

This is the accepted manuscript made available via CHORUS. The article has been published as:

# First detection of cosmic microwave background lensing and Lyman- $\alpha$ forest bispectrum

Cyrille Doux, Emmanuel Schaan, Eric Aubourg, Ken Ganga, Khee-Gan Lee, David N. Spergel, and Julien Tréguer

Phys. Rev. D **94**, 103506 — Published 9 November 2016

DOI: [10.1103/PhysRevD.94.103506](https://doi.org/10.1103/PhysRevD.94.103506)

# First Detection of Cosmic Microwave Background Lensing and Lyman- $\alpha$ Forest Bispectrum

Cyrille Doux,<sup>1,\*</sup> Emmanuel Schaan,<sup>2</sup> Eric Aubourg,<sup>1</sup> Ken Ganga,<sup>1</sup>  
Khee-Gan Lee,<sup>3,4</sup> David N. Spergel,<sup>2</sup> and Julien Tréguer<sup>1</sup>

<sup>1</sup>*AstroParticule et Cosmologie, Université Paris Diderot,  
CNRS, CEA, Observatoire de Paris, Sorbonne Paris Cité*

*Bâtiment Condorcet, 10, rue Alice Domon et Léonie Duquet, F-75205 Paris Cedex 13, France*

<sup>2</sup>*Dept. of Astrophysical Sciences, Peyton Hall, Princeton University, Princeton, NJ USA 08544*

<sup>3</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley CA 94720, USA*

<sup>4</sup>*Hubble Fellow*

We present the first detection of a correlation between the Lyman- $\alpha$  forest and cosmic microwave background (CMB) gravitational lensing. For each Lyman- $\alpha$  forest in SDSS-III/BOSS DR12, we correlate the one-dimensional power spectrum with the CMB lensing convergence on the same line of sight from Planck. This measurement constitutes a position-dependent power spectrum, or a squeezed bispectrum, and quantifies the non-linear response of the Lyman- $\alpha$  forest power spectrum to a large-scale overdensity. The signal is measured at  $5\sigma$  and is consistent with the  $\Lambda$ CDM expectation. **We measure the linear bias of the Lyman- $\alpha$  forest with respect to the dark matter distribution, and constrain a combination of non-linear terms including the non-linear bias.** This new observable provides a consistency check for the Lyman- $\alpha$  forest as a large-scale structure probe and tests our understanding of the relation between intergalactic gas and dark matter. In the future, it could be used to test hydrodynamical simulations and calibrate the relation between the Lyman- $\alpha$  forest and dark matter.

PACS numbers: 98.80.-k, 98.70.Vc, 98.62.Ra

## I. INTRODUCTION

Quasars are bright lanterns that illuminate the dark Universe and probe the distribution of gas. The Lyman- $\alpha$  forest observed in their spectra reveals the presence of intervening neutral hydrogen absorbing light at 1216 Å. It has been used to study the thermal history of intergalactic gas and hydrogen reionization [1–4]. Assuming that the flux transmission in the Lyman- $\alpha$  forest traces the matter density makes it a powerful probe of the large-scale structure of the Universe at intermediate redshifts, on a wide range of scales. The one-dimensional power spectrum along the line of sight probes the matter fluctuations on the smallest scales, constraining the sum of neutrino masses [5, 6], models of warm dark matter [7, 8] and primordial black hole dark matter [9]. Combining different lines of sight enables probing the three-dimensional power spectrum on larger scales [4, 10] and provides a measurement of the baryonic acoustic oscillations (BAO) at high redshift [11–13].

The interpretation of these results rely on the Lyman- $\alpha$  transmission tracing the underlying matter density field. If the hydrogen is in photoionization equilibrium with a uniform UV background, and there is no other sources of entropy, then the relationship is described analytically through variations of the fluctuating Gunn-Peterson approximation (FGPA; [14]) and is evaluated numerically using hydrodynamical simulations [15, 16]. However, the connection between

Lyman- $\alpha$  transmission and the underlying matter density is complex [17] and non-linear. It is affected by the proximity effect on the largest scales [17], by thermal broadening, Jeans smoothing and non-linear gravitational evolution on the smallest scales [18] and by the gas equation of state throughout. For these reasons, and in the light of the tension between BAO measurements from the Lyman- $\alpha$  forest and galaxies [19], consistency checks for the link between Lyman- $\alpha$  transmission and matter density are valuable.

Gravitational lensing of the cosmic microwave background (CMB) is sourced by large-scale structures located between the last scattering surface and the observer, and provides a measurement of the projected density of matter. In this paper, we cross-correlate, for the first time, the power spectrum of the flux transmission in the Lyman- $\alpha$  forest of quasars with CMB lensing to test our understanding of the relation between intergalactic gas and dark matter.

Since both the CMB lensing convergence and the mean Lyman- $\alpha$  transmission probe the mean density on a given line of sight, it is natural to compute their cross-spectrum. However, the mean Lyman- $\alpha$  transmission is strongly affected by continuum fitting in the quasar spectrum, making this a challenging observable. For this reason, we instead correlate the CMB lensing convergence with the small-scale Lyman- $\alpha$  power spectrum on the same line of sight. The origin of this signal is more complex, and corresponds to a position-dependent power spectrum [20, 21], or a squeezed bispectrum of the matter density. Simply put, a positive CMB convergence corresponds to an overdense line of sight; on this line

