

# 1

## Analyse des données et résultats

---

Dans ce chapitre, nous présentons les diverses analyses que nous avons menées sur les données, avec les mocks comme support de référence. Un élément clé à la construction des mocks a été de déterminer quels paramètres  $\text{Ly}\alpha$  nous souhaitions avoir dans nos mocks. Ceci nous a conduit à mener une analyse des données DR16 dans quatre bins en redshift. En produisant cette analyse, nous nous sommes rendus compte que les paramètres  $\text{Ly}\alpha$  obtenus dépendaient fortement de la modélisation des HCD. Nous avons dû faire un choix quant à cette modélisation. Nous présentons donc d’abord l’analyse des données qui a servi de référence pour l’ajustement des paramètres des mocks. Puis, nous discutons la modélisation des HCD et présentons des modélisations alternatives. #prov finir de donner les autres sections

## 1 L’analyse des données DR16

### 1.1 Résultats

L’analyse des données finale d’eBOSS (DR16), dont nous avons déjà parlé et qui est présentée dans **prov** analyse les fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ ,  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ ,  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{QSO}$  et  $\text{Ly}\alpha(\text{Ly}\beta) \times \text{QSO}$ . Ces fonctions de corrélation sont construites sur l’ensemble des données, les paramètres ajustés sont donc donnés uniquement pour le redshift effectif  $z_{\text{eff}} = 2,334$  de la mesure. L’appendice F de **prov** présente cependant l’analyse des données DR16 dans deux bins en redshift. Mais ces deux bins ne sont pas suffisant pour estimer l’évolution des paramètres  $\text{Ly}\alpha$  dans la gamme en redshift  $1,9 < z < 3,6$ . Afin d’estimer  $b_{\text{eff,Ly}\alpha}(z)$  et  $\beta_{\text{Ly}\alpha}(z)$  dans cette gamme, nous avons produit l’analyse des données DR16 dans quatres bins en redshift. De manière à limiter les potentielles systématiques, nous nous limitons à l’analyse de la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (abrégée en  $\text{Ly}\alpha \times \text{Ly}\alpha$  dans la suite de ce chapitre). Pour constituer chacun des bins en redshift, nous pourrions séparer les paires de pixels selon leur redshift effectif. Cependant, à cause de l’ajustement du continuum, cette stratégie induit des corrélations parasites lorsqu’une forêt se trouve dans deux bins en redshift à la fois. Pour palier ce problème, nous divisons l’échantillon de forêts selon le redshift des quasars (voir Appendice B de AGATHE et al. (2019)). Les quatres intervalles choisis pour construire les bins en redshift sont  $[0; 2,35]$ ,  $[2,35; 2,65]$ ,  $[2,65; 3,05]$  et  $[3,05; 10]$ . Dans chacun des bins, nous calculons la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , ainsi que la matrice de distorsion et la matrice des métaux. Enfin, nous procédons à l’ajustement des quatres fonctions de corrélation. Le modèle utilisé pour cet ajustement est le même que celui utilisé pour l’analyse des données finale d’eBOSS (**prov**), il est présenté dans la section ?? . Le modèle est ajusté pour  $10 \leq r \leq 180 h^{-1} \text{ Mpc}$ . Chacune des fonctions de corrélation est ajustée au redshift effectif de la mesure. Ces redshifts sont  $z_1 = 2,136$ ,  $z_2 = 2,276$ ,  $z_3 = 2,551$  et  $z_4 = 2,914$ .

La figure 1.1 présente la fonction de corrélation et le meilleur ajustement du modèle dans chacun des bins en redshift. Les différents graphiques montrent les différents bins en redshift. Dans chaque graphique, la fonction de corrélation est affichée dans plusieurs bins en  $\mu$ . Le tableau 1.1 donne le résultat de l’ajustement dans chacun des bins en redshift. La première section du tableau donne les paramètres ajustés, la deuxième donne le redshift effectif  $z_{\text{eff}}$  et le  $\chi^2$  obtenu. Le nombre de bins dans lesquels la fonction de corrélation est ajustée est  $N_{\text{bin}} = 1590$ , ce qui donne un nombre de degrés de liberté  $N_{\text{d.o.f.}} = 1590 - 13 = 1577$ . Enfin, la troisième colonne donne le biais et le biais effectif du  $\text{Ly}\alpha$ .

Une fois cette analyse produite, et toujours dans le but d’obtenir  $b_{\text{eff,Ly}\alpha}(z)$  et  $\beta_{\text{Ly}\alpha}(z)$  pour  $1,9 < z < 3,6$ , nous ajustons les paramètres  $\text{Ly}\alpha$  mesurés dans les données par une loi de puissance. La figure ?? donne les mesures  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans les quatre bins en redshift, ainsi que l’ajustement fait sur ces quatres points. Pour le biais effectif, nous obtenons  $b_{\text{eff,Ly}\alpha}(z) \propto (1+z)^\gamma$  avec  $\gamma = 3,474 \pm 0,025$ . Pour le paramètre RSD, nous obtenons  $\beta_{\text{Ly}\alpha}(z) \propto (1+z)^\gamma$  avec  $\gamma = -2,32 \pm 1,97$ . Ces deux ajustements sont utilisés comme référence pour l’ajustement des paramètres des mocks (section ??). Ils sont extrapolés de  $z = 1,9$  jusqu’à  $z = 3,6$ .

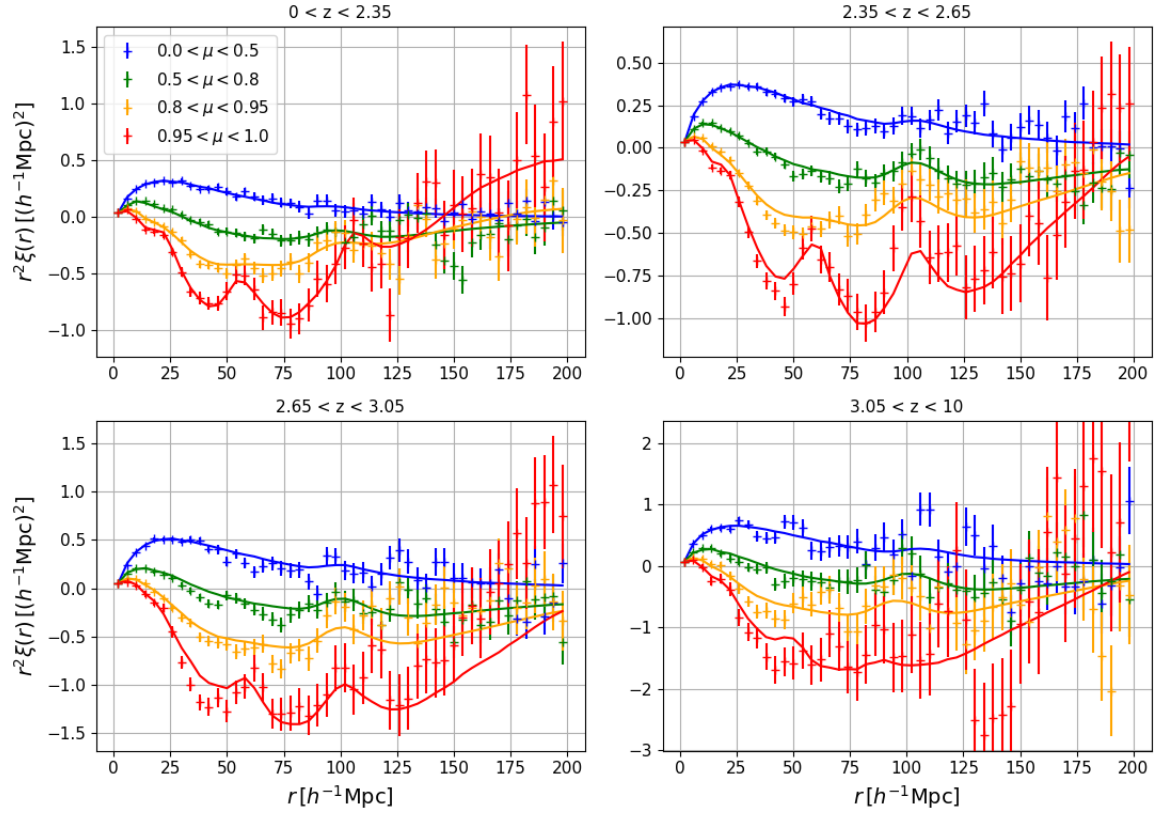


FIGURE 1.1 – Fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  dans chacun des bins en redshift de l'analyse. Les courbes en trait plein donne le meilleur ajustement du modèle obtenu avec `picca`. Chaque graphique correspond à un bin en redshift. Pour chacun des bins, la fonction de corrélation et l'ajustement sont montrés dans quatre bins en  $\mu$ .

