

Lawrence Berkeley National Laboratory

LBL Publications

Title

Weighted FFT estimators for 1D and 3D correlations of the Lyman- α forest

Permalink

<https://escholarship.org/uc/item/6dp1z39b>

Journal

Journal of Cosmology and Astroparticle Physics, 2025(11)

ISSN

1475-7516

Authors

Lokken, Martine

Font-Ribera, Andreu

McDonald, Patrick

Publication Date

2025-11-01

DOI

10.1088/1475-7516/2025/11/070

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Weighted FFT estimators for 1D and 3D correlations of the Lyman- α forest

Martine Lokken^a, Andreu Font-Ribera^a, Patrick McDonald^b

^aInstitut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, 08193 Bellaterra (Barcelona), Spain

^bLawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

E-mail: mlokken@ifae.es, afont@ifae.es

Abstract. Correlations in the Lyman- α (Ly α) forest, both as a function of line of sight separation (1D) and 3D separation, provide a unique window to the distribution of matter at redshifts not accessible by current galaxy surveys. While optimal quadratic estimators have been used to measure 1D correlations, they are computationally expensive and difficult to extend to 3D analyses. On the other hand, estimators based on the Fast Fourier Transform (FFT) are significantly faster, but are affected by missing data in the spectra (masked pixels) and so far have not used pixel weights to reduce the uncertainties in the measurement. In this publication we describe how to compute the window matrix that enables forward-modelling the impact of masked pixels and weights on the FFT-based estimators. We use Gaussian and hydrodynamical simulations with artificially masked pixels to validate the method on the measurement of 1D correlations. Finally, we show that the formalism can be extended to model the impact on 3D correlations, in particular on the cross-spectrum, the correlation of 1D Fourier modes as a function of transverse separation. This work will enable more precise clustering measurements with the Ly α forest dataset recently collected by the Dark Energy Spectroscopic Instrument (DESI).

Contents

1	Introduction	1
2	Masking in FFT estimators of P_{1D}	3
2.1	FFT measurements without pixel weights	3
2.1.1	Discretizing the skewers	3
2.1.2	Discretizing the Fourier transforms	4
2.1.3	FFT estimators in periodic skewers	4
2.1.4	FFT estimators in finite, non-periodic skewers	6
2.2	Modelling the impact of zero-padding, masking, and pixel weights	8
3	Validation with masked mock spectra from hydrodynamic simulations	11
4	Masking in FFT estimators of $P_{\times}$	14
4.1	Overview	14
4.2	$P_{\times}$ from pairs of quasars with different redshifts	15
5	Conclusion	18

1 Introduction

The three-dimensional power spectrum of the Ly α forest contains ample information about the distribution of matter in the high-redshift universe, enabling state-of-the-art constraints on the sum of the neutrino masses, the properties of dark matter particles, and on the physics of the intergalactic medium (IGM). Observations of Ly α from redshifts $z \sim 2 - 4$ are improving rapidly as DESI spectroscopically surveys a large sample of quasars, providing the opportunity to test our physical models at an early era of structure formation that is not accessible through galaxy surveys.

However, until recently, the Ly α spectra were sparsely distributed on the sky, leading the field to focus on measuring only the large-scale baryon acoustic oscillations in 3D [e.g., 1–6]. Small scales have been probed instead by measuring the 1D correlation of spectral pixels within a given quasar spectrum, called P_{1D} [7–13]. Recent expansions in data have also enabled significant measurements of the cross-spectrum $P_{\times}$ [14], which correlates spectra separated by a given transverse distance, incorporating both small-scale line-of-sight information and a range of small-to-intermediate transverse scales. At the time of publication, the Dark Energy Spectroscopic Instrument (DESI) [15, 16] has already observed four times more Ly α quasars than previous surveys [6]. Given the unprecedented precision possible for P_{1D} and $P_{\times}$ measurements from this new data, accurate systematics control and modeling of these statistics is essential to obtain unbiased results on cosmological and IGM parameters.

The quasar spectra are affected in many ways before summary statistics are measured. Among these, particular wavelengths in the spectra are contaminated by nearby absorption (from Galactic gas and the atmosphere) and emission (from the atmosphere). Once the light arrives to the instrument, the spectrographs’ finite resolution smears the signal, the spectra are read out and stored in discrete pixels, and some pixels end up in an untrustworthy state for a variety of reasons. Before measuring the summary statistic, the spectral-line contaminants

and bad pixels are typically removed through masking. The effects of resolution, pixelization, and masking can be described, in combination, in a window matrix which transforms the original correlations to the measured correlations [see, e.g., 17].

Robust inference with P_{1D} and $P_{\times}$ must be able to either remove mask effects from the measurement or include them in the theory model prior to comparing measurement with model. The former approach is achieved well by optimal quadratic estimators, as described in [9, 18, 19], and used in measurements from the DESI Early Data Release (EDR) [13]. These estimators are able to make an unbiased measurement by maintaining the spectra in their native pixel space, down-weighting noisy pixels and removing the contributions from masked pixels. However, the optimal quadratic estimator suffers in speed due to the computationally-expensive algebraic calculations.

The alternative approach is to use fast Fourier transform (FFT)-based estimators, which perform the power spectra measurements by averaging the square of the amplitude of the discrete 1D Fourier modes computed with an FFT. The gain in speed is especially relevant for the purposes of measuring $P_{\times}$, which scales computationally as a function of the number of quasar pairs ($\mathcal{O}(N_q^2)$, where N_q is the number of quasar spectra) rather than the number of quasars like for P_{1D} ($\mathcal{O}(N_q)$). However, if the FFT of spectra are used directly to calculate P_{1D} or $P_{\times}$ without deconvolving the window function, any masking or weighting of pixels will bias the measurement with respect to the truth, as we will demonstrate in this work.

FFT-based estimators have been used to calculate P_{1D} with Ly α data from the Baryon Oscillation Spectroscopic Survey (BOSS, [10]), from its extension eBOSS [11] and from DESI EDR [12], and to make the first $P_{\times}$ measurement on real data from eBOSS [14]; we will generally refer to the approach in those works as the `picca-fft` estimators as the code is currently implemented in the `picca` code package¹. Previous analyses using `picca-fft` have not included per-pixel weights, nor the computation of a window matrix to model impacts of masking. In these analyses, the impact of masking was approximated as a multiplicative correction that is calibrated using synthetic data. Moreover, in order to minimize the impact of masking in the measurements of $P_{\times}$, the implementation in `picca-fft` is limited to pairs of Ly α spectra where both uniformly cover a particular redshift bin.

It is possible to improve the signal-to-noise from FFT-based measurements, avoiding data cuts and using pixel-based weights, if one can accurately describe the weights' impact on the summary statistics using a window function. For robust inference, either the theory model must be then convolved with the window function, or the data measurement deconvolved. Window function computations have been used extensively in inference of a variety of cosmological data for decades [e.g. 17, 20–23]. In this work, we present the window calculations specifically for FFT estimators of Ly α P_{1D} and $P_{\times}$. We describe estimators that can use all the data, include per-pixel weights, and are configured to enable rapid modeling of the window function.

In section 2, we present the FFT estimators and the impacts on them from weights and masked pixels. We present the model and demonstrate its accuracy for P_{1D} using Gaussian simulations of 1D spectra. In section 3, we further validate the model for P_{1D} by applying various forms of mask to mock spectra from hydrodynamic simulations. In section 4, we discuss the application to estimators of $P_{\times}$ and demonstrate a simple example with Gaussian simulations. We conclude in section 5.

¹<https://github.com/igmhub/picca>

2 Masking in FFT estimators of P_{1D}

We start in section 2.1 by introducing our notation, and reviewing the FFT algorithm to measure P_{1D} in the absence of pixel weights or masking. In section 2.1.4 we use simple Gaussian random fields to show the impact of assuming periodic boundary conditions in the FFT algorithm. In section 2.2, we model the impact of masking in the measured P_{1D} , and we use the same framework to describe the impact of working with non-periodic Ly α skewers.

2.1 FFT measurements without pixel weights

In order to properly model the measured P_{1D} , we need to understand the relationship between the continuous field and the discrete, pixelized measurements. This section describes that relationship, for the case of FFT estimators without pixel weights or masking. Let us start with an infinitely long, continuous field $\delta(x)$ describing the fluctuations in the Ly α forest along a line-of-sight (we will refer to this as a Ly α *skewer*). Assuming translational invariance (ignoring redshift evolution), we can introduce the one-dimensional (1D) correlation function as the covariance of two fluctuations separated by a distance r :

$$\xi_{1D}(r) = \langle \delta(x) \delta(x+r) \rangle . \quad (2.1)$$

Let us now define the (one-dimensional) continuous Fourier Transform (FT) of the Ly α skewer and its inverse as:

$$\delta(k) = \int dx e^{-ikx} \delta(x) \quad , \quad \delta(x) = \int \frac{dk}{2\pi} e^{ikx} \delta(k) . \quad (2.2)$$

Using these definitions, the covariance of two Fourier modes can be written as

$$\langle \delta(k) \delta^*(k') \rangle = 2\pi \delta^D(k - k') P_{1D}(k) , \quad (2.3)$$

where $\delta^D(k)$ is the Dirac delta function and $P_{1D}(k)$ is the one-dimensional power spectrum, the FT of $\xi_{1D}(r)$.

