

# 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 pour 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 une meilleure modélisation.

## 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 DU MAS DES BOURBOUX et al. (2020), utilise 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 estimées sur la gamme en redshift complète, les paramètres ajustés sont donc donnés uniquement pour le redshift effectif  $z_{\text{eff}} = 2,334$  de la mesure. L’appendice F de DU MAS DES BOURBOUX et al. (2020) présente cependant l’analyse des données DR16 dans deux bins en redshift. Mais ces deux bins ne sont pas suffisants pour estimer l’évolution des paramètres  $\text{Ly}\alpha$  dans la gamme en redshift  $1,9 < z < 3,6$ . Afin d’estimer  $b_{\text{eff},\text{Ly}\alpha}(z)$  et  $\beta_{\text{Ly}\alpha}(z)$  dans cette gamme, nous avons analysé les données DR16 dans quatre 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 DE SAINTE AGATHE et al. (2019)). Les quatre intervalles choisis pour construire les bins en redshift des quasars 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 quatre fonctions de corrélation. Le modèle utilisé pour cet ajustement est le même que celui utilisé pour l’analyse des données finale d’eBOSS, il est présenté dans la section ???. Le modèle est ajusté pour  $10 \leq r \leq 180 h^{-1}$  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 quatre bins en  $\mu$ . Le tableau 1.1 donne le résultat de l’ajustement dans chacun des bins en redshift, ainsi que l’ajustement fait sur la corrélation moyennée sur ces bins. Une fois cette analyse produite, et toujours dans le but d’obtenir  $b_{\text{eff},\text{Ly}\alpha}(z)$  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 1.2 donne les mesures  $b_{\text{eff},\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans les quatre bins en redshift, ainsi que l’ajustement fait sur ces quatre points. Pour le biais effectif, nous obtenons  $b_{\text{eff},\text{Ly}\alpha}(z) \propto (1+z)^\gamma$  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  $z = 3,6$ .

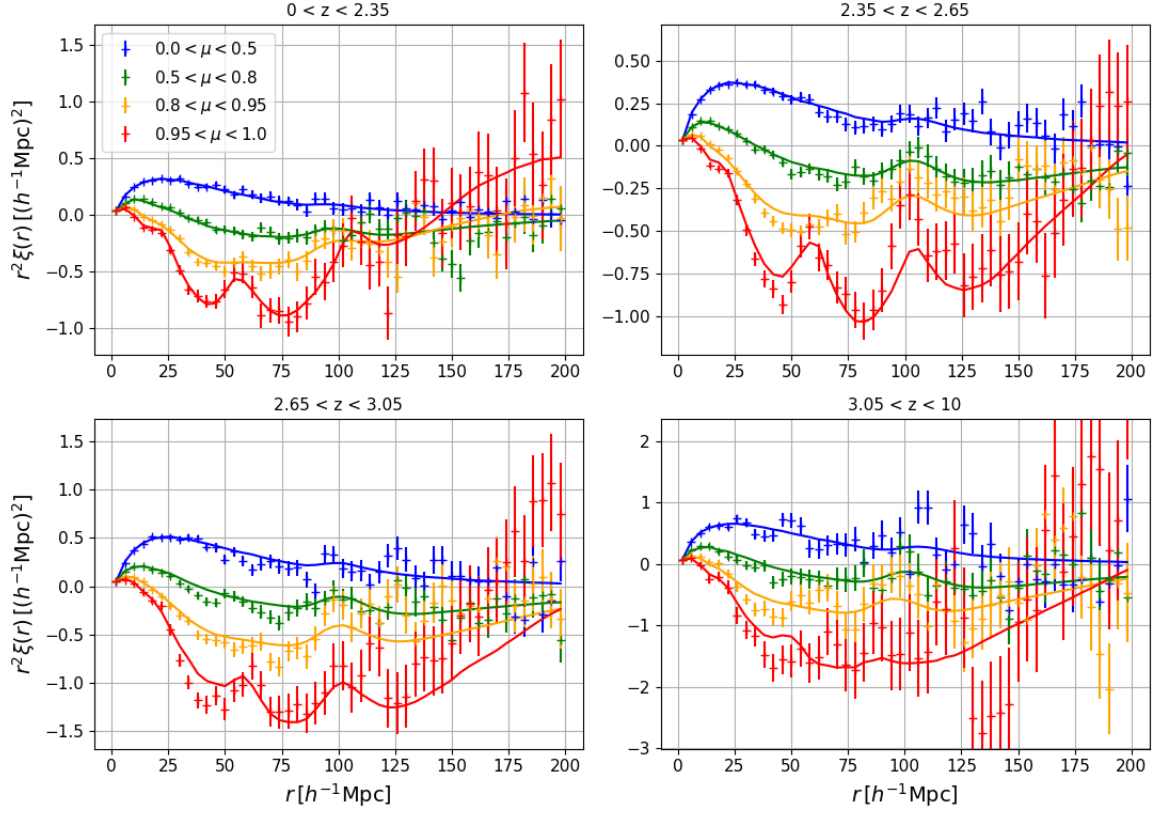


FIGURE 1.1 – Fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  dans chacun des bins en redshift de l’analyse. Les courbes en trait plein donne le meilleur ajustement du modèle obtenu avec *picca*. Chaque graphique correspond à un bin en redshift. Pour chacun des bins, la fonction de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

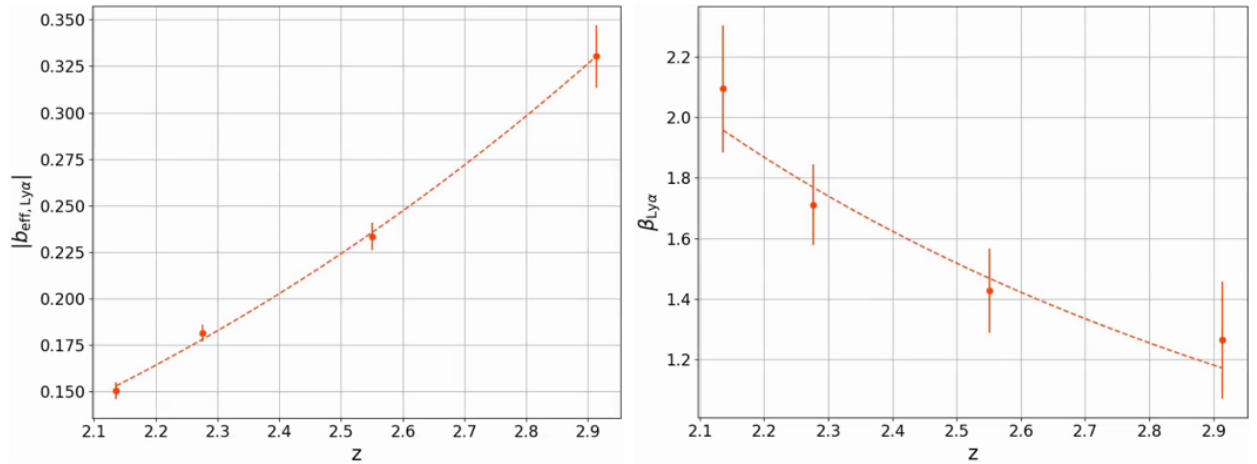


FIGURE 1.2 – Mesure des paramètres  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans les données DR16. Les mesures sont faites dans quatre bins en redshift, indiquées par les points. La ligne en pointillés donne le meilleur ajustement par une loi de puissance.