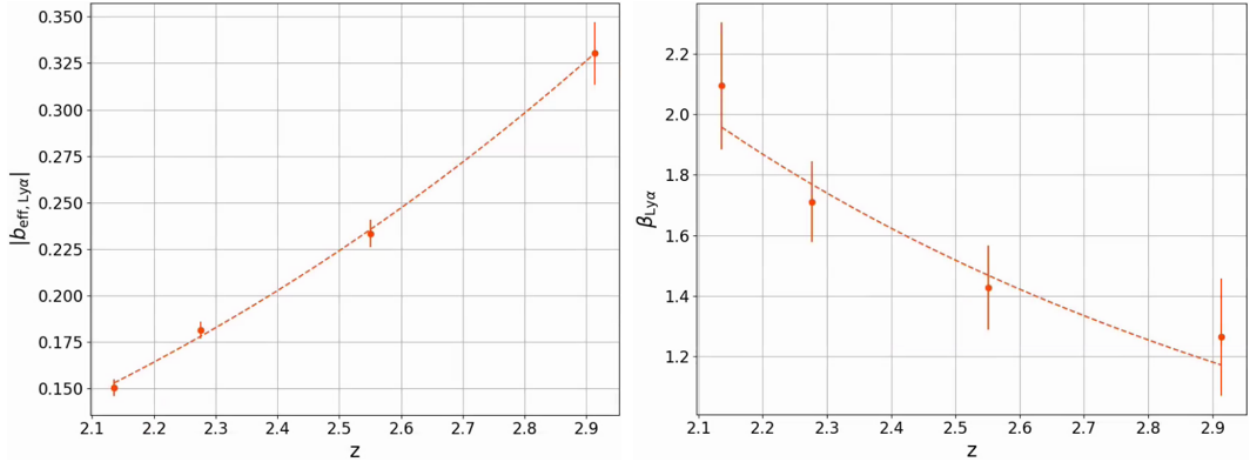


FIGURE 1.2 – Mesure des paramètres  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans les données DR16. Les mesures sont faites dans quatre bins en redshift, indiquées par les points. La ligne en pointillés donne le meilleur ajustement d'une loi de puissance. Cet ajustement donne  $b_{\text{eff,Ly}\alpha}(z) \propto (1+z)^\gamma$  avec  $\gamma = 3,474 \pm 0,025$  et  $\beta_{\text{Ly}\alpha}(z) \propto (1+z)^\gamma$  avec  $\gamma = -2,32 \pm 1,97$ .

TABLE 1.1 – Résultats de l’ajustement fait avec `picca` des fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  calculées sur les données DR16. Chaque colonne donne le résultat de l’ajustement d’un bin en redshift. La dernière colonne donne le résultat de l’ajustement de la corrélation moyennée sur ces quatre bins en redshift. La première section du tableau donne les paramètres du modèle qui sont ajustés. La seconde donne le  $\chi^2$  et le redshift effectif  $z_{\text{eff}}$ . Le nombre de bins sur lesquels le modèle est ajusté est  $N_{\text{bin}} = 1590$ . Le modèle comporte 13 paramètres libres. Enfin, la dernière section donne le biais et le biais effectif du  $\text{Ly}\alpha$ . Ils sont reliés aux paramètres  $b_{\eta, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  par les équations ?? et ??.

Paramètre	$0 < z < 2,35$	$2,35 < z < 2,65$	$2,65 < z < 3,05$	$3,05 < z < 10$	$0 < z < 10$
$\alpha_{\parallel}$	$1.063 \pm 0.066$	$1.019 \pm 0.041$	$1.029 \pm 0.072$	$1.12 \pm 0.081$	$1.047 \pm 0.034$
$\alpha_{\perp}$	$1.063 \pm 0.108$	$0.965 \pm 0.057$	$1.016 \pm 0.058$	$0.926 \pm 0.072$	$0.98 \pm 0.042$
$b_{\eta, \text{Ly}\alpha}$	$-0.1796 \pm 0.0058$	$-0.1938 \pm 0.0053$	$-0.2237 \pm 0.0084$	$-0.2929 \pm 0.0187$	$-0.1998 \pm 0.0039$
$\beta_{\text{Ly}\alpha}$	$2.094 \pm 0.21$	$1.711 \pm 0.133$	$1.427 \pm 0.138$	$1.265 \pm 0.194$	$1.633 \pm 0.087$
$10^3 b_{\eta, \text{SiIII}(1190)}$	$-1.83 \pm 1.1$	$-3.66 \pm 0.68$	$-2.8 \pm 1.01$	$0.36 \pm 1.64$	$-3.02 \pm 0.51$
$10^3 b_{\eta, \text{SiIII}(1193)}$	$-4.83 \pm 1.1$	$-1.94 \pm 0.69$	$-0.79 \pm 0.97$	$-2.13 \pm 1.72$	$-2.07 \pm 0.5$
$10^3 b_{\eta, \text{SiIII}(1260)}$	$-3.38 \pm 1.33$	$-1.97 \pm 0.8$	$-1.32 \pm 1.05$	$0.9 \pm 1.79$	$-2.16 \pm 0.63$
$10^3 b_{\eta, \text{SiIII}(1207)}$	$-7.87 \pm 1.1$	$-4.52 \pm 0.75$	$-2.11 \pm 1.05$	$-2.89 \pm 1.74$	$-4.59 \pm 0.52$
$10^3 b_{\eta, \text{CIV}(\text{eff})}$	$-4.77 \pm 2.54$	$-5.15 \pm 2.64$	$-5.06 \pm 2.62$	$-5.02 \pm 2.61$	$-5.12 \pm 2.63$
$b_{\text{HCD}}$	$-0.0596 \pm 0.007$	$-0.0452 \pm 0.006$	$-0.0665 \pm 0.01$	$-0.0228 \pm 0.0218$	$-0.0521 \pm 0.0045$
$\beta_{\text{HCD}}$	$0.551 \pm 0.086$	$0.56 \pm 0.086$	$0.508 \pm 0.088$	$0.502 \pm 0.09$	$0.603 \pm 0.083$
$10^2 A_{\text{sky}}$	$1.585 \pm 0.098$	$0.87 \pm 0.082$	$0.729 \pm 0.133$	$0.646 \pm 0.338$	$0.947 \pm 0.06$
$\sigma_{\text{sky}}$	$32.5 \pm 1.8$	$31.6 \pm 2.6$	$31.9 \pm 4.3$	$34.1 \pm 16.0$	$31.4 \pm 1.7$
$\chi^2$	1568	1512	1681	1675	1602
$z_{\text{eff}}$	2.136	2.276	2.551	2.914	2.334
$b_{\text{Ly}\alpha}$	$-0.0832 \pm 0.0065$	$-0.1099 \pm 0.0063$	$-0.1521 \pm 0.0103$	$-0.2247 \pm 0.023$	$-0.1187 \pm 0.0046$
$b_{\text{eff}, \text{Ly}\alpha}$	$-0.1506 \pm 0.0046$	$-0.1814 \pm 0.0045$	$-0.2336 \pm 0.0074$	$-0.3305 \pm 0.0168$	$-0.1922 \pm 0.0033$

## 1.2 Stabilité des paramètres $\text{Ly}\alpha$

Une fois avoir produit les ajustements présentés dans la section précédente, nous avons cherché à savoir si la mesure des paramètres  $\text{Ly}\alpha$  était fiable. Nous avons donc d’abord regardé la corrélation des paramètres  $b_{\eta, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  avec les autres paramètres du modèle. Le tableau 1.4 présente ces corrélations. Premièrement, nous pouvons remarquer que les paramètres  $\text{Ly}\alpha$  sont très corrélés entre eux. Ceci vient du fait que le paramètre RSD du  $\text{Ly}\alpha$  est très important : le signal se trouve principalement le long de la ligne de visée. Avec peu de signal perpendiculairement à la ligne de visée, il est difficile de décorréler le biais du paramètre RSD du traceur.