2.1.1 Discretizing the skewers

Let us now assume that all the skewers in our survey have the same length L and the same number of pixels N , centered around positions $x_a = a\Delta x$ with $\Delta x = L/N$. We can define the pixelised fluctuations δ_a as integrals over the continuous fluctuation $\delta(x)$ convolved by a function $R_a(x)$ that takes into account both the pixelisation and the spectral resolution of a given pixel:

$$\delta_a = \int dx R_a(x - x_a) \delta(x) . \quad (2.4)$$

We define R_a with this sign of $x - x_a$, opposite the usual convolution convention, because this is what we expect to have most directly available — a quantification of the contribution to each discrete pixel from the continuous true flux, which we naturally store and plot with the contribution to the pixel from shorter wavelengths on the left, longer on the right.

Using the FT definitions above, and the equivalent for $R_a(x)$, we can write

$$\delta_a = \int \frac{dk}{2\pi} e^{ikx_a} R_a(-k) \delta(k) , \quad (2.5)$$

and the covariance of two pixels can be described as

$$\langle \delta_a \delta_b \rangle = \int \frac{dk}{2\pi} e^{ik(x_a - x_b)} R_a(-k) R_b(k) P_{1D}(k) , \quad (2.6)$$

where we expanded δ_b using the complex conjugate of the integrand of equation (2.5) and used $(R_b(k))^* = R_b(-k)$; both are possible because δ_b and R_b are real.

If the functions $R_a(x)$ are symmetric, then their Fourier transforms are real, and we have $R_a(-k) = R_a(k)$. Moreover, if we also assume that all pixels have the same resolution, we can use a single function $R_a(x) = R(x)$ for all pixels. In this case, equation (2.6) becomes a function of the pixel separation $r_{ab} = |x_b - x_a|$, and it can be expressed as a (smoothed) correlation function: $\langle \delta_a \delta_b \rangle = \xi_{1D}^R(r_{ab})$. The spectroscopic pipeline of DESI [24] provides a *resolution matrix* that is different for each pixel, and not necessarily symmetric. For simplicity in this article, we will ignore this and use a single symmetric $R(x)$ function.

2.1.2 Discretizing the Fourier transforms

The next step towards deriving the FFT estimator is to define the Discrete Fourier Transform (DFT) of the pixelized skewers and its inverse as:

$$\delta_m = \sum_a \delta_a e^{-ik_m x_a} \quad , \quad \delta_a = \frac{1}{N} \sum_m \delta_m e^{ik_m x_a} , \quad (2.7)$$

with $k_m = m\Delta k$, $\Delta k = 2\pi/L$ and $m \in [0, N)$. While the N values of δ_a are real, the N discrete Fourier modes δ_m are complex. However, not all of them are independent, and they satisfy $\delta_{N-m} = \delta_m^*$ ². The covariance of two of the discrete Fourier modes is:

$$\langle \delta_m \delta_n^* \rangle = \sum_a e^{-ik_m x_a} \sum_b e^{ik_n x_b} \langle \delta_a \delta_b \rangle , \quad (2.8)$$

where $a, b \in [0, N)$. At this point we could use $\langle |\delta_m|^2 \rangle$ to define a FFT estimator for P_{1D} . However, equation (2.8) shows that we would need to do N^2 operations in order to compare a given model to the measured power.

2.1.3 FFT estimators in periodic skewers

The situation is considerably simpler when dealing with periodic $\text{Ly}\alpha$ skewers, for instance when working with skewers extracted from hydrodynamic simulations with periodic boundary conditions. This means that we can extend them outside of the original range by defining $\delta_{a+N} = \delta_a$. In this case, the correlation function is also periodic, and because it is also symmetric around zero separation, we have that $\xi_{1D}^R(r_{ab}) = \xi_{1D}^R(L - r_{ab})$. Even though the $\text{Ly}\alpha$ skewers in real observations are not periodic, this is a useful exercise that will allow us to introduce some of the relevant concepts used later on. In section 2.1.4 we will show that assuming periodic conditions on real observations has important consequences that need to be taken into account when comparing measurements and model predictions.

²In order to speed up the computation and save the memory footprint we could use `rf` function of the `numpy.fft` package and work with $N/2 + 1$ independent modes. However, for simplicity in this article we use the `fft` function, the standard `numpy` implementation of the Fast Fourier Transform (FFT) algorithm. Note that this returns the modes with $m \in [-N/2, N/2)$, sorted beginning with $m = 0$ followed by the positive modes, then the negative ones in ascending order, but we can use $\delta_{-m} = \delta_{N-m}$.

With the assumption of periodic skewers, we can define $c = b - a$, $d = c + N$, and rewrite equation (2.8) as:

$$\begin{aligned}
\langle \delta_m \delta_n^* \rangle &= \sum_{a=0}^N e^{-ik_m x_a} \sum_{c=-a}^{N-a} e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle \\
&= \sum_{a=0}^N e^{-ik_m x_a} \left(\sum_{c=0}^{N-a} e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle + \sum_{c=-a}^0 e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle \right) \\
&= \sum_{a=0}^N e^{-ik_m x_a} \left(\sum_{c=0}^{N-a} e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle + \sum_{d=N-a}^N e^{ik_n(x_a+x_d-L)} \langle \delta_a \delta_{a+d-N} \rangle \right) \\
&= \sum_{a=0}^N e^{-ik_m x_a} \left(\sum_{c=0}^{N-a} e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle + \sum_{d=N-a}^N e^{ik_n(x_a+x_d)} \langle \delta_a \delta_{a+d} \rangle \right) \\
&= \sum_{a=0}^N e^{-ik_m x_a} \sum_{c=0}^N e^{ik_n(x_a+x_c)} \langle \delta_a \delta_{a+c} \rangle \\
&= \sum_{a=0}^N e^{-i(k_m-k_n)x_a} \sum_{c=0}^N e^{ik_n x_c} \xi_{1D}^R(x_c) \\
&= N \delta_{mn}^K P_m ,
\end{aligned} \tag{2.9}$$

where we have used the following property of the Kronecker delta function

$$\delta_{mn}^K = \frac{1}{N} \sum_a e^{i(k_n-k_m)x_a} , \tag{2.10}$$

and we have introduced the FFT of the (pixelized) 1D correlation function:

$$P_m = \sum_{c=0}^N e^{-ik_m x_c} \xi_{1D}^R(x_c) . \tag{2.11}$$

In the limit of very long skewers, one can discretize the integral in equation (2.6) and show that P_m (dimensionless) and the actual P_{1D} (with units of length) are related by:

$$\begin{aligned}
P_m &= \sum_c e^{-ik_m x_c} \xi_{1D}^R(x_c) \\
&= \sum_c e^{-ik_m x_c} \int \frac{dk}{2\pi} e^{ik x_c} R^2(k) P_{1D}(k) \\
&\approx \sum_c e^{-ik_m x_c} \frac{\Delta k}{2\pi} \sum_n e^{ik_n x_c} R^2(k_n) P_{1D}(k_n) \\
&= \frac{\Delta k}{2\pi} \sum_n R^2(k_n) P_{1D}(k_n) \sum_c e^{-i(k_m-k_n)x_c} \\
&= \frac{1}{\Delta x} R^2(k_m) P_{1D}(k_m)
\end{aligned} \tag{2.12}$$

Parameter	P_0 [km s ⁻¹]	k_0 [s km ⁻¹]	k_F [s km ⁻¹]	a_{Si}	r_{Si} [km s ⁻¹]
Value	10.0	0.01	0.1	0.05	2271.0

Table 1. Parameters describing the input power of the Gaussian skewers (first three columns), and the contamination by Silicon absorption (last two columns).

Therefore, ignoring instrumental noise, the standard FFT algorithm to estimate $P_{1\text{D}}$ at discrete Fourier mode m is to average over all N_q skewers the square of the amplitude of their δ_m^q values:

$$\hat{P}_{1\text{D},m} \equiv \frac{\Delta x}{N R^2(k_m)} \frac{1}{N_q} \sum_q |\delta_m^q|^2 . \quad (2.13)$$

For the case of periodic Ly α skewers, we can substitute Eq. 2.9 to demonstrate that the expected value of the estimator is unbiased:

$$\langle \hat{P}_{1\text{D},m} \rangle = \frac{\Delta x}{N R^2(k_m)} N P_m = P_{1\text{D}}(k_m) . \quad (2.14)$$

2.1.4 FFT estimators in finite, non-periodic skewers

In order to demonstrate that the FFT estimator in equation (2.13) is unbiased, we had to assume that the skewers were periodic, enabling the rearrangement of indices leading to equation (2.9). Therefore, the estimator is only unbiased in unrealistic scenarios, such as large numerical simulations with very long skewers and periodic boundary conditions. In realistic scenarios with finite, non-periodic skewers, it is biased, and to calculate the expected value of the estimator we must return to equation (2.8) for the δ_m substitution:

$$\begin{aligned} \langle \hat{P}_{1\text{D},m} \rangle &= \frac{\Delta x}{N R^2(k_m)} \langle |\delta_m|^2 \rangle \\ &= \frac{\Delta x}{N R^2(k_m)} \sum_a e^{-ik_m x_a} \sum_b e^{ik_m x_b} \langle \delta_a \delta_b \rangle \\ &= \frac{\Delta x}{N R^2(k_m)} \int \frac{dk}{2\pi} |R(k)|^2 P_{1\text{D}}(k) \sum_a e^{i(k-k_m)x_a} \sum_b e^{-i(k-k_m)x_b} . \end{aligned} \quad (2.15)$$

Here we will use simple synthetic skewers to visualize this bias from non-periodicity. We generate a Gaussian random field with the following input power spectrum:

$$P_{1\text{D}}(k) = P_0 \frac{1 + k/k_0}{1 + (k/k_0)^2} e^{-(k/k_F)^2} , \quad (2.16)$$

with parameter values described in table 1. The scale dependence is a simple attempt to capture the main features of the real $P_{1\text{D}}$: constant power at large scales, slow rise at intermediate scales, followed by a gradual suppression as one goes to smaller scales, before pressure smoothing erases any structure below the Jeans length. As we will see in the next section, the assumption of periodicity will cause a convolution of the true $P_{1\text{D}}$, and therefore any sharp feature will be smoothed. An example of sharp features in $P_{1\text{D}}$ are the oscillations introduced by the contamination by Silicon absorption observed in measurements of $P_{1\text{D}}$ [9]. This contamination is often modeled with a multiplicative correction to the Ly α power, $[1 + a_{\text{Si}}^2 + 2a_{\text{Si}} \cos(kr_{\text{Si}})]$, where a_{Si} sets the amplitude of the oscillations and r_{Si} its frequency. We add this contamination to the Gaussian skewers, using the parameters described in table 1.