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 modèle comporte 13 paramètres libres. Le nombre de bins sur lesquels le modèle est ajusté est  $N_{\text{bin}} = 1590$ , ce qui donne un nombre de degrés de liberté  $n_{d.o.f.} = 1577$ . 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	$0 < z < 2,35$	$2,35 < z < 2,65$	$2,65 < z < 3,05$	$3,05 < z < 10$	$0 < z < 10$
$\alpha_{\parallel}$	$1.063 \pm 0.066$	$1.019 \pm 0.041$	$1.029 \pm 0.072$	$1.12 \pm 0.081$	$1.047 \pm 0.034$
$\alpha_{\perp}$	$1.063 \pm 0.108$	$0.965 \pm 0.057$	$1.016 \pm 0.058$	$0.926 \pm 0.072$	$0.98 \pm 0.042$
$b_{\eta, \text{Ly}\alpha}$	$-0.1796 \pm 0.0058$	$-0.1938 \pm 0.0053$	$-0.2237 \pm 0.0084$	$-0.2929 \pm 0.0187$	$-0.1998 \pm 0.0039$
$\beta_{\text{Ly}\alpha}$	$2.094 \pm 0.21$	$1.711 \pm 0.133$	$1.427 \pm 0.138$	$1.265 \pm 0.194$	$1.633 \pm 0.087$
$10^3 b_{\eta, \text{SiII}(1190)}$	$-1.83 \pm 1.1$	$-3.66 \pm 0.68$	$-2.8 \pm 1.01$	$0.36 \pm 1.64$	$-3.02 \pm 0.51$
$10^3 b_{\eta, \text{SiII}(1193)}$	$-4.83 \pm 1.1$	$-1.94 \pm 0.69$	$-0.79 \pm 0.97$	$-2.13 \pm 1.72$	$-2.07 \pm 0.5$
$10^3 b_{\eta, \text{SiII}(1260)}$	$-3.38 \pm 1.33$	$-1.97 \pm 0.8$	$-1.32 \pm 1.05$	$0.9 \pm 1.79$	$-2.16 \pm 0.63$
$10^3 b_{\eta, \text{SiIII}(1207)}$	$-7.87 \pm 1.1$	$-4.52 \pm 0.75$	$-2.11 \pm 1.05$	$-2.89 \pm 1.74$	$-4.59 \pm 0.52$
$10^3 b_{\eta, \text{CIV}(\text{eff})}$	$-4.77 \pm 2.54$	$-5.15 \pm 2.64$	$-5.06 \pm 2.62$	$-5.02 \pm 2.61$	$-5.12 \pm 2.63$
$b_{\text{HCD}}$	$-0.0596 \pm 0.007$	$-0.0452 \pm 0.006$	$-0.0665 \pm 0.01$	$-0.0228 \pm 0.0218$	$-0.0521 \pm 0.0045$
$\beta_{\text{HCD}}$	$0.551 \pm 0.086$	$0.56 \pm 0.086$	$0.508 \pm 0.088$	$0.502 \pm 0.09$	$0.603 \pm 0.083$
$10^2 A_{\text{sky}}$	$1.585 \pm 0.098$	$0.87 \pm 0.082$	$0.729 \pm 0.133$	$0.646 \pm 0.338$	$0.947 \pm 0.06$
$\sigma_{\text{sky}}$	$32.5 \pm 1.8$	$31.6 \pm 2.6$	$31.9 \pm 4.3$	$34.1 \pm 16.0$	$31.4 \pm 1.7$
$\chi^2$	1568	1512	1681	1675	1602
$z_{\text{eff}}$	2.136	2.276	2.551	2.914	2.334
$b_{\text{Ly}\alpha}$	$-0.0832 \pm 0.0065$	$-0.1099 \pm 0.0063$	$-0.1521 \pm 0.0103$	$-0.2247 \pm 0.023$	$-0.1187 \pm 0.0046$
$b_{\text{eff}, \text{Ly}\alpha}$	$-0.1506 \pm 0.0046$	$-0.1814 \pm 0.0045$	$-0.2336 \pm 0.0074$	$-0.3305 \pm 0.0168$	$-0.1922 \pm 0.0033$

TABLE 1.2 – Corrélations des paramètres du modèle avec les paramètres  $b_{\eta, \text{Ly}\alpha}$  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 la moyenne des fonctions de corrélation calculées dans chaque bin en redshift, soit l’ensemble des données DR16.

Paramètre	$b_{\eta, \text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$
$\alpha_{\parallel}$	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 %

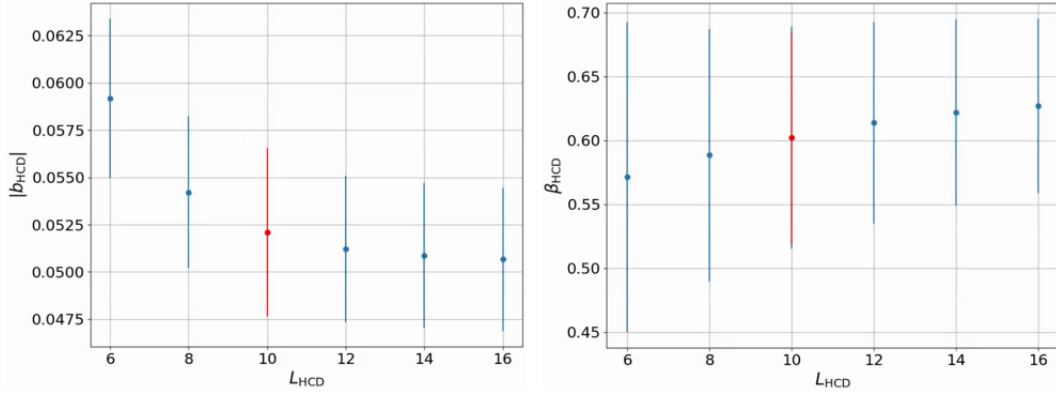


FIGURE 1.3 – Evolution des mesures des paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  en fonction de la valeur  $L_{\text{HCD}}$  choisie pour l’ajustement. Les mesures sont faites sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16 et moyennée sur les quatre bins en redshift.

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

Après 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 d’abord regardé la corrélation des paramètres  $b_{\eta, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  avec les autres paramètres du modèle. La table 1.2 présente ces corrélations. Premièrement, nous pouvons remarquer que les deux paramètres  $\text{Ly}\alpha$  ajustés sont très anticorrélés. 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 DU MAS DES BOURBOUX et al. (2020) 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  $100\%$  (Chabanier et al. (in prep)). 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  $\text{Ly}\alpha$ . Ensuite, le paramètre effectif  $L_{\text{HCD}}$  est fixé à  $10 h^{-1} \text{ Mpc}$  car il est corrélé avec les autres paramètres. Sa valeur, qui dépend de la distribution des HCD non masqués, est difficile à déterminer. Les paramètres des HCD obtenus dépendent de la valeur de  $L_{\text{HCD}}$  choisie, comme illustré sur la figure 1.3. 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  $10 h^{-1} \text{ Mpc}$  et avec une largeur  $\sigma = 1 h^{-1} \text{ Mpc}$ , nous mesurons une corrélation entre  $L_{\text{HCD}}$  et  $\beta_{\text{Ly}\alpha}$  de  $-38\%$ . 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 2.3.

## 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  $\text{Ly}\alpha$  présente, et de comparer cette quantité à ce qui est mesuré par l’ajustement. De plus, nous connaissons le nombre et les distributions

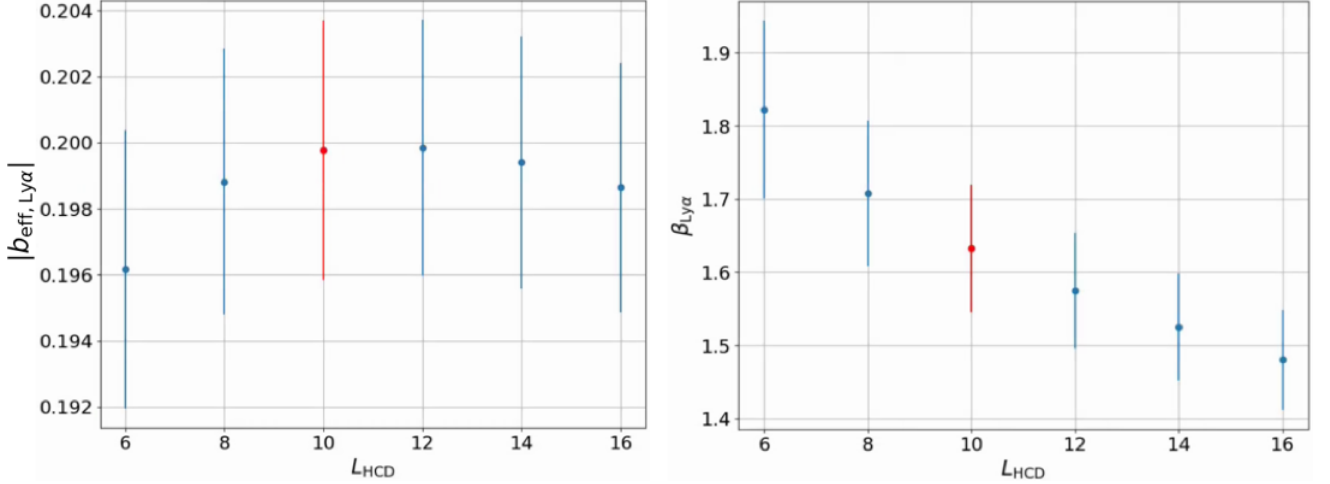


FIGURE 1.4 – Evolution des mesures des paramètres  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  en fonction de la valeur  $L_{\text{HCD}}$  choisie pour l’ajustement. Les mesures sont faites sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16 et moyennée sur les quatre bins en redshift.

en  $z$  et en  $\log n_{\text{HI}}$  des HCD ajoutés dans les mocks, ce qui n’est pas le cas des données. 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). 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  $20 < r < 180 h^{-1}$  Mpc et mesurons les paramètres  $\text{Ly}\alpha$ . 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 DR16. Ceci montre que la procédure d’ajustement des paramètres des mocks que nous avons mise en place est efficace pour obtenir les bons  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$ .

Lorsque nous comparons maintenant les valeurs de  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  mesurées dans les raw mocks à celles mesurées dans les mocks eboss-0.0, nous observons un écart statistiquement significatif. L’effet 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 (de l’ordre de 5 %) et statistiquement significatif. Il semble que la matrice de distorsion ne prenne pas totalement en compte l’ajout du continuum et du bruit par **quickquasars** et l’ajustement du continuum dans l’analyse.

