



**HAL**  
open science

# Mesure de la position du pic d'oscillations acoustiques baryoniques dans les forêts $\text{Ly}\alpha$ et $\text{Ly}\beta$ des spectres des quasars du relevé eBOSS-SDSS IV

Victoria De Sainte Agathe

► **To cite this version:**

Victoria De Sainte Agathe. Mesure de la position du pic d'oscillations acoustiques baryoniques dans les forêts  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  des spectres des quasars du relevé eBOSS-SDSS IV. Cosmologie et astrophysique extra-galactique [astro-ph.CO]. Sorbonne Université, 2019. Français. NNT : 2019SORUS373 . tel-03008641

**HAL Id: tel-03008641**

**<https://theses.hal.science/tel-03008641>**

Submitted on 16 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT  
DE SORBONNE UNIVERSITÉ

*présentée par*

**Victoria de Sainte Agathe**

*Pour obtenir le grade de*

DOCTEUR ÈS SCIENCES  
DE SORBONNE UNIVERSITÉ

*Spécialité : Physique*

École doctorale des Sciences de la Terre et de l'Environnement et Physique de l'Univers, Paris

**Mesure de la position du pic d'oscillations  
acoustiques baryoniques dans les forêts  $\text{Ly}\alpha$  et  $\text{Ly}\beta$   
des spectres des quasars du relevé eBOSS-SDSS IV**

Soutenue le 23 septembre 2019 devant le jury composé de :

Pr	Christophe	BALLAND	Directeur de thèse
Pr	Alain	BLANCHARD	Rapporteur
Pr	Pedro	FERREIRA	Examineur
Dr	Julien	GUY	Invité
Pr	Delphine	HARDIN	Présidente
Pr	Jean-Paul	KNEIB	Examineur
Dr	Vanina	RUHLMANN-KLEIDER	Rapporteur
Dr	Cécile	RENAULT	Examineur



PhD THESIS  
OF SORBONNE UNIVERSITÉ

*presented by*

**Victoria de Sainte Agathe**

*Submitted in fulfillment of the requirements for the degree of*

DOCTEUR ÈS SCIENCES  
DE SORBONNE UNIVERSITÉ

*Spéciality : Physics*

École doctorale des Sciences de la Terre et de l'Environnement et Physique de l'Univers, Paris

**Measuring the position of the baryonic acoustic  
oscillation peak with the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  forests of the  
eBOSS-SDSS IV quasar spectra**

Defended on September 23<sup>rd</sup> 2019 in front of the committee :

Pr	Christophe	BALLAND	Supervisor
Pr	Alain	BLANCHARD	Referee
Pr	Pedro	FERREIRA	Examiner
Dr	Julien	GUY	Guest
Pr	Delphine	HARDIN	President
Pr	Jean-Paul	KNEIB	Examiner
Dr	Vanina	RUHLMANN-KLEIDER	Referee
Dr	Cécile	RENAULT	Examiner





*Whenever life gets you down, Mrs. Brown,  
And things seem hard or tough,  
And people are stupid, obnoxious or daft,  
And you feel that you've had quite eno-o-o-o-ough,*

*Just remember that ...*

*Our universe itself keeps on expanding and expanding,  
In all of the directions it can whiz ;  
As fast as it can go, at the speed of light, you know,  
Twelve million miles a minute and that's the fastest speed there is.  
So remember, when you're feeling very small and insecure,  
How amazingly unlikely is your birth ;  
And pray that there's intelligent life somewhere out in space,  
'Cause there's bugger all down here on Earth !*