Deuxièmement, les paramètres du  $\text{Ly}\alpha$  sont très corrélés avec ceux des HCD, notamment  $\beta_{\text{Ly}\alpha}$  qui est corrélé à  $-75\%$  avec  $b_{\text{HCD}}$ . Ceci pose plusieurs problèmes : d’abord, la modélisation des HCD choisie dans `prov` et utilisée ici consiste à identifier puis masquer les HCD avec  $\log n_{\text{HI}} > 20,3$ , les HCD non masqués étant pris en compte par le terme  $F_{\text{HCD}}$  (voir section ??). Cependant, l’algorithme utilisé ne possède pas une efficacité de  $100\%$  (`#prov` ref papier solene). Des HCD avec une grande densité de colonne ne sont donc pas masqués. Ces HCD produisent des absorptions intenses, non prises en compte par le terme  $F_{\text{HCD}}$ , ce qui a pour effet d’augmenter le biais du  $\text{Ly}\alpha$ . De plus, le paramètre  $L_{\text{HCD}}$  est fixé à  $10 h^{-1}$  Mpc car il est corrélé avec les autres paramètres. Sa valeur, qui dépend de la distribution des HCD non masqués, est difficile à déterminer. Du fait des corrélations avec les autres paramètres, les paramètres des HCD obtenus dépendent de la valeur de  $L_{\text{HCD}}$  choisie. La figure 1.3 montre la dépendance de  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  avec  $L_{\text{HCD}}$ . A cause des corrélations entre les paramètres liés aux HCD et ceux liés au  $\text{Ly}\alpha$ , le fait de changer  $L_{\text{HCD}}$  change aussi les paramètres  $\text{Ly}\alpha$  obtenus. La figure 1.4 montre la dépendance de  $b_{\text{eff}, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  avec  $L_{\text{HCD}}$ . Ainsi, le paramètre RSD  $\beta_{\text{Ly}\alpha}$  est très corrélé avec  $L_{\text{HCD}}$ . Lorsque nous laissons libre  $L_{\text{HCD}}$ , en utilisant un prior gaussien centré sur  $10 h^{-1}$  Mpc et avec une largeur  $\sigma = 1 h^{-1}$  Mpc, nous mesurons une corrélation entre  $L_{\text{HCD}}$  et  $\beta_{\text{Ly}\alpha}$  de

TABLE 1.2 – Corrélations des paramètres du modèle avec le paramètre  $b_{\eta, \text{Ly}\alpha}$  (colonne du milieu) et  $\beta_{\text{Ly}\alpha}$  lors de l’ajustement de la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ . L’ajustement est fait sur l’addition des fonctions de corrélations calculées dans chaque bin en redshift, soit l’ensemble des données DR16.

Paramètre	$b_{\eta, \text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$
$\alpha_{\parallel}$	0 %	0 %
$\alpha_{\perp}$	1 %	−2 %
$b_{\eta, \text{Ly}\alpha}$	100 %	−87 %
$\beta_{\text{Ly}\alpha}$	−87 %	100 %
$10^3 b_{\eta, \text{SiII}(1190)}$	2 %	−8 %
$10^3 b_{\eta, \text{SiII}(1193)}$	3 %	−6 %
$10^3 b_{\eta, \text{SiII}(1260)}$	−1 %	−3 %
$10^3 b_{\eta, \text{SiIII}(1207)}$	6 %	5 %
$10^3 b_{\eta, \text{CIV}(\text{eff})}$	−7 %	−10 %
$b_{\text{HCD}}$	48 %	−75 %
$\beta_{\text{HCD}}$	35 %	−23 %
$10^2 A_{\text{sky}}$	34 %	−19 %
$\sigma_{\text{sky}}$	−10 %	−2 %

−38 %. Lorsque  $L_{\text{HCD}}$  est laissé totalement libre, il est corrélé, en valeur absolue, à plus de 85 % avec les paramètres  $b_{\text{eff}, \text{Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et  $b_{\eta, \text{SiIII}(1207)}$ .

Enfin, le modèle des HCD choisi influence la mesure des paramètres  $\text{Ly}\alpha$ . Toujours dans le but d’avoir une mesure robuste des paramètres  $\text{Ly}\alpha$ , nous avons essayé d’utiliser un autre modèle pour les HCD, développé par Edmond Chaussidon et Julien Guy (modèle C-G). Nous détaillons l’analyse en utilisant ce modèle dans la section ??.

## 2 Etude de la modélisation des HCD

Suite aux différents points énoncés dans la section précédente, nous avons étudié l’effet qu’ont les HCD sur le  $\text{Ly}\alpha$  dans les mocks. En effet, les mocks sont l’outil parfait pour ce genre d’analyse : ils permettent, contrairement aux données, de connaître la quantité de  $\text{Ly}\alpha$  et de HCD présents, et de comparer cette quantité à ce qui est mesuré par l’ajustement. Dans cette section, nous comparons les paramètres  $\text{Ly}\alpha$  mesurés dans les mocks sans HCD (raw mocks, eboss-0.0) et avec HCD (eboss-0.2 et eboss-0.3 #prov). Nous comparons aussi les paramètres  $\text{Ly}\alpha$  mesurés en utilisant différentes modélisations des HCD.

### 2.1 Comparaison des mocks

Comme expliqué dans le chapitre ??, nous avons analysé 30 réalisations des raw mocks, des mocks eboss-0.0 et des mocks eboss-0.2. Dans chacun des cas, nous ajustons le modèle sur  $20 < r < 180 h^{-1} \text{ Mpc}$ , et mesurons les paramètres du  $\text{Ly}\alpha$ . La figure ?? présente les mesures de ces paramètres dans chaque bin en redshift pour chacune des versions des mocks. Nous pouvons remarquer que les valeurs des paramètres  $\text{Ly}\alpha$  mesurés changent selon la version des mocks. Les paramètres mesurés dans les raw mocks sont très proches des paramètres visés, mesurés dans les données. Ceci montre que la procédure d’ajustement des paramètres des mocks que nous avons mise en place est efficace pour obtenir les bons  $b_{\text{eff}, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$ .

Lorsque nous comparons maintenant les valeurs de  $b_{\text{eff}, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  mesurées dans les raw mocks à celle mesurées dans les mocks eboss-0.0, nous observons un écart statistiquement significatif. L’effet

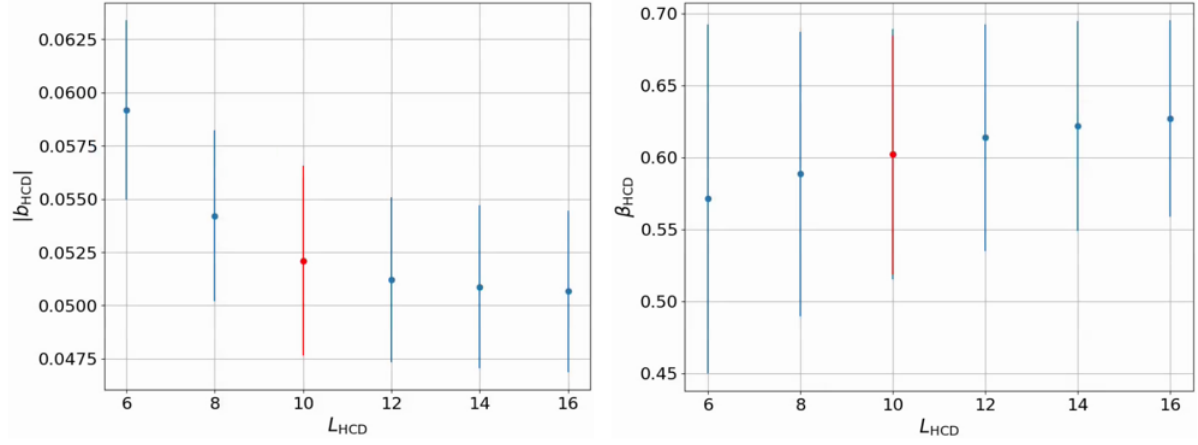


FIGURE 1.3 – Evolution des mesures des paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  en fonction de la valeur  $L_{\text{HCD}}$  choisie pour l’ajustement. Le biais des HCD est corrélé avec la valeur de  $L_{\text{HCD}}$ .

