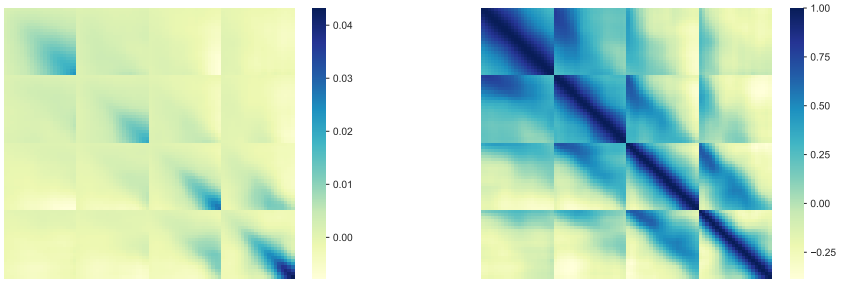

Figure 3. Figure (a) shows the covariance matrix of a sample of 200 mock simulations. Figure (b) show the corresponding correlation matrix of the same sample.

$$C_{ij} = \frac{1}{N-1}\sum_{k=1}^{N}(\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j), \tag{3.1}$$

where N is the total number of samples, i, j represent the spatial bins of the correlation function and $\bar{\xi}_i$ is the mean correlation at the i-th bin.

By construction, the covariance matrix is composed of 88 spatial bins splited into 4

angular bins, each with 22 spatial bins. In the left side of Figure 3 we show the corresponding covariance matrix to our sample of 200 mocks where each of the sub-divisions of the matrix represents one of the wedges. We notice that the matrix is fairly diagonal with some noisy far off-diagonal elements. To have a better understanding of the noise levels we calculated the correlation matrix as:

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \tag{3.2}$$

From the right panel of Figure 3 we observe that we have some correlation between close by wedges, as to be expected, with less correlation from far separated wedges. Having non-zero correlation from widely separated wedges could be problematic as it could bias the parameter inference.

One way to eliminate unwanted correlated data is to smooth the correlation matrix. A tested method is the one used by BOSS [32]. It consists on averaging all elements of the correlation matrix that have the same separation difference $\Delta r_\parallel = r_\parallel - r'_\parallel$ and $\Delta r_\perp = r_\perp - r'_\perp$, and setting to zero those elements whose average is below a set threshold. While this method works well when the correlation is represented in $(r_\parallel, r_\perp)$ coordinates, for the wedge representation the low number of angular bins, which translate to a low number of separations with a large correlation variance, generates a improper smoothed covariance matrix matrix. Another possible way to perform an smoothing is via a Tapering matrix [33]. The methods uses a tapering function of the spatial separation to compute an isotropic covariance matrix that when applied to the precision matrix, smooths the entries far from the diagonal. We performed our statistical analysis using the noise covariance matrix and a smooth the covariance matrix via tapering and found that results of our statistical analysis are very similar in both cases (refer to Appendix **??** for a detailed explanation of the Tapering matrix).

4 Modeling $\xi(r, \mu)$

To model the Lyman-α flux auto-correlation function we used the Kaiser power spectrum for a biased tracer:

$$P_F(k, \mu) = b^2(1 + \beta\mu^2)^2 P_L(k), \tag{4.1}$$

where b is the bias associated to the relation between the DM density fluctuations and the flux fluctuations, δ_F; β is the redshift space distortion parameter; μ is the wedge in which we calculate the power spectrum, and $P_L(k)$ is the linear power spectrum.

Previous measurements of galaxy clustering have found that the description given by equation 4.1 wrongly estimates the amplitude of the BAO signal in the correlation function. This is attributed to a large scale clustering bias caused by tracer pairs shifting from their original position due to their peculiar motions, causing a degradation (seen as a dampening) of the signal at the BAO scale[19], [34],[35]. Such peculiar motions can be described as a Lagrangian displacement field, which in turn can be parametrized up to linear order as:

$$\Sigma(\mu)^2 = \mu^2\Sigma_\parallel^2 + (1 - \mu^2)\Sigma_\perp^2, \tag{4.2}$$

where $\Sigma_\perp = \Sigma_0 G$ and $\Sigma_\parallel = \Sigma_\perp(1 + f)$ describe the displacements parallel and perpendicular to the line of sight respectively. $\Sigma_0 = 12.4\sigma_8/0.9$ Mpch^{-1}; G is the growth function and $f \approx \Omega_m^{0.55}$.

