

6

Analyse des données et résultats

Sommaire

| | |
|----------------------------------------------------------|------------|
| 6.1 L'analyse des données DR16 | 140 |
| 6.1.1 Résultats | 140 |
| 6.1.2 Stabilité des paramètres Ly α | 143 |
| 6.2 Etude de la modélisation des HCD | 144 |
| 6.2.1 Comparaison des mocks | 145 |
| 6.2.2 Effet du masquage des HCD | 147 |
| 6.2.3 Une meilleure modélisation des HCD | 149 |
| 6.2.4 Mieux comprendre les HCD | 151 |
| L'effet des HCD sur la fonction de corrélation | 151 |
| Des HCD avec une même densité de colonne | 152 |

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 Ly α 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 Ly α 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.

6.1 L'analyse des données DR16

6.1.1 Résultats

L'analyse des données finale d'eBOSS (DR16), dont nous avons déjà parlé et qui est présentée dans DU MAS DES BOURBOUX, RICH et al. (2020), utilise les fonctions de corrélation Ly α (Ly α) \times Ly α (Ly α), Ly α (Ly α) \times Ly α (Ly β), Ly α (Ly α) \times QSO et Ly α (Ly β) \times 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, RICH 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 Ly α 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 quatres bins en redshift. De manière à limiter les potentielles systématiques, nous nous limitons à l'analyse de la fonction de corrélation Ly α (Ly α) \times Ly α (Ly α) (abrégée en Ly α \times Ly α 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 quatres 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 Ly α \times Ly α , 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, il est présenté dans la section 5.4.1. 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 6.1 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 μ . Le tableau 6.1 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 Ly α mesurés dans les données par une loi de

entre bins
en z (?)

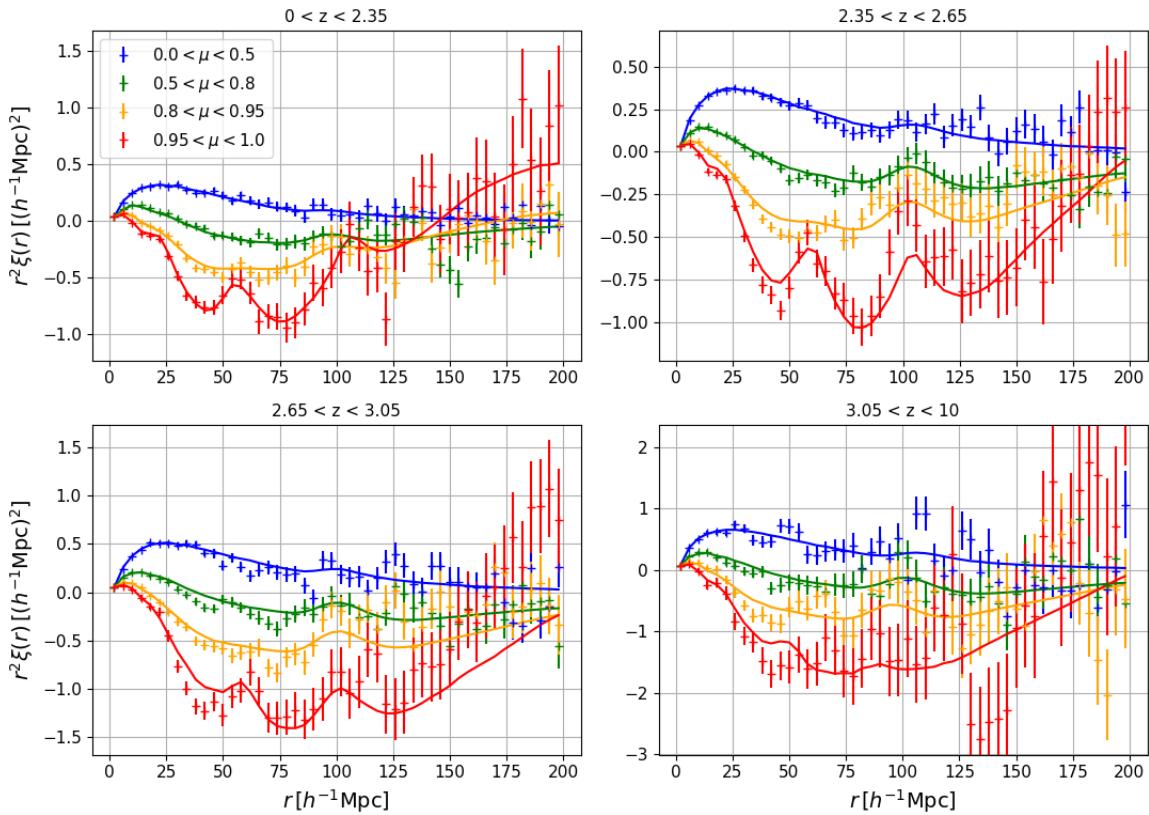


FIGURE 6.1 – Fonctions de corrélation Ly α × Ly α dans chacun des bins en redshift de l'analyse. Les courbes en trait plein donne le meilleur ajustement du modèle obtenu avec `picca`. Chaque graphique correspond à un bin en redshift. Pour chacun des bins, la fonction de corrélation et l'ajustement sont montrés dans quatre bins en μ .

TABLE 6.1 – Résultats de l’ajustement fait avec `picca` des fonctions de corrélation Ly α × Ly α 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 χ^2 et le redshift effectif z_{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 Ly α . Ils sont reliés aux paramètres $b_{\eta, \text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ par les équations 4.66 et 4.67.

| Paramètre | $0 < z < 2,35$ | $2,35 < z < 2,65$ | $2,65 < z < 3,05$ | $3,05 < z < 10$ | $0 < z < 10$ |
|-----------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| α_{\parallel} | 1.063 ± 0.066 | 1.019 ± 0.041 | 1.029 ± 0.072 | 1.120 ± 0.081 | 1.047 ± 0.034 |
| α_{\perp} | 1.063 ± 0.108 | 0.965 ± 0.057 | 1.016 ± 0.058 | 0.926 ± 0.072 | 0.980 ± 0.042 |
| $b_{\eta, \text{Ly}\alpha}$ | -0.1796 ± 0.0058 | -0.1938 ± 0.0053 | -0.2237 ± 0.0084 | -0.2929 ± 0.0187 | -0.1998 ± 0.0039 |
| $\beta_{\text{Ly}\alpha}$ | 2.094 ± 0.210 | 1.711 ± 0.133 | 1.427 ± 0.138 | 1.265 ± 0.194 | 1.633 ± 0.087 |
| $10^3 b_{\eta, \text{SiII}(1190)}$ | -1.83 ± 1.10 | -3.66 ± 0.68 | -2.80 ± 1.01 | 0.36 ± 1.64 | -3.02 ± 0.51 |
| $10^3 b_{\eta, \text{SiII}(1193)}$ | -4.83 ± 1.10 | -1.94 ± 0.69 | -0.79 ± 0.97 | -2.13 ± 1.72 | -2.07 ± 0.50 |
| $10^3 b_{\eta, \text{SiII}(1260)}$ | -3.38 ± 1.33 | -1.97 ± 0.80 | -1.32 ± 1.05 | 0.90 ± 1.79 | -2.16 ± 0.63 |
| $10^3 b_{\eta, \text{SiIII}(1207)}$ | -7.87 ± 1.10 | -4.52 ± 0.75 | -2.11 ± 1.05 | -2.89 ± 1.74 | -4.59 ± 0.52 |
| $10^3 b_{\eta, \text{CIV}(\text{eff})}$ | -4.77 ± 2.54 | -5.15 ± 2.64 | -5.06 ± 2.62 | -5.02 ± 2.61 | -5.12 ± 2.63 |
| b_{HCD} | -0.0596 ± 0.0070 | -0.0452 ± 0.0060 | -0.0665 ± 0.0100 | -0.0228 ± 0.0218 | -0.0521 ± 0.0045 |
| β_{HCD} | 0.551 ± 0.086 | 0.560 ± 0.086 | 0.508 ± 0.088 | 0.502 ± 0.090 | 0.603 ± 0.083 |
| $10^2 A_{\text{sky}}$ | 1.585 ± 0.098 | 0.870 ± 0.082 | 0.729 ± 0.133 | 0.646 ± 0.338 | 0.947 ± 0.060 |
| σ_{sky} | 32.5 ± 1.8 | 31.6 ± 2.6 | 31.9 ± 4.3 | 34.1 ± 16.0 | 31.4 ± 1.7 |
| χ^2 | 1568 | 1512 | 1681 | 1675 | 1602 |
| z_{eff} | 2.136 | 2.276 | 2.551 | 2.914 | 2.334 |
| $b_{\text{Ly}\alpha}$ | -0.0832 ± 0.0065 | -0.1099 ± 0.0063 | -0.1521 ± 0.0103 | -0.2247 ± 0.0230 | -0.1187 ± 0.0046 |
| $b_{\text{eff}, \text{Ly}\alpha}$ | -0.1506 ± 0.0046 | -0.1814 ± 0.0045 | -0.2336 ± 0.0074 | -0.3305 ± 0.0168 | -0.1922 ± 0.0033 |

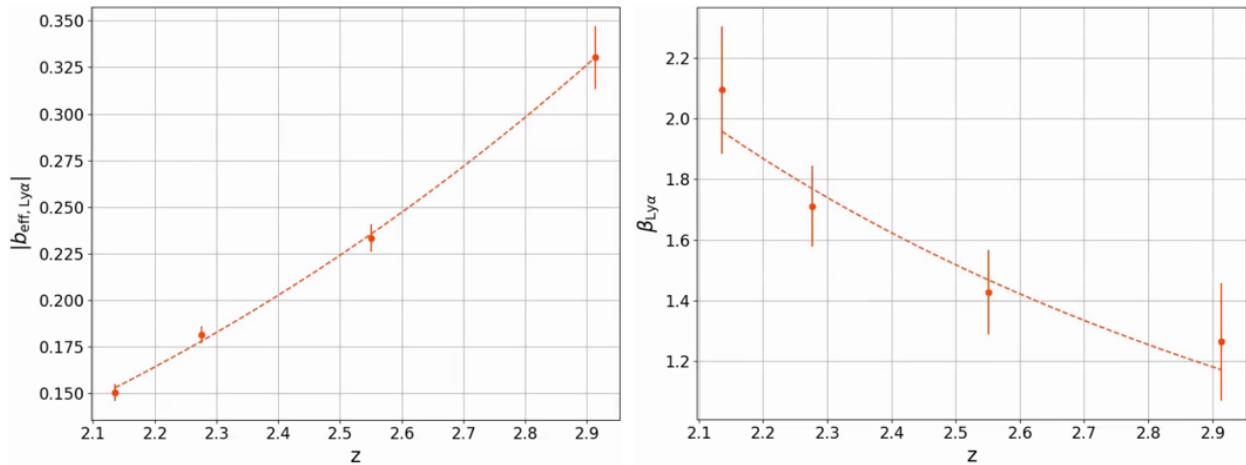


FIGURE 6.2 – Mesure des paramètres $b_{\text{eff}, \text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ dans les données DR16. Les mesures sont faites dans quatre bins en redshift, indiquées par les points. La ligne en pointillés donne le meilleur ajustement par une loi de puissance.