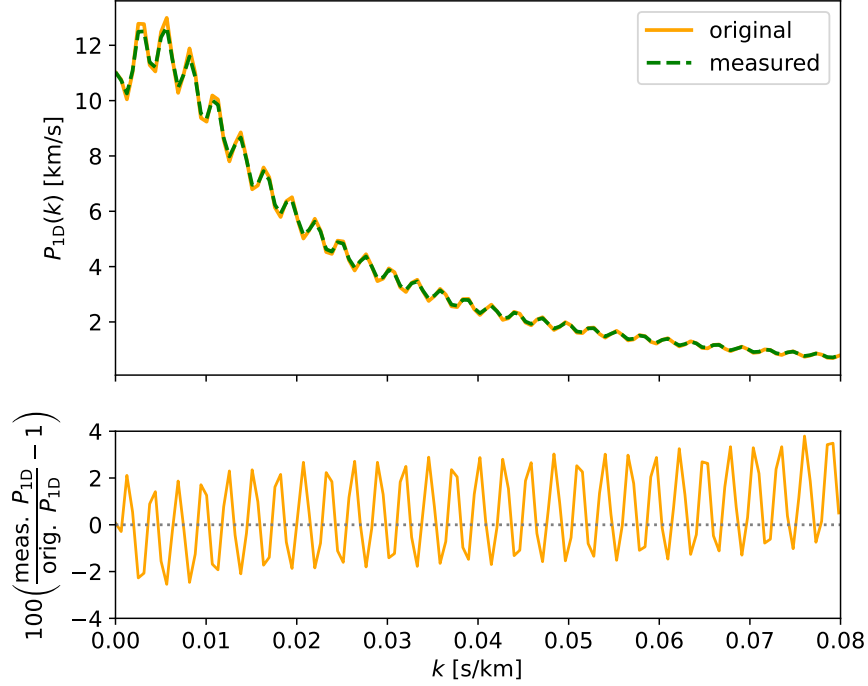


Figure 1. Measurements of P_{1D} in non-periodic Gaussian skewers (green dashed line), compared to the input power (orange). The bias is more clearly seen in the bottom panel, which shows the residuals (measurement over original, minus one) as a percentage. The bias appears as both a distortion of the Silicon oscillations and a broad scale dependence increasing with k .

The latest FFT-based measurement of P_{1D} by the DESI collaboration [12] split each Ly α skewer in three *chunks* of approximately 10^4 km s $^{-1}$ in length. Here we generate a periodic field with 2048 pixels and a total length of 8×10^4 km s $^{-1}$, but we only keep 1/8 of the field to obtain a *non-periodic* skewer with 256 pixels, 10^4 km s $^{-1}$ long. This gives a Fourier mode resolution of $\Delta k = 2\pi/L = 0.0006$ km s $^{-1}$.

In figure 1 we show (in green) the measurement of P_{1D} from 10^7 of these non-periodic skewers, and compare it to the input power used to generate the skewers (in orange). The residuals shown in the bottom panel demonstrate a clear bias in the measurement of P_{1D} . First, the oscillatory features mimicking the contamination by Silicon absorption have been smoothed, causing a bias oscillating by $\sim 4\%$. Second, the “broadband” power is slightly under-predicted at low values of k ($k < 0.015$ km s $^{-1}$) and over-predicted at intermediate and large values of k (reaching the 1% level at $k \sim 0.07$ km s $^{-1}$). This second bias is also present in mock spectra without Silicon contamination.

Conceptually, the two effects arise because the DFT (Eq. 2.7) assumes the signal to be one period of an infinitely repeating signal, and therefore that the final and initial pixel of each spectrum are neighboring. This causes the Fourier signal, and therefore the output power, to become mixed across modes compared to the true input power. The effect is easiest to consider for the correlation function in configuration pixel space: although the initial and final pixels are truly separated by the full length of the spectrum, when the signal is assumed

to be periodic the pair contributes to the correlation function bin at a single-pixel length.

One could use equation (2.15) to forward model the impact of non-periodic boundary conditions, but it would be quite tedious computationally. It is faster and more convenient to model the impacts after *zero-padding* the signal, as shown in the following section. This entails extending the signal by a chain of zeros. After adding zeros, the estimator is still biased³, but it is easier to forward-model the theoretical predictions for the summary statistics.

Ideally, the zero padding should extend the FFT grid used in all computations to at least twice the original length. This is because the discretized computations of P_{1D} introduced in this section perform circular convolution of the real-space signal, which introduces a bias at the overlapping edges that have been wrapped around (when applied to a non-periodic signal). Performing each convolution after padding the array with an equal amount of zeros produces $2N - 1$ non-zero pixels, and thus avoids overlap between the end and beginning of the array which are separated by $2N$. Longer zero-padding will continue to improve the convergence of the discretized calculations toward the continuous integrals in k , but practical limits on memory usage constrain this extension. P_{1D} and $P_{\times}$ analyses are typically less memory-intensive than full three-dimensional analyses, but in any case, the best practice is to perform tests to determine the minimal zero-padding required to reduce bias from discretization and non-periodicity to a satisfactory level given computational constraints.

2.2 Modelling the impact of zero-padding, masking, and pixel weights

Zero-padding is only one source of nonphysical values in spectral data; most analyses of the Ly α forest also implement some level of masking at the pixel level, motivated by instrumental artifacts or sky lines. Practically, masking is typically done by setting a spectral pixel with an unwanted feature to zero; the mask is thus a specific case of pixel weights, where the weights can only take binary values (ones or zeros). In this section we will model the impact of a general array of pixel weights w_a that can take arbitrary values, with the understanding that a masked pixel will have $w_a = 0$.

From now on, we will use N to refer to the total number of pixels in the FFT grid, including the zero padding. We can now define an array of N weighted pixels $f_a = w_a \delta_a$, and compute its FFT:

$$f_m = \sum_a e^{-ik_m x_a} w_a \delta_a = \int \frac{dk}{2\pi} w(k_m - k) R(k) \delta(k) , \quad (2.17)$$

where we have introduced a new function

$$w(k) = \sum_a e^{-ikx_a} w_a . \quad (2.18)$$

Note that we have not defined $w(k)$ as the continuous Fourier transform of a $w(x)$ function, and that it is a complex number and dimensionless.

³Imagine extending the signal by zeros such that the new array is twice as long. As evident from Eq. 2.7, besides an overall amplitude suppression from the the null contribution of zeros to a signal with larger N , the δ_m values at the same discrete modes k_m will be exactly equivalent to the values before zero-padding. However, there are now twice as many k_m modes, with the new discrete modes falling at k values between the previous ones. Thus, for extended bins in k , even after correcting for the scalar amplitude suppression, the padded signal will be subtly different than the original signal. This is consistent with the fact that ξ_{1D} must change.

The expected value of the (dimensionless) variance of the (discrete) Fourier modes is:

$$\langle |f_m|^2 \rangle = \int \frac{dk}{2\pi} |w(k_m - k)|^2 R^2(k) P_{1D}(k) , \quad (2.19)$$

where the integral reminds us of a convolution of the Fourier transform of the weights with the product of P_{1D} and the resolution kernel. In 1D analyses it should be easy to add enough zero-padding to make the skewers as long as needed to have a very good frequency resolution ($\Delta k = 2\pi/L$), such that we can discretize this integral with the same set of k_m wavenumbers:

$$\langle |f_m|^2 \rangle = \frac{\Delta k}{2\pi} \sum_n |w_{m-n}|^2 R^2(k_n) P_{1D}(k_n) = \frac{1}{N} \sum_n |w_{m-n}|^2 P_n , \quad (2.20)$$

where we have introduced $w_m = w(k_m)$, the DFT of w_a , and we have used equation (2.12) above.

Similarly to equation (2.13), we can now derive an alternative estimator of P_{1D} to be used in the presence of weights, masking or zero-padding:

$$\hat{P}_{1D,m} \equiv C_m \frac{\Delta x}{N R^2(k_m)} \frac{1}{N_q} \sum_q |f_m^q|^2 \quad (2.21)$$

(recall f_m is the FT of $f_a = w_a \delta_a$ and N is the number of pixels in the FFT grid). The same normalization factors as in equation (2.13) give dimensions of length⁴ to P_{1D} , while C_m is an extra normalization factor that we introduce for convenience.