To introduce the effect of the dampening in the model we followed the methodology of [36], where the linear power spectrum in equation 4.1 is separated into a broadband component, $P_{\rm smooth}(k)$ and a component with only the contribution of the BAO, $P_{\rm BAO}(k)$ to which an exponential dampening factor is applied such that:

$$P_{\rm NL}(k,\mu) = P_{\rm smooth}(k) + e^{-k^2\Sigma^2(\mu)/2} P_{\rm BAO}(k). \tag{4.3}$$

Following [36], we also account for a shift of the position of the BAO peak by considering scaling parameters applied to the BAO contribution of the correlation function, such that:

$$\xi(r_\parallel, r_\perp) = \xi_{\rm smooth}(r_\parallel, r_\perp) + \xi_{\rm BAO}(\alpha_\parallel r_\parallel, \alpha_\perp r_\perp), \tag{4.4}$$

with

$$\alpha_\parallel = \frac{[r_d H(z)]_{\rm fid}}{[r_d H(z)]},\ \alpha_\perp = \frac{[r_d D_A(z)]_{\rm fid}}{[r_d D_A(z)]}, \tag{4.5}$$

where r_d is the sound horizon scale; $D_A(z)$ is the angular distance function, and the "fid" label stands for the fiducial cosmology presented in section 1.

To compute the correlation function, the linear power spectrum from equation 4.1 was substituted by equation 4.3 and a Fourier transform was applied to obtain the corresponding anisotropic correlation function:

$$\xi(r_\perp, r_\parallel) = \frac{1}{\sqrt{2\pi r_\perp}} \int_0^\infty \left[\sqrt{k_\parallel} P_X(r_\perp, k_\parallel)\right] J_{-1/2}(r_\parallel k_\parallel) r_\parallel dk_\parallel, \tag{4.6}$$

where

$$P_X(r_\perp, k_\parallel) = \frac{1}{2\pi} \int_0^\infty \left[k \perp P_F(k_\perp, k_\parallel)\right] J_0(k_\perp r_\perp) dk_\perp, \tag{4.7}$$

is the cross-power spectrum, which measures the correlation between modes along and transverse to the line of sight; J_α are the Bessel functions of the first kind. To obtain the above integrals we used the relation between Fourier and Hankel transforms and solve them by using mcfit, a python port of the FFTLog [37]. Note that the transform was applied to the "smooth" and "BAO" components of the flux power spectrum separately.

5 Results

5.1 Fitting the data

We used a Markov-Chain Monte Carlo (MCMC) method to find the posterior probability distribution for the free parameters of our model. The likelihood function is given by:

$$\mathrm{Ln}(\mathcal{L}) \propto (\xi_{\rm model}(\mathrm{r},\mu,\Theta) - \xi_{\rm meas}(\mathrm{r},\mu))^{\mathrm{T}} \mathrm{C}^{-1} (\xi_{\rm model}(\mathrm{r},\mu,\Theta) - \xi_{\rm meas}(\mathrm{r},\mu)), \tag{5.1}$$

where $\xi_{\rm meas}(r,\mu)$ represent the average correlation measured from the mocks; $\xi_{\rm model}(r,\mu,\Theta)$ is our model given by equation 4.6 (or eq. 4.1), where Θ is a vector whose elements are the parameters: bias b, redshift distortion parameter β and $\Sigma_\perp$ (for the non-linear model); C is the sample covariance matrix.

To implement the MCMC we choose the multimodal nested sampler MULTINEST [38]. The nested sampler is not only a method to calculate the posterior probability distribution of the model parameters, but it is also a tool to efficiently calculate the bayesian evidence, which