présente

puissance. La figure 6.2 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 quatres 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 4.4). Ils sont extrapolés de $z = 1,9$ jusqu'à $z = 3,6$.

donner les deux paramètres du fit

6.1.2 Stabilité des paramètres Ly α

On masque ceux que l'on détecte et en pratique cela revient à $\log n_{\text{HI}} > 20.3$

Après avoir produit les ajustements présentés dans la section précédente, nous avons cherché à savoir si la mesure des paramètres Ly α 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 6.2 présente ces corrélations. Premièrement, nous pouvons remarquer que les deux paramètres Ly α ajustés sont très anticorrélés. -87% Deuxièmement, les paramètres du Ly α sont très corrélés avec ceux des HCD, notamment $\beta_{\text{Ly}\alpha}$ qui est corrélé à -75 % avec b_{HCD} . Ceci pose plusieurs problèmes : d'abord, la modélisation des HCD choisie dans DU MAS DES BOURBOUX, Rich 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_{HCD} (voir section 5.4.1). 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_{HCD} , ce qui a pour effet d'augmenter le biais du Ly α . Ensuite, le paramètre effectif L_{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_{HCD} choisie, comme illustré sur la figure 6.3. A cause des corrélations entre les paramètres liés aux HCD et ceux liés au Ly α , le fait de changer L_{HCD} change aussi les paramètres Ly α obtenus. La figure 6.4 montre la dépendance de $b_{\text{eff}, \text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ avec L_{HCD} . Ainsi, le paramètre RSD $\beta_{\text{Ly}\alpha}$ est très corrélé avec L_{HCD} . Lorsque nous laissons libre L_{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_{HCD} et $\beta_{\text{Ly}\alpha}$ de -38 %. Lorsque L_{HCD}

Une valeur de $10 h^{-1} \text{ Mpc}$ a été choisie dans l'analyse DR16, de façon un peu arbitraire¹⁴³
dans la mesure ou cela n'affecte pas la mesure de l'échelle BAO.

TABLE 6.2 – 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 | $b_{\eta,\text{Ly}\alpha}$ | $\beta_{\text{Ly}\alpha}$ |
|-----------------------------------|----------------------------|---------------------------|
| α_{\parallel} | 0 % | 0 % |
| α_{\perp} | 1 % | -2 % |
| $b_{\eta,\text{Ly}\alpha}$ | 100 % | -87 % |
| $\beta_{\text{Ly}\alpha}$ | -87 % | 100 % |
| $b_{\eta,\text{SiII}(1190)}$ | 2 % | -8 % |
| $b_{\eta,\text{SiII}(1193)}$ | 3 % | -6 % |
| $b_{\eta,\text{SiII}(1260)}$ | -1 % | -3 % |
| $b_{\eta,\text{SiIII}(1207)}$ | 6 % | 5 % |
| $b_{\eta,\text{CIV}(\text{eff})}$ | -7 % | -10 % |
| b_{HCD} | 48 % | -75 % |
| β_{HCD} | 35 % | -23 % |
| A_{sky} | 34 % | -19 % |
| σ_{sky} | -10 % | -2 % |

indiquer combien vaut alors L_HCD (plus L_HCD petit plus corrélation importante)

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_{HCD} et $b_{\eta,\text{SiIII}(1207)}$.

Enfin, le modèle des HCD choisi influence la mesure des paramètres Ly α . Toujours dans le but d'avoir une mesure robuste des paramètres Ly α , 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 6.2.3.

6.2 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 Ly α dans les mocks. En effet, les mocks sont l'outil parfait pour ce genre d'analyse : ils les paramètres Ly α mesurés

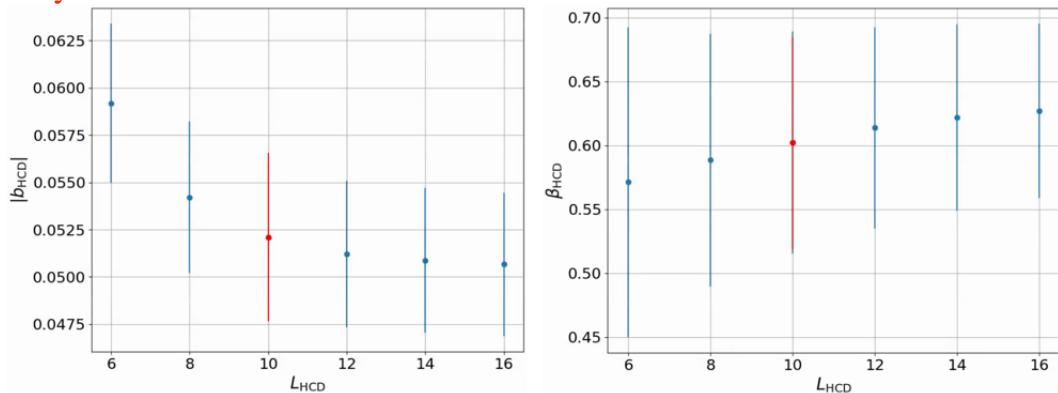


FIGURE 6.3 – Evolution des mesures des paramètres b_{HCD} et β_{HCD} en fonction de la valeur L_{HCD} choisie pour l'ajustement. Les mesures sont faites sur l'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des données DR16 et moyennée sur les quatre bins en redshift.

6.2. ETUDE DE LA MODÉLISATION DES HCD

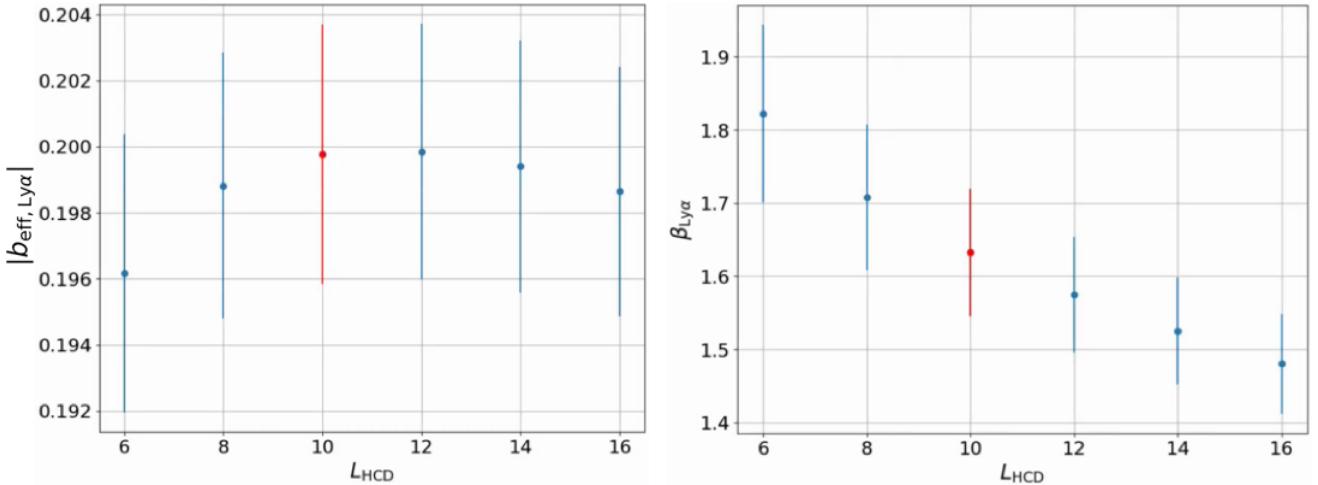


FIGURE 6.4 – Evolution des mesures des paramètres $b_{\text{eff}, \text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ en fonction de la valeur L_{HCD} choisie pour l’ajustement. Les mesures sont faites sur l’auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des données DR16 et moyennée sur les quatre bins en redshift.

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 en redshift (z) et en densité de colonne ($\log n_{\text{HI}}$) 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.

6.2.1 Comparaison des mocks

Comme expliqué dans le chapitre 5, 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 $\text{Ly}\alpha$. La figure 5.16 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 $b_{\text{eff}, \text{Ly}\alpha}$ 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 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 $b_{\text{eff}, \text{Ly}\alpha}$ est important (de l’ordre de 10 %) et statistiquement significatif. Les écarts observés sont dus attribués à la matrice de distorsion, qui ne capture pas l’intégralité des distorsions produites par l’ajustement du continuum. La figure 6.5 présente la comparaison des mocks eboss-0.0, en rouge, avec les mocks eboss-raw multipliés par la matrice de distorsion, en vert. Si la matrice de distorsion était parfaite, les courbes rouge et verte seraient confondues, ce qui n’est pas le cas. Une étude approfondie de la matrice de distorsion et de l’ajout du