The ensemble average of the estimator is related to the original P_{1D} by:

$$\begin{aligned} \langle \hat{P}_{1D,m} \rangle &= C_m \frac{\Delta x}{N R^2(k_m)} \frac{1}{N_q} \sum_q \langle |f_m^q|^2 \rangle \\ &= C_m \frac{\Delta x}{N R^2(k_m)} \frac{1}{N_q} \sum_q \frac{\Delta k}{2\pi} \sum_n |w_{m-n}^q|^2 R^2(k_n) P_{1D}(k_n) \\ &= \frac{C_m}{N^2 R^2(k_m)} \sum_n W_{m-n} R^2(k_n) P_{1D}(k_n) \\ &= \sum_n M_{mn} P_{1D}(k_n) , \end{aligned} \quad (2.22)$$

where we have introduced W_m as the average of the squared FFT of the individual weights:

$$W_m \equiv \frac{1}{N_q} \sum_q |w_m^q|^2 , \quad (2.23)$$

and we have introduced the window matrix of the survey:

$$M_{mn} \equiv \frac{C_m}{N^2 R^2(k_m)} W_{m-n} R^2(k_n) . \quad (2.24)$$

The window matrix M_{mn} specifies how much the undistorted model at a given wavenumber k_n contributes to the distorted prediction at a wavenumber k_m . We can choose to set

⁴In Ly α analyses, the units of such a ‘length’ could be either km/s, Mpc, or Å.

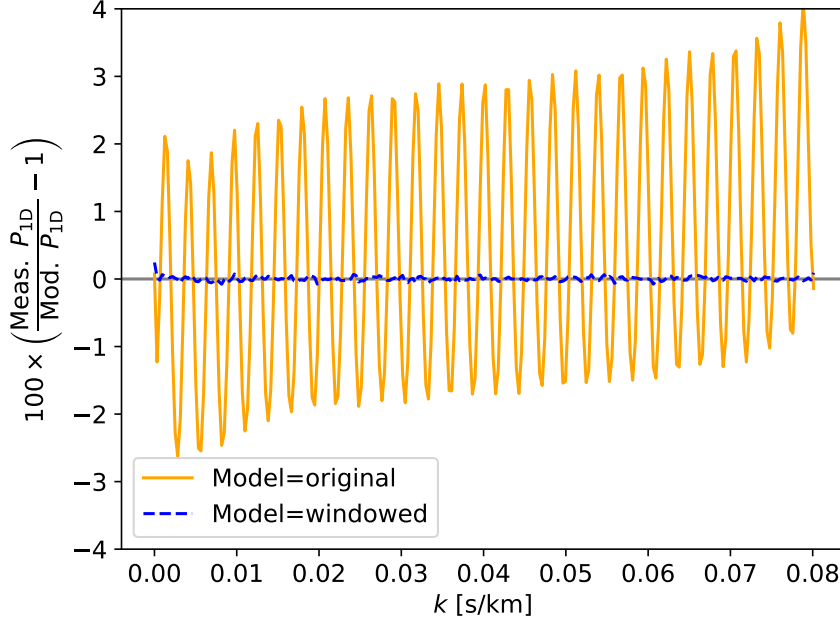


Figure 2. Percentage residuals of the P_{1D} measurement from non-periodic skewers with zero-padding using the estimator in equation (2.26), compared to the original input P_{1D} (orange) and compared to the model computed using the window matrix in equation (2.27) (blue, dashed). Application of the window matrix forward-models the impact of zero-padding on the original input P_{1D} . With this forward modeling, residuals reduce to a negligible level.

the normalization factor C_m such that the each row of the window matrix is normalized to 1, i.e., $1 = \sum_o M_{mo}$. This condition is satisfied if we choose:

$$C_m = \frac{N^2 R^2(k_m)}{\sum_o W_{m-o} R^2(k_o)} . \quad (2.25)$$

If we ignore the impact of the resolution kernel (set $R(k) = 1$), one can see that for simple masking (binary weights), the normalization factor C_m is simply a scalar corresponding to the inverse of the fraction of pixels not masked. So, C_m can be thought of as a factor rescaling the measured P_{1D} amplitudes (typically lower than the original P_{1D} due to masking) to values similar to the original.

Using this normalization factor, the estimator can be rewritten as:

$$\hat{P}_{1D,m} \equiv \frac{L}{\sum_o W_{m-o} R^2(k_o)} \frac{1}{N_q} \sum_q |f_m^q|^2 , \quad (2.26)$$

and the window matrix as:

$$M_{mn} \equiv \frac{W_{m-n} R^2(k_n)}{\sum_o W_{m-o} R^2(k_o)} . \quad (2.27)$$

Thus, to predict $\langle \hat{P}_{1D,m} \rangle$ from a theory model for the original P_{1D} , one needs to compute the window matrix of the survey only once. The process in practice begins with Fourier

transforming the array of real-space weights, then squaring its magnitude and summing over all skewers to get W_m . Next, one calculates the window with equation (2.27); the convolution in the denominator can be done rapidly by multiplying the FFTs of W and R^2 and transforming back via an inverse FFT. The resulting real-valued, $N \times N$ array is matrix-multiplied by the discretized original P_{1D} theory vector to yield the predicted weighted P_{1D} vector.

We make a measurement of P_{1D} from the same 10^7 non-periodic skewers used in figure 1, but this time using the estimator from equation (2.26) and a factor of two zero-padding (we started with the same 256 pixels, but added now the same amount of pixels with $w_a = 0$ such that the FFT grid has $N = 512$ pixels). Because the distortion is caused by an input alteration to the signal in a form that we know exactly, we can use equation (2.22) to forward-model the distortion. In figure 2 we show the residuals from comparing the measurement to the forward model (blue dashed line) versus comparing the measurement to the original input power (orange). The small residuals using the windowed model demonstrate that the model captures both the smoothing of the oscillatory features and the broadband trend.

3 Validation with masked mock spectra from hydrodynamic simulations

To validate the model on a more realistic set of mock skewers, we move to a suite of **MP-Gadget** hydrodynamical simulations first presented in [25]. Each simulation box has periodic boundary conditions with a side length $L = 67.5$ Mpc, with 768^3 particles, and skewers extracted with $N = 1350$ pixels along the line of sight, i.e., $\Delta x = 0.05$ Mpc. This pixelisation sets the natural grid to use in the FFTs, and its corresponding wave-numbers are equispaced by $\Delta k = 2\pi/L = 0.0931 \text{Mpc}^{-1}$, with a Nyquist frequency of $k_{Ny} = \pi/\Delta x = 62.83 \text{Mpc}^{-1}$. We work with the simulation snapshot at $z = 3$.

Because of the periodicity of the hydro boxes, there is no need to use zero-padding; we will apply other forms of realistic masks in this section. As before, we will focus on the effects of these masks on realistic data whose power spectra include localized features in k -space. To incorporate such features, we add Silicon contamination using the model introduced in section 2.1.4 and parameters $a_{Si} = 0.05$, $r_{Si} = 20$ Mpc. Additionally, we include an arbitrary and simple additional source of contamination, a spike in amplitude localized in Fourier space. To do so, we boost the amplitude of a single discrete Fourier mode δ_m for all skewers, with $k_m \sim 0.47 \text{Mpc}^{-1}$, such that the P_{1D} is augmented by 0.5 Mpc for that mode.

We apply five different types of mask to the mock spectra: (a) single emission/absorption line, (b) double emission/absorption line, (c) small DLAs, (d) large DLAs, and (e) random masking to mimic randomly-placed corrupted pixels. In a typical analysis, true $\text{Ly}\alpha$ data would be masked for a combination of these effects, but we separate them to examine their distinct features. We enforce that each case results in an approximately equal total number of masked pixels: $\sim 2\%$ of the total number of pixels summed over all skewers.

The single-line mask (a) sets thirty consecutive pixels to zero at a particular x value in all skewers. The width of the mask (1.5 Mpc, or $\delta\lambda \sim 1.9\text{\AA}$) is chosen to roughly correspond to the masks for sky emission lines and Galactic Calcium absorption lines used in the DESI $\text{Ly}\alpha$ early data release [26], as translated into the rest-frame of a $z = 3$ quasar. The double-line mask (case b) mimics the single-line mask, but sets 15 consecutive pixels to zero in two locations, separated by 125 pixels (6 Mpc / $\Delta\lambda \sim 8\text{\AA}$). For DLA-like masking (c) and (d), we choose a subset of skewers in which to mask one to two DLA-length sets of consecutive pixels. For the large DLA case, we set 20 Mpc (400 consecutive pixels) to zero at a random

