Parameterisation of Lyα forest likelihoods

Chris Pedersen[1, *]

[1] *Department of Physics and Astronomy, University College London, Gower Street, London, United Kingdom*

(Dated: June 16, 2020)

In this short internal note we discuss different sets of likelihood parameters we could use in our emulator, and ways in which we could interface our emulator with other cosmological probes. We propose a likelihood parameterisation around the amplitude and slope of the linear power spectrum in velocity units as presented in ref [1], and propose a test to demonstrate that all the cosmologically relevant information in the Lyα forest is contained within these parameters.

I. INTRODUCTION

One of the novelties in the construction of our emulator when compared with previous approaches is the use of different parameter spaces in different parts of the emulator. A motivation for this is to avoid using parameters that are highly degenerate with one another. The Lyα forest is sensitive to the matter power spectrum in a small redshfit ($2 < z < 5$) and k ($0.1 \lessapprox k \lessapprox 10$ Mpc^{-1}) range, and when considering only this regime, the effects of ΛCDMν parameters are strongly degenerate with one another[2]. The effects of changing these parameters on the flux statistics are also often degenerate with changes to the intergalactic medium (IGM). In this redshift range, the universe is also very close to Einstein de-Sitter, so the dependence of the growth factor and the expansion rate are close to indepenent of cosmological model. Previous studies have put constraints on parameters such as H_0, Σ_ν, Ω_m directly from Lyα forest data[3], however these constraints are achieved through the effect of these parameters on the small scale matter power spectrum. So we instead parameterise our emulator directly on the linear power, using a set of parameters that are more orthogonal to one another and closer to the data which should result in more reliable predictions, and more robust constraints.

In addition to using a reduced and more orthogonal set of parameters, this approach also provides a model-independent emulator that can be used in joint constraints of extended models, such as curvature, non-standard $N_{\rm eff}$, and time-varying dark energy without these models needing to be included in the original emulator. First in section II we cover the emulator parameter space in closer detail, and describe how our emulator can be combined with analysis of the CMB. Then in section III we discuss ways to simplify this procedure so that users do not need to run the emulator code and marginalise over the IGM in order to use Lyα forest results.

II. EMULATOR PARAMETER SPACE

We have opted to train our emulator to predict a $P_{\rm 1D}(k_\parallel)$ as a function of the linear matter power spectrum, which we have parametrised in terms of an amplitude and a slope around a small scale pivot:

$$\Delta_p^2(z) = k^3 P(k,z)/2\pi^2|_{k=k_\star} \tag{1}$$

$$n_p(z) = d\mathrm{lnP(k,z)/dln(k)}|_{\mathrm{k=k_\star}} \tag{2}$$

where $P(k,z)$ is the linear CDM + baryon power spectrum, and we set the pivot scale $k_\star = 0.69$ Mpc^{-1}. We also use 4 nuisance parameters to describe the IGM:

- $\langle F \rangle$, the mean flux

---

* chrisp@star.ucl.ac.uk

- σ_T, the temperature at mean density of the IGM
- γ, quantifying how the temperature scales with density
- k_F, the pressure smoothing scale

We train one Gaussian process on all snapshots from all simulations in a given suite to predict a comoving $P_{1\mathrm{D}}(k_{\parallel})$ as a function of these 6 parameters, which we refer to as the emulator parameters. One benefit of this approach is that the emulator can be used to create theoretical predictions for data at any redshift within the range where we have training data, meaning that the snapshot outputs of our training simulations does not need to match the redshift binning of future or past surveys. Training simulations with different snapshot outputs can also in principle be combined to form bolstered training sets.

In Fig 1 we show how our emulator could be combined with CMB likelihoods. At each cosmological model in the Planck chain, a CAMB object is used to calculate the emulator parameters $\Delta_p^2(z)$ and $n_p(z)$ at each redshift where we have data. This results in N emulator calls, where N is the number of data redshift bins. The emulated $P_{1\mathrm{D}}(k_{\parallel})$ is then converted into velocity units using the factor $H(z)/1+z$ calculated from the CAMB model. The log likelihood for this given model is then calculated from the data. NB that for brevity we do not discuss the details of modelling or marginalisation of the IGM in this document, let's for now assume that the log likelihood presented at the end of this flow diagram is after marginalisation over the nuisance parameters.

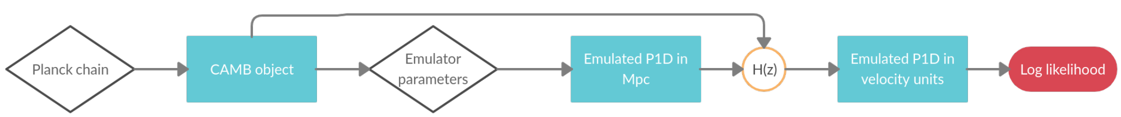

FIG. 1. Procedure for evaluating a Lyα forest likelihood for a given point in the Planck chain. For each cosmological model, a CAMB object is created, and the emulator parameters defined in equations 1 and 2 are calculated, for each redshift z. The $H(z)$ calculated for each model is then used to convert the comoving emulated $P_{1\mathrm{D}}(k_{\parallel})$ to velocity units, and a log likelihood is calculated.