*The Meaning of Life, Monty Python*



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Quelques éléments de cosmologie</b>	<b>3</b>
1.1 Le modèle standard de la cosmologie . . . . .	3
1.1.1 Modèle $\Lambda$ CDM . . . . .	3
1.1.2 Le fond diffus cosmologique . . . . .	4
1.1.3 La mesure des paramètres de $\Lambda$ CDM . . . . .	6
1.2 L’Univers en équations . . . . .	7
1.2.1 Les équations de champ d’Einstein . . . . .	7
1.2.2 Décrire l’espace . . . . .	8
1.2.3 La composition énergétique de l’Univers . . . . .	10
1.2.4 Les équations de Friedmann-Lemaître . . . . .	12
1.3 Les distances en cosmologie . . . . .	14
1.3.1 Le <i>redshift</i> . . . . .	15
1.3.2 La distance comobile . . . . .	15
1.3.3 La distance angulaire . . . . .	16
1.3.4 La distance de Hubble . . . . .	17
1.3.5 La distance de luminosité . . . . .	18
1.4 Mesurer l’expansion de l’Univers . . . . .	19
1.4.1 Les supernovae de type Ia . . . . .	19
1.4.2 Les oscillations acoustiques baryoniques . . . . .	21
<b>2 Fonction de corrélation de la matière et forêts Ly<math>\alpha</math></b>	<b>27</b>
2.1 La fonction de corrélation à deux points de la matière . . . . .	27
2.1.1 Définition . . . . .	27
2.1.2 Effets de l’expansion et de la formation des structures . . . . .	29
2.1.3 Traceur de la matière et biais . . . . .	31
2.1.4 Espace réel et espace des <i>redshifts</i> . . . . .	31
2.1.5 Formule de Kaiser . . . . .	33
2.2 Les forêts Lyman- $\alpha$ . . . . .	34
2.2.1 Quasars . . . . .	34
2.2.2 L’idée de Gunn-Peterson . . . . .	34
2.2.3 Régions Lyman- $\alpha$ et Lyman- $\beta$ . . . . .	37
2.3 Mesure de la position du pic BAO avec les forêts Lyman- $\alpha$ . . . . .	38

<b>3</b>	<b>L'expérience eBOSS du relevé SDSS-IV</b>	<b>41</b>
3.1	Le relevé SDSS . . . . .	41
3.2	L'instrument . . . . .	42
3.2.1	Le télescope . . . . .	42
3.2.2	La caméra . . . . .	45
3.2.3	Le système de spectroscopie . . . . .	48
3.3	La prise de données . . . . .	51
3.3.1	Objectifs de eBOSS . . . . .	51
3.3.2	Phase photométrique . . . . .	53
3.3.3	Sélection des cibles . . . . .	53
3.3.4	Phase spectroscopique . . . . .	56
3.4	Le catalogue des spectres des quasars . . . . .	57
3.4.1	Identification des objets . . . . .	57
3.4.2	Détermination du <i>redshift</i> . . . . .	58
3.4.3	Détection des <i>Broad Absorption Lines</i> . . . . .	61
<b>4</b>	<b>Du spectre au champ d'absorption</b>	<b>63</b>
4.1	Description du modèle . . . . .	64
4.1.1	Définition des régions $\text{Ly}\alpha$ et $\text{Ly}\beta$ . . . . .	64
4.1.2	Modéliser le champ d'absorption . . . . .	64
4.1.3	L'estimation des poids . . . . .	66
4.2	Préparation des spectres . . . . .	67
4.2.1	Sélection des spectres . . . . .	69
4.2.2	Sélection des longueurs d'onde . . . . .	71
4.3	Calcul du champ d'absorption . . . . .	71
4.3.1	Initialisation . . . . .	71
4.3.2	Relation de récurrence . . . . .	72
4.3.3	Extraction du champ de $\delta$ . . . . .	73
4.4	Prise en compte des effets liés à l'ajustement du continuum . . . . .	75
4.4.1	Champ vrai et champ mesuré . . . . .	75
4.4.2	Projecteur $\mathbb{P}$ . . . . .	76
4.4.3	Projection des $\delta_q(\lambda)$ . . . . .	77
4.5	Évolution du champ de $\delta$ en fonction du <i>redshift</i> . . . . .	77
4.5.1	Le spectre de puissance à une dimension . . . . .	78
4.5.2	Ré-étalonnage des poids . . . . .	81
<b>5</b>	<b>Mesure de la fonction de corrélation <math>\text{Ly}\alpha</math></b>	<b>83</b>
5.1	La fonction de corrélation à trois dimensions . . . . .	84
5.1.1	Les types d'absorption dans les régions $\text{Ly}\alpha$ et $\text{Ly}\beta$ . . . . .	84
5.1.2	Les coordonnées . . . . .	85
5.1.3	Séparation entre les absorptions $\text{Ly}\alpha$ et $\text{Ly}\beta$ dans la région $\text{Ly}\beta$ . . . . .	86
5.1.4	Estimateur de la fonction de corrélation . . . . .	88
5.1.5	Calcul de la fonction de corrélation . . . . .	91
5.1.6	Résultats . . . . .	94
5.2	Détection des métaux . . . . .	97
5.2.1	Empilement des DLA . . . . .	97
5.2.2	La fonction de corrélation à une dimension . . . . .	100
5.3	La matrice de covariance . . . . .	100