\* cdoux@apc.in2p3.fr

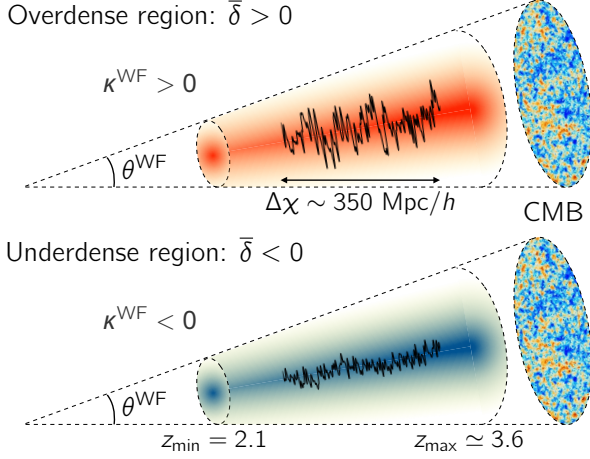


FIG. 1. Schematic of the correlation: overdense regions (respectively underdense regions), in red on the top panel (blue on the bottom panel) have positive (negative) CMB lensing convergence and are expected to produce more (less) small-scale structures under non-linear gravitational evolution, which is detectable in the amplitude of the Lyman- $\alpha$  forest power spectrum. The extent of the aforementioned regions is determined by the angular resolution  $\theta^{\text{WF}}$  of the Wiener-filtered convergence map and depth of the lensing efficiency function. In this analysis, we select Lyman- $\alpha$  forests in the redshift range 2.1 – 3.6.

of sight, the matter power spectrum is enhanced on all scales, due to non-linear evolution under gravity [20–22] (see Fig. 1 for a schematic of this idea). This bispectrum would therefore vanish at linear order in the perturbation theory of the density field, where short and long modes are independent. Instead, for a non-linear density field, this signal probes the response of the Lyman- $\alpha$  power spectrum to a mean overdensity.

This method was proposed in [23–25]. In this paper, we present the first detection of this signal, and propose a new theoretical description of it, based on the response of the matter power spectrum to a mean overdensity.

## II. THEORY

We aim at evaluating the covariance between the one-dimensional power spectrum  $P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel})$  of the Lyman- $\alpha$  forest transmission on one line of sight and the CMB convergence  $\kappa$  on the same line of sight.

The CMB lensing convergence  $\kappa$  probes the large-scale matter distribution along the line of sight. Our estimate  $\hat{\kappa}$  comes from CMB lensing reconstruction [26, 27] and is Wiener-filtered such that  $\hat{\kappa} = \Lambda_{\kappa} * \kappa + \text{noise}$ . As a result, the estimated convergence effectively probes the mass distribution within a “cone”, whose line of sight dimension is determined by the lensing efficiency kernel  $W_{\kappa}(\chi) = \delta\kappa/\delta[\delta(\chi)]$ , and whose angular size is determined by the Wiener filter  $\Lambda_{\kappa}$ . This is depicted in Fig. 1.

We split this cone into thin slices of fixed comoving distance  $\chi$  and thickness  $d\chi$ . The variance of the density field  $\bar{\delta}(\chi)$  averaged over this thin slice is given by  $\text{Var}[\bar{\delta}(\chi)] = \sigma^2(\chi)/d\chi$  (see [21]) with

$$\sigma^2(\chi) = \int \frac{d^2 k_{\perp}}{(2\pi)^2} |\Lambda_{\kappa}(\ell = \chi k_{\perp})|^2 P_{\text{lin}}(k_{\perp}, \chi) \quad (1)$$

where  $P_{\text{lin}}(k_{\perp}, \chi)$  is the linear matter power spectrum at comoving distance  $\chi$ . The covariance of  $\bar{\delta}(\chi)$  with the 1d power spectrum measured on the same slice is [21]

$$\text{Cov}[\bar{\delta}(\chi), P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}, \chi)] = \text{Var}[\bar{\delta}(\chi)] \frac{\partial P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}, \chi)}{\partial \bar{\delta}} + \mathcal{O}(\sigma^4). \quad (2)$$

In other words, the response of the Lyman- $\alpha$  power spectrum to the mean matter overdensity produces a non-zero covariance. This is the signal we wish to detect. For a given Lyman- $\alpha$  forest, measured between  $\chi_{\text{min}}$  and  $\chi_{\text{max}}$ , we define an average power spectrum

$$P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}) = \frac{1}{\Delta\chi} \int_{\chi_{\text{min}}}^{\chi_{\text{max}}} d\chi P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}, \chi) \quad (3)$$

where  $\Delta\chi = \chi_{\text{max}} - \chi_{\text{min}}$ . Since the CMB convergence  $\kappa = \int d\chi W(\chi) \delta(\chi)$  is a weighted average of the matter density field, we perform the same average on Eq. (2) to get an integrated bispectrum between the CMB lensing convergence and fluctuations in the Lyman- $\alpha$  forest

$$\begin{aligned} B_{\kappa, \text{Ly}\alpha}(k_{\parallel}) &\triangleq \text{Cov}[\kappa, P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel})] \\ &= \frac{1}{\Delta\chi} \int d\chi W_{\kappa}(\chi) \frac{\partial P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}, \chi)}{\partial \bar{\delta}} \sigma^2(\chi), \end{aligned} \quad (4)$$

where the integral runs from  $\chi_{\text{min}}$  to  $\chi_{\text{max}}$ . We assumed the various redshift slices are uncorrelated, as in the Limber approximation [28, 29], **because the lensing window function  $W_{\kappa}(\chi)$  varies slowly over the integral width, which in turn, is much larger than the scales where the correlation is expected, of order 0.1 – 1 h/Mpc.**