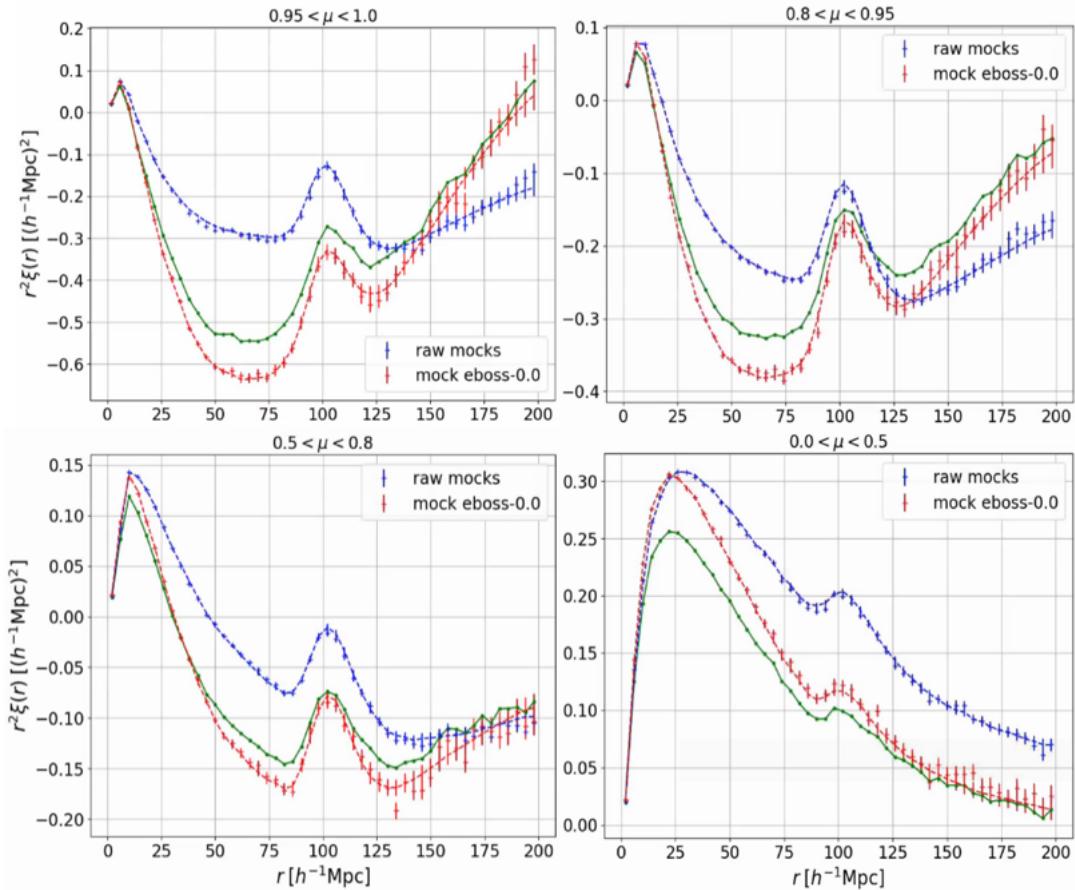


FIGURE 6.5 – Fonction de corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des mocks eboss-raw, en bleu, et des mocks eboss-0.0, en rouge. Les fonctions de corrélation sont estimées à partir de 30 réalisations des mocks. La courbe verte correspond à la corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des mocks eboss-raw et multipliée par la matrice de distorsion.

continuum par `quickquasars` est nécessaire pour comprendre ces différences.

Enfin, nous observons un écart entre les valeurs de $b_{\text{eff},\text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ mesurées dans les mocks eboss-0.0 et eboss-0.2. L'écart mesuré pour $b_{\text{eff},\text{Ly}\alpha}$ 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 6.1.2, les paramètres $\text{Ly}\alpha$ sont très corrélés avec ceux des HCD, en particulier avec L_{HCD} qui est fixé à $10 h^{-1} \text{ Mpc}$. Nous regardons dans un premier temps la corrélation des paramètres $\text{Ly}\alpha$ avec L_{HCD} dans les mocks. La figure 6.6 présente la mesure de $b_{\text{eff},\text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ dans les mocks-0.2 pour différentes valeurs de L_{HCD} . Le corrélation avec L_{HCD} est similaire à celle observée dans les données (figure 6.4). Cette figure suggère qu'il faut utiliser une valeur de L_{HCD} entre 2 et $4 h^{-1} \text{ Mpc}$ afin d'obtenir une mesure des paramètres $\text{Ly}\alpha$ en accord avec les mocks eboss-0.0. Ainsi nous produisons 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 6.2.4). Le résultat de cet ajustement est présenté dans le tableau 6.3. 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_{HCD} utilisée dans l'ajustement des données DR16 est mal choisie. Nous

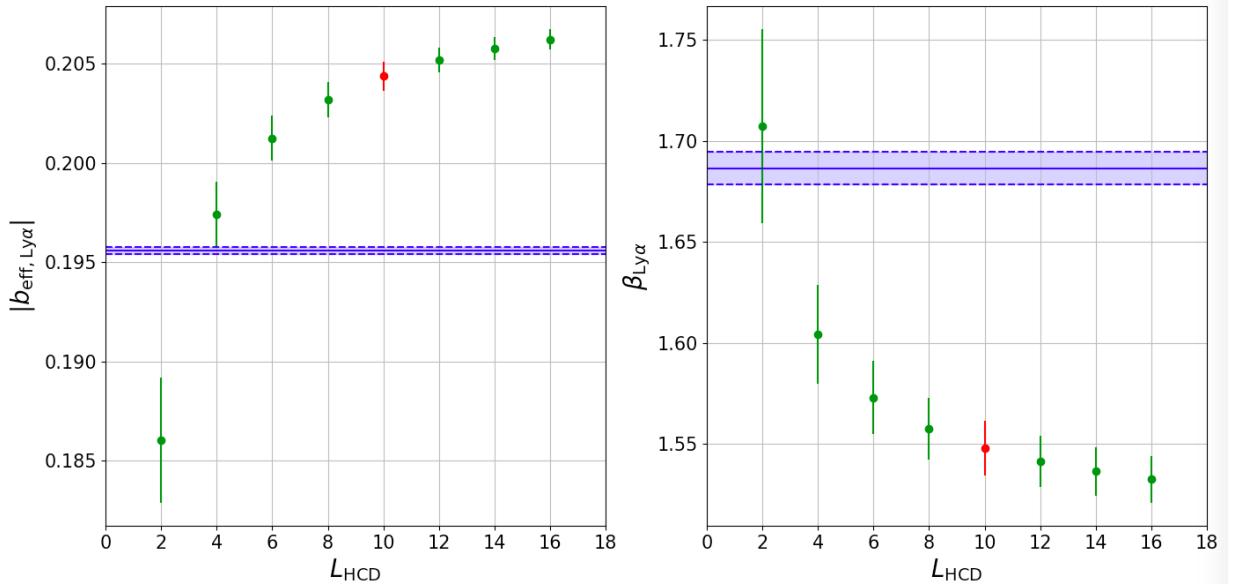


FIGURE 6.6 – Evolution des mesures des paramètres $b_{\text{eff}, \text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$ en fonction de la valeur L_{HCD} choisie pour l’ajustement. Les mesures sont faites sur l’auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir de 30 réalisations des 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 bleues donnent les erreurs à 1σ .

TABLE 6.3 – 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 5.4.2. 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_{HCD} | β_{HCD} | $L_{\text{HCD}} [h^{-1} \text{ Mpc}]$ | $\chi^2 (n_{\text{dof}})$ |
|-----------|-----------------------------------|---------------------------|----------------------|----------------------|---------------------------------------|---------------------------|
| eboss-0.0 | -0.1956 ± 0.0002 | 1.687 ± 0.008 | | | | 1562 (1570) |
| eboss-0.2 | -0.2044 ± 0.0007 | 1.548 ± 0.014 | -0.0080 ± 0.0010 | 0.487 ± 0.089 | 10 | 1573 (1568) |
| eboss-0.2 | -0.1925 ± 0.0023 | 1.647 ± 0.034 | -0.0177 ± 0.0025 | 0.455 ± 0.090 | 2.8 | 1578 (1568) |

investigons le choix de L_{HCD} dans la section 6.2.4.

6.2.2 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 δ_F . Comme expliqué dans la section 6.1.2, 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 δ_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 *eboss-0.2_finder*. La figure 6.7 présente la fonction de corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée

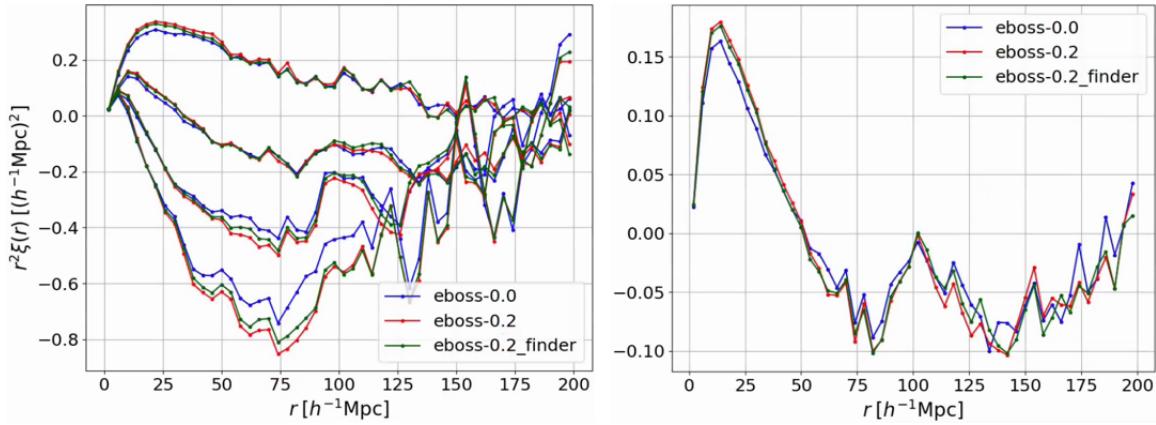


FIGURE 6.7 – Fonctions de corrélation Ly α × Ly α estimées à partir d’une réalisation des mocks eboss-0.0 (bleu), eboss-0.2 (rouge) et eboss-0.2_finder (vert). La version eboss-0.2_finder correspond aux mocks eboss-0.2, dans lesquels les HCD ont été masqués en utilisant le catalogue de HCD produit par l’algorithme d’identification. Le graphique de gauche montre les corrélations dans quatre gammes en μ . 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 $0 < \mu < 1$.

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 α × Ly α , et ainsi d’avoir des mesures de biais comparables. L’effet des HCD (avec une densité de colonne $17,2 < \log n_{\text{HI}} < 20,3$) est visible en comparant la corrélation montrée en rouge à celle montrée en bleu. Comme expliqué dans la section 5.4.1, 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 correctement 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 6.8, qui compare la densité de colonne $\log n_{\text{HI}}$ trouvée par le finder (output) à la “vrai” densité de colonne des HCD des mocks (input). Nous voyons sur cette figure que l’algorithme a tendance à surestimer la densité de colonne $\log n_{\text{HI}}$, 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 6.4 donne les résultats des ajustements des trois corrélations présentées sur la figure 6.7. La statistique d’une seule réalisation n’est pas suffisante pour identifier des potentielles systématiques. Cependant, la précision de la mesure des paramètres Ly α 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

TABLE 6.4 – Résultat de l’ajustement de la corrélation Ly α × Ly α estimée 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_{HCD} | β_{HCD} |
|------------------|-----------------------------------|---------------------------|----------------------|----------------------|
| eboss-0.0 | -0.1970 ± 0.0009 | 1.641 ± 0.039 | | |
| eboss-0.2 | -0.1979 ± 0.0038 | 1.578 ± 0.069 | -0.0201 ± 0.0053 | 0.499 ± 0.090 |
| eboss-0.2_finder | -0.1951 ± 0.0039 | 1.592 ± 0.074 | -0.0186 ± 0.0054 | 0.494 ± 0.091 |

réalisations.