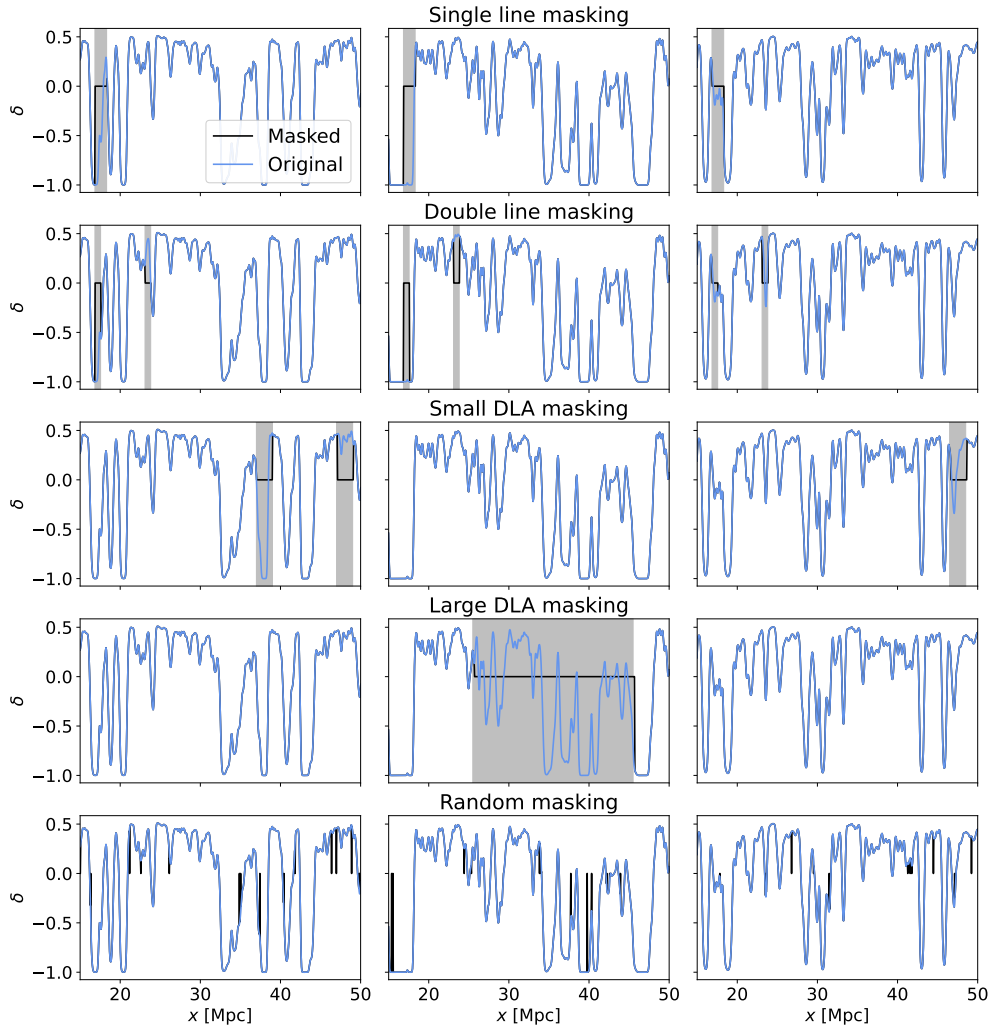


Figure 3. Representation of the five mask schemes (rows) for three mock spectra (columns). The black masked spectra are only visible compared to the over-plotted blue unmasked spectra in the masked regions. Gray shaded regions indicate the masks, except in the case of random masking where in most cases masked pixels are non-consecutive. The mask always appears in the same location in the mock spectra for single- and double-line masking (upper two rows). For the DLA-like masking, not all spectra host a ‘DLA’, and some host multiple. In all cases, the mask is uncorrelated with the original signal.

location along the skewer; for small DLAs, this is reduced to 40 pixels (2 Mpc). We note that the DLA masks are placed randomly, not at absorption peaks. Realistic DLA masking would be correlated with the field, but modeling the impact of correlated masks is beyond the scope of this work; hence we use uncorrelated placement to enable the application of the simple model presented in this paper. Lastly, in the random mask (e), we mask each skewer by a random number of pixels between 0 and 20 in random locations. Visualizations of the four masks are shown in figure 3.

We estimate P_{1D} for each set of masked skewers using equation (2.26), producing a measurement $P_{1D,\text{masked}}$. We repeat the same estimate with the set of original skewers to produce a measurement without masking effects, $P_{1D,\text{orig}}$. As both have very similar cosmic

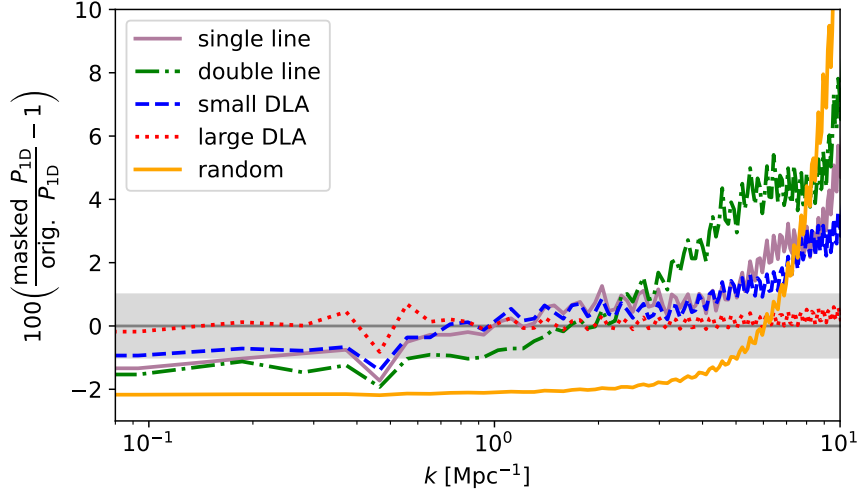


Figure 4. The impacts of the five mask schemes on P_{1D} measured from hydro simulation mock skewers. The P_{1D} estimates are computed by equation (2.26), and the figure shows the percentage difference between the measurements with masking and without (‘orig.’). The ‘large DLA’ case is broadly scale-independent due to the long extent of each DLA mask, but each of the other masking types induce broad k -dependent trends which depend on the extent and spacing of the masked pixels. Additionally, in all cases oscillations are clearly visible due to the mask distorting the measurement of the Silicon-contaminated original power spectrum. Distortions around the spike at $k \sim 0.5 \text{ Mpc}^{-1}$ are also clear for all but the random mask. Overall, the strongest biases occur at very small scales, but even for a reasonable observational range of $k < 3 \text{ Mpc}^{-1}$, all masks except for the large-DLA type cause a bias $> 1\%$ (shaded region). In practice, the amount of bias will depend on the amount of masking.

variance, their residuals are nearly completely dominated by the mask impacts. Figure 4 shows the residuals. The large DLA masking has a broadly scale-independent impact, but due to their small-scale features, the four other types decrease low- k power and increase high- k power. Moreover, all cases show oscillatory features in k space related to the added Silicon-like contamination. Most also cause a distortion at the location of the spike, again demonstrating that the bias from masking is sensitive to the presence of features in the true power spectrum.

Figure 5 reproduces the mask impacts and also shows the performance of the window-convolved model for each case. Here, we calculate the window matrix from equation (2.27) using each set of masks and apply it directly to the unmasked P_{1D} measurement $P_{1D,\text{orig}}$. We choose to use $P_{1D,\text{orig}}$ as the ‘theory model’, rather than a smoothed theory model, in order to remove the effects from cosmic variance in the comparison between model and measurement. Comparing the masked measurements with the original model reveals $> 1\%$ biases within $k < 5 \text{ Mpc}^{-1}$ for all masks with small-scale features (orange lines). However, the model including the window function reproduces the masked signal to within 0.5% in a scale-independent manner (blue dashed lines). Building from the tests with Gaussian skewers in section 2.1.4, these validation tests with hydrodynamic mock data further demonstrate that comparison of a masked measurement with a masked model can yield robust, unbiased inference for a wide range of masking types, even when the original power spectrum has features.