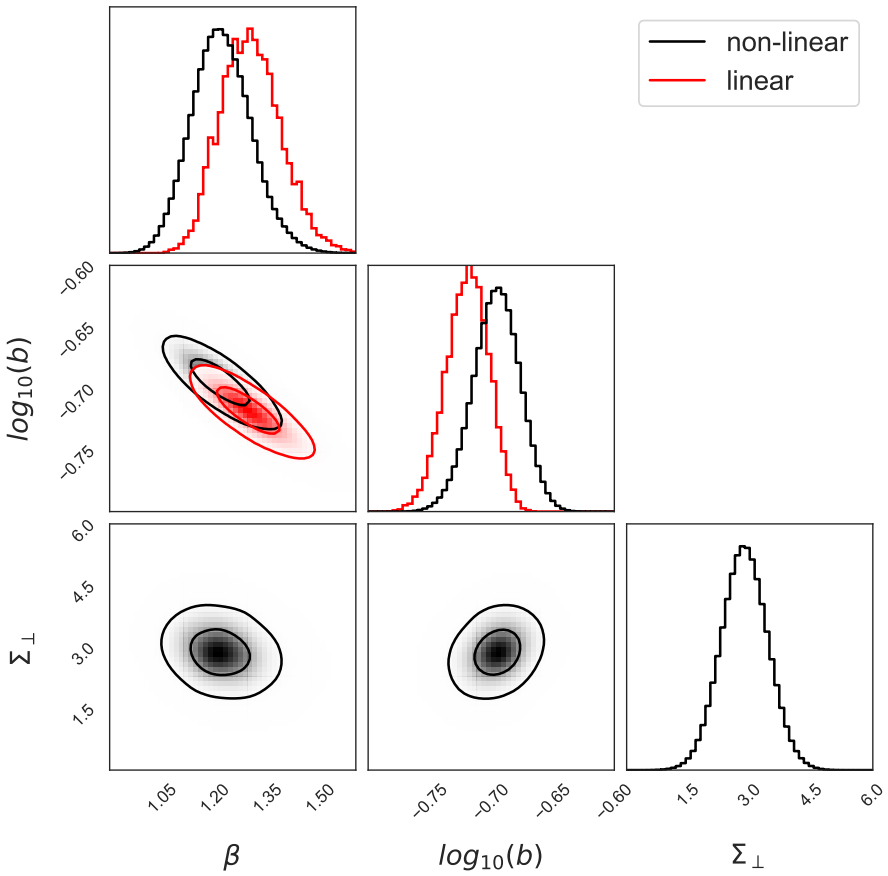

Figure 4. Posterior probability distribution obtained with MULTINEST run with 2000 live points and flat priors. The distributions in black are from the non-linear model, while the red distributions are from the lineal model.

is why MULTINEST is very useful for model selection. We configured the MCMC to be run with 2000 "live" points and used uninformative priors for the free parameters with ranges: $0 < \beta < 5$, $-3 < \log_{10}(b) < 1$ and $0 < \Sigma_\perp/\mathrm{Mpc}\, h^{-1} < 10$. In Figure 4 we show the posteriors for the non-linear model (black line) and a linear model (red line). In Table 5.1 we show the median values from the posteriors and the 2σ confidence interval for both models, as well as the corresponding value of the Akaike Information Criterion, $\mathrm{AIC} = 2k - 2\ln(\hat{\mathcal{L}})$, where k is the number of free parameters in the model and $\hat{\mathcal{L}}$ is the maximum Likelihood. We choose the AIC as an estimator for a preferred model as it penalizes models with higher number of parameters, and since the priors for the parameters are uninformative, the maximum Likelihood is the best estimator of the best fit model (set of parameters).

We notice that the value we report for $\Sigma_\perp$ is lower than the expected value of $\Sigma_\perp = 3.14$ from the cosmology, but still contained at the 2σ level of our distribution. The deviation could be attributed to our calibration of the FGPA formulation or to the non-linear relation between the optical depth and the transmitted flux fraction. To test this we remade our analysis using the unbiased Dark Matter density field. From Figure 6 we notice that median of the posterior probability distribution of $\Sigma_\perp$ matches very well with the predicted value from the cosmology

Model	β	$\log_{10}(b)$	$\Sigma_\perp$	AIC
Linear	1.30 ± 0.17	-0.72 $\pm$ 0.03		34.1
Non-linear	1.22 ± 0.16	-0.69 ± 0.03	2.86 ± 1.0	19.8
Matter Test	1.03 ± 0.2	-0.04 ± 0.05	3.0 ± 1.09	12.8

Table 1. Mean values obtained from the inference for different models. The error reported for all the parameters are at 95% confidence. The AIC values were calculated using the output nested sampler log-evidence.

(red line). This indicates that the transformation from the matter density field to the flux over-densities causes a shift in the distribution of the broadening parameter, but this doesn't mean that our results are wrong as the goal of this paper is to test if such effect are observed in the Lyman-α flux auto-correlation and not to report what would be the measured value for an specific survey.

In Figure 2 we compare the best fits of the linear (green line) and non-linear (orange lines) models. It is observed that the linear model misestimate the measured correlation (blue dots). From Figure 5 we can observe that absolute difference of the linear model respect to the measured correlation is always higher that the non-linear model, more noticeably around the BAO peak ($\sim 100\,\mathrm{Mpc}\,h^{-1}$), suggesting that adding the exponential dampening factor is preferred. This is also backed by the values of the AIC, with a smaller value for the non-linear model, even after penalizing the model for having an extra parameter.

6 Conclusions

In this work we presented a method to generate mocks of the Lyman-α flux using COLA simulations of the dark matter density field. With this method we account for the non-linear evolution of the density field, something that in a Gaussian random field realization has to be tuned in, while compared to to an N-body simulation we are significantly less constrained by the computation resources needed to efficiently generate the large number of simulation required to obtain a good covariance matrix and study the non-linear broadening of the BAO in the Lyman-α flux auto correlation function.

We found that there is a noticeable difference, as seen in Figure 2, between the linear Kaiser model of the two-point correlation function and the same model with an extra parameter accounting for the BAO broadening, with the latter seemingly the preferred model. This is supported by the Bayesian evidence obtained from the parameter inference and the difference between their AIC values. As stated throughout this paper, the presented mocks were not by any means intended to reproduce measurements of the two-point correlation function from surveys such as BOSS or DESI, thus the inferred distribution of $\Sigma_\perp$ from this work suggest that similar analysis should be performed in future measurements as a contrast to fixing the parameter to a fiducial value.