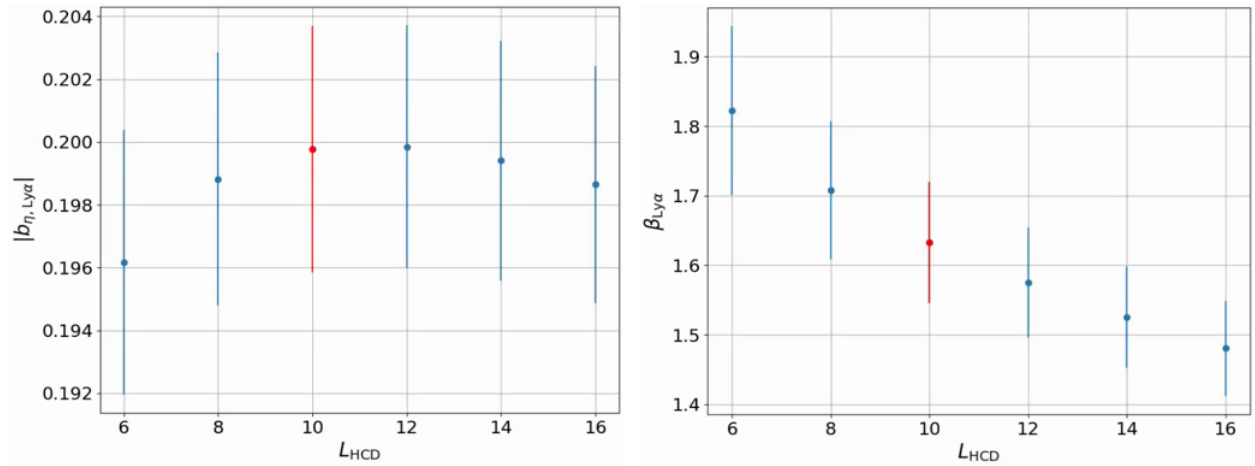


FIGURE 1.4 – Evolution des mesures des paramètres  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  en fonction de la valeur  $L_{\text{HCD}}$  choisie pour l’ajustement. Le paramètre RSD du Ly $\alpha$  est très corrélé avec la valeur de  $L_{\text{HCD}}$ .

sur  $\beta_{\text{Ly}\alpha}$  est faible, et les valeurs de  $\beta_{\text{Ly}\alpha}$  mesurées dans les mocks eboss-0.0 restent compatibles avec les données DR16. Cependant, l’effet sur le biais effectif  $b_{\text{eff,Ly}\alpha}$  est important et statistiquement significatif. Cela implique que l’ajout du continuum et du bruit par quickquasars, l’ajustement du continuum ou la prise en compte des effets liés à cet ajustement par la matrice de distorsion affecte la mesure de  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$ . Nous suspectons la matrice de distorsion de ne pas capturer tous les effets produits par l’ajustement du continuum, ce qui pourrait expliquer l’écart entre les raw mocks et les mocks eboss-0.0.

Enfin, nous observons de nouveau un écart entre les valeurs de  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  mesurées dans les mocks eboss-0.0 et eboss-0.2. L’écart mesuré pour  $b_{\text{eff,Ly}\alpha}$  entre les mocks eboss-0.2 et eboss-0.0 est comparable à celui mesuré entre les mocks eboss-0.0 et les raw mocks. L’écart mesuré sur  $\beta_{\text{Ly}\alpha}$  est plus important. Les valeurs obtenues dans l’ajustement des mocks eboss-0.2 ne sont pas compatibles avec celles mesurées dans les données DR16. Nous suspectons les HCD d’être à l’origine de cet écart. Comme expliqué dans la section 1.2, les paramètres Ly $\alpha$  sont très corrélés avec ceux des HCD. Nous aurions pu par exemple choisir un  $L_{\text{HCD}}$  plus faible dans l’ajustement des mocks eboss-0.2, ce qui aurait permis d’avoir des mesures compatibles de  $\beta_{\text{Ly}\alpha}$  entre les mocks eboss-0.0 et eboss-0.2 (voir figure 1.4).

## 2.2 Effet du masquage des HCD

Dans l’analyse des mocks eboss-0.2, comme pour les données, les HCD pour lesquels  $\log n_{\text{HI}} > 20,3$  sont masqués lors du calcul des  $\delta_F$ . Comme expliqué dans la section 1.2, le masquage des HCD dans les données s’effectue selon le résultat de l’algorithme d’identification, alors que dans les mocks, les HCD sont masqués à partir du “vrai” catalogue. Nous étudions ici l’effet du masquage à partir du catalogue produit par l’algorithme d’identification. Pour ce faire, nous produisons l’analyse d’une réalisation de mock eboss-0.2, pour laquelle nous utilisons l’algorithme d’identification pour créer un catalogue de HCD. Le champ  $\delta_F$  est calculé en masquant les HCD identifiés par l’algorithme, puis la fonction de corrélation Ly $\alpha$ ×Ly $\alpha$  est estimée dans les quatre bins en redshift utilisés jusqu’ici. Nous nommons cette analyse *eboss-0.2\_finder*. La figure 1.5 présente la fonction de corrélation Ly $\alpha$ ×Ly $\alpha$  estimée à partir de la même réalisation des mocks, en version eboss-0.0, eboss-0.2 et eboss-0.2\_finder. Les fonctions de corrélation affichées sont la moyenne des fonctions de corrélation estimées dans les quatre bins en redshift. Dans les trois versions, le code `quickquasars` utilise les mêmes quasars pour produire les spectres synthétiques, et ajoute le même bruit à ces spectres dans les trois cas. Ceci nous permet d’avoir les mêmes fluctuations statistiques dans le calcul de la corrélation Ly $\alpha$ ×Ly $\alpha$ , et ainsi d’avoir des mesures de biais comparables. L’effet des HCD (avec  $17,2 < \log n_{\text{HI}} < 20,3$ ) est visible en comparant la corrélation montrée en bleu à celle montrée en rouge. Comme expliqué dans la section ??, l’effet principal des HCD est d’augmenter le biais effectif. Par ailleurs, le fait que l’effet des HCD soit légèrement plus important sur la corrélation de la version eboss-0.2 que sur celle de la version eboss-0.2\_finder suggère que l’algorithme identifie et masque les HCD pour lesquels  $\log n_{\text{HI}} > 20,3$ , et identifie des HCD avec  $\log n_{\text{HI}} < 20,3$  et les reconstruit avec  $\log n_{\text{HI}} > 20,3$ . Ceci a pour effet de masquer des HCD qui ne possèdent pas une densité de colonne supérieure à 20,3. Ceci semble confirmé par la figure 1.6, qui compare la densité de colonne  $\log n_{\text{HI}}$  trouvée par le finder (output) à la densité de colonne des HCD des mocks (input). Nous voyons sur cette figure que l’algorithme a tendance à surestimer  $\log n_{\text{HI}}$ , ce qui est en accord avec nos observations.

