

# 1

## Validation des mocks

---

Ce chapitre a pour vocation de présenter l’analyse menée sur les mocks, afin de valider leur construction et le choix des différents paramètres. Nous présentons d’abord les estimateurs des diverses fonctions de corrélations, puis les modèles ajustés sur celles-ci. Nous donnons premièrement le modèle ajusté sur les données et présenté dans prov puis le modèle modifié ajusté sur les mocks. Enfin, nous présentons l’analyse des 100 réalisations de mocks produites. Le résultat de l’analyse des données, utilisée pour choisir les paramètres  $\text{Ly}\alpha$  des mocks, est présentée dans le chapitre suivant. Toutes ces analyses sont produites avec le code `picca`. Ce code permet de calculer les champs  $\delta$ , les fonctions de corrélations, les matrices de distorsion, les matrices de covariance ainsi que de réaliser l’ajustement du modèle.

## 1 Les estimateurs

Nous présentons ici les estimateurs utilisés pour calculer la fonction d’auto-corrélation du  $\text{Ly}\alpha$ , la fonction de corrélation croisée  $\text{Ly}\alpha$ -QSO, ainsi que la fonction d’auto-corrélation d’objets ponctuels tels les quasars. Dans le cas des objets ponctuels, il est nécessaire de produire un catalogue aléatoire d’objets qui prend en compte la complétude du relevé. Il est aussi nécessaire d’estimer l’homogénéité de la sélection des cibles, afin de pondérer chaque quasar par un poids qui tient compte de celle-ci. Pour le calcul des fonctions de corrélation impliquant le  $\text{Ly}\alpha$ , il n’est pas utile d’avoir un catalogue aléatoire. En effet, le champ  $\delta_F$  nous donne directement accès, à un biais près, au contraste de densité. Nous pouvons donc calculer la fonction de corrélation avec une formule du type de l’équation ??, ce qui n’est pas possible dans le cas des objets ponctuels. Ceci nous permet aussi de ne pas tenir compte de l’homogénéité de la sélection des cibles, et donc d’utiliser des poids qui servent uniquement à maximiser le rapport signal sur bruit.

Pour le calcul du champ  $\delta_F$ , nous distinguons deux cas : le cas où nous analysons les *raw mocks* et le cas où nous analysons les *cooked mock* ou les données. Les raw mocks (les mocks bruts, non préparés) désignent les mocks avant d’avoir tourné le code `quickquasars`. Les fonctions de corrélation sont alors calculées en utilisant directement la fraction de flux transmise  $F$  donnée dans les fichiers de transmissions. Etant donné que pour chaque forêt, nous avons directement accès à  $F$ , le champ  $\delta_F$  est donné par

$$\delta_F = \frac{F}{\bar{F}(z)} - 1. \quad (1.1)$$

Les cooked mocks quant à eux désignent les mocks après avoir appliqué `quickquasars` à chaque forêt. Dans ce cas, comme pour les données, la fraction de flux transmise  $F$  n’est pas accessible. Il faut alors utiliser la procédure d’ajustement du continuum décrite dans la section ??. Une fois le champ  $\delta_F$  calculé, nous pouvons calculer les différentes fonctions de corrélation.

### 1.1 L’auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$

Comme expliqué dans le chapitre d’introduction, la fonction d’auto-corrélation du champ d’absorption  $F$  du  $\text{Ly}\alpha$  est définie comme

$$\xi_{\text{Ly}\alpha}(\vec{r}) = \langle \delta_F(\vec{r}') \delta_F(\vec{r} + \vec{r}') \rangle. \quad (1.2)$$

Nous pouvons utiliser comme estimateur

$$\hat{\xi}(\vec{r}) = \langle \delta_i \delta_j \rangle, \quad (1.3)$$

où  $\langle . \rangle$  désigne la moyenne sur tous les pixels  $i$  et  $j$  qui vérifient  $\vec{r}_{ij} = \vec{r}$ . Afin d’estimer  $\xi_{\text{Ly}\alpha}$ , une grille en  $(r_{\parallel}, r_{\perp})$  est créée. Les bins mesurent  $4 h^{-1}$  Mpc dans chaque direction. **Cette taille de bin est**

choisie de façon à avoir suffisamment de bins pour résoudre correctement la région du pic BAO, mais aussi de façon à ne pas avoir trop de bins pour le calcul de la matrice de covariance. La taille de bin choisie et la gamme en  $r$  utilisée produisent déjà une matrice de covariance de  $2500 \times 2500$  bins. Une fois cette grille construite, la séparation de chaque paires  $(\Delta\theta, \Delta z)$  est transformée en distance comobile afin d'obtenir la séparation  $(r_{\parallel}, r_{\perp})$  :

$$\begin{cases} r_{\parallel} = [D_C(z_i) - D_C(z_j)] \cos\left(\frac{\Delta\theta}{2}\right), \\ r_{\perp} = [D_M(z_i) - D_M(z_j)] \sin\left(\frac{\Delta\theta}{2}\right), \end{cases} \quad (1.4)$$

où  $D_C$  est la distance comobile le long de la ligne de visée, et  $D_M$  la distance comobile transverse (voir section ??, paragraphe *Les distances*).  $z_i$  et  $z_j$  sont les redshifts des pixels  $i$  et  $j$ . Puis, pour chaque bin  $A$  de la grille en  $(r_{\parallel}, r_{\perp})$ , toutes les paires de pixels dont la distance de séparation se trouve dans ce bin  $A$  sont considérées. La fonction de corrélation dans ce bin est alors donnée par

$$\xi_A = \frac{\sum_{(i,j) \in A} w_i w_j \delta_i \delta_j}{\sum_{(i,j) \in A} w_i w_j}, \quad (1.5)$$

où  $w_i$  est le poids associé au pixel  $i$ . Ces poids sont définis dans l'équation ?? . La correction de la dépendance en redshift du biais du  $\text{Ly}\alpha$  permet de donner plus de poids aux pixels à grand redshift, où l'amplitude de la fonction de corrélation est plus importante. Le calcul de la fonction de corrélation est donc effectué à l'aide d'une double boucle sur les pixels. Afin de réduire le temps de calcul, la fonction de corrélation est calculée uniquement pour les paires de pixels pour lesquelles  $r_{\parallel}$  et  $r_{\perp}$  sont inclus dans  $[0; 200] h^{-1}$  Mpc. La fonction de corrélation est donc calculée dans  $50 \times 50 = 2500$  bins. Les paires formées par des pixels provenant de la même forêt sont exclues du calcul afin d'éviter que les erreurs sur l'ajustement du continuum biaisent la mesure de la fonction de corrélation.

L'analyse  $\text{Ly}\alpha$  des données complète d'eBOSS (**prov**) utilise deux fonctions de corrélation distinctes : les fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . La région  $\text{Ly}\beta$  ne fournit pas suffisamment de pixels pour pouvoir calculer la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\beta) \times \text{Ly}\alpha(\text{Ly}\beta)$  et y détecter le pic BAO. De plus, le gain de statistique que représente cette fonction de corrélation est négligeable. L'utilisation de ces deux fonctions de corrélation permet de mesurer la position du pic BAO avec une plus grande précision. Dans ce manuscrit, nous nous intéressons à la mesure de  $b_{\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  afin de construire correctement nos mocks. Pour limiter les potentielles systématiques, nous considérons donc uniquement la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . Le graphique de droite de la figure ?? présente les distributions pondérées en redshift des paires  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . L'analyse des données DR16 présentée dans ce manuscrit considère donc uniquement la distribution indiquée en orange.

## 1.2 La corrélation croisée $\text{Ly}\alpha \times \text{QSO}$

Nous donnons dans cette section l'estimateur de la fonction de corrélation croisée  $\text{Ly}\alpha \times \text{QSO}$ . De la même manière que précédemment, le calcul s'effectue dans des bins en  $(r_{\parallel}, r_{\perp})$ . L'estimateur de la fonction de corrélation dans le bin  $A$  est donnée par

$$\hat{\xi}_A = \frac{\sum_{(i,j) \in A} w_i w_j \delta_i}{\sum_{(i,j) \in A} w_i w_j}, \quad (1.6)$$

où l'indice  $i$  court sur les pixels des forêts et  $j$  sur les quasars pour lesquels la distance de séparation est comprise dans le bin  $A$ . Les paires  $ij$  où le pixel  $i$  appartient à la forêt du quasar  $j$  sont rejetées.

Similairement au  $\text{Ly}\alpha$ , les poids  $w_j$  associés aux quasars favorisent les quasars à plus grand redshift. Ces poids sont définis comme

$$w_j = \left( \frac{1 + z_j}{1 + 2.25} \right)^{\gamma_{\text{QSO}}}, \quad (1.7)$$

où  $\gamma_{\text{QSO}} = 1.44 \pm 0.08$  (DU MAS DES BOURBOUX et al. 2019). Dans le calcul de la fonction de corrélation croisée  $\text{Ly}\alpha \times \text{QSO}$ , nous nous restreignons aux paires pour lesquelles  $r_\perp \in [0; 200] h^{-1} \text{ Mpc}$ . Contrairement à l'auto-corrélation du  $\text{Ly}\alpha$ , la fonction de corrélation croisée n'est pas symétrique en  $r_\parallel$ . Cette asymétrie est produite par l'erreur systématique sur la mesure du redshift des quasars, et aussi par le rayonnement des quasars qui n'est pas isotrope. La fonction de corrélation croisée est donc calculée dans les bins pour lesquels  $r_\parallel \in [-200; 200] h^{-1} \text{ Mpc}$ . Ceci représente  $100 \times 50 = 5000$  bins.