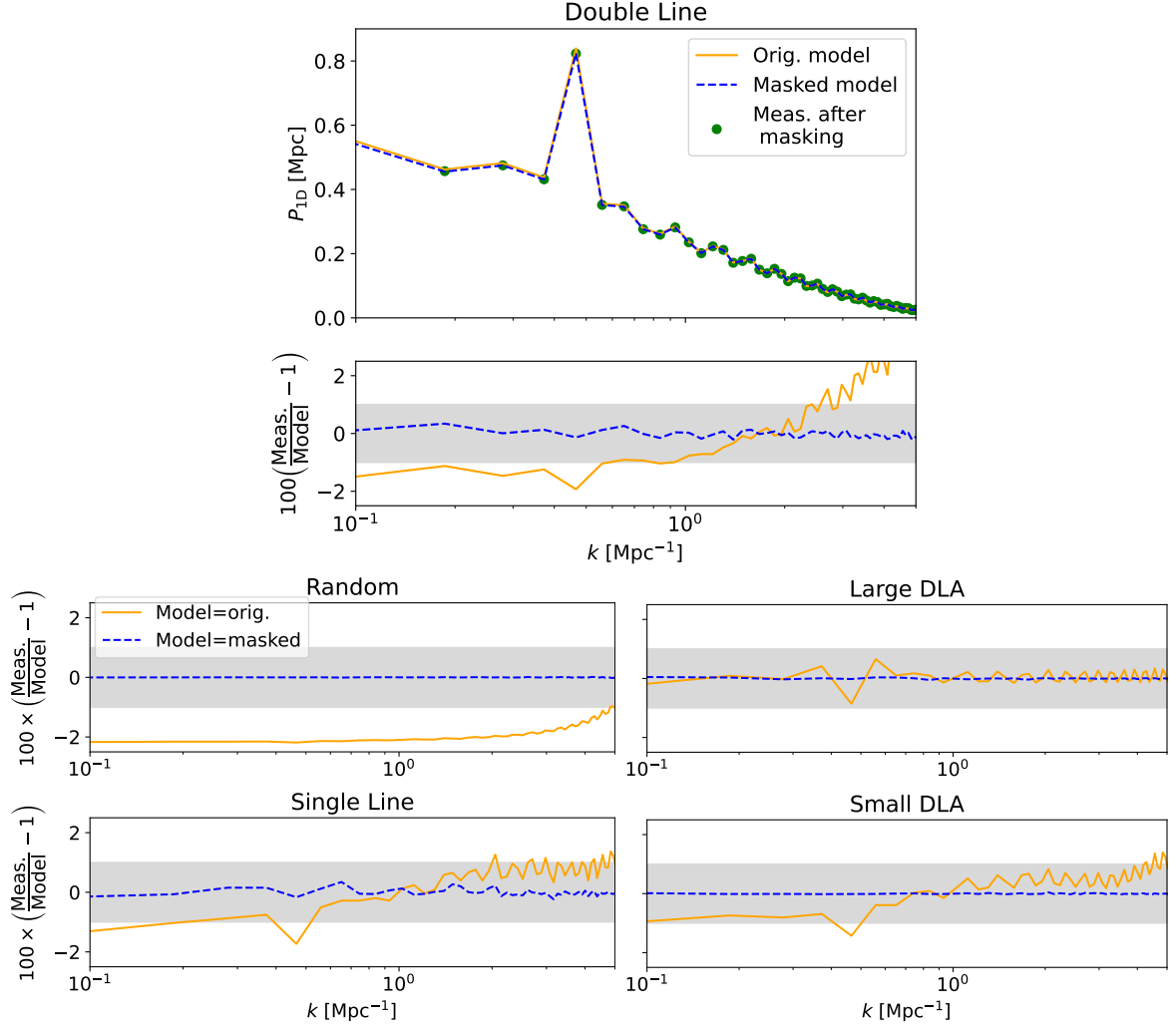


Figure 5. Top: P_{1D} measured from the double-line masked spectra (green points) compared to two models. The orange line shows the measurement from unmasked spectra, which we consider the ‘original’ model, and the blue dashed line shows the model after applying the window matrix to the original model, where the window matrix incorporates the double-line mask. The lower panel of the top figure shows the percentage difference between measurement and model when comparing with the original model (orange) and masked model (blue). The residuals are well within the 1% shaded regions when the mask is modeled, i.e., applying the window function accurately replicates the masking impacts. Bottom: the percentage residuals are displayed for the four other masking cases.

4 Masking in FFT estimators of $P_{\times}$

4.1 Overview

While FFT estimators have been widely used in measurements of $P_{1D}(k_{\parallel})$, the geometry of the Ly α forest surveys makes it difficult to apply them to measurements of the 3D power spectrum, $P_{3D}(k_{\perp}, k_{\parallel})$. Each quasar spectrum gives us information about the Ly α forest fluctuations along the line of sight, with a fairly homogeneous window matrix (except for the occasional masked pixels discussed above). However, the window function is very complicated

in the transverse direction, where we only sample the 3D field whenever we happen to have a background quasar.

This has motivated the use of alternative methods to measure 3D correlations [14, 18, 23, 27, 28]. An interesting idea, first proposed in [29], is to measure the cross-spectrum, $P_{\times}(r_{\perp}, k_{\parallel})$, the correlation of 1D Fourier modes $\delta(k_{\parallel})$ in neighboring quasars separated by a transverse separation $r_{\perp}$. This summary statistic captures the same information contained in P_{3D} , while not being affected by the complex transverse sampling. In the continuous, static (no redshift evolution) limit, it is defined as:

$$\left\langle \delta(x_{\perp}, k_{\parallel}) \delta^*(x'_{\perp}, k'_{\parallel}) \right\rangle = 2\pi \delta^D(k_{\parallel} - k'_{\parallel}) P_{\times}(r_{\perp}, k_{\parallel}), \quad (4.1)$$

with $r_{\perp} = |x'_{\perp} - x_{\perp}|$. It is easy to see that in the limit of $r_{\perp} \rightarrow 0$, $P_{\times}(r_{\perp} = 0, k_{\parallel})$ is equivalent to $P_{1D}(k_{\parallel})$. $P_{\times}$ is related to P_{3D} through the inverse Fourier transform of the perpendicular modes $\mathbf{k}_{\perp} = (k_{\perp,x}, k_{\perp,y})$ for a particular $r_{\perp}$ transverse scale:

$$P_{\times}(k_{\parallel}, r_{\perp}) = \frac{1}{2\pi} \int_0^{\infty} dk_{\perp} k_{\perp} J_0(k_{\perp} r_{\perp}) P_{3D}(k_{\parallel}, k_{\perp}), \quad (4.2)$$

where J_0 is the Bessel function of the first kind. At zero separation, $J_0(r_{\perp} = 0) = 1$, giving the relation between P_{3D} and P_{1D} . We note that equation (4.2) neglects to consider redshift evolution (see [18] for a discussion on how to define $P_{\times}$ in an evolving universe).

The optimal quadratic estimator to measure $P_{\times}$, discussed in [18], is quite complex and computationally expensive. On the other hand, the FFT estimators used to measure P_{1D} can be easily extended to measure $P_{\times}$. Recently, [14] used an FFT estimator to measure $P_{\times}$ from the final Ly α dataset of eBOSS. In this section we show that the methodology described in section 2 can be trivially extended to $P_{\times}$, enabling the use of weights and the possibility to forward model the impact of masked pixels.

4.2 $P_{\times}$ from pairs of quasars with different redshifts

The discussion of zero-padding from Section 2 is particularly relevant to $P_{\times}$, as the measurement of correlation between skewer pairs via the FFT estimator requires that both skewers in each pair share the same FFT grid. Observationally, this is challenging, because the Ly α skewers are not necessarily the same length, and the redshifts of the quasars (marking the start of each skewer) vary. [14] tackled this issue by defining a limited common wavelength range for the $P_{\times}$ measurement, and removing data that falls beyond that range and spectra that only partially covered it. Rather than removing this information and thus depleting the statistical power of the measurement, it is possible to retain all the data by adding zero-padding to each spectrum such that all fall along a common grid in wavelength space, which is then converted into a common grid in Fourier space. In figure 6 we show a visual representation of this change in methodology. As discussed in section 2.1.4, the zero-padding is also necessary to efficiently model the impact of masking, pixel weights, and non-periodic skewers.

Using the method described in section 2, we generate 10^7 pairs of periodic, Gaussian Ly α skewers with Fourier modes $(\delta^A(k), \delta^B(k))$, where k refers to one-dimensional modes, called $k_{\parallel}$ in the previous sub-section). In order for them to have a non-zero $P_{\times}$, we generate a new set of Fourier modes $\delta^C(k)$ as a linear combination of these two, $\delta^C = f_{\times} \delta^A + \sqrt{1 - f_{\times}^2} \delta^B$. Computed this way, P_{1D} of fields δ_A and δ_C is equal to the input power detailed in table 1, and the cross-spectrum of δ^A and δ^C will be equal to $P_{\times} = f_{\times} P_{1D}$. We set $f_{\times} = 0.7$.

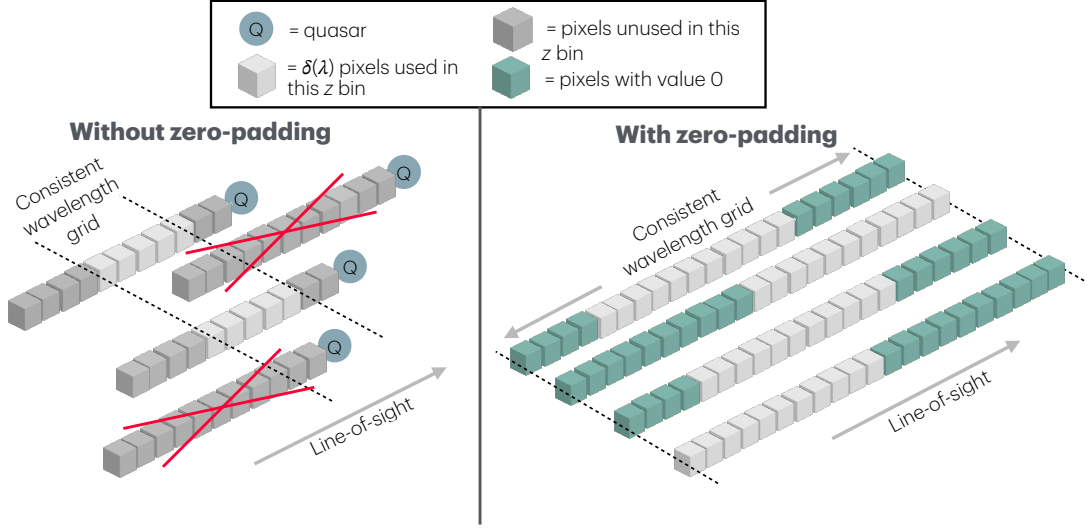


Figure 6. A visualization of the way in which zero-padding can be used to preserve data when measuring $P_{\times}$. At left, a limited wavelength (redshift) grid is defined, and two of the four spectra are removed because they do not fully span the grid. The shaded pixels beyond the grid are also unused in the measurement. At right, a larger wavelength (redshift) range is defined to incorporate more data. The addition of zeros places all spectra into a consistent wavelength grid, corresponding to a consistent FFT grid once Fourier transformed. All pixels are used, so the impacts of zeros must be modeled at the measurement or inference stage.

Following section 2, in order to break the periodicity of the skewers we only keep 256 pixels (1/8 of the original skewer), and add a factor of two zero-padding to obtain an array of 512 pixels, half of them masked ($w_a = 0$). In order to simulate the fact that each line-of-sight covers a different range of observed wavelengths (set by the redshift of the background quasars), each simulated skewer has a different starting point⁵.

For each pair, we then measure the correlation of 1D Fourier modes from the two skewers I and J , with masks w_a^I and w_a^J , respectively, and separated by θ_{IJ} ⁶. We will refer to their discrete, masked Fourier modes as f_m^I and f_m^J , respectively, defined as in equation (2.17). Analogously to equation (2.20), the correlation of these modes is:

$$\langle f_m^I f_m^{J*} \rangle = \frac{\Delta k}{2\pi} \sum_n w_{m-n}^I w_{m-n}^{J*} R^2(k_n) P_{\times}(\theta_{IJ}, k_n), \quad (4.3)$$

and analogously to equation (2.26) we can write an estimator for $P_{\times}$ (in some narrow θ bin

⁵To be precise, we draw a random number between 0 and 256 to be used as a starting point for the 256 pixels that are kept.