To go further, we need to evaluate the response of the Lyman- $\alpha$  power spectrum to the mean overdensity  $\partial P_{\text{Ly}\alpha}^{\text{1d}}/\partial \bar{\delta}$ . The 1d power spectrum is related to the 3d power spectrum as

$$P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel}) = \int \frac{d^2 \vec{k}_{\perp}}{(2\pi)^2} P_{\text{Ly}\alpha}^{\text{3d}}(k_{\parallel}, \vec{k}_{\perp}), \quad (5)$$

which, in turn, is related to the linear matter power spectrum as [18]

$$P_{\text{Ly}\alpha}^{\text{3d}}(k_{\parallel}, \vec{k}_{\perp}) = b_1^2 (1 + \beta \mu^2)^2 D(k, \mu) P_{\text{lin}}(k), \quad (6)$$

where  $\mu$  is the cosine of the angle between  $\vec{k}_{\parallel}$  and  $\vec{k} = \vec{k}_{\parallel} + \vec{k}_{\perp}$ . The term  $b_1$  represents the linear bias of the Lyman- $\alpha$  transmission,  $\beta$  corresponds to linear redshift-space distortions, and  $D$  encapsulates several non-linearities: Jeans smoothing at small scales under gas pressure, thermal broadening of absorption lines due

to local thermal velocity dispersion and finally non-linear structure formation under gravity (it goes to 1 as  $k$  goes to zero and has an exponential cut-off at small scales). This fitting formula was obtained in [18] by comparing hydrodynamical simulations of the intergalactic medium and assumes that the ionizing UV background is homogeneous. It neglects possible fluctuations in this background or in the density-temperature relation arising from inhomogeneous reionization, as well as astrophysical effects such as galactic winds and quasar effects which would most likely affecting only very small volumes (see [18] and references therein).

In the presence of an overdensity  $\delta$ , the linear power spectrum responds as [20]

$$\frac{\partial \ln P_{\text{lin}}}{\partial \delta} = \frac{68}{21} - \frac{1}{3} \frac{\partial \ln k^3 P_{\text{lin}}}{\partial \ln k}. \quad (7)$$

The linear bias term  $b_1$ , the Kaiser term  $\beta$  and the baryonic non-linearities encapsulated in  $D(k, \mu)$  may also respond to a mean overdensity. We characterize the response of these terms by an effective non-linear bias

$$b_2^{\text{eff}}(k, \mu) = \frac{\partial}{\partial \delta} \ln \left( b_1^2 (1 + \beta \mu^2)^2 D(k, \mu) \right). \quad (8)$$

This quantity combines non-linear bias<sup>1</sup> and the response of redshift-space distortions and non-linear clustering of gas. It will be measured from the bispectrum, and can, in principle, be measured from simulations. The response of the Lyman- $\alpha$  forest power spectrum is thus

$$\frac{\partial P_{\text{Ly}\alpha}^{\text{1d}}(k_{\parallel})}{\partial \delta} = \int \frac{d^2 \vec{k}_{\perp}}{(2\pi)^2} P_{\text{Ly}\alpha}^{3d}(\vec{k}) \left( \frac{\partial \ln P_{\text{lin}}}{\partial \delta} + b_2^{\text{eff}}(k, \mu) \right). \quad (9)$$

Combining Eqs (4) and (9), the CMB lensing - Lyman- $\alpha$  bispectrum becomes the sum of two terms: one representing the response of the linear matter power spectrum, and one for the non-linear and baryonic terms.

### III. LYMAN- $\alpha$ FOREST POWER SPECTRUM

We use quasar spectra from the twelfth data release of SDSS-III/BOSS [31–33]. The continuum fitting is performed using a mean-flux-regulated principal component analysis method described in [31] that was applied to the DR9 in [34]. The domain of the Lyman- $\alpha$  forest of a given quasar spectrum is defined by limits on the rest frame wavelength

$$1041\text{\AA} \leq \lambda_{\text{rf}} \leq 1185\text{\AA}, \quad (10)$$

<sup>1</sup> Note that the first-order non-linear bias  $b_2$  is generally defined through the response of the linear bias to an overdensity via  $\partial b_1 / \partial \delta = b_2 - b_1^2$ , e.g. in [20, 30]. Therefore, if we neglect the Kaiser term and the baryonic term, the quantity  $b_2^{\text{eff}}$  defined here is related to the non-linear bias  $b_2$  as  $b_2^{\text{eff}} = 2(b_2 - b_1^2)/b_1$ .

where  $\lambda_{\text{rf}} = \lambda / (1 + z_{\text{QSO}})$  for a quasar at redshift  $z_{\text{QSO}}$ .

Spectra displaying damped Lyman- $\alpha$  absorption systems (DLAs, identified using the technique described in [35]) or broad absorption lines (BALs, identified by visual inspection in [33]) were discarded. We select quasars with a signal-to-noise ratio in the Lyman- $\alpha$  forest, measured by the BOSS pipeline, greater than 1 and a redshift between 2.15 and 4.0. The noise estimation gives poor results close to the spectrograph blue-end, so we additionally cut parts of the forests below  $z = 2.1$  as in [36]. Finally, we discard quasars lying outside of the Planck lensing mask. These cuts select 87,085 quasars out of the 155,002 in the DR12 catalog.

The flux transmission fraction in the Lyman- $\alpha$  forest at redshift  $z = \lambda / \lambda_{\text{Ly}\alpha}$  is

$$F(z) = \frac{f(z)}{C(z)}, \quad (11)$$

where  $f(z)$  is the measured flux and  $C(z)$  is the estimated continuum. We then estimate the normalized transmitted flux fraction as a function of redshift as

$$\delta_i(z) = \frac{F_i(z)}{\langle F(z) \rangle} - 1 \quad (12)$$

where  $\langle F(z) \rangle$  is the mean flux transmission fraction obtained by averaging over quasars and  $i$  stands for the forest index.