5.3.1	Développement de Wick . . . . .	101
5.3.2	Estimation par la méthode de <i>subsampling</i> . . . . .	102
5.3.3	Covariance entre deux fonctions de corrélation . . . . .	106
<b>6</b>	<b>Modélisation et ajustement de la fonction de corrélation mesurée</b>	<b>109</b>
6.1	La fonction de corrélation $\text{Ly}\alpha$ . . . . .	109
6.1.1	Extraction du pic BAO du spectre de puissance linéaire . . . . .	110
6.1.2	Le modèle de Kaiser . . . . .	112
6.1.3	Les effets non-linéaires . . . . .	117
6.1.4	Modéliser l'échantillonnage . . . . .	120
6.1.5	Du spectre de puissance à la fonction de corrélation . . . . .	122
6.2	Les métaux . . . . .	123
6.2.1	Le champ d'absorption des métaux . . . . .	125
6.2.2	La matrice des métaux et la fonction de corrélation métallique . . . . .	127
6.2.3	Contribution totale des métaux . . . . .	128
6.3	La matrice de distorsion . . . . .	128
6.3.1	Effet de l'ajustement du continuum sur la fonction de corrélation . . . . .	128
6.3.2	La projection des $\delta$ et ses effets sur la fonction de corrélation . . . . .	131
6.3.3	Le cas $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . . . . .	132
6.4	Modéliser les erreurs systématiques inconnues . . . . .	133
6.5	L'ajustement . . . . .	136
6.5.1	La fonction de vraisemblance . . . . .	136
6.5.2	Contraintes sur l'ajustement des paramètres . . . . .	137
6.5.3	Les modèles ajustés . . . . .	138
<b>7</b>	<b>Résultats</b>	<b>143</b>
7.1	Résultats des ajustements . . . . .	143
7.1.1	Ajustement simple de $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . . . . .	143
7.1.2	Ajustement combiné de $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ et $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . . . . .	147
7.1.3	Ajustement combiné des fonctions de corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ et $\text{Ly}\alpha$ -quasar . . . . .	155
7.2	Contraintes cosmologiques . . . . .	161
7.2.1	Contraintes sur les paramètres $(D_H(z)/r_d, D_M(z)/r_d, D_V(z)/r_d)$ . . . . .	161
7.2.2	Contraintes sur les paramètres $(\Omega_m, \Omega_\Lambda)$ . . . . .	165
	<b>Conclusion</b>	<b>169</b>
	<b>A Cosmologie fiducielle</b>	<b>171</b>
	<b>B Transitions atomiques utiles</b>	<b>173</b>
	<b>C Module picca</b>	<b>175</b>
C.1	Calcul du champ de $\delta$ . . . . .	175
C.2	Calcul de la fonction de corrélation . . . . .	176
C.3	Calcul de la matrice de distorsion . . . . .	177
C.4	Calcul de la matrice de covariance . . . . .	177
C.5	Calcul de la matrice des métaux . . . . .	177
C.6	L'ajustement . . . . .	178
	<b>D Publication</b>	<b>179</b>



# Introduction

À la fin des années 1990, la découverte de l'accélération de l'expansion de l'univers (Riess et al. 1998, Perlmutter et al. 1999) a bouleversé les conceptions cosmologiques de l'époque. En effet, comment expliquer cette accélération dans un univers composé majoritairement de matière? Deux hypothèses ont alors été proposées :

- soit, il existe une mystérieuse forme d'énergie, l'*énergie noire*, qui a une pression négative et qui agit comme le "carburant" de l'expansion,
- soit, la relativité générale n'est plus valide aux échelles cosmologiques.

Seule la mesure précise de l'évolution de l'expansion avec le *redshift* permettra de déterminer laquelle de ces deux hypothèses est la bonne et de construire un modèle théorique capable de décrire l'univers aux très grandes échelles.

Pour réaliser cet objectif, plusieurs méthodes d'observation doivent être réalisées. D'une part, car cela assure de la validité des résultats et d'autre part, car le *redshift* de la mesure dépend de la technique employée. Par exemple, la mesure des différences de température des photons du fond diffus cosmologique nous renseigne sur l'état de l'univers au moment de la recombinaison ( $z \simeq 1100$ ), alors que la mesure des distances de luminosité des supernovae de type Ia nous donne accès au taux d'expansion aux *redshifts* entre 0.5 et 1.