⁶It is important to note that the two skewers share the same FFT grid, with the quasar redshift determining where the padded zeros should be placed.

A) as:

$$\begin{aligned}
\hat{P}_{\times,m}^A &\equiv C_m^A \frac{\Delta x}{N R^2(k_m)} \frac{1}{N_A} \sum_{\substack{I, J \in A \\ I < J}} (f_m^I f_m^{J*} + f_m^{I*} f_m^J) \\
&= C_m^A \frac{\Delta x}{N R^2(k_m)} \frac{2}{N_A} \sum_{\substack{I, J \in A \\ I < J}} \Re(f_m^I f_m^{J*}) ,
\end{aligned} \tag{4.4}$$

where the sum runs over all pairs of skewers within the transverse separation bin A without double-counting.

The ensemble average of the estimator is related to the original $P_{\times}$ by:

$$\begin{aligned}
\langle \hat{P}_{\times,m}^A \rangle &= C_m^A \frac{\Delta x}{N R^2(k_m)} \frac{2}{N_A} \sum_{\substack{I, J \in A \\ I < J}} \langle \Re(f_m^I f_m^{J*}) \rangle \\
&= \frac{C_m^A}{N^2 R^2(k_m)} \frac{2}{N_A} \sum_{\substack{I, J \in A \\ I < J}} \sum_n \Re(w_{m-n}^I w_{m-n}^{J*}) R^2(k_n) P_{\times}(\theta_A, k_n) \\
&= \frac{C_m^A}{N^2 R^2(k_m)} \sum_n W_{m-n}^A R^2(k_n) P_{\times}(\theta_A, k_n) \\
&= \sum_n M_{mn}^A P_{\times}(\theta_A, k_n) ,
\end{aligned} \tag{4.5}$$

where we have introduced W_m^A as:

$$W_m^A \equiv \frac{2}{N_A} \sum_{\substack{I, J \in A \\ I < J}} \Re(w_m^I w_m^{J*}) \tag{4.6}$$

and we have introduced an equivalent window matrix for $P_{\times}$:

$$M_{mn}^A \equiv \frac{C_m^A}{N^2 R^2(k_m)} W_{m-n}^A R^2(k_n) . \tag{4.7}$$

Here again we can choose the normalization factor C_m^A such that $1 = \sum_o M_{mo}^A$. This condition is satisfied if we choose:

$$C_m^A = \frac{N^2 R^2(k_m)}{\sum_o W_{m-o}^A R^2(k_o)} . \tag{4.8}$$

Using this normalization factor, the estimator can be rewritten as:

$$\hat{P}_{\times,m}^A \equiv \frac{L}{\sum_o W_{m-o}^A R^2(k_o)} \frac{2}{N_A} \sum_{\substack{I, J \in A \\ I < J}} \Re(f_m^I f_m^{J*}) , \tag{4.9}$$

and the window matrix as:

$$M_{mn}^A \equiv \frac{W_{m-n}^A R^2(k_n)}{\sum_o W_{m-o}^A R^2(k_o)} . \tag{4.10}$$

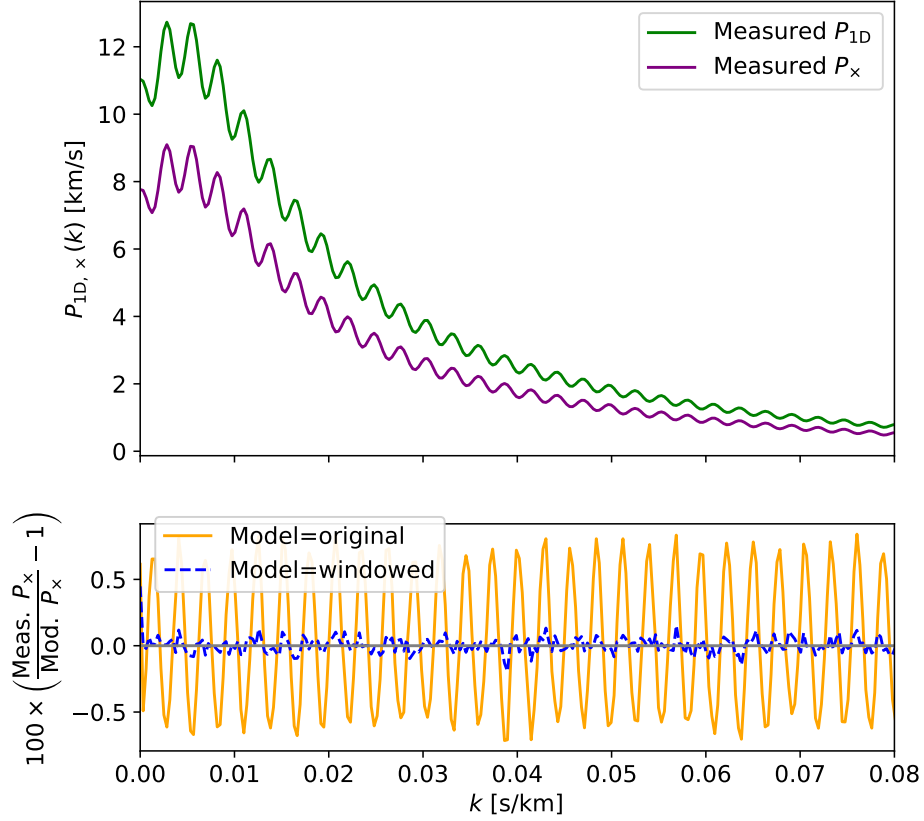


Figure 7. (Top): measured P_{1D} and P_x from 10^7 pairs of Gaussian skewers. (Bottom): residuals of the measured P_x from the same pairs of Ly α skewers compared to the true P_x used to generate the mocks (orange), compared to the prediction that uses the window matrix in equation (4.10) to forward-model the impact of masking and zero padding (blue, dashed).

The measurement of P_{1D} and P_x from these pairs of Gaussian skewers is shown in the upper panel of figure 7. The lower panel compares the P_x measurement to the original model and also to the windowed model. Similarly to the P_{1D} demonstrations of previous sections, application of the window matrix faithfully reproduces the distortions that zero-padding causes to the original power spectrum, reducing the residuals by a factor of ~ 10 . The exercise demonstrates that in cases of different skewer starting locations, it is possible to retain all the data to optimize signal-to-noise for a P_x measurement, by incorporating zero-padding and modeling it in predictions.

5 Conclusion

This work presented an efficient method to estimate and robustly model the P_{1D} and P_x statistics of the Ly α forest with FFTs in the presence of spectral pixel masking and pixel weights. We first demonstrated how the non-periodicity and finite length of realistic spectra induce mode-mixing in the power spectra, which causes scale-dependent bias and smooths features. We described an estimator which zero-pads the spectra by at least a factor of two in order to make the k grid finer and enable the distortions to be modeled rapidly using FFTs. We then translated the calculations of a window matrix given masking—already commonly

used in galaxy surveys, the CMB, and other Ly α contexts to describe sky incompleteness and spectral pixel masking—to the specific domain of FFT estimators for Ly α P_{1D} and $P_{\times}$. Using Gaussian and hydrodynamic simulations, we demonstrated that the model accounts for biases from zero-padding and other varied types of realistic masking at high accuracy. This combined estimator and modeling framework enables precision inference with DESI data and other future surveys.

In practice, the window matrix calculation could be applied in an inference pipeline through two different approaches. In one approach, the theory prediction gets contaminated with realistic models for contaminants, convolved with the same window matrix that affects the data, and then compared to observational measurements. (Subsequently, parameters are iterated through with a sampling technique such as MCMC.) Alternatively, in some settings it might be possible to invert the window matrix in order to directly deconvolve it from the measurement. In that case, one would not need to forward-model the masking or resolution effects on the theory prediction, but rather compare the original prediction directly to the measurement. An additional advantage to this approach is the ability to compare between an optimal quadratic estimator and an FFT estimator. On the other hand, for either estimator, inverting the window requires making unnecessary assumptions about the discretization of the theory, and generally leads to anti-correlated error covariance. E.g., past OQE work advocated an intermediate transformation of the window which left the errors uncorrelated [30]. In any case, given a well-computed window matrix we can easily explore different ways to use it.

An important feature of the weighted FFT estimators discussed in this work is the ability to optimize them using inverse-variance pixel weights. With an accurate model for the window matrix in hand, one can down-weight high-variance pixels to place more importance on well-measured pixels, and model the resulting impacts on P_{1D} or $P_{\times}$ by calculating the window matrix with that array of weights. Pixel weights have thus far not been used in the `picca-fft` estimators, and their inclusion will help maximize the information gained from DESI Ly α spectra.

There are several limitations to the methods recommended in this work. First, this paper does not address the complication of masks that correlate with the signal. Correlated masks will complicate the measurement and modeling of any estimator (not only FFT-based). This is true of DLA masking, where the removal of dense absorbers also removes Ly α forest signal at biased locations. When masking only the largest DLAs, highly extended masks are typically used, removing broad regions of the forest that include a wide range of densities and are thus nearly unbiased on average. However (somewhat counterintuitively) masking small DLAs causes more bias to the signal, as the masks are shorter and therefore remove a shorter and less representative region. Analyses masking only large DLAs can likely ignore the correlation and use the formalism presented here, but must model the bias to the power spectra given by the inclusion of residual small DLAs. Those masking a broader range of DLA sizes take care of the latter bias, but then face the problem of a more complicated masking bias which cannot use the approach of this paper. Such cases likely need to be handled by modeling via hydro simulations, which is beyond the scope of this paper.