The normalized flux fraction is converted from a function of redshift to a function of radial comoving distance  $\chi(z)$ , which is evaluated using cosmological parameters from Planck 2015 (TT,TE,EE+lowP+lensing+ext) [37]. Because the spacing between pixels of the BOSS spectrograph is logarithmic in wavelength, with  $\Delta(\log_{10} \lambda) = 10^{-4}$ , and because sky emission lines are masked (on average 1.2% of pixels), the spacing in distance space is slightly irregular, albeit monotonically growing. Therefore the Fourier transform  $\tilde{\delta}(k)$  of the normalized transmitted flux fraction  $\delta(\chi)$  is computed using the NFFT library [38]. For a forest of length  $\Delta\chi$  and mean pixel spacing  $\bar{\delta}\chi$ , the smallest mode is  $k_{\text{min}} = 2\pi/\Delta\chi$  while the largest mode is  $k_{\text{max}} = 2\pi/\bar{\delta}\chi$ . Forests have a mean length of 420 – 280 Mpc/h with a spacing of order 0.68 – 0.58 Mpc/h for forests in the redshift range 2.1 – 3.6, giving  $k_{\text{max}} \approx 6$  h/Mpc. These scales are highly affected by non-linear clustering and baryonic effects. Moreover, the power spectrum becomes noisier because of the resolution of the spectrograph (see the spectrograph window function in Eq. (15)), so we restrict our analysis to  $k_{\text{max}} = 1.5$  h/Mpc, corresponding to scales of order 4 Mpc/h, consistent with [36]. We note that the large-scale modes over 60 Mpc/h, i.e.  $k \lesssim 0.1$  h/Mpc may be slightly affected by continuum fitting, though in a way uncorrelated with CMB lensing. The raw power spectrum is obtained by

$$\hat{P}_i^{\text{raw}}(k) = \frac{|\tilde{\delta}_i(k)|^2}{\Delta\chi_i} \quad (13)$$

where  $\Delta\chi_i$  is the comoving length of the  $i$ -th. forest.

Multiple observations of the same quasars allow for an estimation of the noise level  $\sigma_{\text{noise}}^2(\chi)$  for each pixel in the forest. We assume the noise to have a white spectrum, which agrees well with [4, 36], and estimate its power spectrum by averaging

$$P_i^{\text{noise}} = \overline{\sigma_{\text{noise},i}^2} \frac{\pi}{\Delta k} \quad (14)$$

where  $\Delta k = k_{\text{max}} - k_{\text{min}}$ .

The resolution of the spectrograph is of order 1 Mpc/ $h$  and varies slowly with wavelength (by about 10% over one forest). Therefore the spectrograph window function is

$$W_{\text{spectro}}(k, R_i) = \exp\left(-\frac{k^2 R_i^2}{2}\right) \times \text{sinc}\left(\frac{k \bar{\delta}\chi_i}{2}\right), \quad (15)$$

where  $R_i$  is the resolution of the spectrograph averaged over the  $i$ -th forest,

$$R_i = \frac{c(1+z)}{H(z)} \delta_{\text{disp}} \Delta \log \lambda, \quad (16)$$

with  $\delta_{\text{disp}}$  being the measured dispersion in units of  $\Delta \log_{10} \lambda$ . The second term, in which  $\text{sinc}(x) = \sin(x)/x$ , accounts for the pixelization. Finally, the estimator of the one-dimensional power spectrum of the Lyman- $\alpha$  forest is given by

$$\hat{P}_{\text{Ly}\alpha}^{\text{1d}}(k, z) = \left\langle \frac{\hat{P}_i^{\text{raw}}(k) - P_i^{\text{noise}}}{W_{\text{spectro}}^2(k, R_i)} \right\rangle_{i \in z}. \quad (17)$$

where the average is over forests falling in the redshift range.

This straightforward measurement of the one-dimensional power spectrum does not correct for subtle instrumental systematics dealt with in [36]. However, these effects are less important in a cross-correlation measurement, and a precise estimation of the power spectrum is beyond the scope of this paper.

#### IV. CMB LENSING

We use the publicly available<sup>2</sup> CMB lensing convergence map from Planck 2015 [39]. We apply a Wiener filter to the map using the provided signal and noise power spectra to get the minimum-variance linear estimator for the CMB convergence  $\kappa^{\text{WF}}(\theta_i)$  in the direction of each Lyman- $\alpha$  forest.

<sup>2</sup> See <http://pla.esac.esa.int/pla/>.

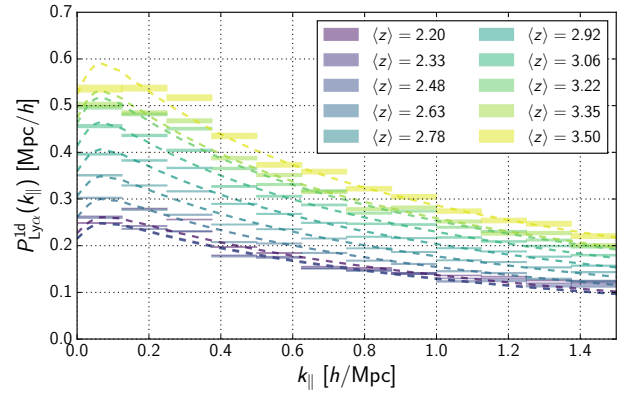


FIG. 2. One-dimensional power spectrum of the Lyman- $\alpha$  forest  $P_{\text{Ly}\alpha}^{\text{1d}}$  in 10 redshift bins. The redshifts indicate the mean value of the middle redshifts of the forests. Error bars are computed from the weighted empirical covariance of the power spectra of different forests. Colored boxes represent the measured spectra averaged over these redshifts bins with their uncertainties and  $k$ -bins width. Dashed lines are theoretical curves with fitted bias  $b_1(z)$ .