Le tableau ?? donne les résultats des ajustements des trois corrélations présentées sur la figure 1.5. La statistique d’une seule réalisation n’est pas suffisante pour identifier des potentielles systématiques. Cependant, la précision de la mesure des paramètres Ly $\alpha$  dans les données DR16 étant comparable à la celle des mocks eboss-0.2, les potentielles systématiques sont inférieures à l’erreur statistique sur cette mesure. Il serait tout de même intéressant de mener cette analyse sur un plus grand nombre de réalisations.

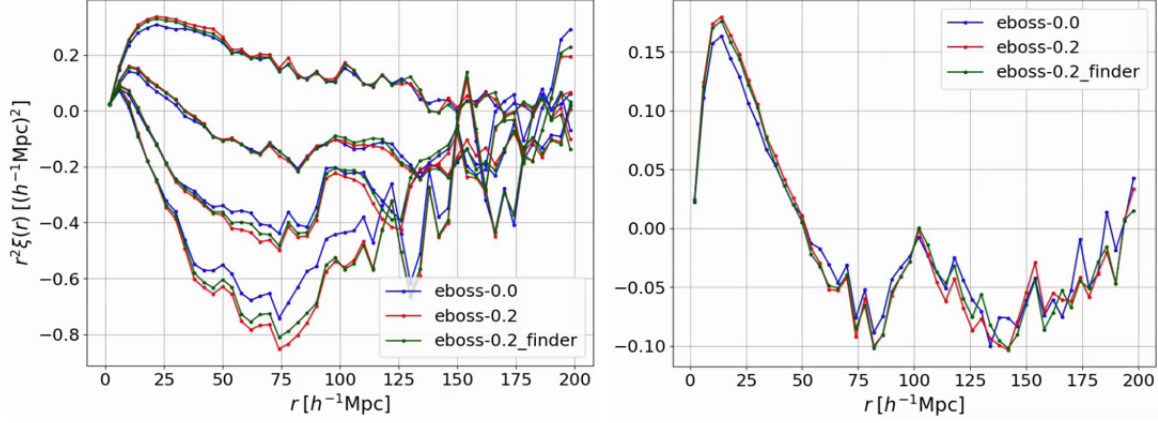


FIGURE 1.5 – Fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimées à partir d’une réalisation des mocks eboss-0.0 (bleu), eboss-0.2 (rouge) et eboss-0.2\_finder (vert). La version eboss-0.2\_finder correspond aux mocks eboss-0.2, dans lesquels les HCD ont été masqués en utilisant le catalogue de HCD produit par l’algorithme d’identification.

TABLE 1.3 – Résultat de l’ajustement de la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimées à partir d’une réalisation des mocks eboss-0.0 (bleu), eboss-0.2 (rouge) et eboss-0.2\_finder (vert). La version eboss-0.2\_finder correspond aux mocks eboss-0.2, dans lesquels les HCD ont été masqués en utilisant le catalogue de HCD produit par l’algorithme d’identification.

version	$b_{\text{eff}, \text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$	$b_{\text{HCD}}$	$\beta_{\text{HCD}}$
eboss-0.0	$-0.1970 \pm 0.0009$	$1.641 \pm 0.039$		
eboss-0.2	$-0.1979 \pm 0.0038$	$1.578 \pm 0.069$	$-0.0201 \pm 0.0053$	$0.499 \pm 0.090$
eboss-0.2_finder	$-0.1951 \pm 0.0039$	$1.592 \pm 0.074$	$-0.0186 \pm 0.0054$	$0.494 \pm 0.091$



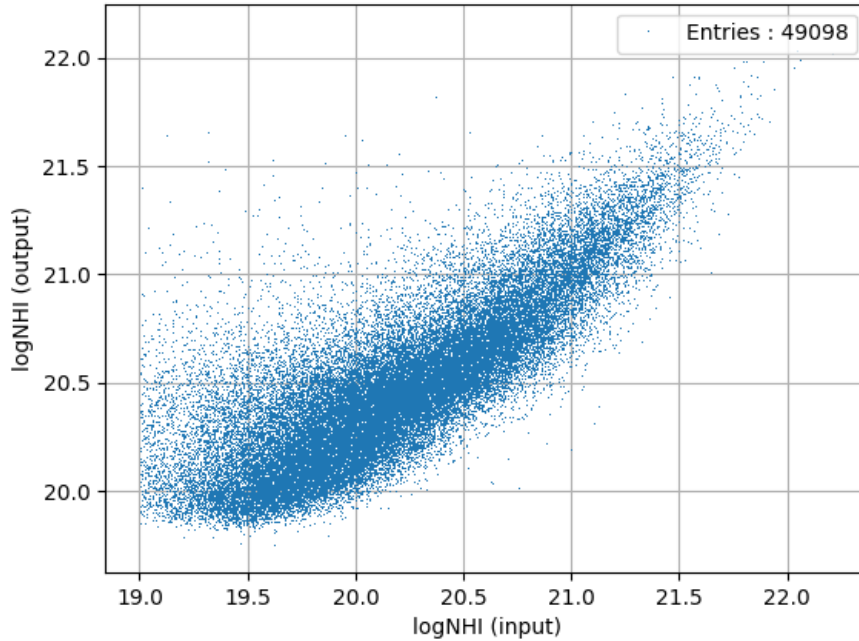


FIGURE 1.6 – Densité de colonne  $\log n_{\text{HI}}$  trouvée par le finder ( $\log\text{NHI}$  output) en fonction de la densité de colonne des HCD des mocks ( $\log\text{NHI}$  input). Cette comparaison est produite par Jim Rich et Solène Chabannier, du groupe cosmologie du CEA, à l’aide des mocks.

### 2.3 Modélisation alternative des HCD

Toujours dans l’optique de tester la robustesse de la mesure des paramètres  $\text{Ly}\alpha$ , nous avons utilisé une modélisation des HCD différente de celle décrite dans la section ?? et utilisée jusqu’ici pour modéliser les mocks et les données. Ce modèle est développé par Edmond Chaussidon et Julien Guy, au sein du groupe  $\text{Ly}\alpha$  de la collaboration DESI. Nous faisons référence à ce modèle via le nom *modèle C-G*. Le modèle des HCD décrit dans la section ?? est dénommé *modèle de Rogers*. Le modèle C-G, contrairement au modèle de Rogers, n’a pas besoin de masquer les HCD identifiés par l’algorithme. Il prend en compte les effets sur les corrélations  $\text{Ly}\alpha \times \text{Ly}\alpha$  et  $\text{Ly}\alpha \times \text{QSO}$  produits par l’ensemble des HCD dans les données. Les deux modèles utilisent la modélisation définie dans l’équation ?. Cependant, dans le cas du modèle C-G, plutôt que de définir  $F_{\text{HCD}}$  comme une fonction exponentielle avec une paramètre  $L_{\text{HCD}}$  qui reflète la taille caractéristique des HCD non masqués (équation ??), la fonction  $F_{\text{HCD}}$  qui est utilisée correspond au profil de Voigt résultant de la distribution en  $\log n_{\text{HI}}$  des HCD présents dans les données. L’avantage de cette méthode est qu’elle n’utilise pas le paramètre effectif  $L_{\text{HCD}}$ . Elle permet donc de modéliser des distributions de HCD avec une plus grande gamme en  $\log n_{\text{HI}}$ , là où le modèle de Rogers, utilisé pour analyser les données DR16, ne fonctionne plus très bien. De plus, le modèle C-G permet de s’affranchir des potentiels systématiques produites par l’utilisation de l’algorithme de détection. Cependant, afin de calculer le profil de Voigt correspondant à la distribution de HCD présents, nous devons justement connaître cette distribution. Si cela est possible pour les mocks, ce n’est pas possible pour les données, car nous ignorons la distribution en  $\log n_{\text{HI}}$  des HCD pour lesquels  $\log n_{\text{HI}} < 20,3$ . Ainsi, lorsque nous analysons les données DR16 avec le modèle C-G, nous supposons que la distribution de HCD dans les données est la même que celle des mocks. Nous utilisons donc le même profil de Voigt.