### 1.3 Le spectre de puissance à une dimension

Comme expliqué dans le chapitre précédent, nous ajustons le spectre de puissance  $P_s$  appliqué à  $\delta_s$  de façon à obtenir un spectre de puissance à une dimension  $P^{1\text{D}}$  en accord avec les données. Nous présentons donc maintenant la mesure de ce spectre de puissance sur les raw mocks. Pour chaque forêt, nous appliquons une transformation de Fourier au champ  $\delta_F$  afin d'obtenir le champ  $\delta_k$ . Puis, le spectre de puissance de cette forêt est obtenu comme

$$\hat{P}^{1\text{D}}(k) = \langle \delta_k^2 \rangle. \quad (1.8)$$

Nous répétons cette procédure pour toutes les forêts, puis le spectre de puissance total est obtenu comme la moyenne du spectre de puissance de chaque forêt.

En ce qui concerne les données et les cooked mocks, la mesure du spectre de puissance à une dimension est plus complexe. Un certain nombre d'effets liés à la mesure doivent être pris en compte. En particulier, le bruit et la résolution doivent être estimés et pris en compte pour ne pas fausser l'estimation du spectre de puissance. Cette analyse est détaillée dans CHABANIER et al. (2018).

### 1.4 La fonction de corrélation à une dimension

Nous présentons maintenant l'estimateur de la fonction de corrélation à une dimension. Similairement au  $P^{1\text{D}}$ , cette fonction de corrélation est calculée en considérant uniquement les paires de pixels appartenant à la même forêt. Nous l'estimons comme

$$\hat{\xi}_A^{1\text{D}} = \frac{\sum_{(i,j) \in A} w_i w_j \delta_i \delta_j}{\sum_{(i,j) \in A} w_i w_j}, \quad (1.9)$$

où  $w_i$  est le poids associé au pixel  $i$ . Ces poids sont définis dans l'équation ?? La fonction de corrélation à une dimension  $\xi^{1\text{D}}$  correspond à la fonction de corrélation à trois dimensions pour  $\mu = 1$ . En considérant uniquement les paires le long de la ligne de visée, elle permet de mettre en évidence les autres absorbeurs présents dans le milieu intergalactique, tels les métaux. La longueur d'onde d'absorption au repos de ces espèces étant différente de celle du  $\text{Ly}\alpha$ , les séparations physiques  $r_\parallel = 0$  sont reconstruite à une séparation  $r_\parallel$  non nulle (voir l'explication de la matrice des métaux, section 4.1). Ceci produit des pics dans la fonction de corrélation à une dimension. Ces pics sont d'autant plus visibles quand la fonction de corrélation est représentée en fonction de  $\lambda_i/\lambda_j$ , où  $\lambda_i$  et  $\lambda_j$  sont les longueurs des deux pixels formant chaque paire.

### 1.5 L'auto-corrélation $\text{QSO} \times \text{QSO}$

# prov To write

## 2 Les matrices de distorsion

L'ajustement du continuum nécessaire au calcul du champ  $\delta_F$  dans les données et les cooked mocks biaise le champ mesuré. Cependant, grâce à la transformation (équation ??) décrite dans la section ??), l'effet sur la fonction de corrélation peut-être pris en compte. L'idée est la suivante : modéliser la distorsion induite par l'ajustement du continuum et par la transformation ?? sur la fonction de corrélation, appliquer cette distorsion au modèle, puis ajuster le modèle "distordu" aux données. L'effet le plus important de l'ajustement du continuum et de la transformation ?? est de forcer la moyenne et la pente de chaque région spectrale à être nulle. Ceci induit des corrélations entre les pixels d'une même région spectrale, et donc le long de la ligne de visée. Ainsi, au premier ordre, nous pouvons considérer que chaque  $\delta_F$  après distorsion d'une forêt est une combinaison linéaire de tous les  $\delta_F$  avant distorsion de cette forêt. La fonction de corrélation distordue dans le bin  $A$  peut alors être reliée à la vraie fonction de corrélation comme

$$\xi_{\text{distorsion}}(A) = \sum_B D_{AB} \xi_{\text{vraie}}(B), \quad (1.10)$$

où  $D_{AB}$  est appelée la *matrice de distorsion*. Celle-ci s'exprime en fonction du terme  $\eta_{ij}^q$ , défini dans l'équation ?. Pour l'auto-corrélation, elle s'exprime comme

$$D_{AB} = \frac{1}{W_A} \sum_{(i,j) \in A} w_i w_j \left( \sum_{(i',j') \in B} \eta_{ii'} \eta_{jj'} \right) \quad (1.11)$$

où  $W_A = \sum_{(i,j) \in A} w_i w_j$  est le poids du bin  $A$ . Pour la corrélation croisée, la matrice de distorsion est donnée par

$$D_{AB} = \frac{1}{W_A} \sum_{(i,j) \in A} w_i w_j \left( \sum_{(i',j') \in B} \eta_{ii'} \right). \quad (1.12)$$

Comme précédemment, les indices  $i$  correspondent aux pixels des forêts, et  $j$  aux quasars. A cause de la double somme, le calcul de la matrice de distorsion est très long. Afin de rendre possible l'estimation de cette dernière, le calcul est fait sur 1 % des paires, tirées au hasard (#prov Bautista 2017 montre que c'est OK avec 5%, est-ce que y a une etude qui montre que c'est ok avec 1% ? Dans DR14, ils utilisent 5% je pense).

La figure ?? (#prov faire la figure) montre l'effet de la distorsion sur l'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  dans les mocks. La fonction de corrélation est présentée dans quatres gammes en  $\mu$ . La courbe noire indique la fonction de corrélation calculée sur les raw mocks, et la courbe bleu la fonction de .... #prov Refaire la figure 11 de Bautista : CF des raw mocks (stack) + CF sur les cooked mocks (stack) avec le fit sur les raw + fit \* DM

La matrice de distorsion est un objet uniquement géométrique. Son calcul est indépendant du champ  $\delta_F$ . Elle ne dépend uniquement de la géométrie du relevé et de la distribution des poids. La figure ?? montre la différence entre ... #prov montrer la difference du fit d'une réa avec sa DM et avec une autre DM et/ou le stack de 10 avec les 10 DM ou dix fois la meme. Est-ce que les DM eboss-0.0 et eboss-0.2 sont différentes ?

Ainsi, il n'est pas nécessaire de calculer la matrice de distorsion pour les 100 réalisations des mocks. Dans l'analyse présentée dans la suite de ce chapitre, nous calculons la matrice de distorsion pour l'auto-corrélation et pour la corrélation croisée une seule fois (#prov ou une fois pour chaque run de quickquasars?). L'ajustement de chaque fonction de corrélation utilise une de ces deux matrices de distorsion.

### 3 Les matrices de covariance

Afin de réaliser l'ajustement de chaque fonction de corrélation, nous avons besoin de calculer les matrices de covariance associées à ces fonctions de corrélation. La covariance de la fonction de corrélation  $\xi$  dans le bin  $A$  et de  $\xi$  dans le bin  $B$  est définie comme

$$C_{AB} = \langle \xi_A \xi_B \rangle - \langle \xi_A \rangle \langle \xi_B \rangle . \quad (1.13)$$

De cette matrice de covariance, la matrice de corrélation est définie comme

$$Corr_{AB} = \frac{C_{AB}}{\sqrt{C_{AA}C_{BB}}} . \quad (1.14)$$

La matrice de corrélation donne la corrélation, comprise dans  $[-1;1]$ , d'un bin  $A$  avec un bin  $B$ . Afin d'estimer la matrice de covariance, le relevé est divisé en pixels HEALPix, en utilisant  $n_{\text{side}} = 16$ . Cette résolution produit des pixels d'une taille sur le ciel de  $3,7 \times 3,7 = 13,4 \text{ deg}^2$ , correspondant à  $250 \times 250 (h^{-1} \text{ Mpc})^2$  à un redshift  $z = 2,33$ . Ces sous-échantillons sont suffisamment grands pour pouvoir négliger les corrélations entre différents pixels HEALPix et ainsi estimer la matrice de covariance comme la variance d'un sous-échantillon à un autre. La matrice de covariance est donc calculée comme

$$C_{AB} = \frac{1}{W_A W_B} \sum_s W_A^s W_B^s (\xi_A^s \xi_B^s - \xi_A \xi_B) , \quad (1.15)$$

où  $s$  est un sous-échantillon, et  $W_A^s$  les poids du bin  $A$  de ce sous-échantillon. Les éléments les plus importants de cette matrice sont les éléments diagonaux : la variance dans chaque bin. Les éléments non-diagonaux, les covariances entre deux bins distincts, sont faibles (#prov donner comme Var\_A est modélisé? Donner comment on modélise les éléments non diag?). Leur estimation est bruitée. La matrice de covariance est donc lissée après avoir été estimée. En ce qui concerne l'auto-corrélation, la matrice de covariance possède  $2500 \times 2500$  bins. Pour la corrélation croisée, elle en possède  $5000 \times 5000$ .