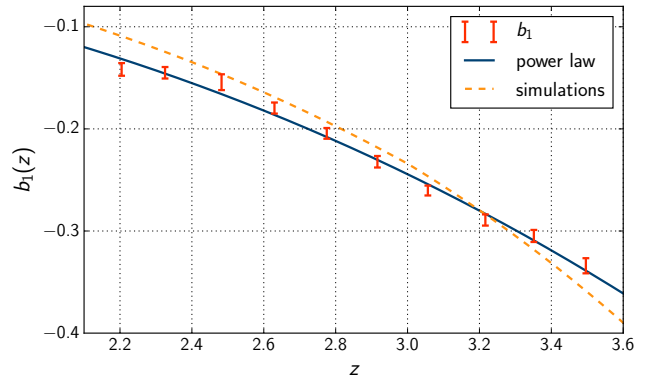


FIG. 3. Lyman- $\alpha$  flux transmission linear bias  $b_1$  as a function of redshift. The data points are depicted with error bars and the power law in  $(1+z)$  (see text) is represented by the blue solid curve. The dashed orange curve represents the linear bias measured from hydrodynamical simulations in [18]. Remark that the transmission is theoretically anti-correlated with the amount of hydrogen, hence the negative sign.

#### V. RESULTS & INTERPRETATION

The first step of our analysis is to measure the one-dimensional power spectrum in order to obtain the linear bias  $b_1(z)$  from Eq. (6) as a function of redshift. The theoretical curves are computed using Eqs. (5) and (6) with parameters, except for the linear bias that we aim at fitting, measured from simulations in [18]. We divide our forest sample into 10 linearly-spaced redshift bins using the central redshift of each forest. Each power spectrum



is given a scale-dependent minimum-variance weight

$$w_i(k) \propto n_i^{\text{pix}} \left( P_{\text{Ly}\alpha}^{\text{1d, fid}}(k, z_i) + \frac{P_i^{\text{noise}}}{W_{\text{spectro}}^2(k, R_i)} \right)^{-2}, \quad (18)$$

where  $n_i^{\text{pix}}$  is the number of pixels of the  $i$ -th forest. The fiducial power spectrum  $P_{\text{Ly}\alpha}^{\text{1d, fid}}(k, z)$  is computed using the linear bias measured from simulations in [18], which will only be used in the weights.

In order to take into account possible wavelength-dependent bias in the noise estimation, we allow for a common rescaling of the estimated noise power spectrum in each redshift bin. Precisely, we introduce a coefficient  $\alpha_z$  in front of  $P_i^{\text{noise}}$  in Eq. (17), common to all spectra in each redshift bin and fit this parameter jointly with the linear bias  $b_1(z)$  in each redshift bin. The estimated one-dimensional power spectrum is shown in Fig. 2 together with theoretical curves with best fit biases. The best fit of the linear bias  $b_1(z)$  is shown in Fig. 3 with error bars including the marginalization over  $\alpha_z$ . We fit this result with a power law in  $(1+z)$  of the form  $b_1(z) = a(1+z)^b$  and find  $a = -0.00507$  and  $b = 2.79$ . It is represented by the solid blue curve in Fig. 3 and is in fairly good agreement with the bias measured in hydrodynamic simulations in [18] (which we only used in the weights).

The next step of our analysis is to compute the weighted unbiased covariance of the lensing convergence and the one-dimensional power spectrum. Quasars have a significant contribution to the lensing of the CMB because the lensing efficiency  $W_\kappa$  peaks at  $z \sim 2$ . Therefore, we expect the mean convergence in the directions of quasars to be positive, and indeed find  $10^4 \times \langle \kappa_i^{\text{WF}} \rangle = 1.35 \pm 0.52$ . This value is consistent with the expected amplitude  $\kappa = (\Lambda_\kappa * \Sigma) / \bar{\rho} \sim 1.5 \times 10^{-4}$  where  $\Sigma$  is the projected density of the haloes hosting the quasars (computed for a NFW profile [40] with a halo mass  $M_h \sim 2 \times 10^{12} M_\odot / h$  and redshift 2.5 [41]) convolved with the Wiener filter and  $\bar{\rho}$  is the mean matter density. With the aim of measuring the correlation between our two probes, we subtract the mean value  $\langle \kappa_i^{\text{WF}} \rangle$  in the computation of the covariance. So as to decrease the effects of noise in this measurement, we also subtract the mean value of the power spectrum in each  $k$ -bin. The estimator for the correlation of CMB lensing and fluctuations in the Lyman- $\alpha$  forest, *i.e.* the CMB lensing – Ly $\alpha$  integrated bispectrum, is defined as

$$\hat{B}_{\kappa, \text{Ly}\alpha}(k_\parallel) \triangleq \text{Cov}_{w(k_\parallel)} [\kappa^{\text{WF}}, P_{\text{Ly}\alpha}^{\text{1d}}(k_\parallel)], \quad (19)$$

where

$$\text{Cov}_w[x, y] = \mathcal{N} \times \sum_i w_i (x_i - \langle x \rangle) (y_i - \langle y \rangle) \quad (20)$$

with the normalization  $\mathcal{N} = \sum_i w_i / \left( (\sum_i w_i)^2 - \sum_i w_i^2 \right)$ . The mean values  $\langle \kappa^{\text{WF}} \rangle$  and  $\langle P^{\text{1d}}(k_\parallel) \rangle$  are computed using the same weights as well. The measured values in each  $k$ -bin are shown in purple in Fig. 4.