A second limitation is that we have elected not to model a full resolution matrix in the development of the model, rather assuming that each pixel has equal resolution which is a symmetric function. Further, we have used perfect resolution ($R = 1$) in the tests with Gaussian and hydro skewers, electing to use the native simulations rather than smoothing to a realistic resolution and accounting for it in the models. In any real observational data,

resolution varies by pixel and by spectrum and the resolution functions generally may be asymmetric. Depending on the size of these impacts, it may be necessary to incorporate the full resolution matrix in models for P_{1D} or $P_{\times}$. This is beyond the scope of this work.

This work presents a key step toward precision modeling of small-scale Ly α forest statistics. With a full pipeline incorporating accurate pure-theory predictions, inclusion of metals and other contaminants, and propagation of the mask, weights, and resolution effects through the estimator, future analyses will be able to maximize the information gained from the vast amounts of incoming spectroscopic data to infer high-redshift cosmology at unprecedented precision.

Acknowledgements

The research of ML and AFR was supported by the European Union’s Horizon Europe research and innovation programme (COSMO-LYA, grant agreement 101044612) and by the Spanish Ministry of Science and Innovation (project PGC2021-123012NB-C41). AFR acknowledges funding by the Spanish Ministry of Science and Innovation under the Ramon y Cajal program (RYC-2018-025210). IF AE is partially funded by the CERCA program of the Generalitat de Catalunya. PM is supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231.

References

- [1] A. Slosar, A. Font-Ribera, M.M. Pieri, J. Rich, J.-M. Le Goff, É. Aubourg et al., *The Lyman- α forest in three dimensions: measurements of large scale flux correlations from BOSS 1st-year data*, *JCAP* **2011** (2011) 001 [[1104.5244](#)].
- [2] N.G. Busca, T. Delubac, J. Rich, S. Bailey, A. Font-Ribera, D. Kirkby et al., *Baryon acoustic oscillations in the Ly α forest of BOSS quasars*, *Astron. Astrophys.* **552** (2013) A96 [[1211.2616](#)].
- [3] A. Font-Ribera, D. Kirkby, N. Busca, J. Miralda-Escudé, N.P. Ross, A. Slosar et al., *Quasar-Lyman α forest cross-correlation from BOSS DR11: Baryon Acoustic Oscillations*, *JCAP* **5** (2014) 27 [[1311.1767](#)].
- [4] H. du Mas des Bourboux, J. Rich, A. Font-Ribera, V. de Sainte Agathe, J. Farr, T. Etourneau et al., *The Completed SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations with Ly α Forests*, *ApJ* **901** (2020) 153 [[2007.08995](#)].
- [5] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander et al., *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, *arXiv e-prints* (2024) [[arXiv:2404.03001](#)] [[2404.03001](#)].
- [6] DESI Collaboration, M. Abdul-Karim, J. Aguilar, S. Ahlen, C. Allende Prieto, O. Alves et al., *DESI DR2 Results I: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, *arXiv e-prints* (2025) [[arXiv:2503.14739](#)] [[2503.14739](#)].
- [7] R.A.C. Croft, D.H. Weinberg, N. Katz and L. Hernquist, *Recovery of the Power Spectrum of Mass Fluctuations from Observations of the Ly α Forest*, *ApJ* **495** (1998) 44 [[astro-ph/9708018](#)].
- [8] P. McDonald, J. Miralda-Escudé, M. Rauch, W.L.W. Sargent, T.A. Barlow, R. Cen et al., *The Observed Probability Distribution Function, Power Spectrum, and Correlation Function of the Transmitted Flux in the Ly α Forest*, *ApJ* **543** (2000) 1 [[astro-ph/9911196](#)].
- [9] P. McDonald, U. Seljak, S. Burles, D.J. Schlegel, D.H. Weinberg, R. Cen et al., *The Ly α Forest Power Spectrum from the Sloan Digital Sky Survey*, *ApJS* **163** (2006) 80 [[astro-ph/0405013](#)].

- [10] N. Palanque-Delabrouille, C. Yèche, A. Borde, J.-M. Le Goff, G. Rossi, M. Viel et al., *The one-dimensional Ly α forest power spectrum from BOSS*, *Astron. Astrophys.* **559** (2013) A85 [[1306.5896](#)].
- [11] S. Chabanier, N. Palanque-Delabrouille, C. Yèche, J.-M. Le Goff, E. Armengaud, J. Bautista et al., *The one-dimensional power spectrum from the SDSS DR14 Ly α forests*, *JCAP* **2019** (2019) 017 [[1812.03554](#)].
- [12] C. Ravoux, M.L. Abdul Karim, E. Armengaud, M. Walther, N.G. Karaçaylı, P. Martini et al., *The Dark Energy Spectroscopic Instrument: one-dimensional power spectrum from first Ly α forest samples with Fast Fourier Transform*, *Mon. Not. Roy. Astron. Soc.* **526** (2023) 5118 [[2306.06311](#)].
- [13] N.G. Karaçaylı, P. Martini, J. Guy, C. Ravoux, M.L.A. Karim, E. Armengaud et al., *Optimal 1D Ly α Forest Power Spectrum Estimation - III. DESI early data*, *Mon. Not. Roy. Astron. Soc.* (2024) [[2306.06316](#)].
- [14] M.L. Abdul Karim, E. Armengaud, G. Mention, S. Chabanier, C. Ravoux and Z. Lukić, *Measurement of the small-scale 3D Lyman- α forest power spectrum*, *JCAP* **2024** (2024) 088 [[2310.09116](#)].
- [15] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein et al., *The DESI Experiment, a whitepaper for Snowmass 2013*, *arXiv e-prints* (2013) arXiv:1308.0847 [[1308.0847](#)].
- [16] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L.E. Allen et al., *The DESI Experiment Part I: Science, Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [[1611.00036](#)].
- [17] L. Hui, S. Burles, U. Seljak, R.E. Rutledge, E. Magnier and D. Tytler, *On Estimating the QSO Transmission Power Spectrum*, *ApJ* **552** (2001) 15 [[astro-ph/0005049](#)].
- [18] A. Font-Ribera, P. McDonald and A. Slosar, *How to estimate the 3D power spectrum of the Lyman- α forest*, *JCAP* **2018** (2018) 003 [[1710.11036](#)].
- [19] N.G. Karaçaylı, A. Font-Ribera and N. Padmanabhan, *Optimal 1D Ly α forest power spectrum estimation - I. DESI-lite spectra*, *Mon. Not. Roy. Astron. Soc.* **497** (2020) 4742 [[2008.06421](#)].
- [20] M.G. Hauser and P.J.E. Peebles, *Statistical Analysis of Catalogs of Extragalactic Objects. II. the Abell Catalog of Rich Clusters*, *ApJ* **185** (1973) 757.
- [21] E. Hivon, K.M. Górski, C.B. Netterfield, B.P. Crill, S. Prunet and F. Hansen, *MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets*, *ApJ* **567** (2002) 2 [[astro-ph/0105302](#)].
- [22] M.J. Wilson, J.A. Peacock, A.N. Taylor and S. de la Torre, *Rapid modelling of the redshift-space power spectrum multipoles for a masked density field*, *Mon. Not. Roy. Astron. Soc.* **464** (2017) 3121 [[1511.07799](#)].
- [23] R. de Belsunce, O.H.E. Philcox, V. Iršič, P. McDonald, J. Guy and N. Palanque-Delabrouille, *The 3D Lyman- α forest power spectrum from eBOSS DR16*, *Mon. Not. Roy. Astron. Soc.* **533** (2024) 3756 [[2403.08241](#)].
- [24] J. Guy, S. Bailey, A. Kremin, S. Alam, D.M. Alexander, C. Allende Prieto et al., *The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument*, *AJ* **165** (2023) 144 [[2209.14482](#)].
- [25] C. Pedersen, A. Font-Ribera, K.K. Rogers, P. McDonald, H.V. Peiris, A. Pontzen et al., *An emulator for the Lyman- α forest in beyond- Λ CDM cosmologies*, *JCAP* **2021** (2021) 033 [[2011.15127](#)].

- [26] C. Ramírez-Pérez, I. Pérez-Ràfols, A. Font-Ribera, M.A. Karim, E. Armengaud, J. Bautista et al., *The Lyman- α forest catalogue from the Dark Energy Spectroscopic Instrument Early Data Release*, *Mon. Not. Roy. Astron. Soc.* **528** (2024) 6666 [[2306.06312](#)].
- [27] B. Horowitz, R. de Belsunce and Z. Lukić, *Maximum a posteriori Ly α estimator (MAPLE): band power and covariance estimation of the 3D Ly α forest power spectrum*, *Mon. Not. Roy. Astron. Soc.* **536** (2025) 845 [[2403.17294](#)].
- [28] N.G. Karaçaylı and C.M. Hirata, *Light in the dark forest – I. An efficient optimal estimator for 3D Lyman-alpha forest power spectrum*, *arXiv e-prints* (2025) [arXiv:2503.15619](#) [[2503.15619](#)].
- [29] L. Hui, A. Stebbins and S. Burles, *A Geometrical Test of the Cosmological Energy Contents Using the Ly α Forest*, *Astrophys. J. Let.* **511** (1999) L5 [[astro-ph/9807190](#)].
- [30] M. Tegmark, M.R. Blanton, M.A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro et al., *The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey*, *ApJ* **606** (2004) 702 [[astro-ph/0310725](#)].