Enfin, nous observons 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  $b_{\text{eff,Ly}\alpha}$  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. Comme expliqué dans la section 1.2, les paramètres  $\text{Ly}\alpha$  sont très corrélés avec ceux des HCD, en particulier avec  $L_{\text{HCD}}$  qui est fixé à  $10 h^{-1}$  Mpc. Nous regardons dans un premier temps la corrélation des paramètres  $\text{Ly}\alpha$  avec  $L_{\text{HCD}}$  dans les mocks. La figure 1.5 présente la mesure de  $b_{\text{eff,Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  dans les mocks-0.2 pour différentes valeurs de  $L_{\text{HCD}}$ . Le corrélation avec  $L_{\text{HCD}}$  est similaire à celle observée dans les données (figure 1.4). Cette figure suggère qu’il faut utiliser une valeur de  $L_{\text{HCD}}$  entre 2 et  $4 h^{-1}$  Mpc afin d’obtenir une mesure des paramètres  $\text{Ly}\alpha$  en accord avec les mocks eboss-0.0. Ainsi nous produisons

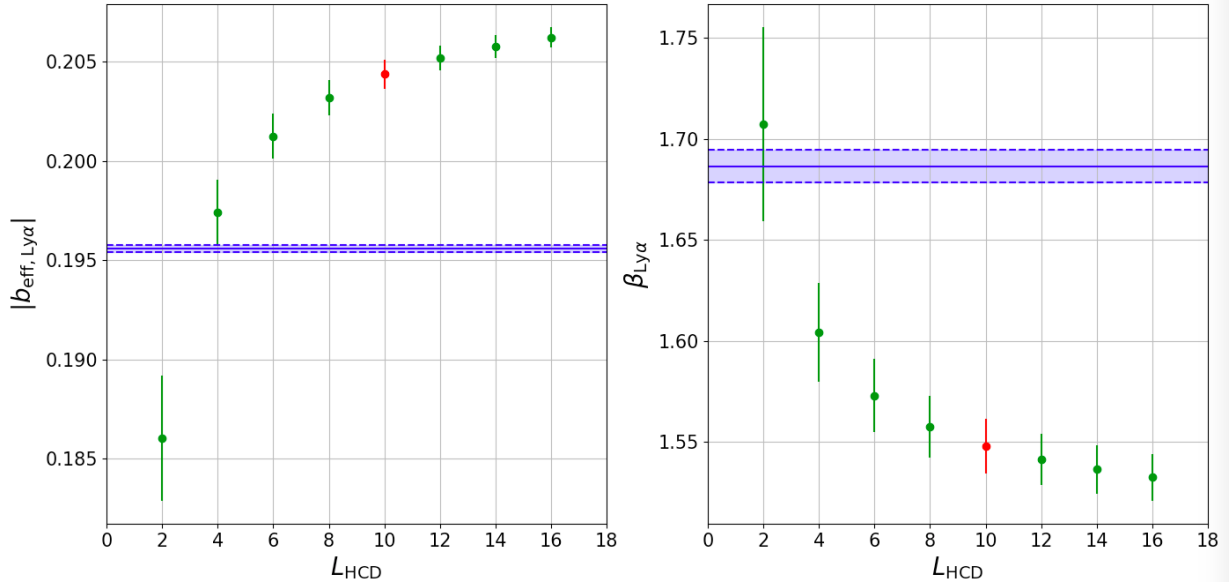


FIGURE 1.5 – Evolution des mesures des paramètres  $b_{\text{eff}, \text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  en fonction de la valeur  $L_{\text{HCD}}$  choisie pour l’ajustement. Les mesures sont faites sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir de 30 réalisations des mocks eboss-0.2 et moyennée sur les 4 bins en redshift. La ligne horizontale bleue donne la mesure faite sur les 30 réalisations des mocks eboss-0.0. Les lignes tiretées donnent les erreurs à  $1\sigma$ .

TABLE 1.3 – Résultats des ajustements de l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée sur 30 réalisations des mocks et moyennée sur les quatre bins en redshift. La première ligne donne l’ajustement des mocks eboss-0.0. La deuxième donne l’ajustement des mocks eboss-0.2 comme décrit dans la section ?? . La troisième ligne donne le même ajustement que la deuxième mais en utilisant  $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ .

version	$b_{\text{eff}, \text{Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$	$b_{\text{HCD}}$	$\beta_{\text{HCD}}$	$L_{\text{HCD}} [h^{-1} \text{ Mpc}]$	$\chi^2 (n_{\text{dof}})$
eboss-0.0	$-0.1956 \pm 0.0002$	$1.687 \pm 0.008$				1562 (1570)
eboss-0.2	$-0.2044 \pm 0.0007$	$1.548 \pm 0.014$	$-0.0080 \pm 0.0010$	$0.487 \pm 0.089$	10	1573 (1568)
eboss-0.2	$-0.1925 \pm 0.0023$	$1.647 \pm 0.034$	$-0.0177 \pm 0.0025$	$0.455 \pm 0.090$	2.8	1578 (1568)

un ajustement des mocks eboss-0.2 avec  $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$  (le choix de cette valeur particulière est expliquée dans la section 2.4). Le résultat de cet ajustement est présenté dans le tableau 1.3. Les paramètres  $\text{Ly}\alpha$  obtenus sont maintenant compatibles à  $1,2\sigma$  et  $1,3\sigma$  avec les mocks eboss-0.0. Ceci suggère que la valeur de  $L_{\text{HCD}}$  utilisée dans l’ajustement des données DR16 est mal choisi. Nous investigons le choix de  $L_{\text{HCD}}$  dans la section 2.4.

## 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. Les HCD identifiés puis reconstruits avec  $\log n_{\text{HI}} > 20,3$  sont masqués. Dans le cas des mocks, les HCD avec  $\log n_{\text{HI}} > 20,3$  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  $\text{Ly}\alpha \times \text{Ly}\alpha$  est estimée dans les quatre bins en redshift utilisés jusqu’ici. Nous nommons cette analyse

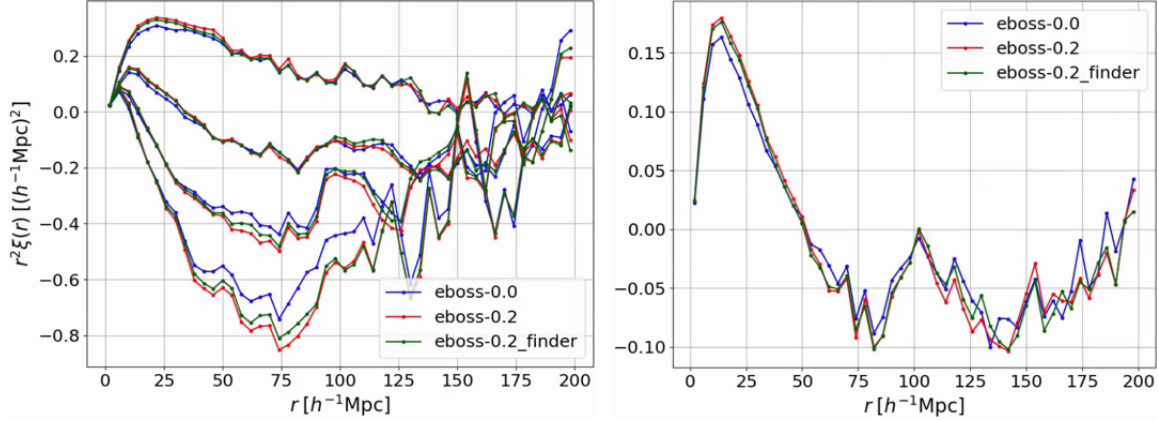


FIGURE 1.6 – Fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimées à partir d’une réalisation des mocks eboss-0.0 (bleu), eboss-0.2 (rouge) et eboss-0.2\_finder (vert). La version eboss-0.2\_finder correspond aux mocks eboss-0.2, dans lesquels les HCD ont été masqués en utilisant le catalogue de HCD produit par l’algorithme d’identification. Le graphique de gauche montre les corrélations dans quatre gammes en  $\mu$ . Ces gammes sont, de haut en bas :  $0 < \mu < 0,5$ ,  $0,5 < \mu < 0,8$ ,  $0,8 < \mu < 0,95$  et  $0,95 < \mu < 1$ . Le graphique de droite montre les corrélations moyennées sur  $0 < \mu < 1$ .

*eboss-0.2\_finder*. La figure 1.6 présente la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  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  $\text{Ly}\alpha \times \text{Ly}\alpha$ , 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 rouge à celle montrée en bleu. 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 une petite partie 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 est confirmé par la figure 1.7, 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. Notons par ailleurs que l’algorithme d’identification possède une efficacité de 90 % pour  $\log n_{\text{HI}} > 20,3$ . L’effet de ces HCD non identifiés avec  $\log n_{\text{HI}} > 20,3$  est donc plus faible que l’effet des HCD dont le  $\log n_{\text{HI}}$  est surestimé. Ceci s’explique par le faible nombre de HCD avec  $\log n_{\text{HI}} > 20,3$ .

Le tableau 1.4 donne les résultats des ajustements des trois corrélations présentées sur la figure 1.6. La statistique d’une seule réalisation n’est pas suffisante pour identifier des potentielles systématiques. Cependant, la précision de la mesure des paramètres  $\text{Ly}\alpha$  dans les données DR16 étant comparable à 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 étude sur un plus grand nombre de réalisations.