6.2.3 Une meilleure modélisation des HCD

Toujours dans l’optique de tester la robustesse de la mesure des paramètres Ly α , nous avons utilisé une modélisation des HCD différente de celle décrite dans la section 5.4.1 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 Ly α de la collaboration DESI. Nous faisons référence à ce modèle via le nom *modèle C-G*. Le modèle des HCD décrit dans la section 5.4.1 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 Ly α × Ly α et Ly α × QSO produits par l’ensemble des HCD dans les données. Les deux modèles utilisent la modélisation définie dans l’équation 5.26. Cependant, dans le cas du modèle C-G, plutôt que de définir F_{HCD} comme une fonction exponentielle avec un paramètre effectif L_{HCD} qui reflète la taille caractéristique des HCD non masqués (équation 5.28), la fonction F_{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_{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_{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 non identifiés, c’est à dire 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 pyigm, utilisée dans les mocks.

Du fait que, pour le modèle C-G, le terme F_{HCD} soit calculé à partir de la distribution en $\log n_{\text{HI}}$ des HCD présents, nous pouvons envisager d’ajuster avec ce modèle, des mocks avec différentes distributions de HCD : les mocks eboss-0.2 où les HCD avec $\log n_{\text{HI}} < 20,3$ sont masqués, les mocks eboss-0.2 où tous les HCD sont présents ($17,2 < \log n_{\text{HI}} < 22,5$), ou n’importe quelle autre distribution.

Afin de tester le modèle C-G, et de le comparer au modèle de Rogers, nous vérifions premièrement que ces deux modèles, lorsqu’ils sont ajustés sur les mocks eboss-0.0, mesurent une quantité de HCD

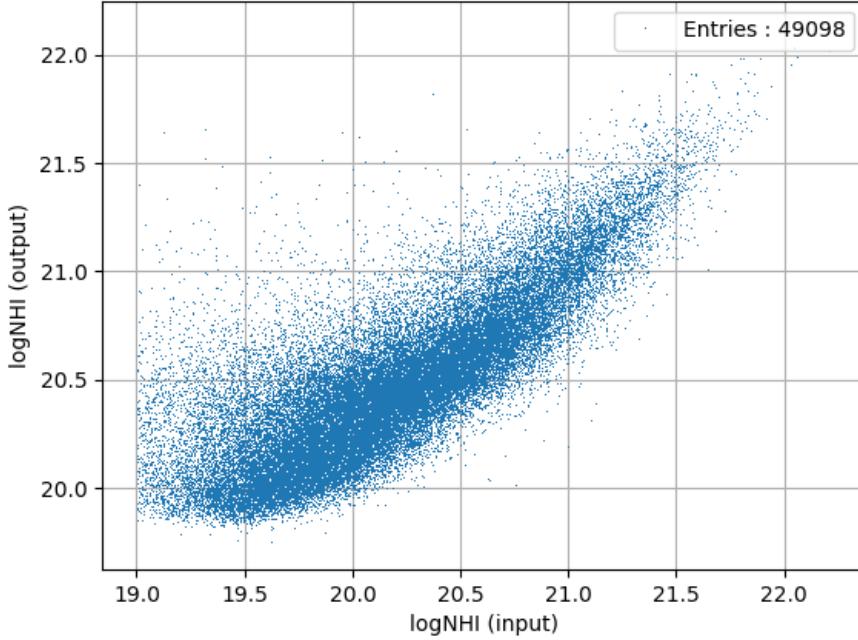


FIGURE 6.8 – Densité de colonne $\log n_{\text{HI}}$ trouvée par le finder ($\log n_{\text{HI}}$ output) en fonction de la “vrai” densité de colonne des HCD des mocks ($\log n_{\text{HI}}$ input). Cette comparaison est produite par Jim Rich et Solène Chabannier, du groupe cosmologie du CEA, à l'aide des mocks (Chabannier et al. (in prep)).

compatible avec 0.

Par ailleurs, dans une étude préliminaire au choix des paramètres Ly α à utiliser pour construire les mocks, nous étudions comment ces deux modèles se comportent lorsque nous triplons la quantité de HCD dans les mocks. 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 et ajoutons le même bruit pour produire les spectres synthétiques puis calculer la corrélation Ly $\alpha \times$ Ly α à partir de ces deux versions. Puis, nous ajustons la fonction de corrélation Ly $\alpha \times$ Ly α , estimées à partir de ces deux versions, avec le modèle de Rogers et le modèle C-G. Le paramètre b_{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 6.5 résume ces mesures. Pour une même version, les mesures de b_{HCD} faites avec le modèle de Rogers ou le modèle C-G ne sont pas compatibles. Comme nous le verrons dans la section 6.2.4, cela est dû à la valeur de L_{HCD} mal choisie pour le modèle de Rogers. De plus, pour un même modèle, nous remarquons que b_{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 a pour effet de réduire la mesure de $\beta_{\text{Ly}\alpha}$. Nous pouvons toutefois noter que le modèle C-G produit des mesures de $b_{\text{eff},\text{Ly}\alpha}$ compatibles dans ces deux versions, ce qui n'est pas le cas du modèle de Rogers.

TABLE 6.5 – Mesures des paramètres $b_{\text{eff},\text{Ly}\alpha}$, $\beta_{\text{Ly}\alpha}$, b_{HCD} et β_{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 | $b_{\text{eff},\text{Ly}\alpha}$ | $\beta_{\text{Ly}\alpha}$ | b_{HCD} | β_{HCD} |
|---------------------------------------------------------|----------------------------------|---------------------------|----------------------|----------------------|
| Rogers | -0.1910 ± 0.0020 | 1.561 ± 0.039 | -0.0167 ± 0.0030 | 0.509 ± 0.089 |
| Rogers & $3 \times \text{HCD}$ | -0.2095 ± 0.0025 | 1.295 ± 0.039 | -0.0296 ± 0.0037 | 0.514 ± 0.087 |
| Rogers & $3 \times \text{HCD}$ & $\text{Ly}\alpha$ fixé | -0.1910 | 1.561 | -0.0555 ± 0.0017 | 0.465 ± 0.061 |
| C-G | -0.1704 ± 0.0058 | 1.752 ± 0.096 | -0.0327 ± 0.0062 | 0.487 ± 0.089 |
| C-G & $3 \times \text{HCD}$ | -0.1713 ± 0.0071 | 1.561 ± 0.110 | -0.0598 ± 0.0077 | 0.477 ± 0.087 |
| C-G & $3 \times \text{HCD}$ & $\text{Ly}\alpha$ fixé | -0.1704 | 1.752 | -0.0651 ± 0.0014 | 0.277 ± 0.040 |

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 6.5 (voir les lignes “ $3 \times \text{HCD}$ & $\text{Ly}\alpha$ 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 β_{HCD} compense les paramètres $\text{Ly}\alpha$ fixés, et la valeur de b_{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 β_{HCD} , nous retrouvons le bon b_{HCD} .

Afin d’étudier davantage le modèle C-G, nous comparons maintenant la mesure des paramètres $\text{Ly}\alpha$ faite avec ce modèle sur les mocks eboss-0.2 où les HCD avec une densité de colonne $\log n_{\text{HI}} > 20,3$ sont masqués, à la mesure faite sur les mocks eboss-0.2 où aucun HCD n’est masqué. Dans chacun des cas, nous utilisons le terme F_{HCD} correspondant à la distribution de HCD considérée. La comparaison est présentée sur la figure 6.9. Nous remarquons que la mesure des paramètres $\text{Ly}\alpha$ faite avec le modèle C-G sur les mocks eboss-0.2 où les HCD sont masqués (orange) est compatible avec la mesure faite sur les mocks eboss-0.2 où les HCD ne sont pas masqués (rouge). Ceci est encourageant car cela signifie que le modèle C-G est cohérent avec lui-même. Lorsque nous comparons les mesures faites avec le modèle C-G aux mesures faites avec le modèle de Rogers, nous pouvons identifier une différence statistiquement significative, en particulier pour $b_{\text{eff},\text{Ly}\alpha}$. Nous étudions la cause de ces différences dans la section suivante. Enfin, les mesures de $b_{\text{eff},\text{Ly}\alpha}$ faites avec le modèle C-G sur les mocks eboss-0.2 sont en meilleur accord avec la mesure faite sur les mocks eboss-0.0 que ne l’est la mesure faite avec le modèle de Rogers sur les mocks eboss-0.2.

6.2.4 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 du 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 6.10 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 5.4.1, il utilise la modélisation de Rogers avec $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$. La figure 6.11 donnent ces

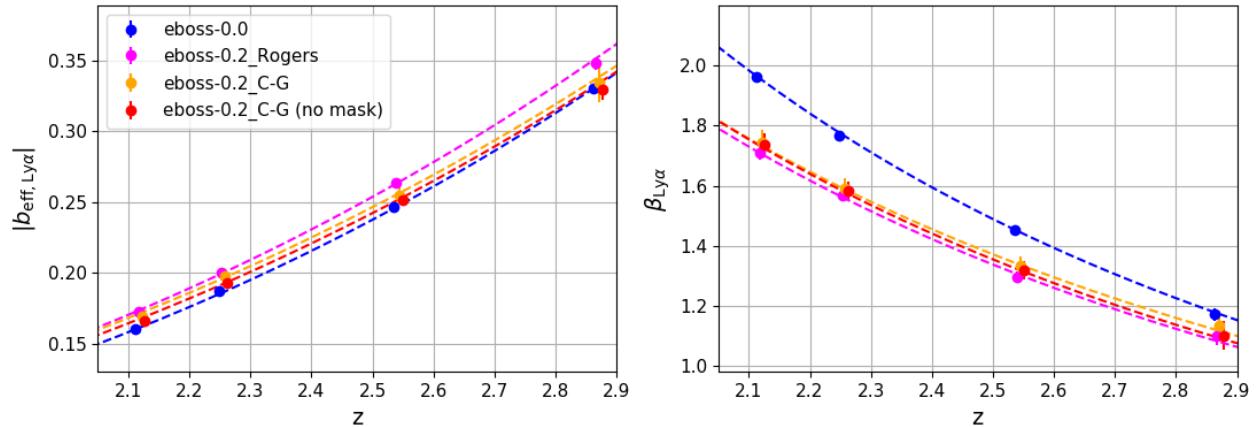


FIGURE 6.9 – Mesure des paramètres $b_{\text{eff}, \text{Ly}\alpha}$ (gauche) et $\beta_{\text{Ly}\alpha}$ (droite) sur l’auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des mocks (30 réalisations utilisées pour chaque mesure). Les points bleus présentent la mesure faite sur les mocks eboss-0.0. Les points magentas la mesure faite sur les mocks eboss-0.2, où les HCD avec $\log n_{\text{HI}} > 20,3$ sont masqués dans le calcul des fonctions de corrélation, et ajustés avec le modèle de Rogers. Les points orange la mesure faite sur ces mêmes fonctions de corrélation, mais ajustées avec le modèle C-G. Enfin, les points rouges présentent la mesure faite sur les mocks eboss-0.2 où aucun HCD n’a été masqué, ajustés avec le modèle C-G. Les tirets correspondent à l’ajustement de chaque jeu de données par une loi de puissance. Les points bleus sont décalés horizontalement de $\Delta z = -1 \times 10^{-2}$, les magentas de $\Delta z = -5 \times 10^{-3}$ et les rouges de $\Delta z = 5 \times 10^{-3}$ pour des raisons de visibilité.