Finally, in the future this mocks can potentially be adapted to include redshift evolution of the flux over-density, δ_F, and to reproduce a given survey geometry and volume by means of lightcone simulations. Also the mocks can be used to study higher order clustering statistics such as the biepectrum and three-point correlation function.

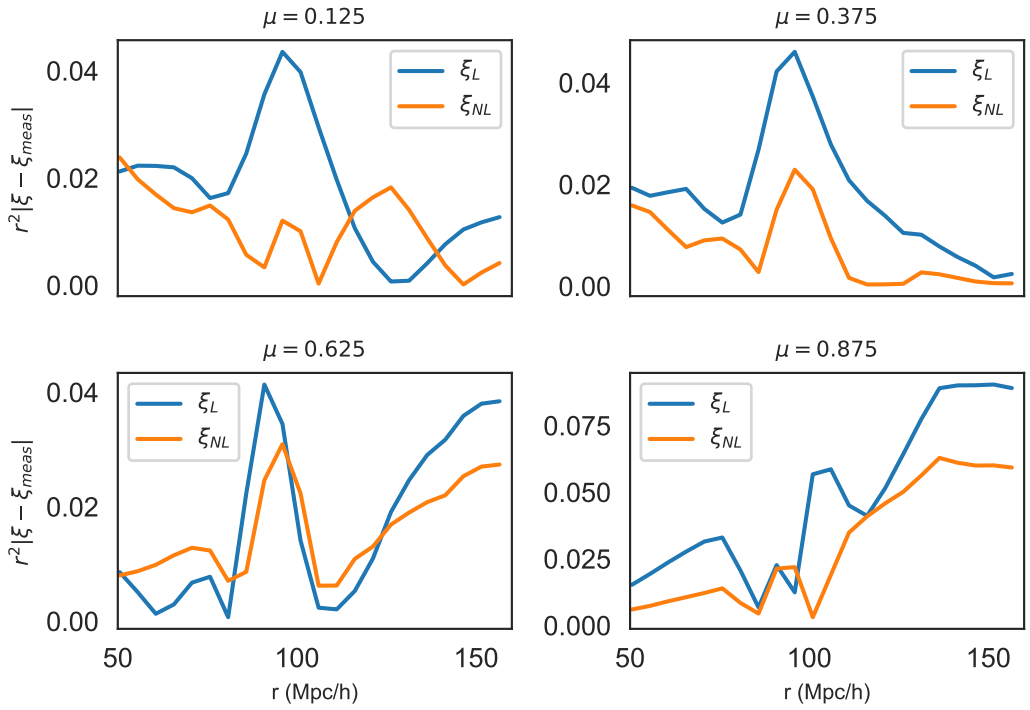

Figure 5. Absolute difference between the models with respect to the measured correlation from the mocks.

A Matter Test

To test that the MCMC that we are using to find the best fit values for the parameters is finding the correct posteriors we make use of the dark matter density field of our mocks to check if we can extract the unbiased distributions. We ran the MCMC with the same conditions as the previous cases: the likelihood function is given by equation 5.1, were the covariance matrix was recomputed using 200 DM only simulation boxes; we used the same uninformative priors for the three parameters and the same number of steps as in the analysis presented in section 5.

In Fig. 6 we show the posterior distributions found applying our analysis on the DM field only; the mean values of the distributions and is errors at a 95% confidence are presented in Table 5.1. Given that this field is unbiased, we expect to find a distribution for the bias parameter centered on one (zero in our case due to the logarithmic scale) but we find that the distribution is shifted almost 2σ to the left. Now, if we examine the distribution of β we see that it is shifted, as expected, by the slight degeneration between b and β; the red line represent the expected value computed using the relation $\beta \approx \Omega_m^{0.55}/b$ where we fixed the value of b to one (the unbiased case), while if we use the mean value of the bias distribution we find that $\beta = 1.06$, at the center of the distribution. Last, the distribution found for $\Sigma_\perp$ is very close to what we find in the analysis of the Ly-α flux field which is a very good sign that our findings are correct. The difference between the mean of the distribution and the value predicted by [19] is less than 1σ apart.