### 2.3 Une meilleure modélisation 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



TABLE 1.4 – 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, eboss-0.2 et eboss-0.2\_finder. 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	$-0.1951 \pm 0.0039$	$1.592 \pm 0.074$	$-0.0186 \pm 0.0054$	$0.494 \pm 0.091$

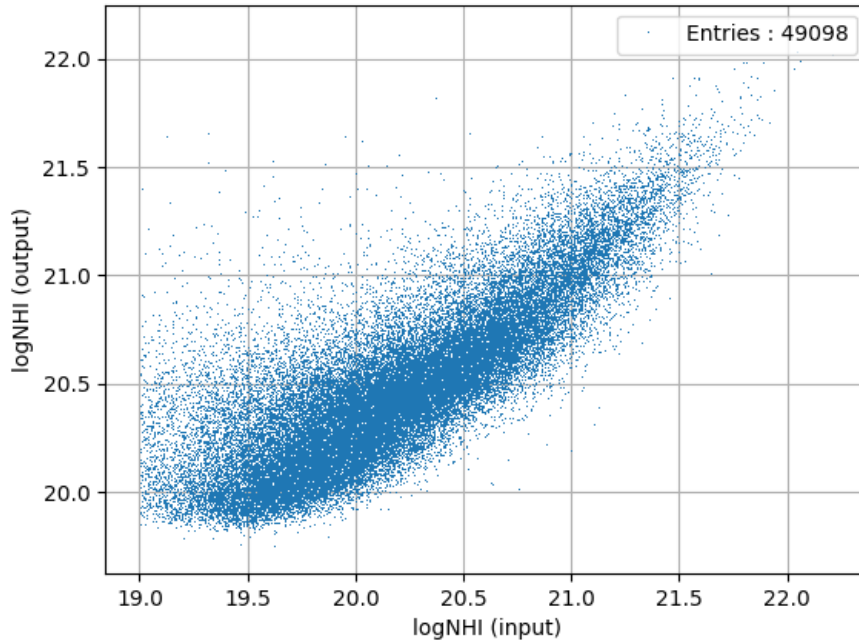


FIGURE 1.7 – Densité de colonne  $\log n_{\text{HI}}$  trouvée par le finder ( $\log\text{NHI}$  output) en fonction de la densité de colonne des HCD des mocks ( $\log\text{NHI}$  input). Cette comparaison est produite par Jim Rich et Solène Chabannier, du groupe cosmologie du CEA, à l’aide des mocks (Chabannier et al. (in prep)).

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 un paramètre effectif  $L_{\text{HCD}}$  qui reflète la taille caractéristique des HCD non masqués (équation ?), la fonction  $F_{\text{HCD}}$  est calculée en prenant en compte 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 induites par l'utilisation de l'algorithme de détection. Cependant, afin de calculer le terme  $F_{\text{HCD}}$  correspondant à la distribution de HCD présents, nous devons justement connaître cette distribution. Si cela est possible pour les mocks, cela ne l'est pas 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 celle du modèle *pygmy*, utilisée dans les mocks.

Nous comparons ici les résultats de l'ajustement des fonctions de corrélation obtenues à partir 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 terme  $F_{\text{HCD}}$  que nous utilisons ici est le terme  $F_{\text{HCD}}$  calculé à partir de la distribution des HCD des mocks pour lesquels  $\log n_{\text{HI}} < 20,3$ . La figure 1.8 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 proches. 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 sont 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 pensons que les mesures faites avec le modèle de Rogers ne sont pas correctes, à cause de  $L_{\text{HCD}}$  qui est mal choisi. Si nous supposons que la distribution des HCD dans les données est semblable à celle des mocks, le tableau 1.3 et la figure 1.5 nous suggèrent d'utiliser un  $L_{\text{HCD}}$  de l'ordre de  $3 h^{-1} \text{ Mpc}$ . Ceci donnerait un  $b_{\text{eff,Ly}\alpha}$  plus faible et un  $\beta_{\text{Ly}\alpha}$  plus grand, et donc peut-être un accord avec les mesures faites avec le modèle C-G. Ceci est dans notre liste des choses à faire.

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. Pour ce faire, nous estimons la corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  sur 10 réalisations eboss-0.2 où les HCD ne sont pas masqués lors du calcul du champ  $\delta_F$ . La figure 1.9 présente la comparaison des deux modèles 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  $b_{\text{eff,Ly}\alpha}$  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 Rogers (jaune) et le modèle C-G (rouge) est un peu plus important que dans la comparaison faite sur la figure 1.8.

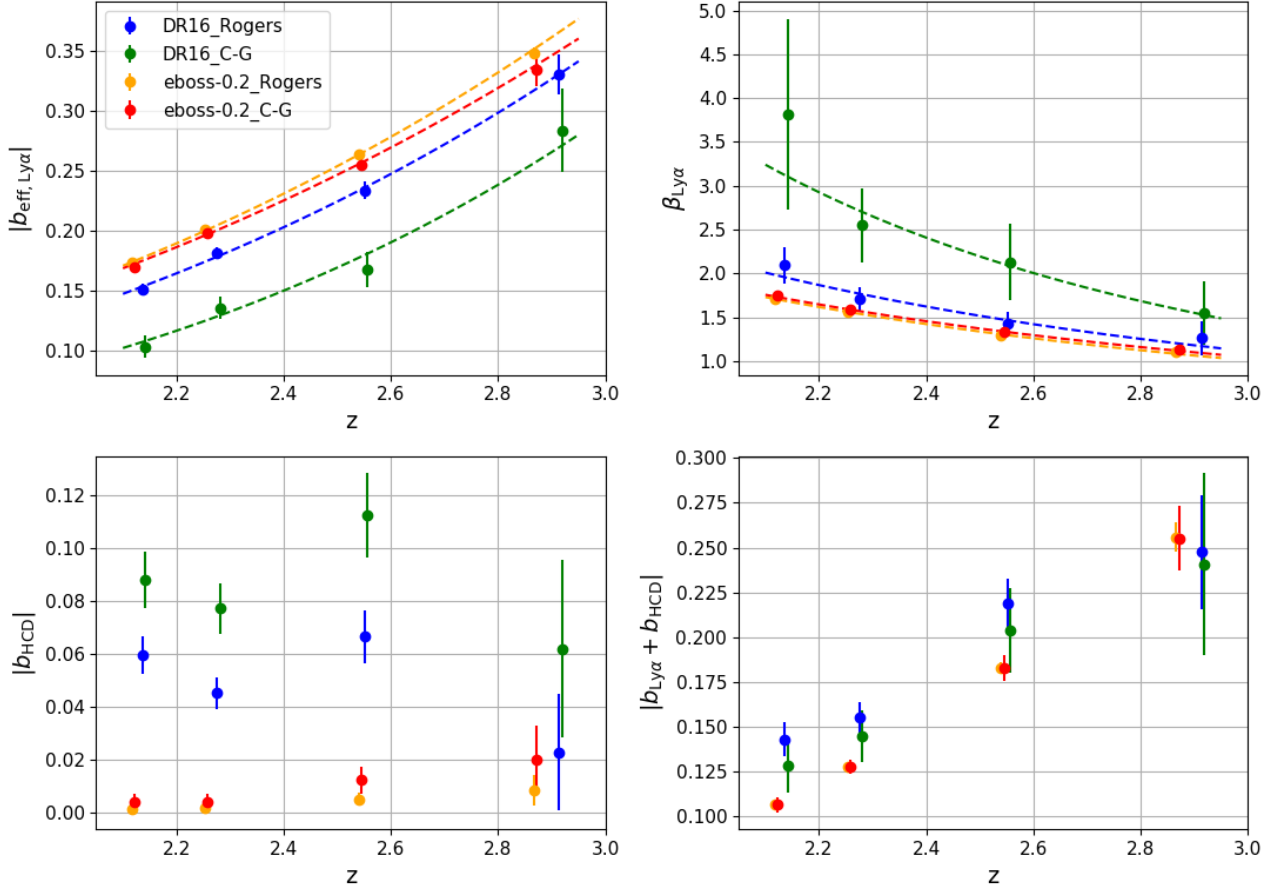


FIGURE 1.8 – Mesure des paramètres  $b_{\text{eff}, \text{Ly}\alpha}$  (haut gauche),  $\beta_{\text{Ly}\alpha}$  (haut droite),  $b_{\text{HCD}}$  (bas gauche) et  $b_{\text{Ly}\alpha} + b_{\text{HCD}}$  (bas droite) sur l'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16 et des mocks eboss-0.2. Les HCD avec  $\log n_{\text{HI}} > 20,3$  sont masqués dans le calcul des fonctions de corrélation. 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). Les points verts sont décalés des points bleus de  $\Delta z = 5 \times 10^{-3}$  pour des raisons de visibilité. Idem pour les points rouges et jaunes.

TABLE 1.5 – Corrélations des paramètres du modèle avec le paramètre  $b_{\eta, \text{Ly}\alpha}$  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 %
$10^3 b_{\eta, \text{SiII}(1207)}$	4 %	−28 %
$10^3 b_{\eta, \text{CIV}(\text{eff})}$	−5 %	−7 %
$b_{\text{HCD}}$	35 %	−91 %
$\beta_{\text{HCD}}$	78 %	−49 %
$10^2 A_{sky}$	29 %	−12 %
$\sigma_{sky}$	−15 %	−7 %

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 avec le modèle C-G (rouge) sont presque compatibles avec les mesures faites sur les mocks eboss-0.0 (gris). En ce qui concerne les données, l’écart entre le modèle C-G (vert) et le modèle de Rogers (bleu) est moins important que dans la comparaison faite sur la figure 1.8. Encore une fois, nous pensons que la mesure faite avec le modèle de Rogers n’est pas correcte car le  $L_{\text{HCD}}$  choisi n’est pas correct. Nous pouvons remarquer que les mesures faites avec le modèle C-G sur la corrélation où les HCD avec  $\log n_{\text{HI}} < 20,3$  sont masqués (en vert sur la figure 1.8) ne sont pas compatibles avec les mesures faites avec le modèle C-G sur la corrélation où les HCD ne sont pas masqués (en vert sur la figure 1.9). Ceci peut-être dû au fait que la distribution des HCD dans les données est différente de celle des mocks, et donc que  $F_{\text{HCD}}$  est mal estimé. Cela peut-être aussi dû à un effet inconnu dans les données qui est pris en compte par la modélisation des HCD.