L'utilisation des oscillations acoustiques baryoniques pour la mesure du taux d'expansion de l'univers est une technique relativement récente qui présente l'avantage d'être peu affectée par des incertitudes astrophysiques, comparativement à d'autres méthodes comme les supernovae. La mesure de l'échelle acoustique, à un *redshift*  $z$ , parallèlement et perpendiculairement à la ligne de visée donne accès, respectivement, aux rapports  $D_M(z)/r_d$  et  $D_H(z)/r_d$  où  $D_H$  est la distance de Hubble,  $D_M$  la distance angulaire comobile et  $r_d$  l'horizon acoustique, *i.e.* l'échelle acoustique à la recombinaison (voir chapitre 1). L'empreinte des oscillations acoustiques baryoniques est détectée sous la forme d'un pic à  $r_d \simeq 100 h^{-1}\text{Mpc}$  dans la fonction de corrélation à deux points de la matière. Puisque l'échelle acoustique est très grande, sa mesure est peu influencée par la formation non-linéaire des structures mais nécessite que le volume du relevé soit suffisamment large pour pouvoir être réalisée. On peut étudier la répartition de la matière à l'aide de différents traceurs. Pour les *redshifts* inférieurs à 1, la répartition des galaxies donne accès à la position du pic BAO. L'expérience BOSS a permis d'en déduire le taux d'expansion de l'univers aux *redshifts* 0.38, 0.51 et 0.61 avec une précision entre 2.2 et 2.7 %. À plus haut *redshift*, la densité de galaxies observées n'est plus suffisante pour détecter l'empreinte BAO. Dans cette thèse, je décrirai l'utilisation des absorptions Lyman- $\alpha$  visibles dans les région Lyman- $\alpha$  et

Lyman- $\beta$  des spectres des quasars. La variation du flux mesuré dépend directement de la densité des nuages d'hydrogène neutre présents sur la ligne de visée du quasar. Ainsi, chaque spectre donne accès aux positions d'une centaine d'absorbants Lyman- $\alpha$ , ce qui permet de mesurer le taux d'expansion à des *redshifts* élevés ( $z \simeq 2.3$ ), même si la densité de quasars observés est faible (voir chapitre 2). La précédente mesure utilisant les absorptions Lyman- $\alpha$  des spectres du relevé BOSS a donné une précision de 3.6% sur la mesure du taux d'expansion à  $z = 2.33$ .

Dans ce manuscrit, je détaille le processus d'analyse qui conduit à la mesure des rapports  $D_H(z)/r_d$  et  $D_M(z)/r_d$ , à  $z = 2.34$ , en utilisant les régions Lyman- $\alpha$  et Lyman- $\beta$  des spectres des quasars du relevé eBOSS-SDSS IV. Dans le chapitre 3, je décris l'instrument d'observation ainsi que la chaîne d'extraction qui aboutit à un échantillon d'environ 200 000 spectres de quasars à haut *redshift* ( $z > 2.0$ ). Dans la première phase de mon travail, j'extrais de chacun de ces spectres le champ d'absorption qui est créé par les absorbants Lyman- $\alpha$  présents sur la ligne de visée (chapitre 4). Puis je calcule la fonction de corrélation Lyman- $\alpha$  à deux points en corrélant deux à deux les champs d'absorption (chapitre 5).

La mesure de l'échelle acoustique est réalisée en ajustant un modèle sur la fonction de corrélation mesurée. Ce modèle, décrit dans le chapitre 6, inclut, entre autres, les effets de distorsion liés au *redshift*, de contamination par les métaux du milieu intergalactique et de distorsion créés par la procédure d'extraction des champs d'absorption. Les résultats de l'ajustement sont décrits dans le chapitre 7. Dans un premier temps je présente la mesure de  $D_H(z)/r_d$  et  $D_M(z)/r_d$  dans le cas de l'ajustement simple de la fonction de corrélation Lyman- $\alpha$ , puis je réalise un ajustement combiné en ajoutant la fonction de corrélation croisée entre les absorptions Lyman- $\alpha$  et les quasars, qui conduit aux précisions 2.5% et 3.4% sur la mesure de l'échelle acoustique respectivement parallèlement et perpendiculairement à la ligne de visée. En combinant ces résultats avec des mesures de l'échelle acoustique à différents *redshifts*, j'en déduis des contraintes sur les paramètres cosmologiques  $\Omega_m$  et  $\Omega_\Lambda$  dans le cadre du modèle  $\Lambda$ CDM. Ces contraintes sont, actuellement, les plus fortes sur ces paramètres à bas *redshift*.