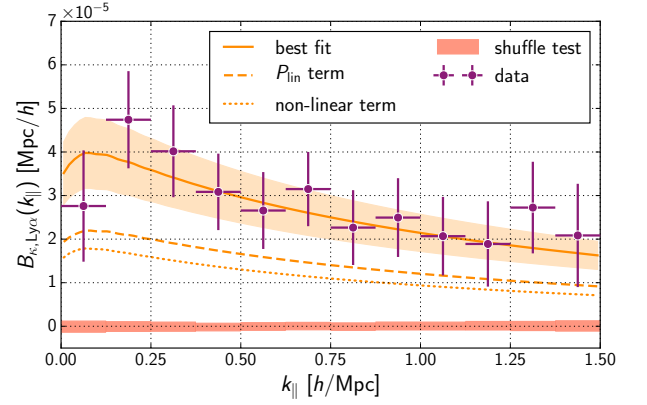


FIG. 4. Integrated bispectrum of CMB lensing and fluctuations in the Lyman- $\alpha$  forest. The Wiener-filtered CMB lensing is measured in the direction of quasars for which we measure the Lyman- $\alpha$  forest one-dimensional power spectrum in the range  $k_\parallel \sim 0.1 - 1.5 \, h/\text{Mpc}$ . Data points (in purple) show a signal measured at  $5 \sigma$ . The theoretical curve (solid orange) is the sum of two terms: the response of the linear matter power spectrum (dashed), and the response of the non-linear terms in the Lyman- $\alpha$  power spectrum (non-linear bias  $b_2$ , Kaiser term and baryonic non-linear term  $D$ ) (dotted). While the first involves no free parameter, the latter has an amplitude characterized by the effective non-linear bias  $b_2^{\text{eff}} = 1.16 \pm 0.53$ , see Eq. (8). The orange area represents the  $1 \sigma$  uncertainty on this non-linearity amplitude. We test that our estimator is coherent with zero in the case of no correlation by a shuffling method (thin red boxes, expanded 10 times for visibility).

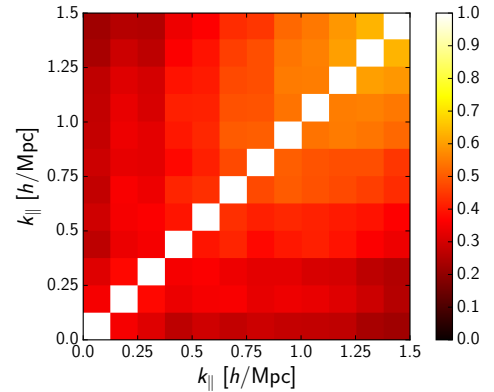


FIG. 5. Correlation matrix of the data-points between  $k$ -bins computed by shuffling the indices of one of the variables. It shows an important correlation ranging from 20% up to almost 65% for the large  $k$  modes.

To compute the covariance matrix for the various  $k$ -bins, we proceed by computing the signal repeatedly with shuffled indices in  $\kappa_i^{\text{WF}}$ . More precisely, for a given random permutation  $\sigma$  of the quasar indices, we compute  $\text{Cov} [\kappa_{\sigma(i)}^{\text{WF}}, P_{\text{Ly}\alpha, i}^{\text{1d}}(k_\parallel)]$  and repeat  $N = 10,000$  times. We then estimate the mean value (thin red boxes on Fig. 4)

and the empirical covariance. The corresponding matrix of correlation coefficients is shown in Fig. 5.

Finally, we aim at comparing our theoretical model and fitting a value of the effective non-linear bias  $b_2^{\text{eff}}$  defined in Eq. (8). We measure a single number, *i.e.* a scale and redshift averaged non-linear bias **integrated over  $\mu$** , characterizing the non-linear response in our sample. **This parametrization is incomplete, but sufficient given the signal-to-noise ratio.** For each line of sight, we evaluate the expected signal using Eq. (4) given the redshift range  $[z_{\text{min}}, z_{\text{max}}]$  of the forest, and the linear bias  $b_1(z)$  from the power law best fit. We then weight the theoretical expected value by the weights in Eq. (18). The best fit value is  $b_2^{\text{eff}} = 1.16 \pm 0.53$ . The theoretical curve (in orange in Fig. 4) is the sum of two contributions, one from the linear power spectrum (dashed line) and the other from the non-linear terms (dotted line).

Using the covariance matrix obtained by our shuffling method and the measured data points, we find a  $\chi^2$  value for the null hypothesis  $\chi_{\text{null}}^2 = 30.1$  for 12 data points. The probability to exceed is 0.27% and the null hypothesis is therefore rejected at a significance of  $3.0 \sigma$ . For the best fit in  $b_2$ , we find  $\chi_{\text{best-fit}}^2 = 5.37$ , a small value that could be explained by over-estimated error bars, which would lead to a better detection. The signal-to-noise ratio for the detection is  $\text{SNR} = \sqrt{\chi_{\text{null}}^2 - \chi_{\text{best-fit}}^2} = 4.97$ , hence this constitutes a  $5 \sigma$  detection of the non-linear response of the Lyman- $\alpha$  power spectrum.

## VI. NULL TESTS

In order to assess the cosmological nature of this signal, we proceed to a number of null tests.

First, we make sure that the correlation estimator is consistent with zero in the case of no correlation. To do so, we compute the mean of the values of the signal measured with shuffled indices in  $\kappa_i^{WF}$  as we expect different lines to be uncorrelated. The result is consistent with zero, as shown on Fig. 4 by the thin red boxes.