## 4 Modélisation des fonctions de corrélation

Dans cette section, nous présentons les modèles utilisés pour ajuster les fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  et  $\text{Ly}\alpha \times \text{QSO}$ . Nous présentons d'abord les modèles utilisés dans l'analyse des données DR16, puis nous donnons les modèles utilisés pour analyser les mocks.

### 4.1 Modélisation des données

Pour l'analyse des données DR16, dont les résultats sont présentés dans le chapitre suivant, nous utilisons le modèle décrit dans **prov** L'analyse décrite dans cette étude est une analyse BAO : l'auto-corrélation et la corrélation croisée sont modélisées de façon à mesurer au mieux les paramètres BAO  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ . Pour ce faire, le modèle est séparé en deux composantes. La première,  $\xi_{\text{smooth}}$ , correspond à la forme globale de la fonction de corrélation. La seconde,  $\xi_{\text{peak}}$ , correspond au pic BAO. C'est cette seconde composante qui dépend des paramètres BAO :

$$\xi(r_{\parallel}, r_{\perp}, \alpha_{\parallel}, \alpha_{\perp}) = \xi_{\text{smooth}}(r_{\parallel}, r_{\perp}) + \xi_{\text{peak}}(\alpha_{\parallel} r_{\parallel}, \alpha_{\perp} r_{\perp}) . \quad (1.16)$$

Cette distinction entre les composantes *smooth* et *peak* est faite lors du calcul du spectre de puissance modèle (voir paragraphe suivant). Le modèle de la fonction de corrélation est ensuite obtenue à l'aide d'une transformation de Fourier inverse de ce spectre de puissance.

Le spectre de puissance utilisé dans la modélisation de la fonction de corrélation des traceurs  $i$  et  $j$  s'exprime comme

$$P(\vec{k}) = b_i b_j (1 + \beta_i \mu_k^2) (1 + \beta_j \mu_k^2) P_{\text{QL}}(\vec{k}) F_{\text{NL}}(\vec{k}) G(\vec{k}). \quad (1.17)$$

Les termes  $b_i(1 + \beta_i \mu_k^2)$  et  $b_j(1 + \beta_j \mu_k^2)$  sont les facteurs de Kaiser (équation ??) relatifs aux traceurs  $i$  et  $j$ .  $P_{\text{QL}}$  est le spectre de puissance *quasi-linéaire*. Il est découpé en deux composantes  $P_{\text{smooth}}$  et  $P_{\text{peak}}$  comme

$$P_{\text{QL}}(\vec{k}, z) = P_{\text{smooth}}(k, z) + \exp\left(-\frac{k_{\parallel}^2 \Sigma_{\parallel}^2 + k_{\perp}^2 \Sigma_{\perp}^2}{2}\right) P_{\text{peak}}(k, z). \quad (1.18)$$

$P_{\text{smooth}}$  est le spectre de puissance linéaire, sans les BAO. Il est construit à partir du spectre de puissance linéaire  $P_{\text{L}}$  donné par Camb, puis les BAO sont retirées en utilisant la technique *side-band* décrite dans KIRKBY et al. (2013). Le spectre de puissance  $P_{\text{peak}}$  est alors obtenu comme la différence  $P_{\text{L}} - P_{\text{smooth}}$  : il contient uniquement les oscillations dues aux BAO présentes dans le spectre de puissance linéaire. Le terme exponentiel devant  $P_{\text{peak}}$ , paramétré par  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$ , prend en compte l'élargissement non linéaire du pic BAO. Nous utilisons  $\Sigma_{\parallel} = 6,42 h^{-1} \text{ Mpc}$  et  $\Sigma_{\perp} = 3,26 h^{-1} \text{ Mpc}$  (EISENSTEIN, SEO et WHITE 2007). Le terme  $F_{\text{NL}}$  prend en compte les non linéarités aux petites échelles. Nous distinguons  $F_{\text{NL}}^{\text{auto}}$  et  $F_{\text{NL}}^{\text{cross}}$ . Pour l'auto-corrélation, les effets non linéaires proviennent de l'élargissement thermique, des vitesses particulières et de la croissance des structures non linéaire. Comme lors de la modélisation du  $P^{1\text{D}}$ , nous utilisons le modèle décrit dans ARINYO-I-PRATS et al. (2015). Nous avons donc  $F_{\text{NL}}^{\text{auto}}(k, \mu) = D(k, \mu)$ , où  $D$  est défini dans l'équation ?. Les paramètres utilisés sont une interpolation à  $z = 2,334$  de ceux donnés dans la section "Planck" de la table 7 de ARINYO-I-PRATS et al. (2015). Pour la corrélation croisée, l'effet dominant est dû aux vitesses non linéaires des quasars. Cet effet est modélisé par une lorentzienne :

$$F_{\text{NL}}^{\text{cross}}(k_{\parallel}) = \frac{1}{1 + (k_{\parallel} \sigma_v)^2}, \quad (1.19)$$

où l'inverse de la demi-largeur à mi-hauteur  $\sigma_v$  est un paramètre libre. L'effet dû aux erreurs statistiques sur la mesure du redshift des quasars étant confondu avec l'effet des vitesses non linéaires des quasars, il est aussi pris en compte par le terme  $F_{\text{NL}}^{\text{cross}}$ . Enfin, le terme  $G(\vec{k})$  prend en compte l'effet du binning utilisé lors du calcul de la fonction de corrélation. Il est défini comme le produit des transformés de Fourier de la fonction porte :

$$G(\vec{k}) = \text{sinc}\left(\frac{k_{\parallel} R_{\parallel}}{2}\right) \text{sinc}\left(\frac{k_{\perp} R_{\perp}}{2}\right), \quad (1.20)$$

avec  $R_{\parallel}$  et  $R_{\perp}$  la largeur des bins, soit  $4 h^{-1} \text{ Mpc}$ .

Afin de modéliser la corrélation croisée  $\text{Ly}\alpha \times \text{QSO}$ , nous ajoutons le paramètre  $\Delta_{r_{\parallel}, \text{QSO}}$ , inclus comme

$$r_{\parallel} = r_{\parallel, \text{measure}} + \Delta_{r_{\parallel}, \text{QSO}}, \quad (1.21)$$

où  $r_{\parallel, \text{measure}}$  est la séparation mesurée des paires  $(i, j)$ , et  $r_{\parallel}$  est la séparation utilisée dans le modèle de la corrélation croisée. L'ajout de ce paramètre permet de prendre en compte les erreurs systématiques sur la mesure du redshift des quasars, qui rendent asymétrique la fonction de corrélation  $\text{Ly}\alpha \times \text{QSO}$ .

La présence des facteurs de Kaiser dans l'équation 1.17 permet de mesurer le biais et le paramètre RSD de nos traceurs. En ce qui concerne l'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , la fonction de corrélation est proportionnelle à  $b_{\text{Ly}\alpha}^2 (1 + \beta_{\text{Ly}\alpha} \mu^2)^2$ . Cependant, la présence de HCD dans les données modifie le biais et le paramètre RSD du  $\text{Ly}\alpha$ . En effet, l'efficacité de l'algorithme de détection n'étant pas de 100 %, il subsiste des DLA non identifiés dans les forêts. De plus, les HCD avec  $\log n_{\text{HI}} < 20,3$  ne sont pas

identifiés. Ces deux effets participent à augmenter le biais mesuré du Ly $\alpha$  significativement. Nous utilisons les paramètres effectifs

$$b'_{\text{Ly}\alpha} = b_{\text{Ly}\alpha} + b_{\text{HCD}} F_{\text{HCD}}(k_{\parallel}), \quad (1.22)$$

$$b'_{\text{Ly}\alpha} \beta'_{\text{Ly}\alpha} = b_{\text{Ly}\alpha} \beta_{\text{Ly}\alpha} + b_{\text{HCD}} \beta_{\text{HCD}} F_{\text{HCD}}(k_{\parallel}), \quad (1.23)$$

où  $F_{\text{HCD}}$  est une fonction qui dépend de la distribution en  $z$  et en  $\log n_{\text{HI}}$  des HCD (FONT-RIBERA et MIRALDA-ESCUDE 2012). Cette fonction est estimée sur des simulations hydrodynamiques (ROGERS et al. 2017). Elle est modélisée comme

$$F_{\text{HCD}}(k_{\parallel}) = \exp(-L_{\text{HCD}} k_{\parallel}), \quad (1.24)$$

où  $L_{\text{HCD}}$  est la taille typique des HCD non masqués. La résolution du spectrographe d'eBOSS rend possible l'identification des DLAs dont la largeur est supérieure à 2 nm, correspondant à une taille d'environ  $14 h^{-1}$  Mpc au redshift effectif de la mesure. Par ailleurs,  $L_{\text{HCD}}$  est très dégénéré avec les autres paramètres du modèle, comme le biais des HCD ou les paramètres du Ly $\alpha$ . Nous fixons donc sa valeur à  $10 h^{-1}$  Mpc dans l'ajustement du modèle.