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 étudions 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 produisons 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 ajustons 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 1.6 résume ces mesures. 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. Cela est dû à la valeur de  $L_{\text{HCD}}$  mal choisie pour le modèle de Rogers. 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

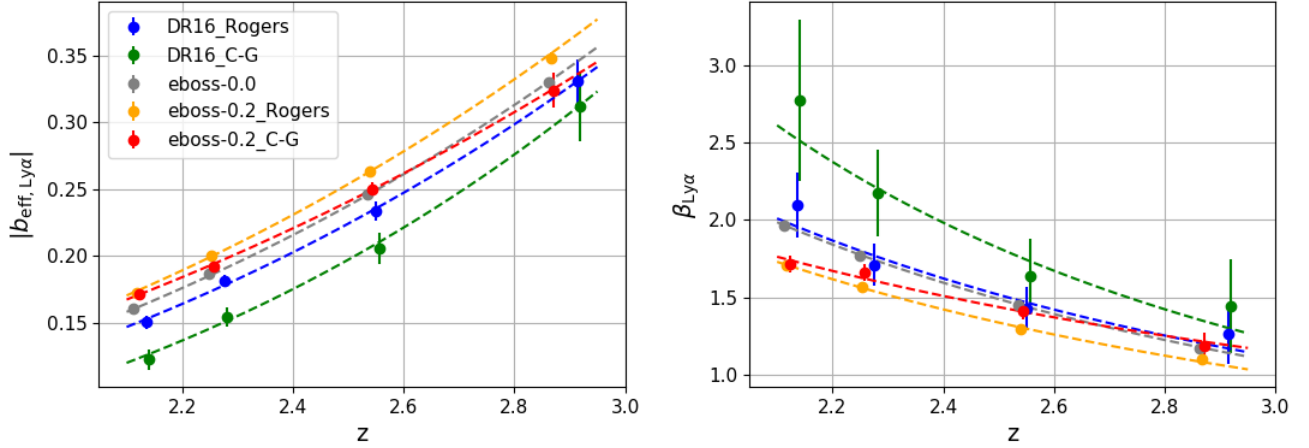


FIGURE 1.9 – Mesure des paramètres  $b_{\text{eff,Ly}\alpha}$  (gauche) et  $\beta_{\text{Ly}\alpha}$  (droite) sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ . Les données représentées sont : les données DR16 avec les HCD masqués et ajustées avec le modèle de Rogers (bleu), les données 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. Les points jaunes et gris sont mesurés sur la corrélation moyennée sur 30 réalisations. Les points rouges sont mesurés sur la corrélation moyennée sur 10 réalisations. Les points verts sont décalés des points bleus de  $\Delta z = 5 \times 10^{-3}$  pour des raisons de visibilité. Les points gris, jaunes et rouges sont décalés de la même manière.

HCD mesuré dans l’autre version. Le biais des HCD mesuré dans la version contenant 3 fois plus de HCD est sous-estimé. Ceci a pour effet de réduire la mesure de  $\beta_{\text{Ly}\alpha}$ . Nous pouvons tout de même 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.

Enfin, nous regardons si en fixant les paramètres  $\text{Ly}\alpha$ , nous obtenons un biais des HCD compatible avec 3 fois le biais des HCD mesuré dans la version standard. Nous faisons ce test pour le modèle de Rogers et pour le modèle C-G. Le résultat de ces ajustements est donné dans le tableau 1.6 (voir les lignes “3×HCD &  $\text{Ly}\alpha$  fixé”). Pour le modèle de Rogers, le fait de fixer les paramètres  $\text{Ly}\alpha$  résout bien les tensions observées : nous attendons  $b_{\text{HCD}} = 0,0501 \pm 0,0090$  et mesurons  $b_{\text{HCD}} = -0,0555 \pm 0,0017$ . Cependant, dans le cas du modèle C-G, le paramètre  $\beta_{\text{HCD}}$  compense les paramètres  $\text{Ly}\alpha$  fixés, et la valeur de  $b_{\text{HCD}}$  obtenue n’est pas compatible avec 3 fois celle mesurée dans la version standard. Il serait intéressant de vérifier si en fixant aussi la valeur de  $\beta_{\text{HCD}}$ , nous retrouvons le bon  $b_{\text{HCD}}$ .

TABLE 1.6 – Mesures des paramètres  $b_{\text{eff,Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  faites 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. Le détail de chaque ajustement est donné dans le texte.

Version	$b_{\text{eff,Ly}\alpha}$	$\beta_{\text{Ly}\alpha}$	$b_{\text{HCD}}$	$\beta_{\text{HCD}}$
Rogers	$-0.1910 \pm 0.0020$	$1.561 \pm 0.039$	$-0.0167 \pm 0.0030$	$0.509 \pm 0.089$
Rogers & 3×HCD	$-0.2095 \pm 0.0025$	$1.295 \pm 0.039$	$-0.0296 \pm 0.0037$	$0.514 \pm 0.087$
Rogers & 3×HCD & $\text{Ly}\alpha$ fixé	$-0.1910$	$1.561$	$-0.0555 \pm 0.0017$	$0.465 \pm 0.061$
C-G	$-0.1704 \pm 0.0058$	$1.752 \pm 0.096$	$-0.0327 \pm 0.0062$	$0.487 \pm 0.089$
C-G & 3×HCD	$-0.1713 \pm 0.0071$	$1.561 \pm 0.110$	$-0.0598 \pm 0.0077$	$0.477 \pm 0.087$
C-G & 3×HCD & $\text{Ly}\alpha$ fixé	$-0.1704$	$1.752$	$-0.0651 \pm 0.0014$	$0.277 \pm 0.040$

## 2.4 Mieux comprendre les HCD

### L'effet des HCD sur la fonction de corrélation

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. La figure 1.10 présente les différentes composantes du modèle ajusté sur l'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16. Ce modèle est présenté dans la section ??, il utilise la modélisation de Rogers avec  $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$ . **La figure 1.11 donnent ces mêmes modèles non multipliés par la matrice de distorsion. En comparant les courbes oranges aux courbes vertes, nous pouvons remarquer sur ces figures que l'effet des HCD est assez semblable à celui d'une augmentation du biais. Nous pouvons cependant distinguer l'effet des HCD de celui d'une augmentation du biais en regardant le long de la ligne de visée : les HCD ajoutent une corrélation positive pour les petites échelles, et les courbes vertes et jaune ne se coupent pas au niveau de l'axe  $y = 0$ . Ceci est davantage visible lorsque les modèles ne sont pas multipliés par la matrice de distorsion (figure 1.11). Si nous regardons maintenant la gamme  $0 < \mu < 0,5$ , l'effet des HCD est très semblable à l'effet d'un facteur multiplicatif appliqué à la courbe verte. La distinction entre une augmentation du biais et l'effet des HCD est d'autant plus facile du fait que les modèles utilisent une valeur de  $L_{\text{HCD}}$  importante ( $10 h^{-1} \text{ Mpc}$ ). Lorsque  $L_{\text{HCD}}$  est plus faible, le terme  $F_{\text{HCD}}$  est plus proche de 1, et donc l'effet des HCD ressemble davantage à celui d'une augmentation de biais (équation ??). Dans un tel cas, les paramètres  $\text{Ly}\alpha$  et ceux des HCD sont très dégénérés.**

### Des HCD avec une même densité de colonne

Nous avons ensuite regardé l'effet des différentes gammes de  $\log n_{\text{HI}}$  des HCD sur la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ . Pour ce faire, à partir de la même réalisation des mocks, nous avons produit cinq versions eboss-0.2. Dans chacune de ces versions, nous ajoutons les HCD avec une valeur de  $\log n_{\text{HI}}$  fixe. Ces valeurs pour les cinq versions sont  $\log n_{\text{HI}} \in [17,6; 18,2; 18,8; 19,4; 20]$ . Les nombres relatifs de HCD entre ces versions suivent la distribution en  $\log n_{\text{HI}}$  utilisée pour construire le catalogue standard de HCD (présentée sur la figure ??). La distribution en  $z$  utilisée est trois fois plus importante que dans le catalogue standard afin d'avoir suffisamment de HCD.

La figure 1.12 présente les fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  pour les différentes versions des mocks décrites précédemment. Pour simplifier les comparaisons, nous n'utilisons pas de bins en redshift, les fonctions de corrélation sont donc estimées à partir de l'ensemble des forêts. Nous pouvons voir sur la figure 1.12 que les HCD qui ont le plus grand effet sur la fonction de corrélation sont les HCD avec une grande densité de colonne, malgré leur nombre plus restreint. L'effet causé par les HCD avec une densité de colonne de 17,6, 18,2 et 18,8 est similaire : la faible densité de colonne est compensée par le nombre plus important de HCD.