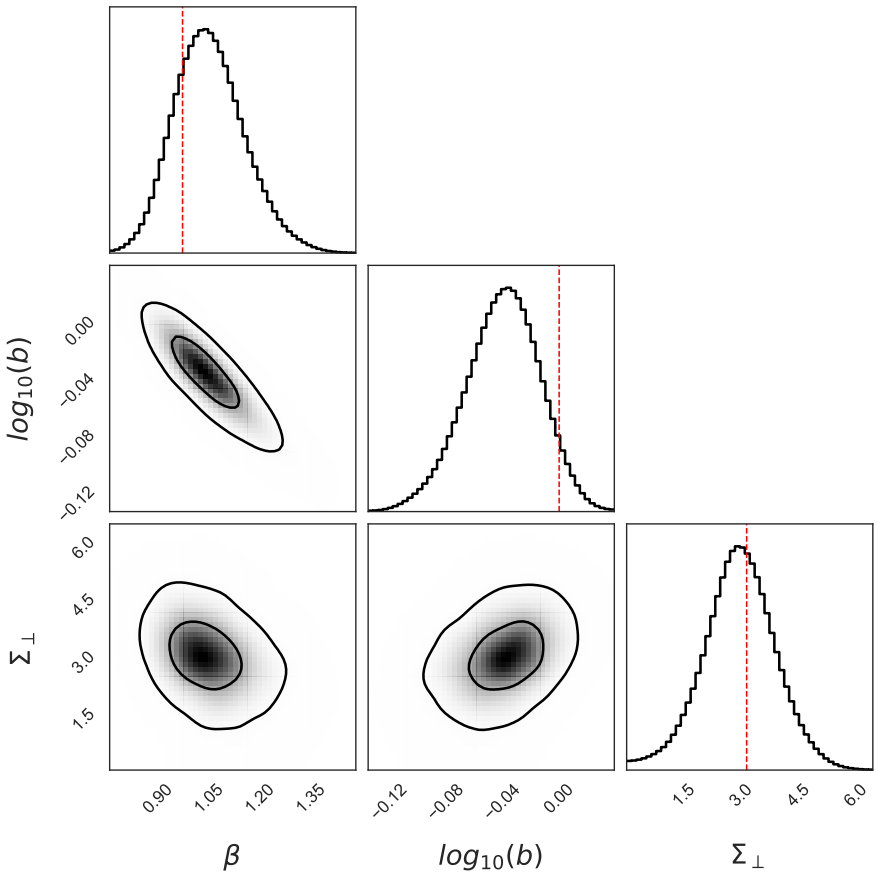

Figure 6. Posterior distribution for the free parameters using only the DM density field. This distribution where found using a MCMC with 200 sample mocks. The red lines show the "true" values of the parameter calculated from the theory.

B Tapering matrix

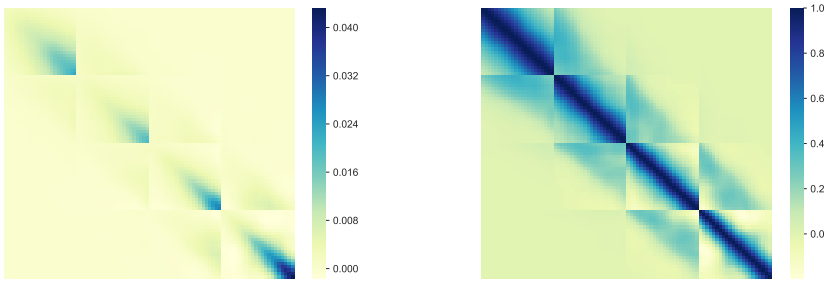

Figure 7. Tapering resutls. Figure (a) shows the covariance matrix of a sample of 200 mock simulations. Figure (b) show the corresponding correlation matrix of the same sample.

C Comoving Lagrangian Acceleration method

The main goal of a N-body code is to solve the equation of motion of the dark matter particles:

$$\frac{d^2x}{d\tau^2} + H(\tau)\frac{dx}{d\tau} + \nabla\phi = 0, \tag{C.1}$$

where τ is the conformal time; $H(\tau)$ is the conformal Hubble parameter; ϕ is the gravitational potential, and x is the particle position that is related to a Lagrangian comoving coordinate q by:

$$x(\tau) = q + \Psi(q, \tau), \tag{C.2}$$

where $\Psi(q, \tau)$ is the Lagrangian displacement field. Taking the divergence of equation C.2 and using the Poisson equation we get:

$$\nabla\left(\frac{d^2x}{d\tau^2} + H(\tau)\frac{dx}{d\tau}\right) = -\frac{3}{2}\Omega_{0,m}H(\tau)\delta(x), \tag{C.3}$$

where $\delta(x)$ is the density contrast in the Eulerian frame.

Using the Jacobian of the transformation $J = (\delta_{ij} + \Psi_{i,j})^{-1}$ and the fact that the density is the same in either of the frames, the equation of motion can be written as:

$$\nabla\left(\frac{d^2\Psi_{i,j}}{d\tau^2} + H(\tau)\frac{d\Psi_{i,j}}{d\tau}\right) = -\frac{3}{2}\Omega_{0,m}H(\tau)\delta(q), \tag{C.4}$$

where ∇ is now in the Lagrangian coordinates as well as the density contrast. $\Psi_{i,j} = \frac{\partial\Psi_i}{\partial x_j}$. We solve the previous equation perturbately such that

$$\Psi = \Psi^{(1)} + \Psi^{(2)} + ... \tag{C.5}$$

The solution of equation C.4, at linear order, is:

$$\nabla\cdot\Psi^{(1)} = -D_1(\tau)\delta_1(q), \tag{C.6}$$

where we have decoupled the the dynamic variable into $D_1(\tau)$, the linear growth factor.