Afin de pouvoir ajuster le même modèle sur tous les bins  $(r_{\parallel}, r_{\perp})$ , la dépendance en redshift de  $\delta_F$  est prise en compte. En considérant que  $\beta_{\text{Ly}\alpha}$  est constant avec le redshift, nous avons  $\delta_F(z) \propto G(z) b_{\text{Ly}\alpha}(z)$ , avec  $G(z) \propto (1+z)^{-1}$ . Concernant le biais du Ly $\alpha$ , nous utilisons la même dépendance que celle choisie lors du calcul des poids (équation ??), c'est à dire  $b_{\text{Ly}\alpha} \propto (1+z)^{\gamma_{\text{Ly}\alpha}}$ , avec  $\gamma_{\text{Ly}\alpha} = 2.9$  (MCDONALD et al. 2004). En ce qui concerne  $\beta_{\text{Ly}\alpha}$ , nous considérons lors de l'ajustement du modèle qu'il est indépendant du redshift. Comme montré dans l'analyse présentée dans le chapitre ??,  $\beta_{\text{Ly}\alpha}$  n'est pas indépendant du redshift. Cependant, le redshift moyen dans chaque bin  $(r_{\parallel}, r_{\perp})$  varie peu. Lors de l'analyse de l'ensemble des données DR16, il varie dans la gamme  $2.31 < z < 2.39$ . Cette variation correspond à une variation de  $\beta_{\text{Ly}\alpha}$  de moins de 5 %. De plus, elle est d'autant plus faible lorsque l'analyse est faite dans différents bins en redshift.

La fonction de corrélation croisée Ly $\alpha$ ×QSO est sensible au produit  $b_{\text{Ly}\alpha}(1 + \beta_{\text{Ly}\alpha}\mu^2)b_{\text{QSO}}(1 + \beta_{\text{QSO}}\mu^2)$ . L'ajustement de cette seule fonction de corrélation ne permet donc pas de lever la dégénérescence des paramètres Ly $\alpha$  et des paramètres des QSO. Nous fixons donc  $b_{\text{QSO}}$  et  $\beta_{\text{QSO}}$ . Concernant le biais des quasars, comme pour le Ly $\alpha$  nous prenant en compte l'évolution avec le redshift. Il est paramétrisé comme

$$b_{\text{QSO}}(z) = 3.77 \left( \frac{1+z}{1+2.334} \right)^{1.44}. \quad (1.25)$$

Cette paramétrisation est celle choisie dans **prov**. Nous gardons cette paramétrisation pour notre analyse des données, présentée dans le chapitre suivant. Cependant, lorsque nous analysons les mocks, nous utilisons la paramétrisation adoptée pour construire les mocks. Elle est donnée dans l'équation ??.  $\beta_{\text{QSO}}$  est choisi constant et vaut  $\beta_{\text{QSO}} = f/b_{\text{QSO}}(z_{\text{eff}})$ , où  $z_{\text{eff}}$  est le redshift effectif de la mesure. Le taux de croissance  $f$  est lui aussi évalué au redshift effectif de la mesure.

Une fois les composantes multiplicatives incluses au modèle, nous pouvons transformer le  $P(\vec{k})$  modèle défini dans l'équation 1.17 en fonction de corrélation  $\xi(r, \mu)$ . Pour ce faire, la fonction de corrélation est décomposée en polynômes de Legendre  $P_l$  jusqu'à  $l_{\text{max}} = 6$ . Pour chaque  $l \in [0; 2; 4; 6]$ , une transformation de Fourier est appliquée au spectre de puissance (**prov** en fait c'est plutôt une intégrale avec une fonction de Bessel. Donner une ref du papier d'Hamilton ? C'est aussi dans le papier Kirkby, eq 2.7 - 2.9). Etant donné que la transformation  $P(\vec{k}) \rightarrow \xi(r, \mu)$  est faite à chaque étape de la minimisation lors de l'ajustement du modèle, il est important que cette transformation se fasse très rapidement. Nous utilisons donc l'algorithme FFTLog (HAMILTON 1999), qui apporte à la fois rapidité et précision. Ainsi, en suivant cette procédure, nous obtenons les fonctions de corrélation  $\xi_{\text{Ly}\alpha \times \text{Ly}\alpha}$  en choisissant  $i = j = \text{Ly}\alpha$ , et  $\xi_{\text{Ly}\alpha \times \text{QSO}}$  en choisissant  $i = \text{Ly}\alpha$  et  $j = \text{QSO}$ . Afin de pouvoir



ajuster les fonctions de corrélation calculées avec les données, nous devons prendre en compte dans nos modèles les corrélations parasites. Ces corrélations s'ajoutent au  $\xi$  modèle calculé précédemment. A ce stade, nous distinguons le modèle utilisé pour l'auto-corrélation et la corrélation croisée. Le modèle de l'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  est défini comme

$$\xi = \xi_{\text{Ly}\alpha \times \text{Ly}\alpha} + \sum_{m,n} \tilde{\xi}_{m \times n} + \xi_{\text{ciel}} , \quad (1.26)$$

où  $\tilde{\xi}_{m \times n}$  est la corrélation de l'absorbeur  $m$  avec l'absorbeur  $n$ , interprétés comme des absorptions  $\text{Ly}\alpha$ . Nous verrons plus tard comment relier la corrélation  $\tilde{\xi}_{m \times n}$  à la corrélation physique  $\xi_{m \times n}$ . Ces absorbeurs peuvent être du  $\text{Ly}\alpha$  ou des métaux ( $m$  et  $n$  ne peuvent pas être tous les deux du  $\text{Ly}\alpha$ ). Le tableau 1.1 liste les métaux ajustés dans les données.  $\xi_{\text{ciel}}$  est la corrélation induite par la soustraction du fond de ciel. Ces termes sont décrits dans les prochains paragraphes. Le modèle de la corrélation croisée  $\text{Ly}\alpha \times \text{QSO}$  est défini comme

$$\xi = \xi_{\text{Ly}\alpha \times \text{QSO}} + \sum_m \tilde{\xi}_{m \times \text{QSO}} + \xi_{\text{prox}} , \quad (1.27)$$

où  $\tilde{\xi}_{m \times \text{QSO}}$  est la corrélation du métal  $m$  (interprété comme du  $\text{Ly}\alpha$ ) avec les quasars, et  $\xi_{\text{prox}}$  donne la corrélation induite par l'effet de proximité des quasars. Ces termes sont décrits dans les prochains paragraphes.

Afin de modéliser la corrélation des métaux, nous utilisons le modèle défini précédemment, utilisé pour décrire les corrélations  $\xi_{\text{Ly}\alpha \times \text{Ly}\alpha}$  et  $\xi_{\text{Ly}\alpha \times \text{QSO}}$ . Comme expliqué dans la section ??, toutes les absorptions sont supposées être des absorptions  $\text{Ly}\alpha$ . Les absorptions causées par les métaux sont donc reconstruites à un mauvais redshift. Ceci résulte dans un décalage de la fonction de corrélation le long de la ligne de visée. **Considérons deux absorbeurs  $m$  et  $n$  ( $m \neq n$ ). La fonction de corrélation  $\xi_{m \times n}$  de ces deux absorbeurs est maximale pour les séparations  $r = 0$ . Cependant, en supposant que ces deux absorptions sont causées par le  $\text{Ly}\alpha$ , cette séparation  $r = 0$  est reconstruite à  $r_{\perp} = 0$  et  $r_{\parallel} \sim (1+z)D_H(z)(\lambda_m - \lambda_n)/\lambda_{\text{Ly}\alpha}$ , où  $z$  est le redshift moyen des deux absorbeurs. Ainsi, l'excès de corrélation observé n'est pas situé à  $r = 0$ , mais se trouve décalé le long de la ligne de visée.** Le tableau 1.1 donne les séparations associées aux corrélations entre le  $\text{Ly}\alpha$  et les métaux ajustés sur les données. Les métaux étant beaucoup moins présents que l'hydrogène dans le milieu intergalactique, les corrélations mettant en jeu deux métaux sont beaucoup moins importantes que les corrélations mettant en jeu un métal et le  $\text{Ly}\alpha$ . L'effet principal vient donc des corrélations  $\xi_{\text{Ly}\alpha \times m}$ . En ce qui concerne les corrélations  $\xi_{m \times m}$ , le décalage est d'origine différente. Une séparation physique  $r = 0$  correspond bien à une reconstruction  $r_{\parallel} = r_{\perp} = 0$ . Cependant, le redshift de la paire est mal estimé. Chaque séparation physique  $(r_{\parallel}, r_{\perp})$  est donc reconstruite à  $(D_H(z_m)/D_H(z))r_{\parallel}$  et  $(D_M(z_m)/D_M(z))r_{\perp}$ . Pour la corrélation  $\xi_{m \times \text{QSO}}$ , le décalage est le même que dans le cas  $\xi_{m \times n}$ , en prenant  $z_n = z_{\text{QSO}}$ .