Nous comparons ici les résultats de l’ajustement des fonctions de corrélation obtenues à partir

TABLE 1.4 – Corrélations des paramètres du modèle avec le paramètre  $b_{\text{eff,Ly}\alpha}$  (colonne du milieu) et  $\beta_{\text{Ly}\alpha}$  lors de l’ajustement de la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  en utilisant le modèle C-G. L’ajustement est fait sur l’addition des fonctions de corrélation calculées dans chaque bin en redshift, soit l’ensemble des données DR16.

Paramètre	$b_{\eta,\text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$
$\alpha_{\parallel}$	−1 %	1 %
$\alpha_{\perp}$	1 %	−2 %
$b_{\eta,\text{Ly}\alpha}$	100 %	−61 %
$\beta_{\text{Ly}\alpha}$	−61 %	100 %
$10^3 b_{\eta,\text{SiII}(1190)}$	−2 %	−21 %
$10^3 b_{\eta,\text{SiII}(1193)}$	−2 %	−16 %
$10^3 b_{\eta,\text{SiII}(1260)}$	−2 %	−11 %
$10^3 b_{\eta,\text{SiIII}(1207)}$	4 %	−28 %
$10^3 b_{\eta,\text{CIV}(\text{eff})}$	−5 %	−7 %
$b_{\text{HCD}}$	35 %	−91 %
$\beta_{\text{HCD}}$	78 %	−49 %
$10^2 A_{\text{sky}}$	29 %	−12 %
$\sigma_{\text{sky}}$	−15 %	−7 %

des données et des mocks en utilisant la modélisation des HCD de Rogers, aux résultats obtenus en utilisant le modèle C-G. Premièrement, nous comparons l’ajustement des fonctions de corrélation estimées à partir du champ  $\delta_F$  où les HCD ont été masqués. Ceci nous permet de faire une comparaison directe des deux modèles ajustés sur les mêmes fonctions de corrélation. Le profile de Voigt que nous utilisons ici est le profile de Voigt calculé à partir de la distribution des HCD des mocks pour lesquels  $\log n_{\text{HI}} < 20,3$ . La figure 1.7 montre la mesure des paramètres  $b_{\text{eff,Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et la somme  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  sur l’auto-corrélation issues des données et des mocks eboss-0.2, en utilisant le modèle de Rogers et le modèle C-G. Sur cette figure, nous pouvons remarquer que l’ajustement des mocks avec le modèle de Rogers et l’ajustement avec le modèle C-G donnent des mesures de  $b_{\text{eff,Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$  et  $b_{\text{HCD}}$  très similaires. De plus, la somme  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  est identique. Ceci est rassurant car la somme  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  donne l’amplitude de la fonction de corrélation. La corrélation ajustée étant la même, nous nous attendons à mesurer la même somme  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$ . En ce qui concerne les données, les mesures de  $b_{\text{eff,Ly}\alpha}$  et  $b_{\text{HCD}}$  en utilisant le modèle de Rogers ou le modèle C-G ne sont pas compatibles. Les sommes  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  dans ces deux cas ne sont pas identiques mais restent compatibles. Ainsi, les proportions de HCD et de  $\text{Ly}\alpha$  mesurés avec le modèle de Rogers et le modèle C-G ne sont pas les mêmes. Nous pouvons aussi remarquer que les barres d’erreurs liées à la mesure des paramètres, notamment celles de  $\beta_{\text{Ly}\alpha}$ , sont bien plus importantes dans le cas du modèle C-G. Ceci provient de la corrélation des paramètres  $\text{Ly}\alpha$ , notamment  $\beta_{\text{Ly}\alpha}$ , avec les autres paramètres qui est plus importante dans le modèle C-G que dans le modèle de Rogers.

Le fait que l’ajustement des mocks avec le modèle de Rogers et le modèle C-G donne des résultats compatibles est rassurant, car dans les deux cas, c’est la même corrélation qui est ajustée. Cependant, le fait que les résultats des ajustements dans le cas des données ne soient pas compatibles suggère qu’il y a quelque chose de non identifié dans les données, et qui est pris en parti en compte par la modélisation des HCD, notamment par le modèle C-G. Ce dernier point est appuyé par le fait que le  $\chi^2$  dans l’ajustement de la corrélation moyennée sur les quatres bins en redshift avec le modèle C-G est 15 points plus bas que celui obtenu dans l’ajustement avec le modèle de Roger.

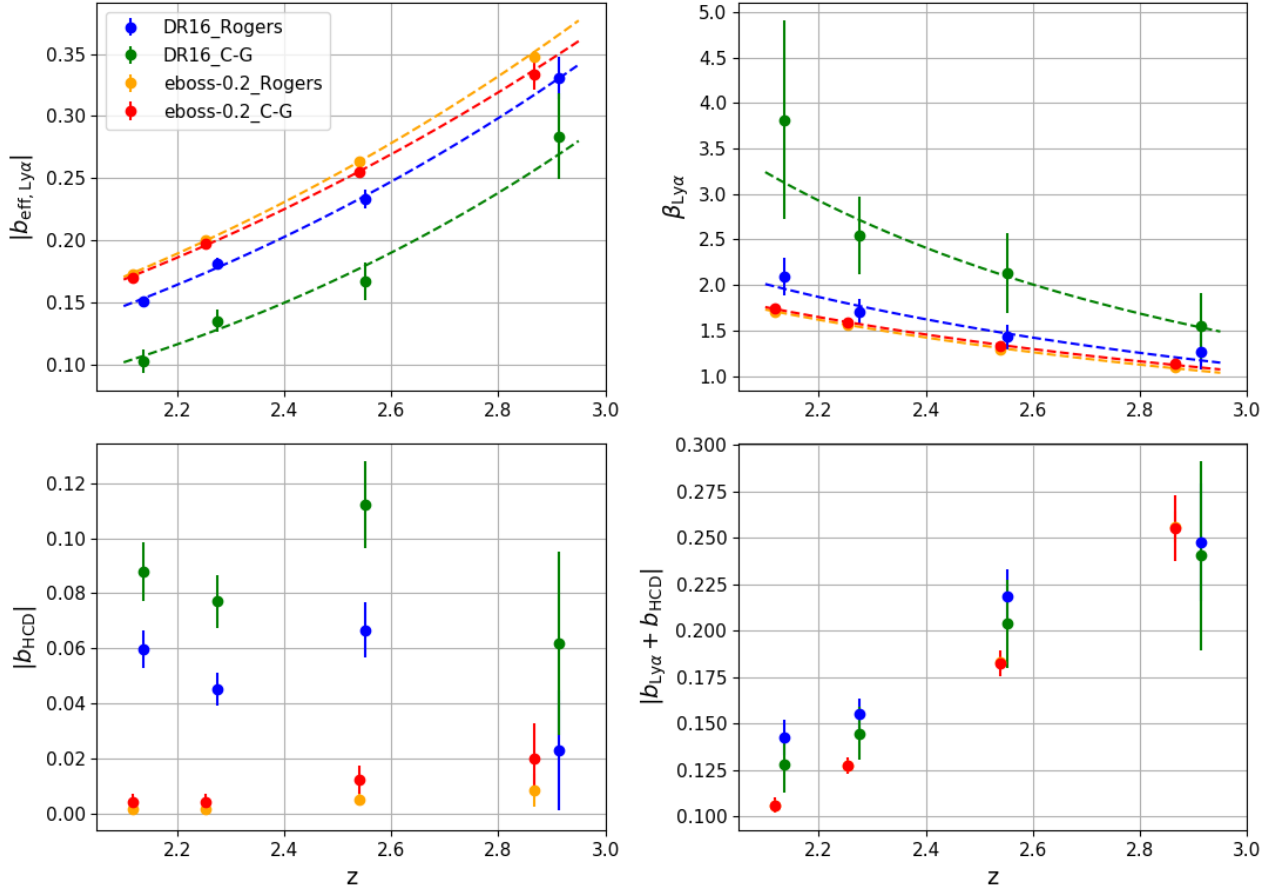


FIGURE 1.7 – Mesure des paramètres  $b_{\text{eff}, \text{Ly}\alpha}$  (haut gauche),  $\beta_{\text{Ly}\alpha}$  (haut droite),  $b_{\text{HCD}}$  (bas gauche) et  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  (bas droite) sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16 et des mocks eboss-0.2. Les données sont ajustées avec le modèle de Rogers (orange) et le modèle C-G (rouge). Les mocks sont ajustés avec le modèle de Rogers (bleu) et le modèle C-G (vert).