At second order the procedure is the same but introduces corrections to the linear case such that:

$$\nabla\cdot\Psi^{(2)} = \frac{1}{2}D_2(\tau)\sum_{i\neq j}(\Psi^{(1)}_{i,i}\Psi^{(1)}_{j,j} - \Psi^{(1)}_{i,j}\Psi^{(1)}_{j,i}), \tag{C.7}$$

where in a flat ΛCDM universe $D_2(\tau) \approx -3D_1^2(\tau)\Omega_{0,m}^{-1/143}/7$.

The Comoving Lagrangian Acceleration (COLA) method developed by Tassev et al. (2013), makes use of the solution of 2LPT to decouple the small and large scales of an N-body code. Because at large scales (larger than 100 Mpc/h) the N-body codes solve for linear theory only, we can use the exact solutions for the linear growth factor given by perturbation theory. With this, we can rewrite the equation of motions for the dark matter particles as

$$\partial_\tau^2\Psi_{res} + \partial_\tau^2 D_1(\tau)\Psi^{(1)} + \partial_\tau^2 D_2(\tau)\Psi^{(2)} + \nabla\phi = 0, \tag{C.8}$$

where

$$\partial_\tau^2 X = \frac{d^2 X}{d\tau^2} + H(\tau)\frac{dX}{d\tau} \tag{C.9}$$

and Ψ_{res} is the resultant from removing the perturbative solutions of the field from the complete displacement field $\Psi_{res} = \Psi - \Psi^{(1)} - \Psi^{(2)}$. To complete solve the equations of motion we need to input a power spectrum at redshift $z = 0$ to compute the initial displacement field that we can shift back in time using the know solutions from 2LPT and only use the N-body solver for Ψ_{res}

References

[1] Edwin Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proc.Nat.Acad.Sci.*, 15:168–173, 1929.

[2] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038, 1998.

[3] T. Abbott et al. The dark energy survey. 2005.

[4] Alexandra Abate et al. Large Synoptic Survey Telescope: Dark Energy Science Collaboration. 2012.

[5] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi, T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, W. Moos, S. Perlmutter, M. Postman, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, J. Centrella, W. Traub, C. Baltay, J. Colbert, D. Bennett, A. Kiessling, B. Macintosh, J. Merten, M. Mortonson, M. Penny, E. Rozo, D. Savransky, K. Stapelfeldt, Y. Zu, C. Baker, E. Cheng, D. Content, J. Dooley, M. Foote, R. Goullioud, K. Grady, C. Jackson, J. Kruk, M. Levine, M. Melton, C. Peddie, J. Ruffa, and S. Shaklan. Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report. *arXiv e-prints*, page arXiv:1305.5422, May 2013.

[6] Kyle S. Dawson, David J. Schlegel, Christopher P. Ahn, Scott F. Anderson, Éric Aubourg, Stephen Bailey, Robert H. Barkhouser, Julian E. Bautista, Alessand ra Beifiori, Andreas A. Berlind, Vaishali Bhardwaj, Dmitry Bizyaev, Cullen H. Blake, Michael R. Blanton, Michael Blomqvist, Adam S. Bolton, Arnaud Borde, Jo Bovy, W. N. Brandt, Howard Brewington, Jon Brinkmann, Peter J. Brown, Joel R. Brownstein, Kevin Bundy, N. G. Busca, William Carithers, Aurelio R. Carnero, Michael A. Carr, Yanmei Chen, Johan Comparat, Natalia Connolly, Frances Cope, Rupert A. C. Croft, Antonio J. Cuesta, Luiz N. da Costa, James R. A. Davenport, Timothée Delubac, Roland de Putter, Saurav Dhital, Anne Ealet, Garrett L. Ebelke, Daniel J. Eisenstein, S. Escoffier, Xiaohui Fan, N. Filiz Ak, Hayley Finley, Andreu Font-Ribera, R. Génova-Santos, James E. Gunn, Hong Guo, Daryl Haggard, Patrick B. Hall, Jean-Christophe Hamilton, Ben Harris, David W. Harris, Shirley Ho, David W. Hogg, Diana Holder, Klaus Honscheid, Joe Huehnerhoff, Beatrice Jordan, Wendell P. Jordan, Guinevere Kauffmann, Eyal A. Kazin, David Kirkby, Mark A. Klaene, Jean-Paul Kneib, Jean-Marc Le Goff, Khee-Gan Lee, Daniel C. Long, Craig P. Loomis, Britt Lundgren, Robert H. Lupton, Marcio A. G. Maia, Martin Makler, Elena Malanushenko, Viktor Malanushenko, Rachel Mandelbaum, Marc Manera, Claudia Maraston, Daniel Margala, Karen L. Masters, Cameron K. McBride, Patrick McDonald, Ian D. McGreer, Richard G. McMahon, Olga Mena, Jordi Miralda-Escudé, Antonio D. Montero-Dorta, Francesco Montesano, Demitri Muna, Adam D. Myers, Tracy Naugle, Robert C. Nichol, Pasquier Noterdaeme, Sebastián E. Nuza, Matthew D. Olmstead, Audrey Oravetz, Daniel J. Oravetz, Russell Owen, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, John K. Parejko, Isabelle Pâris, Will J. Percival, Ismael Pérez-Fournon, Ignasi Pérez-Ràfols, Patrick Petitjean, Robert Pfaffenberger, Janine Pforr, Matthew M. Pieri, Francisco Prada, Adrian M. Price-Whelan, M. Jordan Raddick, Rafael Rebolo, James Rich, Gordon T. Richards, Constance M. Rockosi,

Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, J. A. Rubiño-Martin, Lado Samushia, Ariel G. Sánchez, Conor Sayres, Sarah J. Schmidt, Donald P. Schneider, C. G. Scóccola, Hee-Jong Seo, Alaina Shelden, Erin Sheldon, Yue Shen, Yiping Shu, Anže Slosar, Stephen A. Smee, Stephanie A. Snedden, Fritz Stauffer, Oliver Steele, Michael A. Strauss, Alina Streblyanska, Nao Suzuki, Molly E. C. Swanson, Tomer Tal, Masayuki Tanaka, Daniel Thomas, Jeremy L. Tinker, Rita Tojeiro, Christy A. Tremonti, M. Vargas Magaña, Licia Verde, Matteo Viel, David A. Wake, Mike Watson, Benjamin A. Weaver, David H. Weinberg, Benjamin J. Weiner, Andrew A. West, Martin White, W. M. Wood-Vasey, Christophe Yeche, Idit Zehavi, Gong-Bo Zhao, and Zheng Zheng. The Baryon Oscillation Spectroscopic Survey of SDSS-III. , 145(1):10, January 2013.

[7] D. Heath Jones, Mike A. Read, Will Saunders, Matthew Colless, Tom Jarrett, Quentin A. Parker, Anthony P. Fairall, Thomas Mauch, Elaine M. Sadler, Fred G. Watson, Donna Burton, Lachlan A. Campbell, Paul Cass, Scott M. Croom, John Dawe, Kristin Fiegert, Leela Frankcombe, Malcolm Hartley, John Huchra, Dionne James, Emma Kirby, Ofer Lahav, John Lucey, Gary A. Mamon, Lesa Moore, Bruce A. Peterson, Sayuri Prior, Dominique Proust, Ken Russell, Vicky Safouris, Ken-Ichi Wakamatsu, Eduard Westra, and Mary Williams. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. , 399(2):683–698, October 2009.

[8] Amir Aghamousa et al. The DESI Experiment Part I: Science,Targeting, and Survey Design. 2016.

[9] Michael E. Levi et al. The Dark Energy Spectroscopic Instrument (DESI). 2019.

[10] A. D. Miller, R. Caldwell, M. J. Devlin, W. B. Dorwart, T. Herbig, M. R. Nolta, L. A. Page, J. Puchalla, E. Torbet, and H. T. Tran. A measurement of the angular power spectrum of the cmb from l = 100 to 400. *Astrophys. J.*, 524:L1–L4, 1999.

[11] P. de Bernardis et al. A Flat universe from high resolution maps of the cosmic microwave background radiation. *Nature*, 404:955–959, 2000.

[12] Shaun Cole et al. The 2dF Galaxy Redshift Survey: Power-spectrum analysis of the final dataset and cosmological implications. *Mon. Not. Roy. Astron. Soc.*, 362:505–534, 2005.

[13] Daniel J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophys. J.*, 633:560–574, 2005.

[14] Metin Ata et al. The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: first measurement of baryon acoustic oscillations between redshift 0.8 and 2.2. *Mon. Not. Roy. Astron. Soc.*, 473(4):4773–4794, 2018.

[15] Patrick McDonald and Jordi Miralda-Escudé. Measuring the Cosmological Geometry from the Lyα Forest along Parallel Lines of Sight. , 518(1):24–31, June 1999.

[16] Lam Hui, Albert Stebbins, and Scott Burles. A Geometrical Test of the Cosmological Energy Contents Using the Lyα Forest. , 511(1):L5–L8, January 1999.