Nous étudions maintenant comment le modèle de Rogers et le modèle C-G se comportent lorsque nous ajustons les mocks contenant des HCD avec une densité de colonne fixe. Dans chacun des ajustements, nous devons utiliser pour chaque modèle un terme  $F_{\text{HCD}}$  adéquat aux HCD présents. En ce qui concerne le modèle de C-G, le code `picca` permet de calculer le terme  $F_{\text{HCD}}$  pour une distribution en  $\log n_{\text{HI}}$  données. Nous calculons donc  $F_{\text{HCD}}$  pour chacune des valeurs de  $\log n_{\text{HI}}$ . Pour le modèle de Rogers, la forme de  $F_{\text{HCD}}$  est fixée (voir équation ??). Il nous suffit, pour chaque valeur de  $\log n_{\text{HI}}$ , de trouver la valeur de  $L_{\text{HCD}}$  à utiliser. Mais ceci n'est pas chose aisée. Dans l'analyse DR16,  $L_{\text{HCD}}$  est choisie de manière approximative à  $10 h^{-1} \text{ Mpc}$ , car il influence peu la mesure des paramètres BAO. Mais comme nous l'avons vu à plusieurs reprises,  $L_{\text{HCD}}$  influence grandement la mesure des paramètres  $\text{Ly}\alpha$ . Nous devons donc le choisir avec précaution.

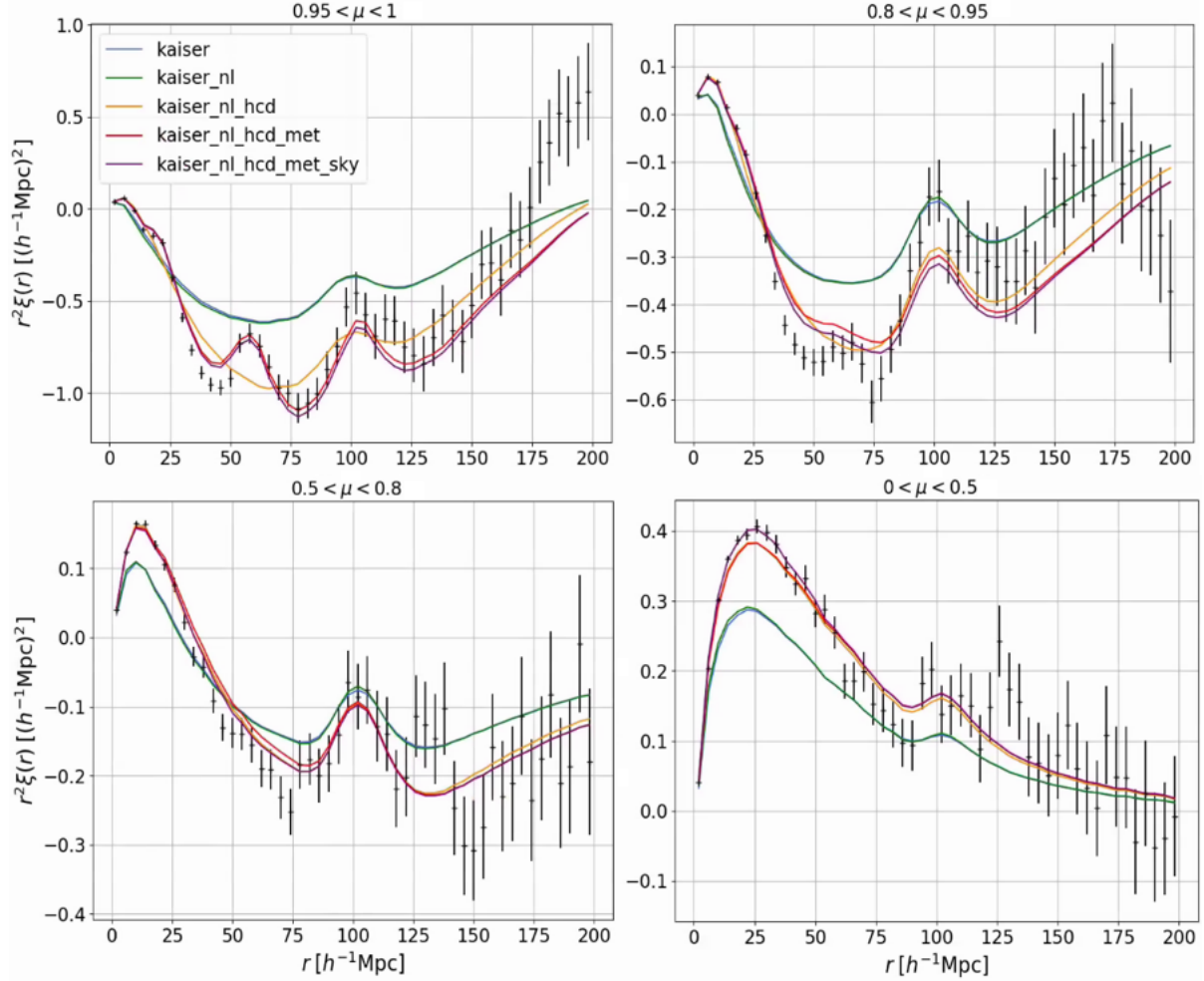


FIGURE 1.10 – Les différentes composantes du modèle ajusté sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  des données DR16. La ligne bleue donne le modèle de Kaiser, c’est à dire la fonction de corrélation obtenue à partir du spectre de puissance  $P_{\text{QL}}$  (équation ??) et multiplié par le facteur de Kaiser  $b_{\text{eff,Ly}\alpha}^2(1 + \beta_{\text{Ly}\alpha}\mu_k^2)$ . La ligne verte donne le modèle de Kaiser multiplié par le terme  $F_{\text{NL}}^{\text{auto}}$  qui prend en compte les non-linéarités aux petites échelles. La ligne orange donne le modèle précédent auquel la contribution des HCD a été ajoutée (équation ??). La ligne rouge donne le modèle précédent plus la contribution des métaux et la ligne violette donne le modèle complet. L’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimée à partir des données DR16 et moyennée dans les quatre bins en redshift est représentée par les points noirs. Les quatre graphiques présentent 4 gammes en  $\mu$  différentes.

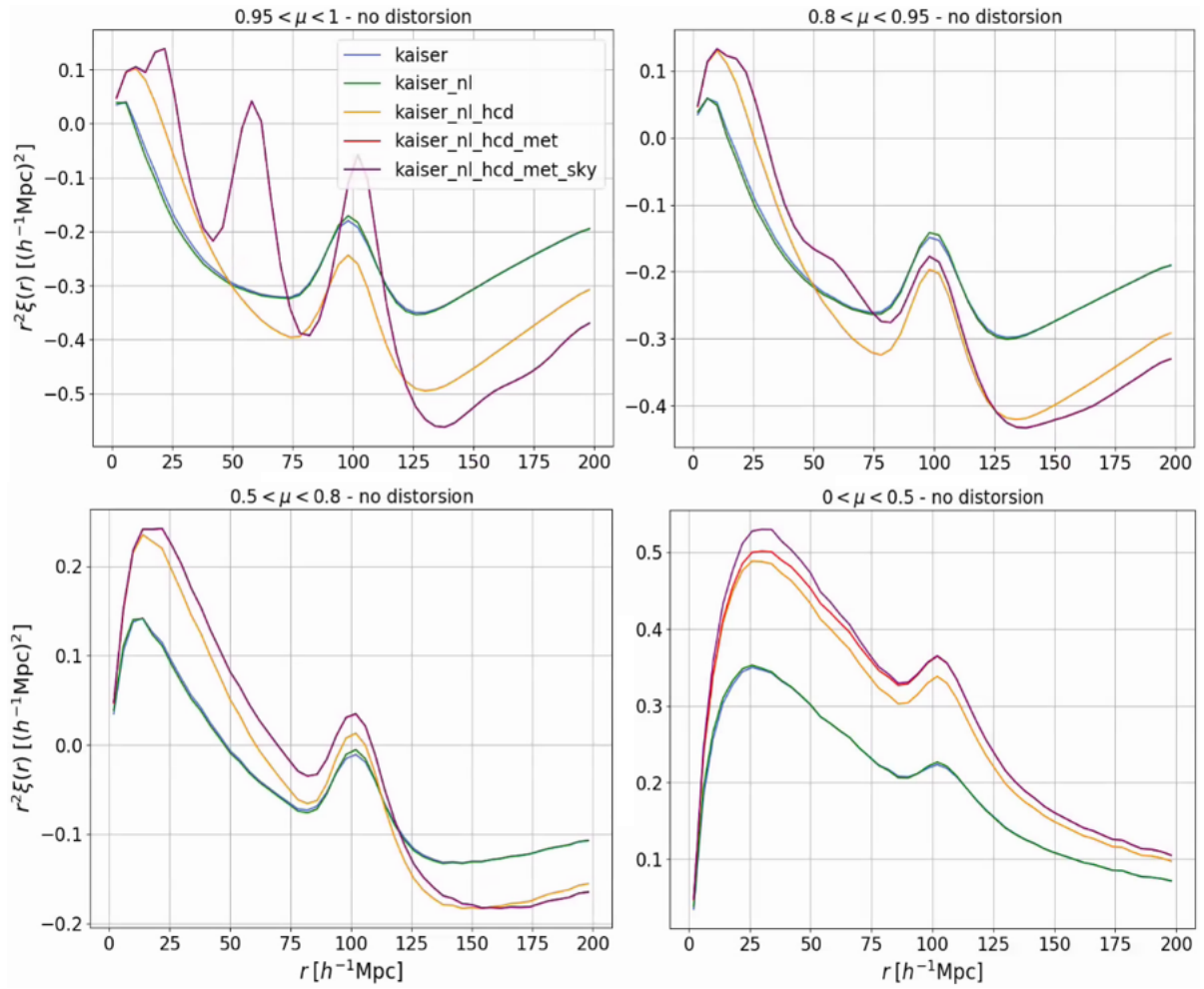


FIGURE 1.11 – Les différentes composantes du modèle ajusté sur l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  des données DR16. Elles sont décrites dans la légende de la figure ???. Les composantes présentées ici n’incluent pas les distorsions produites par l’ajustement du continuum. Les quatre graphiques présentent ces composantes dans 4 gammes en  $\mu$ .