Pour chaque couple  $(m, n)$ , le décalage de la fonction de corrélation  $\xi_{m \times n}$  est pris en compte par la matrice des métaux  $M_{AB}^{m \times n}$ . Nous ne détaillons pas son calcul ici, mais il est donné dans la thèse **CITE:Victoria** La matrice des métaux permet donc de relier la fonction de corrélation  $\xi_{m \times n}$  à la fonction de corrélation  $\tilde{\xi}_{m \times n}$ , où les absorbeurs  $m$  et  $n$  sont interprétés comme des absorptions  $\text{Ly}\alpha$ , utilisée comme modèle dans l'équation 1.26 et 1.27 :

$$\tilde{\xi}_{m \times n}(A) = \sum_B M_{AB}^{m \times n} \xi_{m \times n}(B) , \quad (1.28)$$

Les bins  $A$  correspondent aux séparations calculées en supposant une absorption  $\text{Ly}\alpha$ . Les bins  $B$  correspondent aux séparations physiques, calculées en utilisant les redshifts  $z_m$  et  $z_n$  des absorbeurs. Ainsi, pour chaque couple  $(m, n)$ , la matrice des métaux  $M_{AB}^{m \times n}$  est calculée, puis la fonction de

TABLE 1.1 – Liste des métaux inclus dans le modèle ajusté aux données. La 3<sup>e</sup> colonne donne la séparation reconstruite pour une séparation réelle  $r = 0$ . CIV(eff) indique la raie effective du carbon IV : la résolution du spectrographe d’eBOSS étant trop faible pour distinguer le doublet du CIV, nous ajustons la combinaison des deux raies. La séparation liée au CIV est bien supérieure à  $200 h^{-1}$  Mpc, la corrélation  $\text{Ly}\alpha \times \text{CIV}$  n’a donc pas d’effet sur nos mesures. Nous modélisons cependant l’effet lié à l’auto-corrélation  $\text{CIV} \times \text{CIV}$ .

Raie	$\lambda_m [\text{\AA}]$	$r_{\parallel} [h^{-1} \text{ Mpc}]$
SiII(1190)	1190,4158	−64
SiII(1193)	1193,2897	−56
SiIII(1207)	1206,500	−21
SiII(1260)	1260,4221	+111
CIV(eff)	1549,06	> 200

corrélation  $\tilde{\xi}_{m \times n}$  est estimée et ajoutée à la fonction de corrélation  $\xi_{\text{Ly}\alpha \times \text{Ly}\alpha}$  (équation 1.26). Dans le cas de la corrélation croisée, la matrice est calculée pour tous les couples  $(m, \text{QSO})$ , puis la fonction de corrélation  $\tilde{\xi}_{m \times \text{QSO}}$  est estimée et ajoutée à la fonction de corrélation  $\xi_{\text{Ly}\alpha \times \text{QSO}}$  (équation 1.27).

Le modèle utilisé pour construire les fonctions de corrélation  $\xi_{m \times n}$  et  $\xi_{m \times \text{QSO}}$  et le même que celui utilisé pour construire les fonctions de corrélation  $\xi_{\text{Ly}\alpha \times \text{Ly}\alpha}$  et  $\xi_{\text{Ly}\alpha \times \text{QSO}}$  (équation 1.17) mais avec les paramètres  $b_i$ ,  $b_j$ ,  $\beta_i$  et  $\beta_j$  qui sont ceux des métaux. Du fait que les métaux sont mesurables principalement le long de la ligne de visée, il est difficile d’ajuster à la fois le biais et le paramètre RSD de chaque métal. Nous ajustons donc uniquement le biais de chaque métal, le paramètre RSD étant fixé à  $\beta_m = 0,5$ . **De plus, comme la quantité de métaux dans le milieu intergalactique est bien moins importante que celle de l’hydrogène, il n’existe pas ou peu de structures denses constituées de métaux. Nous n’incluons donc pas la modélisation des HCD dans les fonctions de corrélations  $\xi_{m \times n}$  et  $\xi_{m \times \text{QSO}}$ .**

Le terme additionnel suivant est le terme  $\xi_{ciel}$ . Ce terme prend en compte les corrélations induites par la soustraction du fond de ciel. Lors de la réduction des données, décrite dans la section ??, le spectre du fond de ciel est soustrait à tous les spectres d’une même demi-plaque. Ceci induit alors des corrélations entre tous ces spectres pour  $r_{\parallel} = 0$ . A cause de la distorsion induite par l’ajustement du continuum, cette effet ne se limite pas à  $r_{\parallel} = 0$ . L’effet est modélisé par une fonction gaussienne de  $r_{\perp}$  :

$$\xi_{ciel}(r_{\parallel}, r_{\perp}) = \begin{cases} \frac{A_{sky}}{\sqrt{2\pi\sigma_{sky}^2}} \exp\left(-\frac{r_{\perp}^2}{2\sigma_{sky}^2}\right) & , \text{ si } r_{\parallel} = 0 \\ 0 & , \text{ si } r_{\parallel} \neq 0 \end{cases} . \quad (1.29)$$

Les paramètres  $A_{sky}$  et  $\sigma_{sky}$  sont laissés libres lors de l’ajustement des données. Ils donnent l’amplitude et la largeur de la gaussienne. Le terme  $\xi_{ciel}$  n’est présent que dans l’ajustement de l’auto-corrélation, car ces corrélations parasites ne sont induites que lorsqu’on corréle des pixels d’absorption issus de deux spectres présents sur la même demi-plaque. Cet effet n’a donc pas lieu d’être pour la corrélation croisée.

Enfin, le dernier terme additionnel est le terme  $\xi_{prox}$ . Ce terme n’est présent que dans la fonction de corrélation croisée  $\xi_{\text{Ly}\alpha \times \text{QSO}}$ . Il prend en compte l’effet du rayonnement produit par les quasars sur l’hydrogène environnant. En effet, à cause de leur grande luminosité, et en particulier dans la direction de leur jet, les quasars ionisent le gaz qui les entoure. Ceci réduit donc la fraction d’hydrogène neutre et donc la profondeur optique au voisinage de chaque quasar, ce qui induit des corrélations

supplémentaire entre le champ d'absorption Ly $\alpha$  et la position des quasars. Cet effet est modélisé comme (FONT-RIBERA, ARNAU et al. 2013) :

$$\xi_{prox} = \xi_{0,prox} \left( \frac{1 h^{-1} \text{Mpc}}{r} \right)^2 \exp \left( -\frac{r}{\lambda_{UV}} \right), \quad (1.30)$$

où  $\xi_{0,prox}$  donne l'amplitude de l'effet. L'émission est supposée isotrope, et le paramètre  $\lambda_{UV}$  est fixé à  $300 h^{-1} \text{Mpc}$ .

Une fois tous ces termes inclus, nous obtenons un modèle qui décrit la fonction de corrélation de la matière dans l'espace des redshifts, à un biais près, et qui prend en compte les différents effets astrophysiques ou instrumentaux qui affectent les données. Afin de pouvoir correctement comparer notre modèle aux fonctions de corrélation calculées avec les données, nous devons prendre en compte la distorsion due à l'ajustement du continuum. Ceci est fait, comme décrit dans la section 2, grâce à la matrice de distorsion. Le modèle distordu est alors donné par

$$\xi_{\text{distorsion}}(A) = \sum_B D_{AB} \xi(B), \quad (1.31)$$

où  $D_{AB}$  est la matrice de distorsion, et  $\xi$  est le modèle construit précédemment. Ainsi, le modèle qui est ajusté aux données est  $\xi_{\text{distorsion}}(A)$ .

## 4.2 Modélisation des mocks

Afin d'analyser les fonctions de corrélation Ly $\alpha$ ×Ly $\alpha$  et Ly $\alpha$ ×QSO des mocks, nous utilisons les modèles ajustés sur les données et décrits dans la section précédente. Cependant, un certain nombre d'effets modélisés dans les données ne sont pas présents dans les mocks. Nous modifions donc légèrement les modèles décrits précédemment.

Premièrement, nous n'incluons pas l'élargissement non linéaire du pic BAO dans les mocks. Les non-linéarités prises en compte par le terme  $F_{NL}$  ne sont pas non plus présentes dans nos mocks. Ainsi, le spectre de puissance défini dans l'équation 1.17 et utilisé comme modèle pour l'auto-corrélation ne contient pas le terme  $F_{NL}^{\text{auto}}$ . Cependant, le code `quickquasars` ajoute une vitesse particulière à chaque quasar. Ceci a pour effet d'ajouter une erreur statistique sur la mesure du redshift des quasars. Nous gardons donc le terme  $F_{NL}^{\text{cross}}$  lorsque nous ajustons la corrélation Ly $\alpha$ ×QSO issues des mocks avec `quickquasars`. Ce terme n'est pas présent dans le modèle utilisé pour ajuster les raw mocks.

L'effet instrumental causé par la soustraction du fond de ciel sur l'auto-corrélation n'est pas modélisé par `quickquasars`. Nous n'ajoutons donc pas le terme  $\xi_{\text{ciel}}$  à  $\xi_{\text{Ly}\alpha \times \text{Ly}\alpha}$  dans l'ajustement de l'auto-corrélation. L'effet de proximité des quasars sur le champ Ly $\alpha$  environnant n'est ajouté ni dans les mocks, ni dans `quickquasars`. Le terme  $\xi_{\text{prox}}$  n'est donc pas inclu dans le modèle de la corrélation croisée.

En ce qui concerne l'ajustement des HCD et des métaux, cela dépend de la version des mocks analysée. Pour les raw mocks comme pour les mocks eboss-0.0, ni les HCD ni les métaux ne sont présents. Nous n'incluons donc pas leur modélisation dans ces versions des mocks. Les versions eboss-0.2 incluent les HCD, nous modélisons donc leur présence comme décrit dans la section précédente. Les versions eboss-0.3 incluent à la fois les HCD et les métaux. L'ajustement des fonctions de corrélation issues de ces mocks contient donc les paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$ . De plus, nous calculons la matrice des métaux pour ces versions et ajoutons au modèle les termes  $\xi_{m \times n}$  et  $\xi_{m \times \text{QSO}}$ .