mêmes composantes non multipliées par la matrice de distorsion. En comparant les courbes orange 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 6.11). 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 que les modèles utilisent une valeur de L_{HCD} importante ($10 h^{-1} \text{ Mpc}$). Lorsque L_{HCD} est plus faible, le terme F_{HCD} est plus proche de 1, et donc l’effet des HCD ressemble davantage à celui d’une augmentation de biais (voir équation 5.26). Dans un tel cas, les paramètres Ly α 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 densité de colonne des HCD sur la fonction de corrélation Ly $\alpha \times$ Ly α . 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 densité de colonne 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 4.14). La distribution en z utilisée est trois fois plus importante que dans le catalogue standard afin d’avoir suffisamment de HCD dans chaque

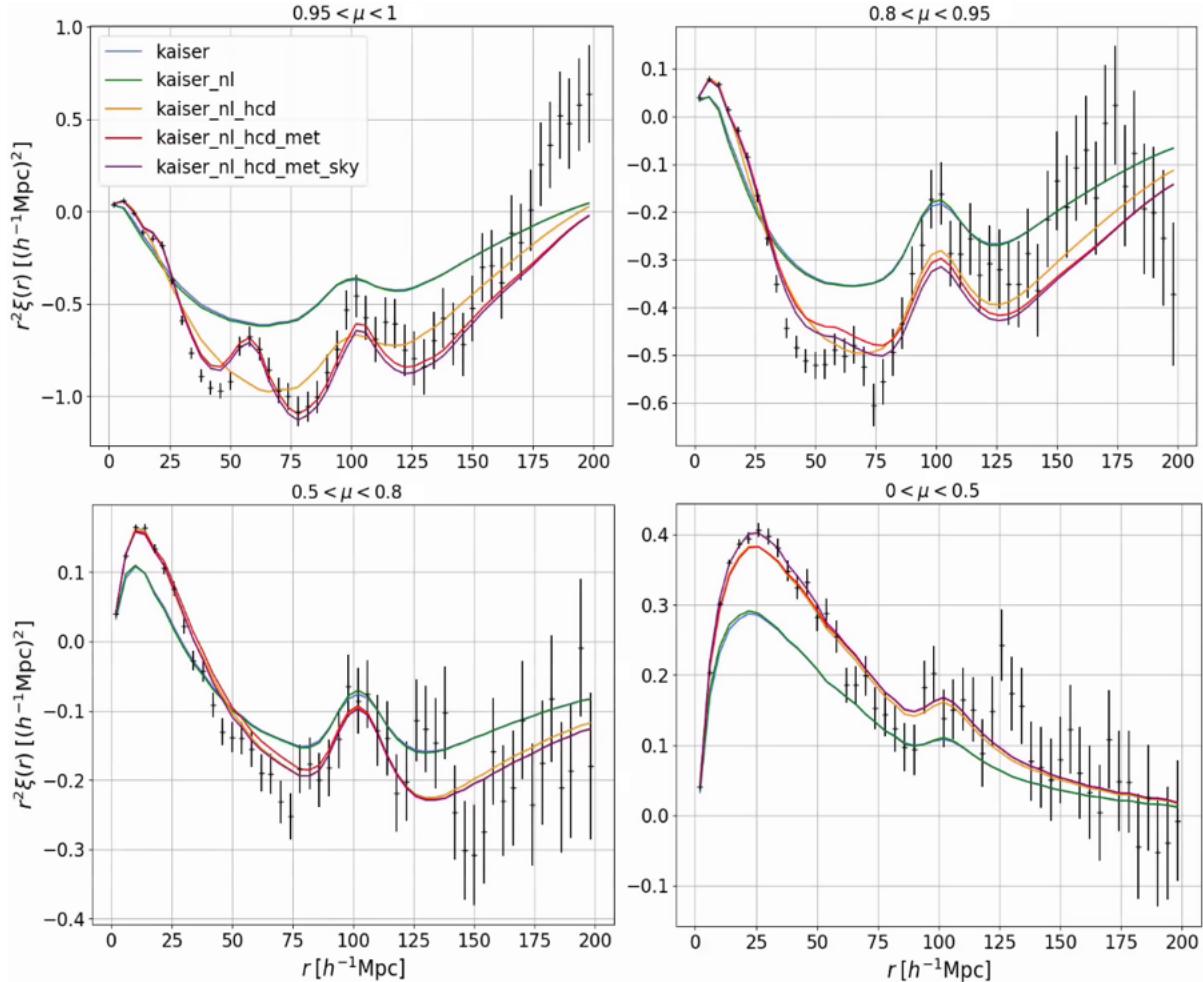


FIGURE 6.10 – Les différentes composantes du modèle ajusté sur l’auto-corrélation Ly α ×Ly α des données DR16. La ligne bleue donne le modèle de Kaiser, c’est à dire la fonction de corrélation obtenue à partir du spectre de puissance P_{QL} (équation 5.22) multiplié par le facteur de Kaiser $b_{\text{eff},\text{Ly}\alpha}^2(1 + \beta_{\text{Ly}\alpha}\mu_k^2)$. 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 5.26). La ligne rouge donne le modèle précédent plus la contribution des métaux et la ligne violette donne le modèle complet. L’auto-corrélation Ly α ×Ly α 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 μ différentes.

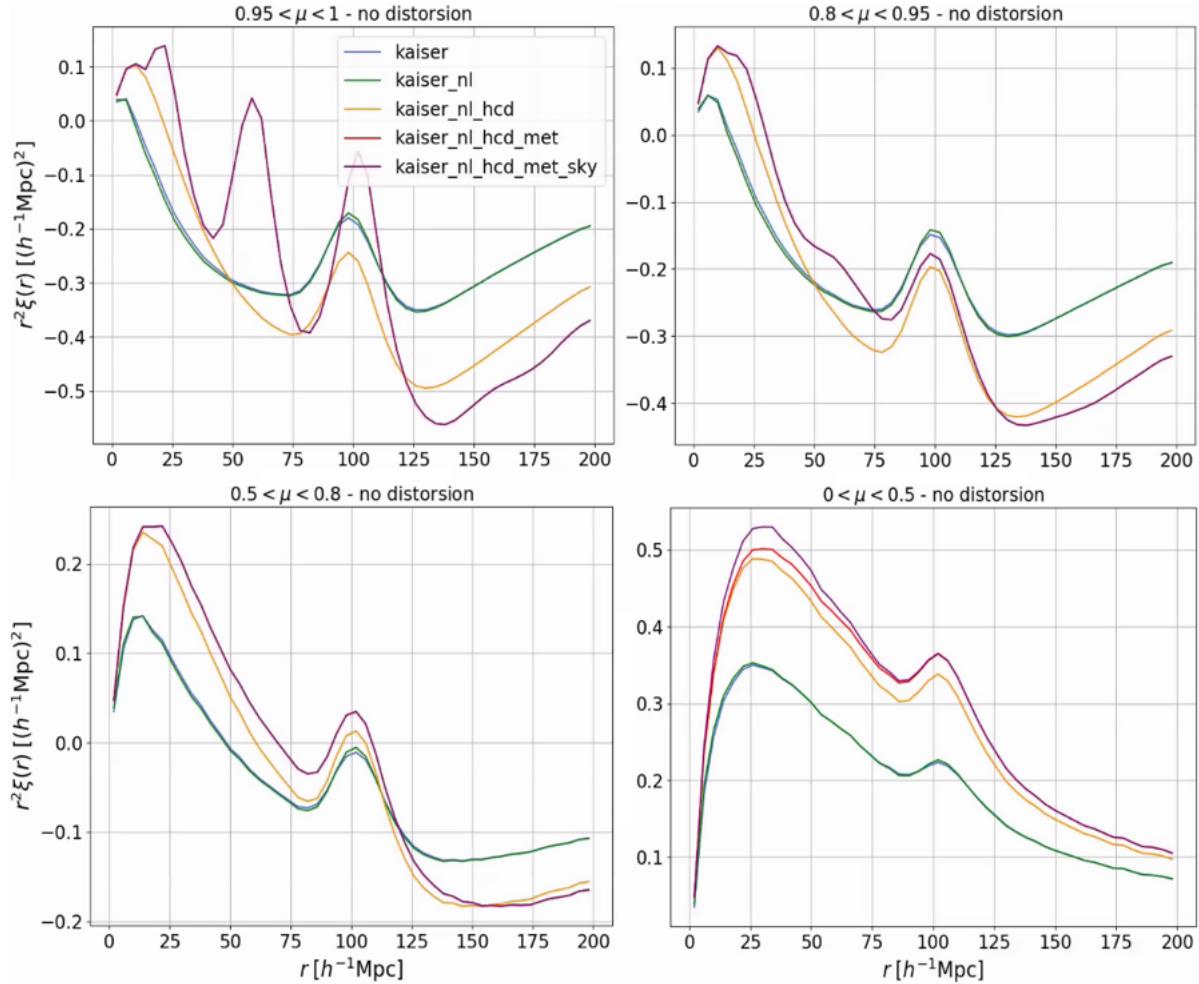


FIGURE 6.11 – Les différentes composantes du modèle ajusté sur l’auto-corrélation Ly α × Ly α des données DR16. Elles sont décrites dans la légende de la figure 6.10. Les composantes présentées ici n’incluent pas les distorsions produites par l’ajustement du continuum. Les quatre graphiques présentent ces composantes dans 4 gammes en μ .