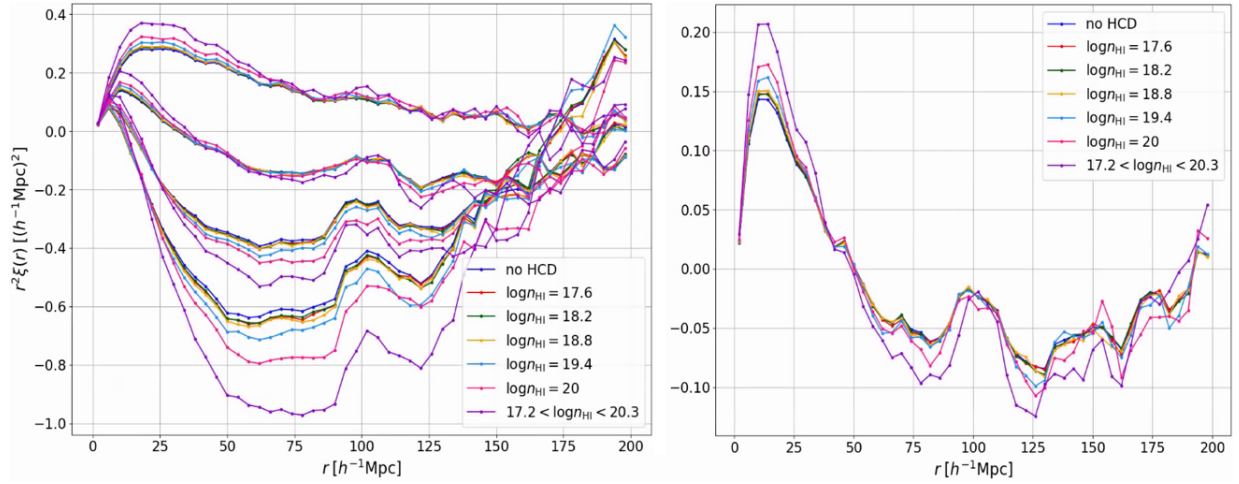


FIGURE 1.12 – Fonctions de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  estimées à partir d’une réalisation. La ligne bleue foncée donne la corrélation pour les mocks eboss-0.0 (sans HCD). La ligne violette donne la corrélation pour les mocks eboss-0.2 incluant les HCD avec  $17,2 < \log n_{\text{HI}} < 20,3$ . Les autres couleurs donnent les corrélations pour les mocks eboss-0.2 incluant des HCD avec une valeur fixe en  $\log n_{\text{HI}}$ . Le graphique de gauche montre les corrélations dans quatre gammes en  $\mu$ . Ces gammes sont, de haut en bas :  $0 < \mu < 0,5$ ,  $0,5 < \mu < 0,8$ ,  $0,8 < \mu < 0,95$  et  $0,95 < \mu < 1$ . Le graphique de droite montre les corrélations moyennées sur  $0 < \mu < 1$ .

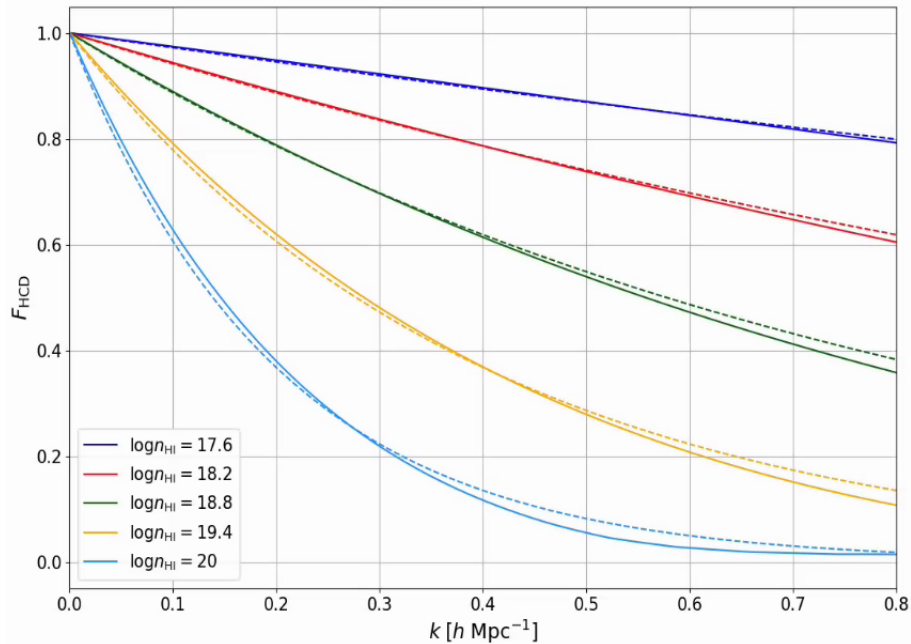


FIGURE 1.13 – Les fonctions  $F_{\text{HCD}}$  utilisées dans le cadre du modèle C-G (lignes continues) ou du modèle de Rogers (lignes tiretées) pour différentes valeurs de  $\log n_{\text{HI}}$ . Les valeurs de  $L_{\text{HCD}}$  utilisées pour le modèle de Rogers sont données dans le texte.

Nous expliquons ici la méthode utilisée pour déterminer  $L_{\text{HCD}}$ . Pour chaque valeur de  $\log n_{\text{HI}}$ , nous choisissons une valeur de  $L_{\text{HCD}}$  qui reproduise au mieux le comportement de  $F_{\text{HCD}}(k)$  dans la gamme  $k \lesssim 0,8 h \text{ Mpc}^{-1}$ . Nous ne nous intéressons pas aux  $k$  plus grands que  $0,8 h \text{ Mpc}^{-1}$  car la fonction de corrélation est estimée dans des bins de  $4 h^{-1} \text{ Mpc}$ , les échelles plus petites que  $k = \pi/4 \sim 0,8 h \text{ Mpc}^{-1}$  ne sont donc pas accessibles. Ainsi, les valeurs de  $L_{\text{HCD}}$  que nous obtenons pour  $\log n_{\text{HI}} \in [17,6; 18,2; 18,8; 19,4; 20]$  sont  $[0,28; 0,60; 1,2; 2,5; 5]$ . La figure 1.13 présente les fonctions  $F_{\text{HCD}}(k)$  pour le modèle C-G et pour le modèle de Rogers en utilisant les valeurs de  $L_{\text{HCD}}$  obtenues précédemment. Pour chaque valeur de  $\log n_{\text{HI}}$ , les fonctions  $F_{\text{HCD}}$  sont similaires dans les deux modèles.

Nous ajustons maintenant chacune des versions des mocks avec les deux modèles. Pour le modèle de Rogers, nous utilisons les valeurs de  $L_{\text{HCD}}$  déterminées précédemment. Pour chacune des versions des mocks, les ajustements faits avec ces deux modèles produisent des résultats très similaires. Les différences sont inférieures au dixième de  $\sigma$ . Ceci est rassurant car les fonctions  $F_{\text{HCD}}$  utilisées dans ces deux modèles sont similaires. Par ailleurs, pour les faibles valeurs de  $\log n_{\text{HI}}$ , les fonctions  $F_{\text{HCD}}$  sont très proches de 1 (voir figure 1.13). Les paramètres  $\text{Ly}\alpha$  et HCD sont quasiment complètement dégénérés. Le paramètre  $\beta_{\text{Ly}\alpha}$ , en particulier, n'est pas contraint. Lorsque  $\log n_{\text{HI}}$  est plus grand (supérieur à 18,8), les ajustements parviennent à mesurer les paramètres du  $\text{Ly}\alpha$ . Cependant ces mesures ne sont pas en accord avec les mesures faites sur les mocks eboss-0.0. Si ces modèles ne parviennent pas à distinguer l'effet du  $\text{Ly}\alpha$  de l'effet des HCD lorsque ces derniers possèdent une densité de colonne fixe, alors il y a peu d'espoirs qu'ils parviennent à le faire lorsque les HCD sont distribués dans une large gamme en  $\log n_{\text{HI}}$ .