Pour l'ajustement des corrélations croisées Ly $\alpha$ ×QSO, nous gardons le paramètre  $\Delta_{r_{\parallel}, \text{QSO}}$ . Même si `quickquasars` ajoute une erreur statistique sur les redshifts des quasars, cette erreur est nulle en moyenne. Nous nous attendons donc à obtenir  $\Delta_{r_{\parallel}, \text{QSO}} = 0$  dans l'ajustement des différentes versions des mocks. Ainsi,  $\Delta_{r_{\parallel}, \text{QSO}}$  sert de test de la construction des mocks, pour vérifier par exemple que les lignes de visées sont placées correctement à partir de chaque quasar.

Comme pour les données, les fonctions de corrélations sont évaluées sur une grille de séparation d'intervalle  $4 h^{-1}$  Mpc. Nous gardons donc le terme  $G(\vec{k})$  dans la modélisation des mocks. Dans le cas des mocks, nous incluons un terme supplémentaire, qui prend en compte le lissage gaussien appliqué au champ  $\delta_l$  interpolé. Ce terme est donné par

$$W(k_{\parallel}, k_{\perp}) = \exp\left(-\frac{k^2 \sigma_{\text{smooth}}^2}{2}\right). \quad (1.32)$$

**Comme montré sur la figure ??, l'effet du lissage gaussien sur la fonction de corrélation est important surtout à petit  $r$ . Les mocks n'étant pas en très bon accord avec le modèle à petit  $r$ , la modélisation du lissage gaussien a tendance à corriger ces écarts lorsque le paramètre  $\sigma_{\text{smooth}}$  est laissé libre. De plus,  $\sigma_{\text{smooth}}$  est dégénéré avec d'autres paramètres du modèle. Nous fixons donc  $\sigma_{\text{smooth}}$  à la taille des voxels, soit  $2,19 h^{-1}$  Mpc lors des ajustements. Le spectre de puissance modèle  $P(k_{\parallel}, k_{\perp})$  est ainsi multiplié par  $W^2(k_{\parallel}, k_{\perp})$ .**

Enfin, similairement à la modélisation des données, le modèle ajusté sur les mocks est multiplié par la matrice de distorsion  $D_{AB}$  (équation 1.31). Cependant, dans le cas des raw mocks, les fonctions de corrélation ne sont pas affectées par la distorsion due à l'ajustement du continuum puisque nous avons accès directement au champ  $\delta_F$ . Dans ce cas, la matrice de distorsion  $D_{AB}$  vaut la matrice identité.

## 5 Analyse des mocks

Comme expliqué dans la section ??, nous produisons différentes versions des mocks : les raw mocks, pour lesquels le champ  $\delta_F$  est obtenu directement à partir des vrais transmissions, et les mocks après l'utilisation de `quickquasars` : `eboss-0.0`, `eboss-0.2` et `eboss-0.3`. Nous présentons dans cette section l'analyse des différentes versions des mocks.

### 5.1 Analyse des raw mocks

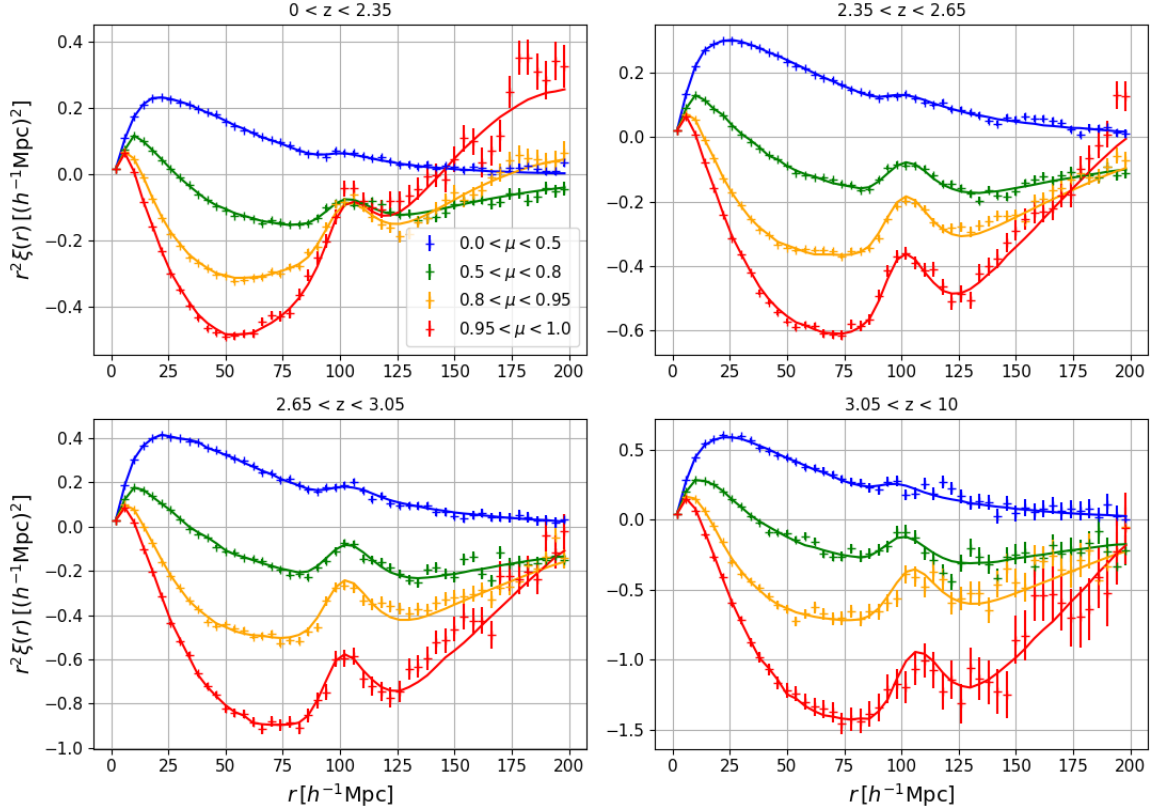
Pour chacune des cent réalisations produites, nous avons calculé les fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  et  $\text{Ly}\alpha \times \text{QSO}$  des raw mocks. L'analyse de ces fonctions de corrélation permet d'identifier plus facilement les problèmes qui peuvent exister au niveau de la construction des mocks, car les effets astrophysiques et instrumentaux introduits par `quickquasars` ne sont pas présents. Nous commençons donc par valider la construction des mocks, via l'étude des raw mocks, puis nous présenterons l'analyse des versions des mocks avec `quickquasars`.

Toutes les fonctions de corrélation présentées ici, à l'exception de la fonction de corrélation à une dimension, sont estimées dans quatre bins en redshift. Ces bins sont les mêmes que ceux choisis pour analyser les données et déterminer  $b_{\text{Ly}\alpha}(z)$  et  $\beta_{\text{Ly}\alpha}(z)$  à utiliser pour la construction des mocks (voir chapitre ??). Ces bins sont :  $[0; 2,35]$ ,  $[2,35; 2,65]$ ,  $[2,65; 3,05]$  et  $[3,05; 10]$ . Nous utilisons les mêmes bins pour faciliter la comparaison entre les mocks et les données.

#### L'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$

La figure 1.1 donne l'addition des cent fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  des raw mocks dans chaque bin en redshift. Pour chaque bin en redshift, la fonction de corrélation est montrée dans quatre bins en  $\mu$  différents. Le bin  $0,95 < \mu < 1$  correspond aux paires avec une séparation le long de la ligne de visée. Le bin  $0 < \mu < 0,5$  correspond aux paires perpendiculaires à la ligne de visée. Les lignes continues donnent le meilleur ajustement du modèle produit par `picca`.

#prov Est-ce qu'il faut pas que le paragraphe qui suit soit pour l'ajustement combiné de la CF et XCF ? Sachant que le biais et beta  $\text{Ly}\alpha$  sont déterminés sur le fit en 4 bins de la CF des données, c'est

FIGURE 1.1 – L’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  calculée sur les raw mocks. La figure

peut-être mieux de faire uniquement sur la CF. Je peux ensuite montrer l’évolution du biais des QSO en fitant la XCF, avec les paramètres  $\text{Ly}\alpha$  fixés à ce que je trouve dans le fit de la CF ? Ou alors je fais le fit combiné, mais dans ce cas là si une partie du  $\text{Ly}\alpha$  passe dans les QSO (ou inversement), j’aurais pas le bon biais QSO.