III. COMPRESSED LIKELIHOOD PARAMETERS

The approach shown in Fig 1 requires the user to run the emulator code and marginalise over the IGM. This requires expertise and is computationally slow, and so while this is appropriate for our first tests of the emulator, it is worthwhile to consider more practical ways which would allow a wider set of users to be able to incorporate results Lyα forest results in their constraints. We propose to follow the approach presented in [1], and parameterise our likelihood in terms of the small scale linear power in velocity units:

$$\Delta_\star^2 = q^3 P(q,z)/2\pi^2|_{q=q_\star, z=z_\star} \tag{3}$$

$$n_\star = d\mathrm{lnP(q,z)/dln(k)}|_{\mathrm{q=q_\star, z=z_\star}} \tag{4}$$

where q refers to a wavenumber in velocity units, and we set a pivot redshift of $z_\star = 3$ and $q_\star = 0.009$ s/km.[1] We refer to these parameters as the compressed likelihood parameters. We note here that when using only these two parameters we are fixing the growth rate and expansion history to a fiducial model. One could then present marginalised likelihoods for the parameters, which could be included in packages like MontePython and CosmoMC. In this case the Lyα forest contribution to a joint likelihood would

[1] Note that this pivot scale can be chosen independently to $k_\star$ used in emulator parameter space.

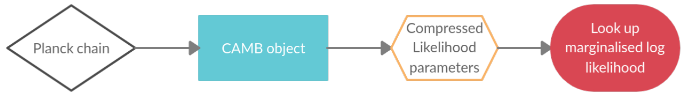

FIG. 2. Once we have calculated marginalised likelihoods for the compressed likelihood parameters, the Lyα forest likelihood for a given cosmological model can simply be looked up.

be calculated using the procedure shown in Fig 2. Here for a given cosmological model, the compressed likelihood parameters are calculated, and the Lyα forest likelihood for this model is looked up in some form of table.

It is important to demonstrate that there is no significant loss of information when compressing the information in the Lyα forest down to the compressed likelihood parameters. The way we propose to show this is through the approach show in Fig 3, which is very similar to the first procedure in Fig 1, except that the emulator calls are made having first gone through the compressed parameters. We intend to compare the combined CMB + Lyα forest posteriors when calculating the Lyα forest likelihood using these two different approaches. If the posteriors are unchanged whether or not the compressed parameters are used, we will be able to conclude that all the cosmologically relevant information in the Lyα forest is contained within the compressed parameters.

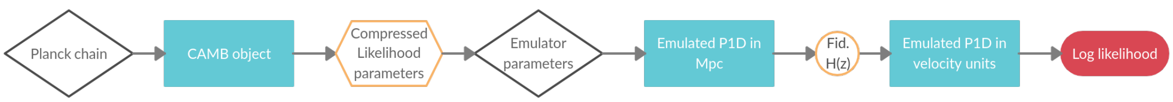

FIG. 3. Here we evaluate each Lyα forest likelihood contribution going through the compressed likelihood parameters. Instead of calculating the emulator calls directly from the cosmological model, the compressed likelihood parameters are calculated first, and these are then used to find the emulator parameters. A fiducial expansion history is then used for the conversion of the emulated $P_{1\mathrm{D}}(k_\parallel)$ to velocity units.

To what extent our approximations of a fixed fiducial expansion history and growth rate are valid is a question we are currently working on. There is a physical argument suggesting we are largely insensitive to these, however we do not know at what threshold the precision in the data and emulator will break this insensitivity. We are testing this by introducing some parameters that add freedom in these two dimensions, which we can go into more detail on in a future discussion. As a final note it is also trivial to add a running of the spectral index to the parameter spaces above, and it is worth discussing whether or not we want to include this in the first paper.

[1] P. McDonald, U. Seljak, R. Cen, D. Shih, D. H. Weinberg, S. Burles, D. P. Schneider, D. J. Schlegel, N. A. Bahcall, J. W. Briggs, J. Brinkmann, M. Fukugita, Ž. Ivezić, S. Kent, and D. E. Vanden Berk, *The Linear Theory Power Spectrum from the Lyα Forest in the Sloan Digital Sky Survey*, *Astrophys. J.* **635** (Dec., 2005) 761–783.

[2] C. Pedersen, A. Font-Ribera, T. D. Kitching, P. McDonald, S. Bird, A. Slosar, K. K. Rogers, and A. Pontzen, *Massive neutrinos and degeneracies in lyman-alpha forest simulations*, *Journal of Cosmology and Astroparticle Physics* **2020** (apr, 2020) 025–025.

[3] C. Yèche, N. Palanque-Delabrouille, J. Baur, and H. du Mas des Bourboux, *Constraints on neutrino masses from Lyman-alpha forest power spectrum with BOSS and XQ-100*, *JCAP* **6** (June, 2017) 047, [`arXiv:1702.0331`].