[17] Anze Slosar et al. The Lyman-alpha forest in three dimensions: measurements of large scale flux correlations from BOSS 1st-year data. *JCAP*, 1109:001, 2011.

[18] Anse Slosar et al. Measurement of Baryon Acoustic Oscillations in the Lyman-alpha Forest Fluctuations in BOSS Data Release 9. *JCAP*, 1304:026, 2013.

[19] Daniel J. Eisenstein, Hee-jong Seo, and Martin J. White. On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter. *Astrophys. J.*, 664:660–674, 2007.

[20] David Kirkby et al. Fitting Methods for Baryon Acoustic Oscillations in the Lyman-$\alpha Forest Fluctuations in BOSS Data Release 9. JCAP, 1303 : 024, 2013.$

[21] Svetlin Tassev, Matias Zaldarriaga, and Daniel Eisenstein. Solving Large Scale Structure in Ten Easy Steps with COLA. *JCAP*, 06:036, 2013.

[22] Andreu Font-Ribera, Patrick McDonald, and Jordi Miralda-Escudé. Generating mock data sets for large-scale Lyman-α forest correlation measurements. , 2012(1):001, Jan 2012.

[23] Julian E. Bautista et al. Mock Quasar-Lyman-α forest data-sets for the SDSS-III Baryon Oscillation Spectroscopic Survey. *JCAP*, 1505(05):060, 2015.

[24] Aniket Agrawal, Ryu Makiya, Chi-Ting Chiang, Donghui Jeong, Shun Saito, and Eiichiro Komatsu. Generating Log-normal Mock Catalog of Galaxies in Redshift Space. *JCAP*, 1710(10):003, 2017.

[25] James Farr et al. LyaCoLoRe: Synthetic Datasets for Current and Future Lyman-α Forest BAO Surveys. 2019.

[26] Cullan Howlett, Marc Manera, and Will J. Percival. L-PICOLA: A parallel code for fast dark matter simulation. *Astron. Comput.*, 12:109–126, 2015.

[27] Albert Izard, Martin Crocce, and Pablo Fosalba. ICE-COLA: Towards fast and accurate synthetic galaxy catalogues optimizing a quasi N-body method. *Mon. Not. Roy. Astron. Soc.*, 459(3):2327–2341, 2016.

[28] Albert Izard, Pablo Fosalba, and Martin Crocce. ICE-COLA: fast simulations for weak lensing observables. *Mon. Not. Roy. Astron. Soc.*, 473(3):3051–3061, 2018.

[29] Diego Blas, Julien Lesgourgues, and Thomas Tram. The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes. *JCAP*, 1107:034, 2011.

[30] Patrick McDonald. Toward a measurement of the cosmological geometry at Z 2: predicting lyman-alpha forest correlation in three dimensions, and the potential of future data sets. *Astrophys. J.*, 585:34–51, 2003.

[31] Nick Hand, Yu Feng, Florian Beutler, Yin Li, Chirag Modi, Uros Seljak, and Zachary Slepian. nbodykit: an open-source, massively parallel toolkit for large-scale structure. *Astron. J.*, 156(4):160, 2018.

[32] TimothÃľe Delubac et al. Baryon acoustic oscillations in the LyÎś forest of BOSS DR11 quasars. *Astron. Astrophys.*, 574:A59, 2015.

[33] Dante J. Paz and Ariel G. Sanchez. Improving the precision matrix for precision cosmology. *Mon. Not. Roy. Astron. Soc.*, 454(4):4326–4334, 2015.

[34] H.-J. Seo and D. J. Eisenstein. Probing Dark Energy with Baryonic Acoustic Oscillations from Future Large Galaxy Redshift Surveys. , 598:720–740, December 2003.

[35] Hee-Jong Seo and Daniel J. Eisenstein. Improved forecasts for the baryon acoustic oscillations and cosmological distance scale. *Astrophys. J.*, 665:14–24, 2007.

[36] David Kirkby, Daniel Margala, Anže Slosar, Stephen Bailey, Nicolás G. Busca, Timothée Delubac, James Rich, Julian E. Bautista, Michael Blomqvist, Joel R. Brownstein, Bill Carithers, Rupert A. C. Croft, Kyle S. Dawson, Andreu Font-Ribera, Jordi Miralda-Escudé, Adam D. Myers, Robert C. Nichol, Nathalie Palanque-Delabrouille, Isabelle Pâris, Patrick Petitjean, Graziano Rossi, David J. Schlegel, Donald P. Schneider, Matteo Viel, David H. Weinberg, and Christophe Yèche. Fitting methods for baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9. , 2013(3):024, Mar 2013.

[37] A. J. S. Hamilton. Uncorrelated modes of the nonlinear power spectrum. *Mon. Not. Roy. Astron. Soc.*, 312:257–284, 2000.

[38] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398:1601–1614, 2009.