Second, we want to verify that the signal we measure does not originate in a possible correlation of the lensing convergence in the directions of quasars with their intrinsic properties. We split our sample in two equal parts according to the median values of various quasar parameters. For each of these parameters, we measure the signal in the two sub-samples using Eq. (19) and compute the difference. The results are shown in Fig. 6 and the meaning of the tested parameters' names are detailed in Table I. We test for galactic latitude of the quasars and galactic HI column density in their directions to verify that the signal is not related to galactic foregrounds. We test for quasar redshift and find no significant variation of the signal. We also test for various intrinsic properties of the quasars linked to their masses: colors (PSF magnitudes in the  $g$  and  $i$  bands), near and far UV fluxes (from GALEX [42]) and quasar spectral index. We also test for

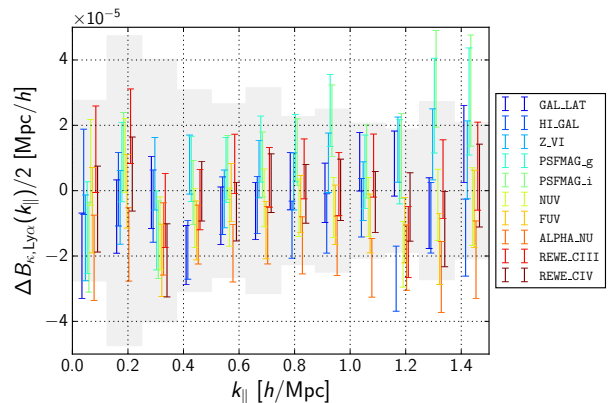


FIG. 6. Null tests for various quasar properties. For each parameter, the sample is split in two parts according to the median value. We then compute the CMB lensing – Ly $\alpha$  integrated bispectrum for these two sub-samples, using Eq. (19) and compute the difference (divided by 2 in order to have the same error bars as the signal itself). The greyed area is delimited by the absolute level of the measured bispectrum for the full sample (as in Fig. 4).

TABLE I. Results of the null tests seeking for correlation between the lensing signal and intrinsic quasars properties.

Label	Description	Null test
GAL_LAT	Galactic latitude	$1.6 \sigma$
HI_GAL	log of galactic HI column density	$1.1 \sigma$
Z_VI	Quasar redshift from visual inspection	$0.2 \sigma$
PSFMAG_g	PSF magnitude (flux in $g$ band)	$1.0 \sigma$
PSFMAG_i	PSF magnitude (flux in $i$ band)	$1.4 \sigma$
NUV	Near UV flux (from GALEX)	$0.7 \sigma$
FUV	Far UV flux (from GALEX)	$0.5 \sigma$
ALPHA_NU	Quasar spectral index	$0.5 \sigma$
REWE_CIII	Rest equivalent width of CIII emission	$1.1 \sigma$
REWE_CIV	Rest equivalent width of CIV emission	$0.3 \sigma$

contamination by carbon lines using the rest equivalent width of the emissions of CIII and CIV. We find that all tests are consistent with zero at the  $0.2 - 1.6 \sigma$  level. Lack of data prevents us from testing the contamination from the SiIV line; however, it is at most a 5% effect according to [36].

## VII. DISCUSSION

We have presented the first detection of a cross-correlation between the Lyman- $\alpha$  forest of quasars and the gravitational lensing of the CMB. Our understanding of this correlation is based on the response of small-scale fluctuations in the matter density, measured by the one-dimensional power spectrum of the transmission in the

Lyman- $\alpha$  forest, to large-scale overdensities probed by the convergence of CMB lensing. This signal corresponds to a bispectrum in the squeezed limit configuration where the two small-scale modes are of order  $k \sim 0.1 - 1$  Mpc/ $h$ . This is the first measurement of the CMB lensing – Ly $\alpha$  integrated bispectrum, and it measures the non-linearity in the Lyman- $\alpha$  forest. Finally, this new observable tests our understanding of the relation between neutral hydrogen and dark matter.

We measured the one-dimensional power spectrum and the linear bias of the Lyman- $\alpha$  forest, finding values consistent with hydrodynamical simulations. Even though the power spectrum is sensitive to a number of systematic effects, these are much less important in a cross-correlation measurement like the integrated bispectrum that we computed. The theoretical bispectrum is the sum of two contributions: the response of the linear matter power spectrum, theoretically well-understood and involving no free parameters, and the response of the bias and non-linear terms, computed up to an effective non-linear bias  $b_2^{\text{eff}}$  which we have fitted. We believe this model provides a reasonable explanation of the observed signal.

However, we notice that our interpretation of the measured bispectrum is limited by theoretical uncertainties mainly related to baryonic physics. That is, the term  $D(k, \mu)$ , taken from [18], encodes a number of effects that are significant at very small scales (of order  $k \sim 60$   $h$ /Mpc, see [43]), but the integral of the three-dimensional power spectrum in Eq. (5) gets contributions from  $k$ -modes greater than 10  $h$ /Mpc, and we cannot neglect these effects. Moreover, this term and the redshift-space distortion term  $\beta$  may also respond to large-scale overdensities. Therefore, the effective non-linear bias term  $b_2^{\text{eff}}$  encompasses several uncertain contributions: it could be compared with simulations, providing both a valuable check for the simulation assumptions while shedding light on the relation between Lyman- $\alpha$  and dark matter.

Another uncertainty arises from the fact that the ionizing UV flux of the quasars reduces the amount of neutral hydrogen around them, a phenomenon known as the proximity effect. Because overdense regions radiate more, the bias of neutral hydrogen  $b_{\text{H I}}(k)$  becomes negative at scales larger than  $k \sim 0.01$   $h$ /Mpc [17]. This impacts the H I power spectrum for scales  $k \lesssim 0.1$   $h$ /Mpc and may also affect our measurement in the lowest  $k$ -bin.

In the future, hydrodynamical simulations could be compared to this new observable and used to model the dependence of the various non-linear terms in the one-dimensional power spectrum on the mean overdensity. **In particular, the anisotropic linear bias of the Lyman- $\alpha$  forest, i.e. its dependence on angle  $\mu$ , could be explored.** Meanwhile, this measurement can inform

simulations and help constrain the relation between intergalactic gas and dark matter. **Other avenues of exploration are to study the dependence of the signal on redshift and perpendicular separation  $r_{\perp}$ . The upcoming SDSS-IV/eBOSS [44] data, covering a broad redshift range, will improve the signal-to-noise ratio and help measure the redshift dependence. Combined with a high signal-to-noise ratio CMB lensing map, it could allow for a measurement of the angular correlation between the Lyman- $\alpha$  forest and large-scale overdensities. Finally, because this bispectrum is sensitive to small scales observed in the Lyman- $\alpha$  forest, it could also provide additional constraints on the total mass of neutrinos and be used as a tool to study alternative models of dark matter predicting small-scale cut-offs.**

## ACKNOWLEDGMENTS

The authors warmly thank the referee for interesting and useful questions and comments.