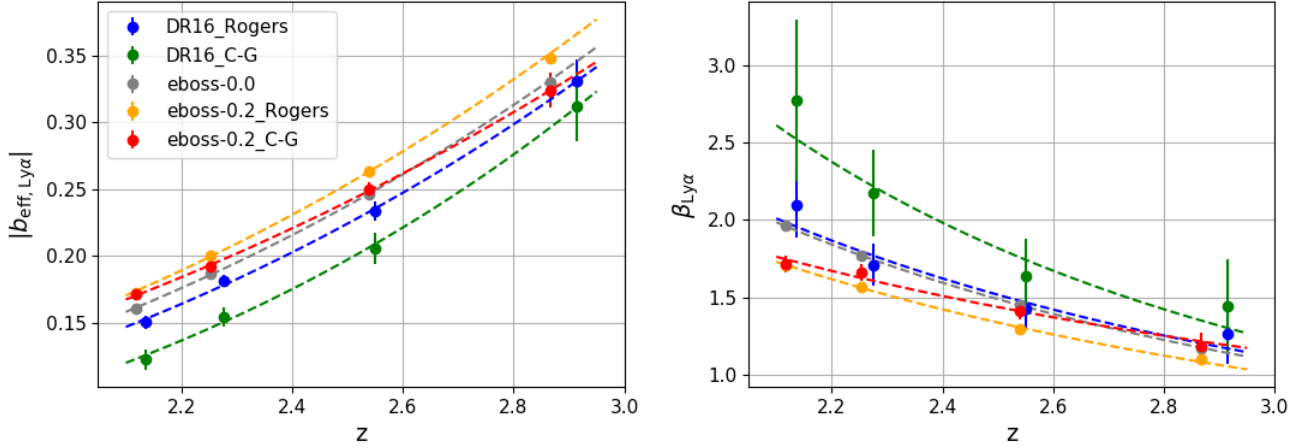


FIGURE 1.8 – Mesure des paramètres  $b_{\text{eff},\text{Ly}\alpha}$  (gauche) et  $\beta_{\text{Ly}\alpha}$  (droite) sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ . Les données représentées sont : les données DR16 avec les HCD masqués et ajustées avec le modèle de Rogers (bleu), les données DR16 sans masquage des HCD et ajustées avec le modèle C-G (vert), les mocks eboss-0.2 avec masquages des HCD et ajustés avec le modèle de Rogers (jaune) et les mocks eboss-0.2 sans masquage des HCD et ajustés avec le modèle C-G (rouge). A titre de comparaison, la mesure sur les mocks eboss-0.0 est représentée en gris.

Nous comparons maintenant la mesure des paramètres  $\text{Ly}\alpha$  faite sur la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , où les HCD sont masqués, ajustée avec le modèle de Rogers, à la mesure faite sur la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , où aucun HCD n’est masqué, ajustée avec le modèle C-G. La figure 1.8 présente cette comparaison pour les données et les mocks eboss-0.2. Sur cette figure, nous avons ajouté la mesure faite sur les mocks eboss-0.0, afin de pouvoir comparer les valeurs de  $b_{\text{eff},\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  obtenues dans l’ajustement des mocks eboss-0.2 à la valeur obtenue sans HCD. Nous pouvons noter que, dans le cas des mocks, l’écart entre le modèle de Roger et le modèle C-G est un peu plus important que dans la comparaison faite sur la figure 1.7. De ce fait, les mesures de  $b_{\text{eff},\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  faites sur les mocks eboss-0.2 sans masquage des HCD et avec le modèle C-G sont presque compatibles avec les mesures faites sur les mocks eboss-0.0. En ce qui concerne les données, paradoxalement l’écart est moins important que dans la comparaison faite sur la figure 1.7. Nous ne comprenons pas vraiment cela.

Nous avons aussi pu vérifier que ces deux modèles, lorsqu’ils sont ajustés sur les mocks eboss-0.0, mesurent une quantité de HCD compatible avec 0. Par ailleurs, dans une étude préliminaire au choix des paramètres  $\text{Ly}\alpha$  à utiliser pour construire les mocks, nous avons étudié comment ces deux modèles se comportent lorsque nous augmentons la quantité de HCD. Pour ce faire, à partir d’une même réalisation des mocks, nous avons produit deux versions de quickquasars eboss-0.2 : l’une avec le catalogue de HCD standard, l’autre avec un catalogue contenant 3 fois plus de HCD. Dans ces deux versions, les HCD qui vérifient  $\log n_{\text{HI}} > 20,3$  sont masqués. Afin de faciliter la comparaison des résultats, nous utilisons le même sous-échantillon de quasars pour produire les spectres synthétiques et calculer la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  à partir de ces deux versions. Puis, nous avons ajusté la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , estimées à partir de ces deux versions, avec le modèle de Rogers et le modèle C-G. Le paramètre  $b_{\text{HCD}}$  étant proportionnel au nombre de HCD, nous nous attendons à mesurer un biais trois fois plus grand dans la réalisation possédant trois fois plus de HCD. Cependant, ce n’est pas ce que nous observons. Le tableau ?? résume ces mesures. Comme observé précédemment, pour une même version, les mesures de  $b_{\text{HCD}}$  faites avec le modèle de Rogers ou le modèle C-G ne sont pas compatibles. De plus, pour un même modèle, nous remarquons que  $b_{\text{HCD}}$  mesuré dans la version contenant 3 fois plus de HCD n’est pas compatible avec 3 fois le biais des HCD mesuré dans l’autre version. Le biais des HCD mesuré dans la version contenant 3 fois plus de HCD est sous-estimé. Ceci

TABLE 1.5 – Mesures des paramètres  $b_{\text{eff,Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$  et  $b_{\text{HCD}}$  à partir de l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée sur une réalisation des mocks eboss-0.2, où les HCD avec  $\log n_{\text{HI}} > 20,3$  ont été masqués. Les deux premières lignes donnent les mesures faites avec le modèle de Rogers et le modèle C-G. Les deux dernières lignes donnent les mesures faites avec ces deux modèles en ajustant la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée sur la même réalisation des mocks mais avec un catalogue contenant 3 fois plus de HCD. Afin que les mesures soient comparables, nous utilisons le même sous-échantillon de quasars pour mesurer la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ .

Version	$b_{\text{eff,Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$	$b_{\text{HCD}}$
modèle Rogers	$-0.1912 \pm 0.0020$	$1.558 \pm 0.039$	$-0.0165 \pm 0.0030$
modèle C-G	$-0.1708 \pm 0.0058$	$1.746 \pm 0.095$	$-0.0323 \pm 0.0062$
modèle Rogers et 3 fois plus de HCD	$-0.2095 \pm 0.0025$	$1.295 \pm 0.039$	$-0.0296 \pm 0.0037$
modèle C-G et 3 fois plus de HCD	$-0.1713 \pm 0.0071$	$1.561 \pm 0.110$	$-0.0598 \pm 0.0077$

a pour effet de réduire la mesure de  $\beta_{\text{Ly}\alpha}$ . Nous pouvons tout fois noter que le modèle C-G produit des mesures de  $b_{\text{eff,Ly}\alpha}$  compatibles dans ces deux versions, ce qui n’est pas le cas du modèle de Rogers.

## 2.4 Mieux comprendre les HCD

Afin de comprendre pourquoi les deux modèles discutés dans la section précédente ont beaucoup de mal à distinguer la contribution des HCD de celle du  $\text{Ly}\alpha$ , nous avons essayé de comprendre l’effet des HCD sur les fonctions de corrélation.