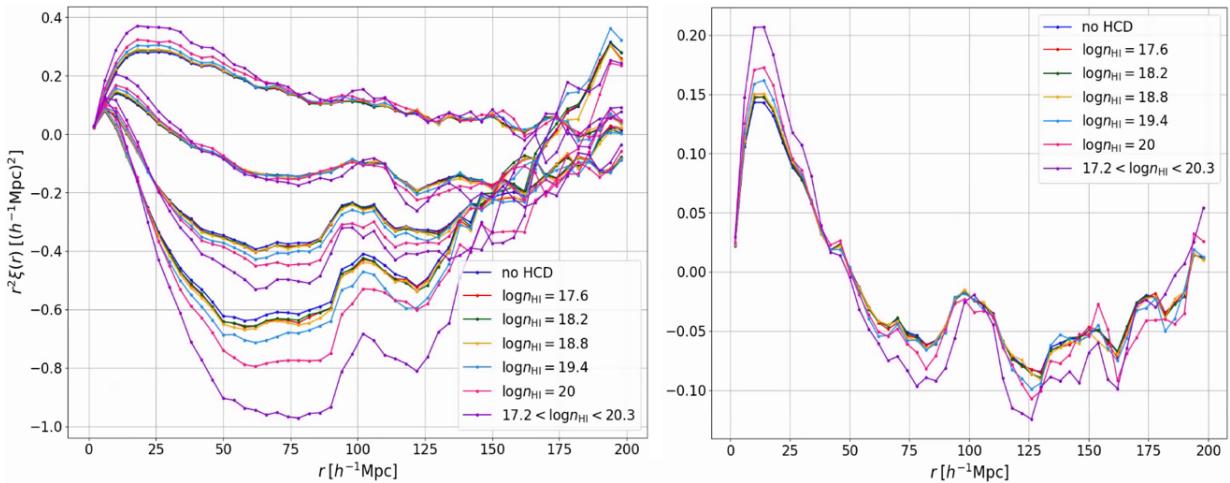


FIGURE 6.12 – Fonctions de corrélation Ly α ×Ly α estimées à partir d’une réalisation. La ligne bleue foncée donne la corrélation pour les mocks eboss-0.0 (sans HCD). La ligne violette donne la corrélation pour les mocks eboss-0.2 incluant les HCD avec $17,2 < \log n_{\text{HI}} < 20,3$. Les autres couleurs 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 μ . 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 $0 < \mu < 1$.

gamme.

La figure 6.12 présente les fonctions de corrélation Ly α ×Ly α 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 6.12 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.

Afin de comprendre les différences entre le modèle de Rogers et le modèle C-G observées sur la figure 6.9, nous étudions maintenant comment ces deux modèles 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_{HCD} adéquat aux HCD présents. En ce qui concerne le modèle de C-G, le code `picca` permet de calculer le terme F_{HCD} pour une distribution en $\log n_{\text{HI}}$ données. Nous calculons donc F_{HCD} pour chacune des valeurs de $\log n_{\text{HI}}$. Pour le modèle de Rogers, la forme de F_{HCD} est fixée (voir équation 5.28). Il nous suffit, pour chaque valeur de $\log n_{\text{HI}}$, de trouver la valeur de L_{HCD} à utiliser. Mais ceci n’est pas chose aisée. Dans l’analyse DR16, L_{HCD} est choisie de manière approximative à $10 h^{-1}$ Mpc, car il influence très peu la mesure des paramètres BAO. Mais comme nous l’avons vu à plusieurs reprises, L_{HCD} influence grandement la mesure des paramètres Ly α . Nous devons donc le choisir avec précaution.

Nous expliquons ici la méthode utilisée pour déterminer L_{HCD} . Pour chaque valeur de $\log n_{\text{HI}}$, nous choisissons une valeur de L_{HCD} qui reproduise au mieux le comportement de $F_{\text{HCD}}(k)$ dans

calculé dans le modèle CG

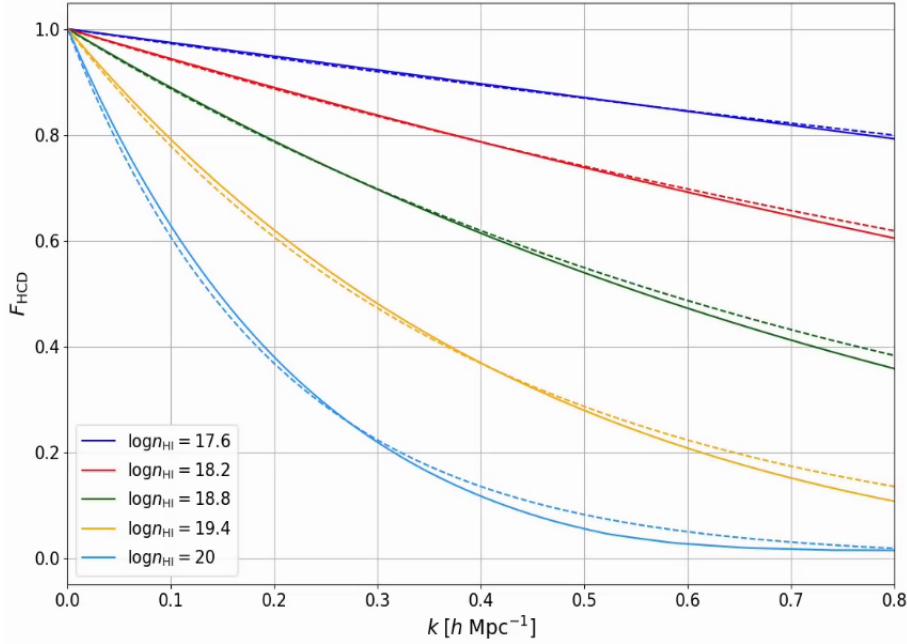


FIGURE 6.13 – Les fonctions F_{HCD} utilisées dans le cadre du modèle C-G (lignes continues) ou du modèle de Rogers (lignes tiretées) pour différentes valeurs de $\log n_{\text{HI}}$. Les valeurs de L_{HCD} utilisées pour le modèle de Rogers sont données dans le texte.

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_{HCD} que nous obtenons pour $\log n_{\text{HI}} \in [17,6; 18,2; 18,8; 19,4; 20]$ sont $L_{\text{HCD}} \in [0,28; 0,60; 1,2; 2,5; 5]$. La figure 6.13 présente les fonctions $F_{\text{HCD}}(k)$ pour le modèle C-G (lignes continues) et pour le modèle de Rogers (lignes tiretées) en utilisant les valeurs de L_{HCD} obtenues précédemment. Pour chaque valeur de $\log n_{\text{HI}}$, les fonctions F_{HCD} sont très similaires ~~dans~~ entre 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_{HCD} déterminées précédemment. Pour chacune des versions des mocks, les ajustements faits avec le modèle de Rogers et le modèle C-G produisent des résultats très similaires. Les différences sont inférieures au dizième de σ . Ceci est rassurant car les fonctions F_{HCD} utilisées dans ces deux modèles sont très similaires. Par ailleurs, pour les faibles valeurs de $\log n_{\text{HI}}$, les fonctions F_{HCD} sont très proches de 1 (voir figure 6.13). Les paramètres Ly α et HCD sont alors 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 Ly α . Cependant ces mesures ne sont pas compatibles entre elles et ne sont pas non plus compatibles avec les mesures faites sur les mocks eboss-0.0. Si ces modèles ne parviennent pas à distinguer l'effet du Ly α 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 densité de colonne.

Les chiffres de cette analyse ne sont pas donnés ?

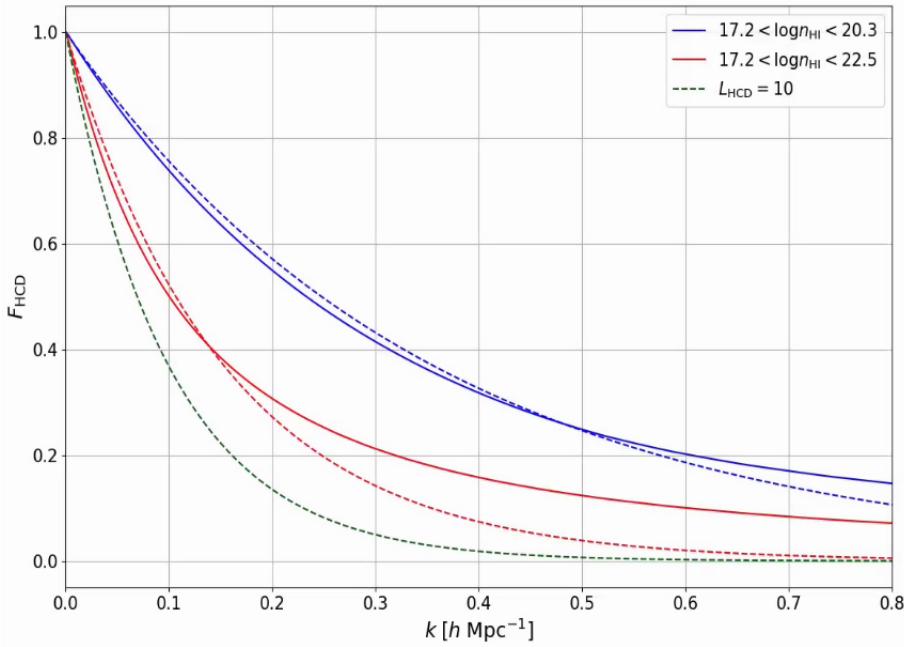


FIGURE 6.14 – Les fonctions F_{HCD} utilisées dans le cadre du modèle C-G (lignes continues) ou du modèle de Rogers (lignes tiretées). Les courbes bleues correspondent à la gamme $17,2 < \log n_{\text{HI}} < 20,3$. Les courbes rouges correspondent à la gamme $17,2 < \log n_{\text{HI}} < 22,5$. La courbe verte donne F_{HCD} utilisé dans le modèle de Rogers avec $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$.