C.D., E.S. and E.A. acknowledge support from CNRS grant PICS APC-Princeton. C.D., E.S. and D.N.S. acknowledge support from NSF grant no. AST-1311756. K.-G.L. acknowledges support for this work by NASA through Hubble Fellowship grant HF2-51361 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555..

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.



- (2015).
- [3] K.-G. Lee *et al.*, *ApJ* **799**, 196 (2015).
  - [4] M. McQuinn and M. White, *Mon. Not. R. Astron. Soc.* **415**, 2257 (2011).
  - [5] N. Palanque-Delabrouille *et al.*, *J. Cosmol. Astropart. Phys.* **11**, 011 (2015).
  - [6] N. Palanque-Delabrouille *et al.*, *J. Cosmol. Astropart. Phys.* **02**, 045 (2015).
  - [7] M. Viel, G. D. Becker, J. S. Bolton, and M. G. Haehnelt, *Phys. Rev. D* **88**, 043502 (2013).
  - [8] U. Seljak, A. Makarov, P. McDonald, and H. Trac, *Phys. Rev. Lett.* **97**, 191303 (2006).
  - [9] N. Afshordi, P. McDonald, and D. N. Spergel, *ApJ* **594**, L71 (2003).
  - [10] A. Slosar *et al.*, *J. Cosmol. Astropart. Phys.* **09**, 001 (2011).
  - [11] A. Slosar *et al.*, *J. Cosmol. Astropart. Phys.* **2013**, 026 (2013).
  - [12] N. G. Busca *et al.*, *A&A* **552**, A96 (2013).
  - [13] T. Delubac *et al.*, *A&A* **574**, A59 (2015).
  - [14] R. A. C. Croft, D. H. Weinberg, N. Katz, and L. Hernquist, *ApJ* **495**, 44 (1998).
  - [15] A. Borde *et al.*, *J. Cosmol. Astropart. Phys.* **07**, 005 (2014).
  - [16] G. Rossi *et al.*, *A&A* **567**, A79 (2014).
  - [17] A. Pontzen, *Phys. Rev. D* **89**, 083010 (2014).
  - [18] A. Arinyo-i Prats, J. Miralda-Escudé, M. Viel, and R. Cen, *J. Cosmol. Astropart. Phys.* **12**, 017 (2015).
  - [19] E. Aubourg *et al.*, *Phys. Rev. D* **92**, 123516 (2015).
  - [20] C.-T. Chiang, C. Wagner, F. Schmidt, and E. Komatsu, *J. Cosmol. Astropart. Phys.* **05**, 048 (2014).
  - [21] E. Schaan, M. Takada, and D. N. Spergel, *Phys. Rev. D* **90**, 123523 (2014).
  - [22] Y. Li, W. Hu, and M. Takada, *Phys. Rev. D* **90**, 103530 (2014).
  - [23] M. Zaldarriaga, U. Seljak, and L. Hui, *ApJ* **551**, 48 (2001).
  - [24] A. Vallinotto, S. Das, D. N. Spergel, and M. Viel, *Phys. Rev. Lett.* **103**, 091304 (2009).
  - [25] A. Vallinotto, M. Viel, S. Das, and D. N. Spergel, *ApJ* **735**, 38 (2011).
  - [26] W. Hu and T. Okamoto, *ApJ* **574**, 566 (2002).
  - [27] T. Okamoto and W. Hu, *Phys. Rev. D* **67**, 083002 (2003).
  - [28] D. N. Limber, *ApJ* **117**, 134 (1953).
  - [29] M. LoVerde and N. Afshordi, *Phys. Rev. D* **78**, 123506 (2008).
  - [30] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, *J. Cosmol. Astropart. Phys.* **09**, 007 (2016).
  - [31] K.-G. Lee, N. Suzuki, and D. N. Spergel, *Astron. J.* **143**, 51 (2012).
  - [32] K. S. Dawson *et al.*, *Astron. J.* **145**, 10 (2013).
  - [33] I. Pâris *et al.*, (2016).
  - [34] K.-G. Lee *et al.*, *Astron. J.* **145**, 69 (2013).
  - [35] P. Noterdaeme *et al.*, *A&A* **547**, L1 (2012).
  - [36] N. Palanque-Delabrouille *et al.*, *A&A* **559**, A85 (2013).
  - [37] Planck Collaboration *et al.*, *arXiv* arXiv:1502.01589 (2015).
  - [38] J. Keiner, S. Kunis, and D. Potts, *ACM Trans. Math. Softw.* **36**, 1 (2009).
  - [39] Planck Collaboration *et al.*, *arXiv* arXiv:1502.01591 (2015).
  - [40] J. F. Navarro, C. S. Frenk, and S. D. M. White, *ApJ* **462**, 563 (1996).
  - [41] M. White *et al.*, *Mon. Not. R. Astron. Soc.* **424**, 933 (2012).
  - [42] A. S. G. Robotham and S. P. Driver, *Mon. Not. R. Astron. Soc.* **413**, 2570 (2011).
  - [43] A. M. Cieplak and A. Slosar, *J. Cosmol. Astropart. Phys.* **03**, 016 (2016).
  - [44] K. S. Dawson *et al.*, *Astron. J.* **151**, 44 (2016).