Nous comparons maintenant les fonctions  $F_{\text{HCD}}$  correspondante à une large gamme en  $\log n_{\text{HI}}$ . Nous considérons premièrement la gamme  $\log n_{\text{HI}}$  des HCD présents dans les mocks eboss-0.2, c'est à dire  $17,2 < \log n_{\text{HI}} < 20,3$ . Comme précédemment, afin de déterminer le  $L_{\text{HCD}}$  à utiliser dans le modèle de Rogers, nous cherchons à reproduire le comportement de  $F_{\text{HCD}}(k)$  dans la gamme  $k \lesssim 0,8 h \text{ Mpc}^{-1}$ . Pour ce faire, nous devons choisir  $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ . La fonction  $F_{\text{HCD}}$  utilisée dans le modèle de C-G pour cette gamme en  $\log n_{\text{HI}}$  est calculée par `picca`. La figure 1.14 présente les fonctions  $F_{\text{HCD}}$  ainsi obtenues. Nous pouvons remarquer que pour une large gamme en  $\log n_{\text{HI}}$ , il est plus difficile d'obtenir un accord entre le modèle de Rogers et le modèle C-G. La figure 1.14 présente aussi, à titre de comparaison, la fonction  $F_{\text{HCD}}$  obtenue avec  $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$ . La différence entre la courbe continue bleue et la courbe tiretée verte est très importante. Nous sommes étonnés<sup>1</sup> que les ajustements faits sur les mocks eboss-0.2 avec le modèle de Rogers et le modèle C-G produisent des résultats si similaires (voir la figure 1.8) compte tenu de ces différences sur  $F_{\text{HCD}}$ . Dans tous les cas, nous pensons que la mesure faite avec  $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$  n'est pas fiable. La figure 1.14 présente aussi les fonctions  $F_{\text{HCD}}$  obtenues pour la gamme  $17,2 < \log n_{\text{HI}} < 22,5$ . La valeur de  $L_{\text{HCD}}$  choisie est  $6,5 h^{-1} \text{ Mpc}$ . Les fonctions  $F_{\text{HCD}}$  obtenues pour cette gamme dans le cadre du modèle C-G et du modèle de Rogers (courbes rouges) sont différentes. Le modèle de Rogers atteint ici ses limites. Ainsi, si nous choisissons d'ajuster les mocks ou les données sans masquer au préalable les HCD, nous devons utiliser le modèle C-G.

La figure 1.15 présente les profils dans l'espace réel correspondant aux fonctions  $F_{\text{HCD}}$ . Ces profils sont obtenus à l'aide d'une transformée de Fourier des fonctions  $F_{\text{HCD}}(k)$ . Comme observé sur la figure 1.14, le profil obtenu avec le modèle de Rogers a une forme très différente du profil obtenu avec le modèle C-G lorsque nous considérons toute la gamme en  $\log n_{\text{HI}}$  (courbes rouges). Les profils obtenus pour  $17,2 < \log n_{\text{HI}} < 20,3$  sont assez semblables (courbes bleues). Il serait intéressant de reproduire l'analyse des mocks-0.2 et des données en utilisant ce profil, c'est à dire le modèle de Rogers avec  $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ .

1. Nous pensons que cela vient du biais des HCD qui est environ 2,2 fois supérieur dans la mesure faite avec le modèle C-G. Nous pensons que le biais des HCD plus bas obtenu avec le modèle de Rogers compense les mesures des paramètres  $\text{Ly}\alpha$ .

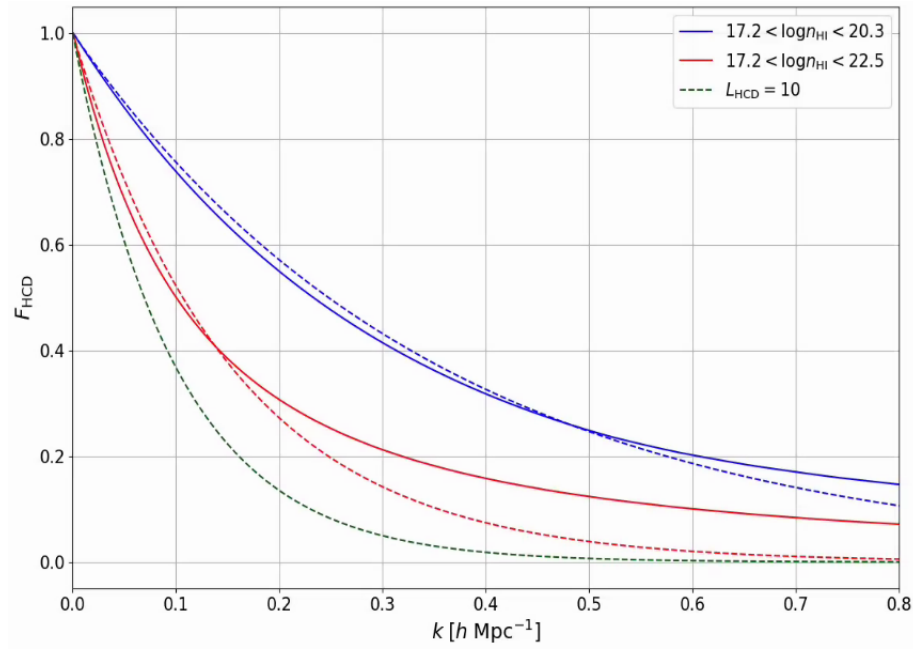


FIGURE 1.14 – Les fonctions  $F_{\text{HCD}}$  utilisées dans le cadre du modèle C-G (lignes continues) ou du modèle de Rogers (lignes tiretées). Les courbes bleues correspondent à la gamme  $17,2 < \log n_{\text{HI}} < 20,3$ . Les courbes rouges correspondent à la gamme  $17,2 < \log n_{\text{HI}} < 22,5$ . La courbe verte donne  $F_{\text{HCD}}$  utilisé dans le modèle de Rogers avec  $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$ .

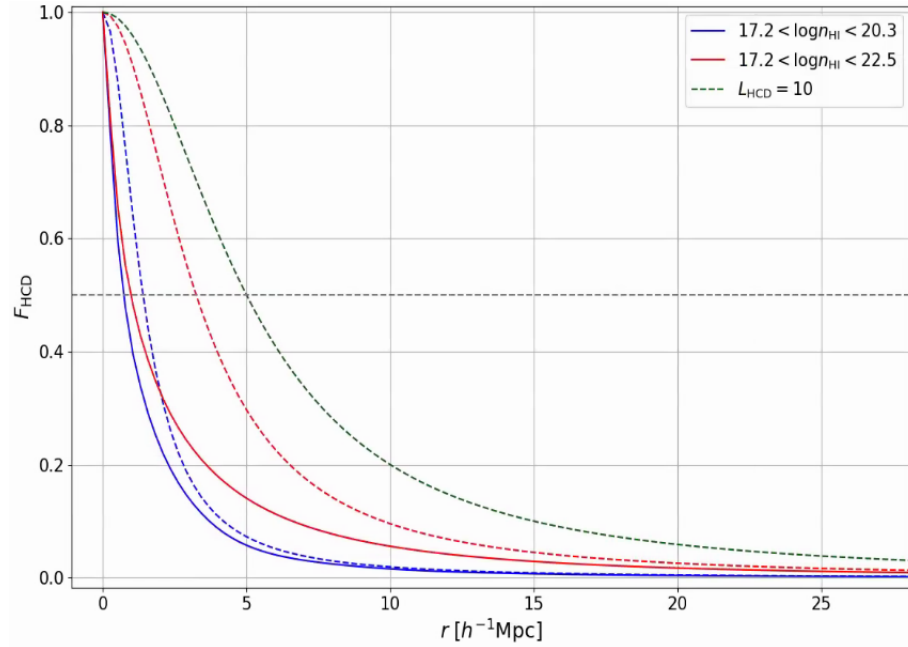


FIGURE 1.15 – Les fonctions  $F_{\text{HCD}}$  dans l'espace réel. Celles-ci sont obtenues avec une transformation de Fourier des fonctions  $F_{\text{HCD}}$  dans l'espace  $k$ , utilisées pour modéliser les HCD. La légende est la même que celle de la figure 1.14.

Les analyses présentées dans ce chapitre sont présentées dans un ordre chronologique : nous nous sommes rendus compte très tardivement que le modèle utilisé pour modéliser les HCD, à savoir le modèle de Rogers avec  $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$ , n'était pas correct. Nous n'avons malheureusement pas eu le temps de reproduire les analyses des mocks et des données en utilisant  $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ . Il serait intéressant de produire ces analyses, et de les comparer aux analyses faites avec le modèle C-G dans le cas où les HCD avec  $\log n_{\text{HI}} > 20,3$  sont masqués, et dans le cas où aucun HCD n'est masqué. Nous imaginons qu'en calculant correctement les fonctions  $F_{\text{HCD}}$  dans ces deux cas, et en estimant correctement  $L_{\text{HCD}}$ , nous pouvons arriver à des mesures des paramètres  $\text{Ly}\alpha$  compatibles entre ces trois analyses.

# Bibliographie

---

- DE SAINTE AGATHE, Victoria 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.
- DU MAS DES BOURBOUX, Hélion et al. (2020). « THE COMPLETED SDSS-IV EXTENDED BARYON OSCILLATION SPECTROSCOPIC SURVEY: BARYON ACOUSTIC OSCILLATIONS WITH LYMAN-ALPHA FORESTS ». In :