Nous comparons maintenant les fonctions F_{HCD} correspondant à une large gamme en densité de colonne. Nous considérons premièrement la gamme en densité de colonne des HCD présents dans les mocks eboss-0.2 lorsque les HCD avec $\log n_{\text{HI}} > 20,3$ sont masqués, c'est à dire la gamme $17,2 < \log n_{\text{HI}} < 20,3$. La fonction F_{HCD} utilisée dans le modèle de C-G pour cette gamme en $\log n_{\text{HI}}$ est calculée par **picca**. Comme précédemment, afin de déterminer le L_{HCD} à utiliser dans le modèle de Rogers, nous cherchons à reproduire le comportement de $F_{\text{HCD}}(k)$ utilisé dans le modèle de C-G 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 figure 6.14 présente les fonctions F_{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 6.14 présente aussi, à titre de comparaison, la fonction F_{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 étonnés¹ que les ajustements faits sur les mocks eboss-0.2 avec le modèle de Rogers (avec $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$) et le modèle C-G produisent des résultats si similaires (voir la figure 6.9) compte tenu de ces différences sur F_{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 6.14 présente aussi les fonctions F_{HCD} obtenues pour la gamme $17,2 < \log n_{\text{HI}} < 22,5$. La valeur de L_{HCD} choisie est $6,5 h^{-1} \text{ Mpc}$. Les fonctions F_{HCD} obtenues pour cette gamme dans le cadre du modèle C-G et du modèle de Rogers (courbes rouges) sont très différentes. Le modèle de Rogers atteint ici ses limites.

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 Ly α . **faible**

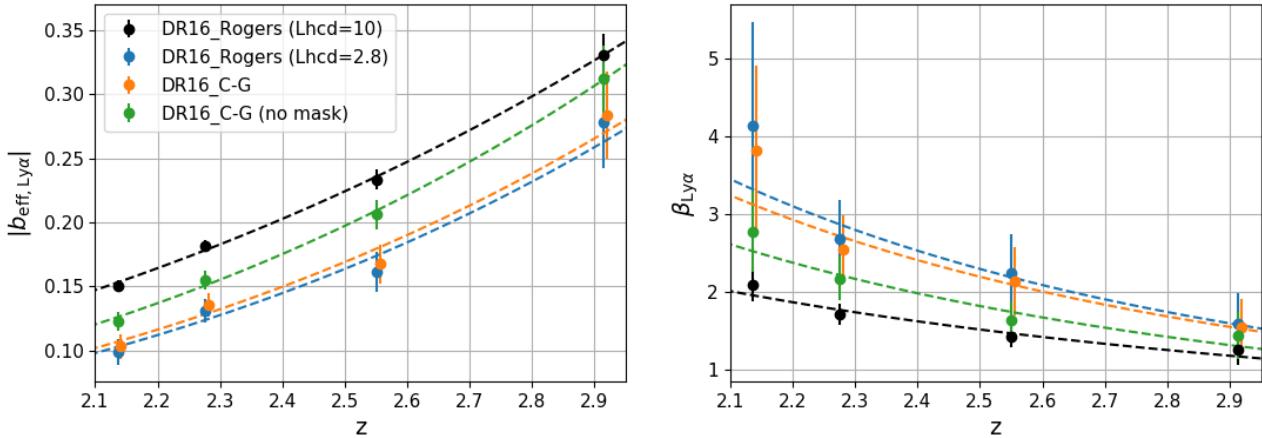


FIGURE 6.15 – Mesure des paramètres $b_{\text{eff}, \text{Ly}\alpha}$ (gauche) et $\beta_{\text{Ly}\alpha}$ (droite) sur l’auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ estimée à partir des données DR16. Les points bleus présentent la mesure faite avec le modèle de Rogers en utilisant $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$. Les points orange la mesure faite avec le modèle C-G lorsque les HCD avec $\log n_{\text{HI}} > 20,3$ sont masqués. Les points verts la mesure faite avec le modèle C-G lorsqu’aucun HCD n’est masqué. Enfin, les points noirs présentent la mesure faite avec le modèle de Rogers en utilisant $L_{\text{HCD}} = 10 h^{-1} \text{ Mpc}$. Les tirets correspondent à l’ajustement de chaque jeu de données par une loi de puissance. Les points orange sont décalés horizontalement de $\Delta z = 5 \times 10^{-3}$ pour des raisons de visibilité.

Les courbes tiretées

Suite à ces observations, nous reproduisons l’analyse des mocks eboss-0.2 où les HCD avec une densité de colonne $\log n_{\text{HI}} > 20,3$ sont masqués. Nous analysons les fonctions de corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ en utilisant le modèle de Rogers, avec $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$. Ainsi, nous vérifions que les résultats de l’ajustement sont très similaires à ceux obtenus avec le modèle C-G (les résultats varient d’un dixième de σ). Ceci est attendu car les fonctions F_{HCD} sont très similaires. De plus, comme nous l’avons vu dans la section 6.2.3, la mesure des paramètres $\text{Ly}\alpha$ faite sur les mocks eboss-0.2 où les HCD avec une densité de colonne $\log n_{\text{HI}} > 20,3$ sont masqués et ajustés avec le modèle C-G, est compatible avec la mesure faite sur les mocks eboss-0.2 où aucun HCD n’est masqué et ajustés avec le modèle C-G. Ainsi les trois mesures des paramètres $\text{Ly}\alpha$ que nous faisons (modèle C-G avec ou sans HCD masqués et modèle de Rogers avec $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$) sur les mocks eboss-0.2 sont compatibles. Cependant il reste toujours un léger écart entre la mesure faite sur les mocks eboss-0.0 et celle faite sur les mocks eboss-0.2 (environ 5 %, voir figure 6.9).

Après avoir fait ces vérifications sur les mocks, nous analysons les données DR16 avec les modèles de HCD étudiés précédemment. La figure 6.15 présente les mesures des paramètres $\text{Ly}\alpha$ faites sur les données DR16 en utilisant ces différents modèles de HCD. Nous pouvons premièrement remarquer que, comme pour les mocks, la mesure faite avec le modèle C-G lorsque les HCD avec une densité de colonne $\log n_{\text{HI}} > 20,3$ sont masqués (orange) est très similaires à la mesure faite avec le modèle de Rogers lorsque nous utilisons $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ (bleu). Cependant, contrairement aux mocks, ces deux mesures ne sont pas compatibles avec la mesure faite avec le modèle C-G lorsqu’aucun HCD n’est masqué (vert). Le fait que les deux mesures faites avec le modèle C-G (HCD masqués ou non) ne soient pas compatibles, alors que les mêmes mesures faites sur les mocks le sont (voir figure 6.9), pose un sérieux problème. Cela signifie que quelque chose dans les données ne se comporte pas de la

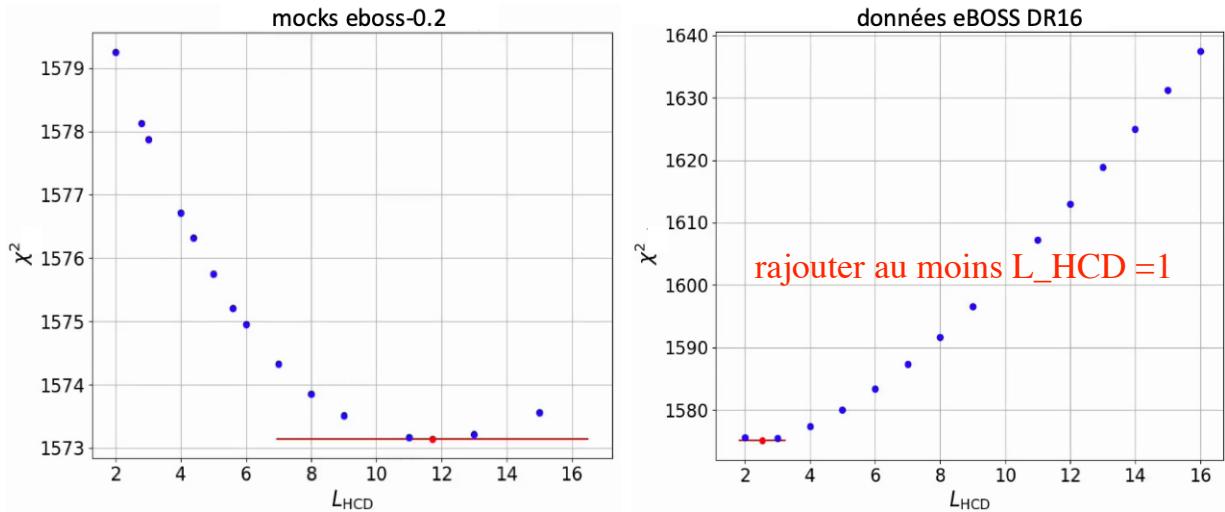


FIGURE 6.16 – Evolution de la valeur de χ^2 de l’ajustement en fonction du paramètre L_{HCD} . La figure de gauche montre l’évolution pour les mocks eboss-0.2 où les HCD avec une densité de colonne $\log n_{HI} > 20,3$ sont masqués. Celle de droite l’évolution pour les données DR16 où les HCD sont aussi masqués. Sur chaque graphique, le point rouge présente la valeur, ainsi que la barre d’erreur qui l’accompagne, de L_{HCD} qui minimise le χ^2 .

même manière que dans les mocks. Il se pourrait qu’un effet inconnu, présent dans les données, soit pris en compte par la modélisation des HCD. Cependant nous n’avons pour l’instant pas d’intuition concernant la nature de cet effet.

La dernière analyse que nous effectuons est d’étudier comment se comporte la valeur de χ^2 de l’ajustement en fonction du paramètre L_{HCD} . Nous effectuons cette étude à la fois pour les mocks et pour les données. La figure 6.16 présente la valeur de χ^2 de l’ajustement obtenue pour différentes valeurs de L_{HCD} . Sur les données DR16 (graphique de droite), nous mesurons $L_{HCD} = 2,54 \pm 0,73 h^{-1}$ Mpc, ce qui est en accord avec la valeur suggérée par l’analyse présentée dans les paragraphes précédents. Cependant, lorsque nous faisons le même exercice sur les mocks (graphique de gauche), nous mesurons $L_{HCD} = 11,73 \pm 4,75 h^{-1}$ Mpc. Cette mesure est peu contrainte sur les mocks, et en tension ($\sim 2\sigma$) avec la valeur $L_{HCD} = 2,8 h^{-1}$ Mpc suggérée par les analyses présentées précédemment.

Ici encore, nous sommes face à une incohérence : lorsque nous étudions la distribution des HCD présents dans les mocks (avec une densité de colonne comprise dans la gamme $17,2 < \log n_{HI} < 20,3$), le modèle C-G nous suggère d’utiliser $L_{HCD} = 2,8 h^{-1}$ Mpc. Pourtant lorsque L_{HCD} est laissé libre, la valeur obtenue est en tension avec $L_{HCD} = 2,8 h^{-1}$ Mpc. Quand nous fixons les paramètres Ly α aux valeurs mesurées sur les mocks eboss-0.0 et laissons libre L_{HCD} , nous mesurons $L_{HCD} = 4,38 \pm 0,47 h^{-1}$ Mpc. Ceci a pour effet de réduire les tensions. Mais la mesure produite n’est tout de même pas compatible ($\sim 3\sigma$) avec ce que suggère le modèle C-G.

on passe de 2 à 3 sigma !