Nous présentons maintenant le résultat de l’analyse en quatre bins en redshift des raw mocks. Pour chacune des réalisations, les fonctions de corrélation ont été calculées dans quatre bins en redshift distincts. Une fois toutes ces fonctions de corrélation produites, nous avons calculé l’addition des cent fonctions de corrélation dans chacun des quatres bins en redshift. Puis, nous avons ajustés ces quatre fonctions de corrélation séparément avec `picca`. Le tableau ?? (#prov faire le tableau) donne le résultat de l’ajustement dans chaque bin en redshift. La figure ?? (#prov faire la figure) montre le biais et le paramètre RSD du  $\text{Ly}\alpha$  obtenus dans l’ajustement des cent raw mocks (couleur), ainsi que ceux obtenus dans l’ajustement des données (couleur). L’analyse en quatre bins en redshift des données est décrite dans la section ?? . Pour chaque bin en redshift, nous générons la prédiction au redshift effectif du bin. Les biais et paramètres RSD mesurés sur la prédiction sont montrés en (couleur). Enfin, pour chacun des jeux des données montrés sur la figure ?? , nous ajustons une fonction puissance  $y(z) = a(1+z)^\gamma$ . Le paramètre  $\gamma$  nous donne la dépendance en redshift de  $b_{\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans chaque cas. Les différents  $\gamma$  obtenus sont résumés dans le tableau ?? (#prov faire le tableau). #prov conclusions

#prov montrer  $\langle F \rangle(z)$  ou au moment du tuning

### La corrélation croisée $\text{Ly}\alpha \times \text{QSO}$

Nous présentons maintenant la même analyse que précédemment, mais en considérant cette fois ci les cent fonctions de corrélation  $\text{Ly}\alpha \times \text{QSO}$  calculées sur les raw mocks. La figure ?? (#prov faire la figure) montre l’addition de ces cent fonctions de corrélation croisées. La fonction de corrélation

est montrée dans quatre bins en redshift. La ligne continue (couleur) donne l'ajustement du modèle produit par `picca`. La prédiction est montrée en (couleur). `#prov conclusion`

Comme précédemment, nous présentons l'analyse en quatre bins en redshift. Le tableau ?? (`#prov` faire le tableau) donne le résultat de l'ajustement dans chaque bin en redshift. Pour cet ajustement, nous avons fixé les paramètres  $\text{Ly}\alpha$  à ce que nous avons obtenu dans la section précédente. Ceci nous permet d'obtenir une mesure des paramètres des quasars moins corrélée avec la mesure des paramètres  $\text{Ly}\alpha$ . (`#prov` donner les corrélations de  $b\text{Ly}\alpha$  et  $\beta\text{Ly}\alpha$  avec les autres paramètres?). La figure ?? (`#prov` faire la figure) montre le biais et le paramètres RSD des quasars obtenus avec un tel ajustement. La ligne en pointillés donne le biais  $b_{\text{QSO}}$  utilisé pour construire le champ des quasars (voir section ??). `#prov conclusion`

### Le spectre de puissance à une dimension

`#prov` montrer le P1D des raw mocks? Plutôt celui obtenu avec `quickquasars`? oui

### L'auto-corrélation $\text{QSO} \times \text{QSO}$

Nous présentons ici la mesure de l'auto-corrélation  $\text{QSO} \times \text{QSO}$ , estimée grâce à l'équation ?? . Comme pour les corrélations  $\text{Ly}\alpha \times \text{Ly}\alpha$  et  $\text{Ly}\alpha \times \text{QSO}$ , la fonction de corrélation  $\text{QSO} \times \text{QSO}$  est estimée puis ajustée dans les quatre bins en redshift  $[0; 2,35]$ ,  $[2,35; 2,65]$ ,  $[2,65; 3,05]$  et  $[3,05; 10]$ . Nous estimons la corrélation de dix réalisations des raw mocks, puis nous calculons et ajustons l'addition de ces dix fonctions de corrélation. La figure 1.2 montre l'addition des dix fonctions de corrélation et leur modèle ajusté dans chaque bin en redshift. Pour chacun des bins, la fonction de corrélation est montrée dans trois bins en  $\mu$ . Le modèle utilisé pour ajuster la corrélation  $\text{QSO} \times \text{QSO}$  est le même que celui utilisé pour ajuster la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , à la différence que nous n'incluons pas le terme représentant le lissage gaussien. Dans le cas de la corrélation  $\text{QSO} \times \text{QSO}$ , contrairement au cas du  $\text{Ly}\alpha$ , nous avons la relation  $b_{\text{QSO}}\beta_{\text{QSO}} = f$ , où  $f$  est le taux de croissance des structures. Lors de l'ajustement, nous fixons donc le paramètre  $b_{\eta, \text{QSO}} = 1$ , et nous ajustons les paramètres  $\beta_{\text{QSO}}$  et  $f$ .

Sur la figure 1.2, nous pouvons remarquer que les mocks sont en très bon accord avec le modèle ajusté par `picca`. Une légère différence est visible à petit  $r$ . Cette différence dépend de  $\mu$  : l'amplitude est trop faible le long de la ligne de visée, et trop importante perpendiculairement à cette dernière. Ceci semble provenir des RSD aux petites échelles. La figure 1.3 montre  $b_{\text{QSO}}$  et  $\beta_{\text{QSO}}$  obtenus avec l'ajustement de l'addition des dix fonctions de corrélation. La ligne bleue donne la paramétrisation utilisée dans les mocks (équation ??). Pour les petits redshifts,  $b_{\text{QSO}}$  est trop faible et  $\beta_{\text{QSO}}$  est trop grand. Cependant, les valeurs ajustées de  $\beta_{\text{QSO}}$  et  $f$  sont corrélées à plus de 99%. Ceci vient du fait que, pour les quasars, le paramètre RSD est faible. Il y a donc peu de différences entre la corrélation le long de la ligne de visée et perpendiculairement à cette dernière. Lorsque nous fixons  $f$  à ce qui est donné par la cosmologie utilisée (équation ??), nous obtenons un biais et un paramètre RSD en très bon accord avec la paramétrisation utilisée. La figure 1.4 montre  $b_{\text{QSO}}$  et  $\beta_{\text{QSO}}$  obtenus avec un tel ajustement.

### La corrélation croisée $\text{Ly}\alpha \times \text{DLA}$

`#prov` montrer cette corrélation? Dans le cas des raw mocks? Le faire avec `quickquasars`, et superposer ce qu'on obtient avec les données? a voir peut être plutôt l'auto  $\text{DLA} \times \text{DLA}$ , pour vérifier le biais qu'on obtient