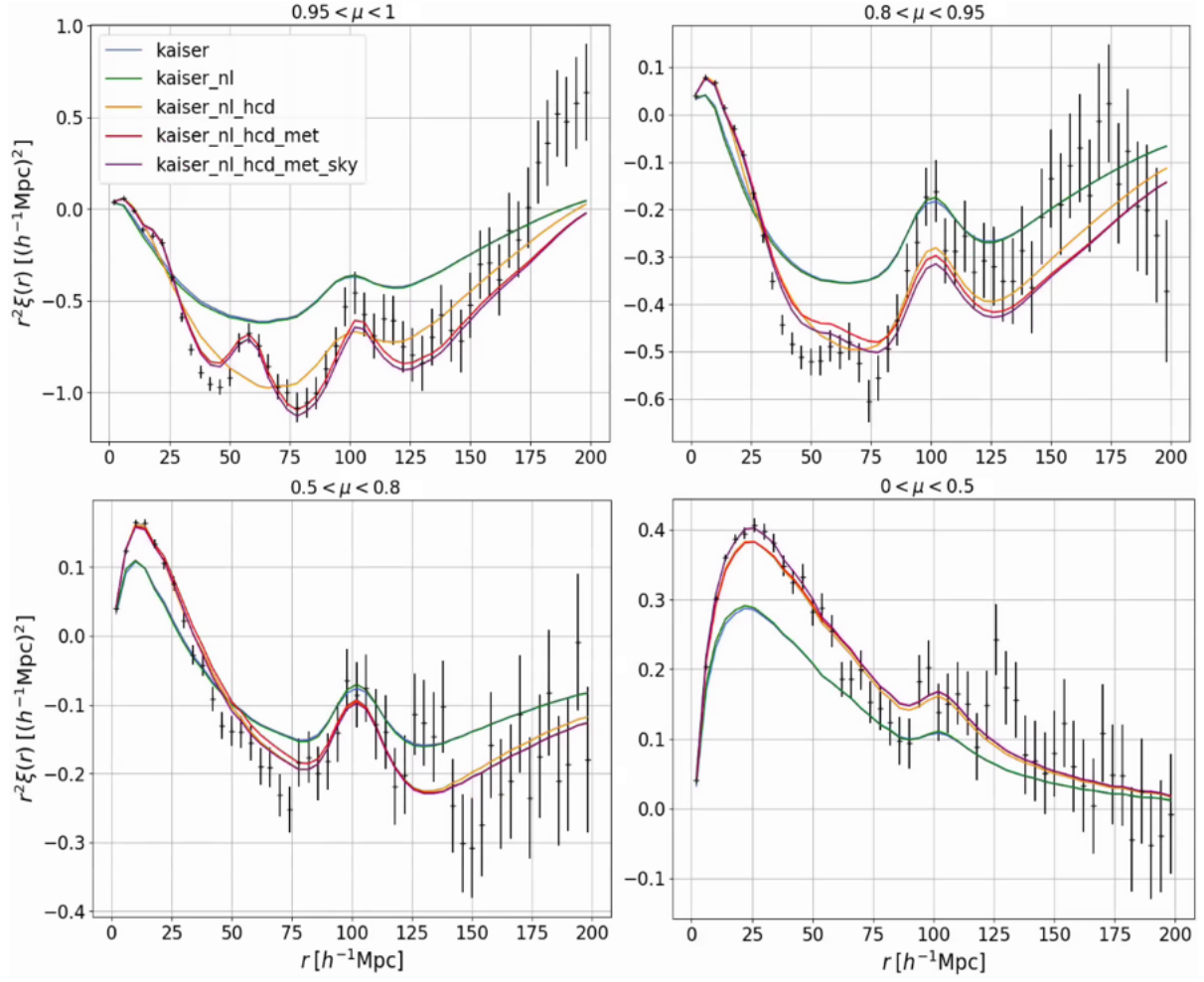


FIGURE 1.9 – bla

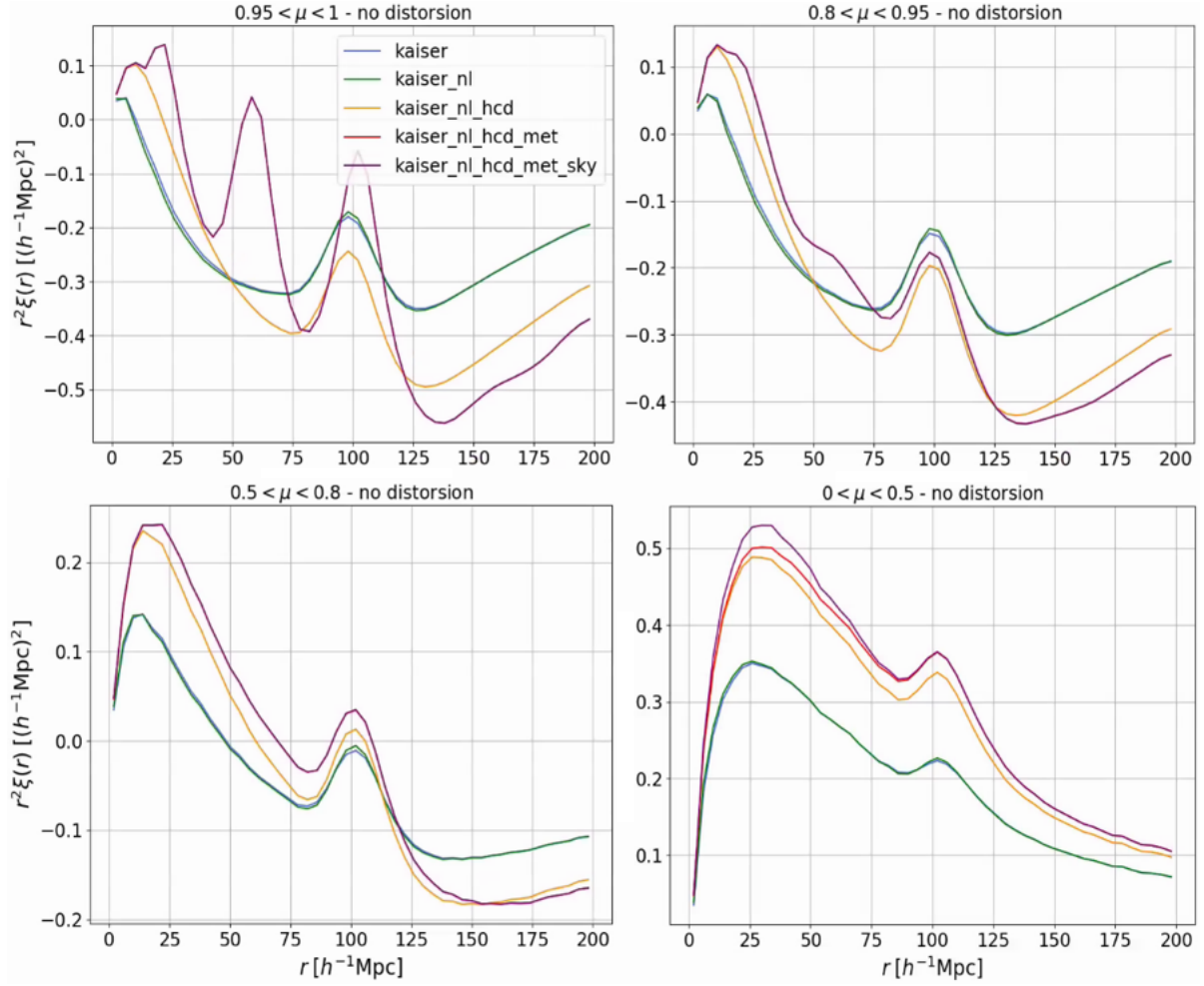


FIGURE 1.10 – bla





# Bibliographie

---

AGATHE, Victoria de Sainte et al. (2019). « Baryon acoustic oscillations at  $z = 2.34$  from the correlations of Lyman alpha absorption in eBOSS DR14 ». In : DOI : 10.1051/0004-6361/201935638. arXiv : 1904.03400.