Après ces nombreuses études, nous ne sommes pas convaincus quant au modèle de HCD à utiliser pour ajuster les données. Le fait que la mesure du paramètre L_{HCD} , quand il est laissé libre, soit compatible avec ce que suggère le modèle C-G (grâce à l’étude des fonctions F_{HCD}) encourage à

utiliser le modèle de Rogers avec $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$, où de manière équivalente le modèle C-G lorsque les HCD avec une densité de colonne $\log n_{\text{HI}} > 20,3$ sont masqués. Cependant, le fait que ce résultat ne soit pas également obtenu sur les mocks, c'est à dire que la valeur du paramètre L_{HCD} obtenu lorsqu'il est laissé libre dans l'ajustement des mocks eboss-0.2 (avec les HCD masqués) ne soit pas compatible avec la valeur suggérée par l'étude des fonctions F_{HCD} avec le modèle C-G, ne nous encourage pas à utiliser le modèle de Rogers avec $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$.

De plus, le fait que les ajustements effectués sur les données DR16 avec le modèle C-G, lorsque les HCD sont masqués ou ne sont pas masqués, ne soient pas compatibles ne nous conforte pas dans l'idée d'utiliser aveuglément ce modèle. Il est probable que des effets systématiques inconnus soient présents dans les données, et que les différents modèles de HCD prennent ces effets en compte.

Etant donné l'état de connaissance actuel de la modélisation des HCD et des effets systématiques qui pourraient affecter les données, nous pensons qu'il est préférable d'utiliser le modèle C-G dans le cas où les HCD ne sont pas masqués dans les données. Premièrement, le fait de ne pas masquer les HCD permet d'avoir plus de pixel d'absorptions dans l'analyse, et donc de disposer de plus de statistique pour effectuer la mesure de la position du pic BAO. De plus, lorsque le nombre de HCD présents dans les données est plus important, la fonction F_{HCD} est plus importante est donc les paramètres Ly α et ceux des HCD sont moins corrélés. Deuxièmement, le fait de masquer les HCD consiste à retirer une partie non aléatoire des données, qui plus est qui correspond à des régions denses de l'univers. Un tel processus peut introduire des effets systématiques dans l'analyse. Pour ces raisons, nous pensons qu'il est préférable de ne pas masquer les HCD, et donc d'utiliser le modèle C-G.

7

Conclusion

Dans ce manuscrit, j'ai présenté le résultat de mes recherches effectuées au sein du groupe de cosmologie du CEA Saclay, dans le cadre de la préparation de ma thèse de doctorat. Durant ces trois ans, j'ai développé des simulations qui imitent les données Lyman alpha ($\text{Ly}\alpha$) acquises par eBOSS et bientôt par DESI. Ces simulations produisent un relevé de quasars, dont les forêts $\text{Ly}\alpha$ présentes dans les spectres synthétiques possèdent les bonnes fonctions de corrélation. Pour ce faire, elles utilisent les champs aléatoires gaussiens. D'abord, nous générions un champ aléatoire qui représente le champ de densité. Nous donnons à ce champ aléatoire la bonne fonction de corrélation à l'aide des transformations de Fourier. Puis, grâce à une transformation log-normale, nous tirons un relevé de quasars. Ensuite, à partir de chaque quasar, nous reconstruisons la densité le long de la ligne de visée en interpolant ce champ. Enfin, nous obtenons la fraction de flux transmise en appliquant l'approximation de Gunn-Peterson (FGPA) au champ interpolé le long de la ligne de visée. Ceci nous permet d'obtenir la forêt d'absorption $\text{Ly}\alpha$ pour chaque quasar de notre relevé. Aussi, nous calculons la prédiction des fonctions de corrélation que nous devons obtenir avec ces simulations. Ceci nous permet de vérifier leur construction en comparant la fonction de corrélation prédite à la fonction de corrélation mesurée dans nos simulations. Enfin, l'analyse des fonctions de corrélation issues des mocks (présentée dans le chapitre 5) montre que les mocks sont en très bon accord avec la prédiction, et qu'ils sont très bien décrits par les modèles utilisés pour ajuster les données.

Ces simulations ont été utilisées dans l'analyse qui mesure l'échelle des oscillations acoustiques de baryon (BAO) à partir des données finales d'eBOSS DR16 (DU MAS DES BOURBOUX, RICH et al. 2020) afin de vérifier que cette analyse produisait des mesures non biaisées de l'échelle BAO. Elles seront aussi utilisées par la collaboration DESI afin de vérifier les analyses $\text{Ly}\alpha$. Par ailleurs, ces simulations ont été utilisées par RAVOUX et al. (2020) pour produire une carte tomographique du champ de matière à grande échelle à partir des données d'eBOSS. Elles ont permis de valider l'algorithme utilisé pour reconstruire ce champ de matière.

Nos simulations, avec celles présentées dans FARR et al. (2019), sont à la pointe des programmes qui simulent la distribution d'absorption $\text{Ly}\alpha$ à grande échelle. Ces simulations sont un atout majeur et indispensable aux analyses de mesure de l'échelle BAO et de mesure des distorsions dans l'espace des redshifts (RSD) avec le $\text{Ly}\alpha$. Les simulations que nous présentons dans ce manuscrit ont l'avantage de posséder une fonction de corrélation prédictible, ce qui profite beaucoup aux analyses RSD. Cette

prédition nous permet d'ajuster efficacement les valeurs du biais $b_{\text{Ly}\alpha}$, du paramètre RSD $\beta_{\text{Ly}\alpha}$, de la transmission moyenne \bar{F} et du spectre de puissance $P^{1D}(k)$ à chaque redshift. De plus, ces simulations bénéficient aussi de distributions de quasars et HCD (High Column Density) très proches de celles observées dans les données, ce qui permet d'obtenir des fonctions de corrélation croisées entre les différents éléments des simulations très correctes. Afin d'améliorer encore ces simulations, il conviendrait d'ajouter les absorptions causées par les HCD et les métaux directement lors de la construction des forêts Ly α , avant l'ajout du continuum de chaque quasar.

Grâce à l'étude approfondie de nos simulations (~~voir les chapitres 5 et 6~~), nous montrons que l'analyse des données finales d'eBOSS DR16, présentée dans DU MAS DES BOURBOUX, RICH et al. (2020), produit une mesure des paramètres BAO α_{\parallel} et α_{\perp} non ~~baisée~~ à 1% près. Nous montrons aussi que nous comprenons la forme de la fonction de corrélation Ly α \times Ly α , c'est à dire la mesure du biais $b_{\text{Ly}\alpha}$ et du paramètre RSD $\beta_{\text{Ly}\alpha}$ du Ly α , à environ 20%. Pour affiner cette compréhension, nous devons comprendre pourquoi les distorsions induites sur la fonction de corrélation Ly α \times Ly α par l'ajustement du continuum des quasars ne sont pas complètement prises en compte par la matrice de distorsion. Ceci doit passer par une comparaison approfondie des simulations brutes (les raw mocks) et des simulations avec continuum (~~les mocks eboss-0.0~~). Afin de simplifier les comparaisons, il est prévu ~~d'étudier les~~ mocks avec continuum sans ajouter le bruit de mesure et sans ~~considérer~~ inclure la diversité des quasars. **de produire des mesures**

Lors de l'étude des simulations avec HCD (~~voir le chapitre 6~~), nous montrons que la modélisation des HCD n'est pas parfaite, et qu'elle mène à ~~une mesure~~ du biais $b_{\text{Ly}\alpha}$ et du paramètre RSD $\beta_{\text{Ly}\alpha}$ du Ly α biaisée. Premièrement, la valeur du paramètre L_{HCD} dans l'analyse des données d'eBOSS (présentée dans DU MAS DES BOURBOUX, RICH et al. (2020)), fixée à $10 h^{-1} \text{ Mpc}$, est mal choisie. Nous suggérons d'utiliser $L_{\text{HCD}} = 2,8 h^{-1} \text{ Mpc}$ si le modèle de Rogers est utilisé. Deuxièmement, nous étudions une modélisation alternative des HCD, que nous appelons le modèle C-G (~~voir le chapitre 6~~). Ce modèle produit des mesures compatibles sur les mocks. Cependant, des différences restent à prendre lorsque nous ajustons les données, notamment lorsque nous ajustons les données avec HCD masqués ou sans masquer les HCD. Suite à ces observations, nous privilégions la piste d'un effet inconnu, présent dans les données (et non dans les mocks) et étant à l'origine de ces différences. Tant que notre compréhension des HCD et de leur modélisation n'aura pas évolué, nous pensons qu'il est préférable d'utiliser le modèle C-G, sans masquer les HCD. Ceci pour deux raisons : le fait de ne pas masquer les HCD permet de disposer d'une plus grande statistique pour mesurer la position du pic BAO et les paramètres Ly α . De plus, les paramètres Ly α sont moins corrélés dans un tel cas. La seconde raison est que le fait de masquer les HCD peut potentiellement introduire des systématiques dans la fonction de corrélation : les pixels d'absorptions identifiés comme HCD et masqués dans l'analyse ne correspondent pas à une partie aléatoire des données, mais à des régions très denses du ciel. Les masquer peut introduire des systématiques dans l'analyse.

Avec l'arrivée du relevé DESI, une grande quantité de données, et donc une grande puissance statistique, sera mise à la disposition de la communauté scientifique. Cette puissance statistique va permettre d'effectuer des mesures BAO avec le Ly α à différents redshifts, et peut-être de mener une première

à

analyse RSD avec le Ly α . Dans ce but, un effort doit être fait pour comprendre les problèmes liés à la matrice de distorsion et comprendre la modélisation des HCD. Un effort doit aussi être fait pour comprendre les différences qu'il y a entre les mesures faites sur l'auto-corrélation Ly α × Ly α et les mesures faites sur la corrélation croisée Ly α × QSO (~~voir le chapitre 5~~). Nous montrons dans ce manuscrit que les simulations que nous avons développées permettent d'étudier ces effets et ainsi de comprendre les potentielles systématiques qui affectent ces mesures. Ces simulations sont donc un atout majeur pour préparer les analyses Ly α du relevé DESI.