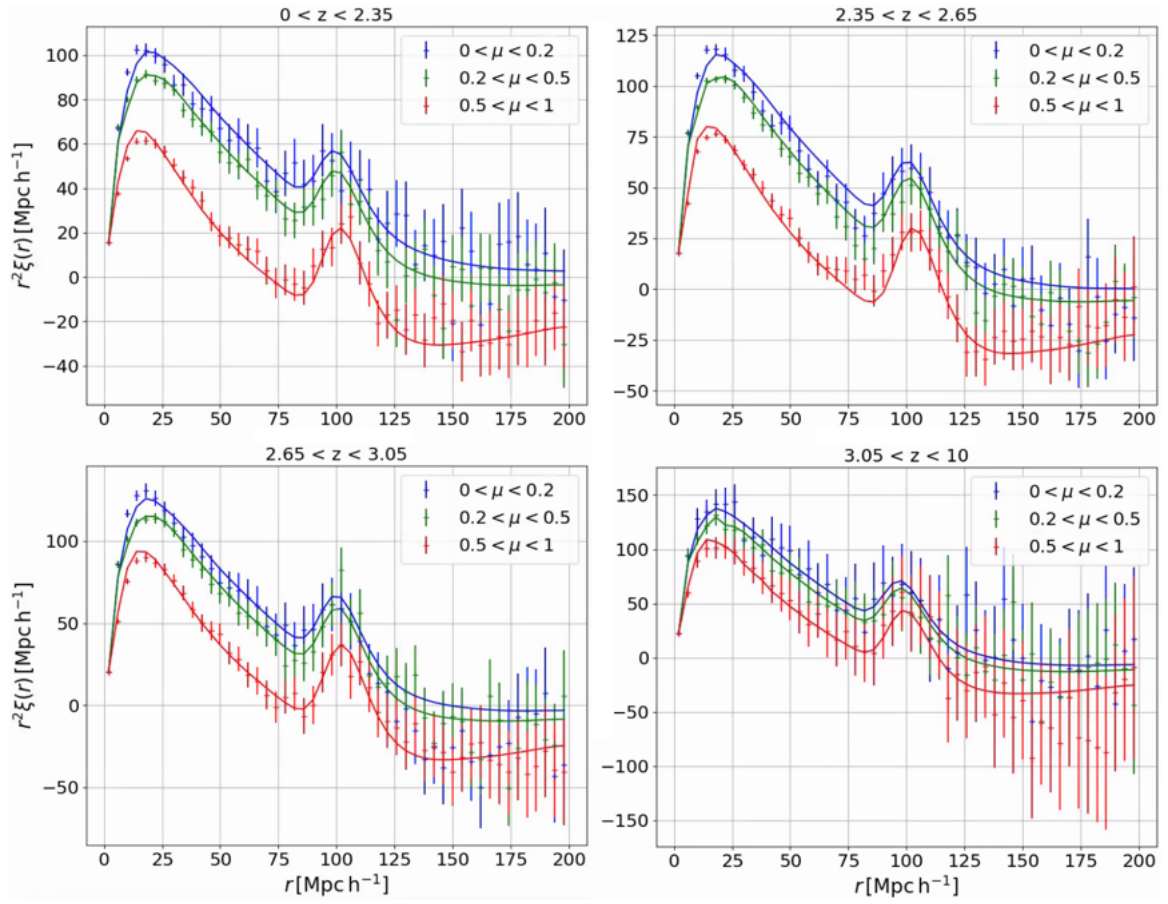


FIGURE 1.2 – bla

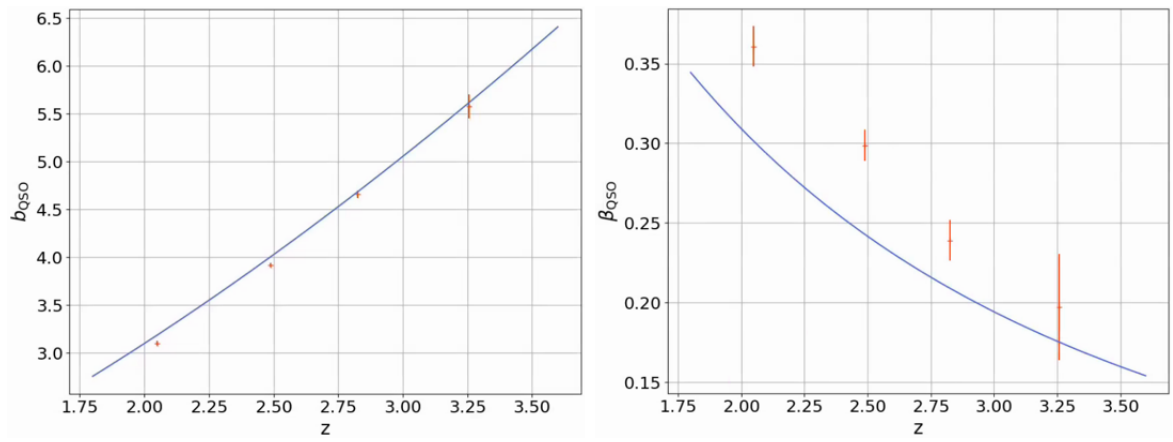


FIGURE 1.3 – bla

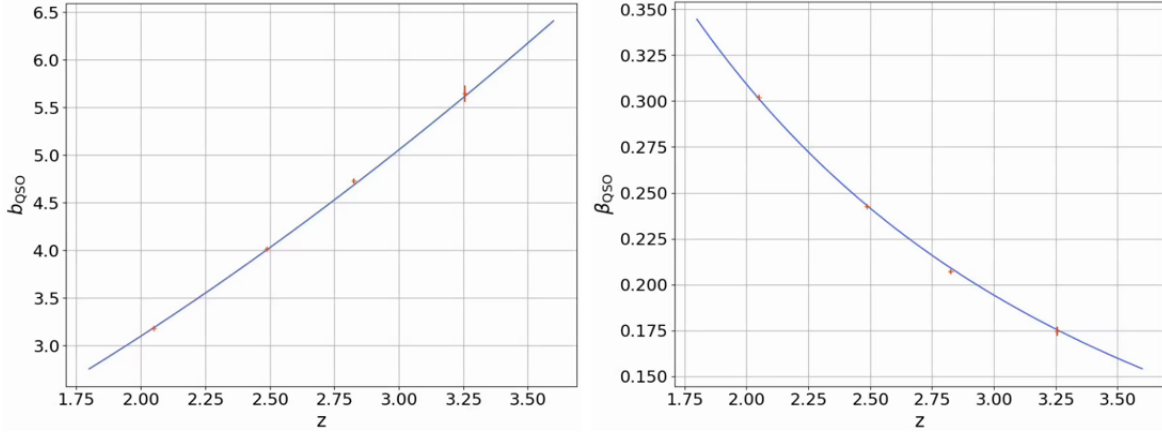


FIGURE 1.4 – bla

## 5.2 Analyse des mocks eboss-0.0

Maintenant que nous avons vérifié que les raw mocks possédaient les bonnes fonctions de corrélation, nous pouvons analyser les mocks après avoir appliqué **quickquasars**. Nous commençons par présenter l'analyse des mocks eboss-0.0. Contrairement aux raw mocks, nous avons recours à l'ajustement du continuum pour calculer le champ  $\delta_F$ . Ainsi, les fonctions de corrélation possèdent les distortions liées à cette ajustement, et les modèles sont multipliés par les matrices de distortions (équation 1.31). De plus, les mocks issus de **quickquasars** contiennent du bruit instrumental. Les fonctions de corrélation sont donc plus bruitées que celles calculées sur les raw mocks.

### L'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$

#prov figure du stack des CF dans les 4 bins en mu + modèle picca lya only ou prédiction \* dmat  
 + tableau qui donne le résultat du fit ?  
 + figure qui montre l'évolution avec  $z$  du biais et beta ? Le plot de la section précédente + le b et beta obtenu pour eboss-0.0 ; ou alors je fais un seul plot avec toutes les versions : raw mocks, eboss-0.0, eboss-0.2, eboss-0.3, prediction, data

### L'auto-corrélation $\text{Ly}\alpha \times \text{QSO}$

#prov figure du stack des XCF dans les 4 bins en mu.  
 + tableau qui donne le résultat du fit ? fit avec les parametres lya fixés ?  
 + figure qui montre l'évolution avec  $z$  du biais et beta ? Comme pour le lya, je met un plot à chaque fois, ou alors un seul plot qui résume tout ?

### Le spectre de puissance à une dimension

#prov montrer le P1D ? bof, ca teste surtout l'analyse de michael

## 5.3 Analyse des mocks eboss-0.2

Nous analysons à présent les mocks eboss-0.2. Ces mocks sont obtenus comme les mocks eboss-0.0, analysés précédemment, à la différence que le code **quickquasars** inclue les HCD dans les spectres synthétiques. Nous modélisons donc les HCD lors de l'ajustement des fonctions de corrélation. #prov je me limite ici aux CF obtenues en masquant les DLA à partir du true catalogue ? Et je parlerai des différences dans le chapitre suivant ?



**L'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$** 

- #prov figure du stack des CF dans les 4 bins en mu + modele picca lya + hcd
- + tableau qui donne le résultat du fit ?
- + figure qui montre l'évolution avec z du biais et beta ?

**L'auto-corrélation  $\text{Ly}\alpha \times \text{QSO}$** 

- #prov figure du stack des XCF dans les 4 bins en mu.
- + tableau qui donne le résultat du fit ? fit avec les parametres lya fixés ?
- + figure qui montre l'évolution avec z du biais et beta ?

**Le spectre de puissance à une dimension**

- #prov montrer le P1D ? bof, pareil

**5.4 Analyse des mocks eboss-0.3**

Cette section présente l'analyse des mocks eboss-0.3. Ces mocks sont obtenus comme les mocks eboss-0.2, analysés précédemment, à la différence que le code `quickquasars` ajoute les métaux dans les spectres synthétiques. Nous incluons donc les termes  $\xi_{m \times n}$  et  $\xi_{m \times \text{QSO}}$  dans la modélisation des fonctions de corrélation.

**L'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$** 

- #prov figure du stack des CF dans les 4 bins en mu + modele picca lya + hcd + met
- + tableau qui donne le résultat du fit ?
- + figure qui montre l'évolution avec z du biais et beta ?

**L'auto-corrélation  $\text{Ly}\alpha \times \text{QSO}$** 

- #prov figure du stack des XCF dans les 4 bins en mu.
- + tableau qui donne le résultat du fit ? fit avec les parametres lya fixés ?
- + figure qui montre l'évolution avec z du biais et beta ?

**Le spectre de puissance à une dimension**

- #prov montrer le P1D ? Pourquoi pas, pour montrer les metaux ? Ou alors le x1d

**La fonction de corrélation à une dimension**

- Comparer les mocks et les données



# Bibliographie

---

- ARINYO-I-PRATS, Andreu et al. (2015). « The Non-Linear Power Spectrum of the Lyman Alpha Forest ». In : DOI : 10.1088/1475-7516/2015/12/017. arXiv : 1506.04519.
- CHABANIER, Solène et al. (2018). « The one-dimensional power spectrum from the SDSS DR14 Lyman-alpha forests ». In : DOI : 10.1088/1475-7516/2019/07/017. arXiv : 1812.03554.
- DU MAS DES BOURBOUX, Héliot et al. (2019). « The extended Baryon Oscillation Spectroscopic Survey: measuring the cross-correlation between the MgII flux transmission field and quasars and galaxies at  $z=0.59$  ». In : arXiv : 1901.01950.
- EISENSTEIN, D J, H.-J. SEO et M WHITE (2007). « On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter ». In : *ApJ* 664, p. 660–674. DOI : 10.1086/518755.
- FONT-RIBERA, Andreu, Eduard ARNAU et al. (2013). « The large-scale Quasar-Lyman alpha Forest Cross-Correlation from BOSS ». In : DOI : 10.1088/1475-7516/2013/05/018. arXiv : 1303.1937.
- FONT-RIBERA, Andreu et Jordi MIRALDA-ESCUDE (2012). « The Effect of High Column Density Systems on the Measurement of the Lyman alpha Forest Correlation Function ». In : DOI : 10.1088/1475-7516/2012/07/028. arXiv : 1205.2018.
- HAMILTON, A. J. S. (1999). « Uncorrelated Modes of the Nonlinear Power Spectrum ». In : DOI : 10.1046/j.1365-8711.2000.03071.x. arXiv : 9905191 [astro-ph].
- KIRKBY, David et al. (2013). « Fitting Methods for Baryon Acoustic Oscillations in the Lyman-alpha Forest Fluctuations in BOSS Data Release 9 ». In : DOI : 10.1088/1475-7516/2013/03/024. arXiv : 1301.3456.
- MCDONALD, Patrick et al. (2004). « The Lyman-alpha Forest Power Spectrum from the Sloan Digital Sky Survey ». In : DOI : 10.1086/444361. arXiv : 0405013 [astro-ph].
- ROGERS, Keir K. et al. (2017). « Correlations in the three-dimensional Lyman-alpha forest contaminated by high column density absorbers ». In : arXiv : 1711.06275.