# Chapitre 1

## Quelques éléments de cosmologie

Dans ce chapitre, j'exposerai quelques notions de cosmologie nécessaires à la compréhension des chapitres suivants. Dans un premier temps (section 1.1), je décrirai brièvement le modèle  $\Lambda$ CDM, qui est le modèle standard utilisé pour décrire la physique de l'univers aux grandes échelles. Puis je présenterai la version mathématique du modèle  $\Lambda$ CDM, avec les équations de champ d'Einstein et les deux équations de Friedmann-Lemaître (section 1.2). Dans l'univers en expansion, la notion de distance peut être ambiguë. Dans la section 1.3, j'exposerai les différentes distances cosmologiques utilisées dans ce manuscrit. Enfin, je décrirai deux sondes cosmologiques, les supernovae de type Ia et les oscillations acoustiques baryoniques, utilisées pour mesurer l'expansion de l'univers (section 1.4).

### 1.1 Le modèle standard de la cosmologie

#### 1.1.1 Modèle $\Lambda$ CDM

Un modèle cosmologique permet de décrire, avec un jeu de paramètres indépendants, les propriétés physiques de l'univers aux grandes échelles. Depuis la découverte de l'accélération de l'expansion (Riess et al. 1998, Perlmutter et al. 1999), le modèle standard de la cosmologie est  $\Lambda$ CDM,  $\Lambda$ -Cold Dark Matter, qui suppose l'existence de deux composantes en plus de la matière ordinaire et du rayonnement :

- $\Lambda$ , appelée *constante cosmologique*. On verra dans la section 1.2 qu'elle peut être considérée comme une propriété intrinsèque de l'espace temps ou bien comme une nouvelle forme d'énergie, l'*énergie noire*, dont la densité est constante.
- CDM, *Cold Dark Matter* (*Matière Noire Froide* en français), qui désigne de la matière non visible (donc noire) et dont les particules ne sont pas relativistes.

Ces deux composantes n'ont pas été directement observées. Leur existence est supposée car elles permettent d'expliquer certains phénomènes physiques observés : dans le cas de l'énergie noire, l'accélération de l'expansion ; dans le cas de la matière noire, entre autres, la formation des grandes structures et la vitesse de rotation de certaines galaxies (Zwicky 1933, Zwicky 1937).

Le modèle  $\Lambda$ CDM se construit en deux étapes. D'abord, on décrit la géométrie de l'univers (voir sous-section 1.2.2), puis on décrit son contenu énergétique (voir sous-section 1.2.3). Les

deux sont liés via les équations de champ d'Einstein (voir sous-section 1.2.1).

La force du modèle  $\Lambda$ CDM est de pouvoir expliquer la plupart des observations actuelles (*e.g.*, l'existence et la forme du fond diffus cosmologique, la distribution des grandes structures, l'accélération de l'expansion ...) avec seulement six paramètres indépendants, qui seront décrits dans la sous-section 1.1.3. Ces paramètres sont ajustés sur les mesures du fond diffus cosmologique (Planck Collaboration et al. 2016), qui fournit des informations très précises sur l'état de l'univers au moment de la recombinaison. Dans la sous-section suivante, je vais expliquer d'où vient le fond diffus cosmologique et comment le mesurer.

### 1.1.2 Le fond diffus cosmologique

Dans l'univers primordial, les photons et les baryons<sup>1</sup> sont couplés et forment un plasma. Lorsque l'univers a suffisamment refroidi pour permettre le découplage des deux espèces ( $z \simeq 1100$ ), les premiers atomes neutres se créent et les photons sont alors libres de diffuser. Cette époque est appelée la recombinaison.

Le *fond diffus cosmologique* (*Cosmic Microwave Background* en anglais, CMB dans la suite du manuscrit) est le rayonnement électromagnétique associé aux photons qui nous parviennent de la recombinaison. Puisque l'univers était opaque avant cette époque, le CMB constitue donc la première observation de l'univers. La mesure du spectre des fluctuations de température du CMB nous donne des informations sur la distribution du plasma au moment de la recombinaison. En effet, les photons qui viennent d'une zone de surdensité ont perdu de l'énergie en s'échappant du potentiel gravitationnel local, ils sont décalés vers le rouge et donc plus froids. A contrario, les photons provenant des zones de sous-densités sont plus bleus et plus chauds.

La figure 1.1 montre une projection sur le ciel du spectre des fluctuations de températures du CMB, mesuré par le satellite Planck (Planck Collaboration et al. 2016). Les zones de surdensité, plus froides, sont en bleu, et les zones de sous-densité, en rouge. Cette carte nous indique que l'univers est quasiment homogène au moment de la recombinaison. On n'y voit que de très petites inhomogénéités, mais qui sont à l'origine des structures à grande échelle. En effet, l'excès de potentiel gravitationnel dans les zones de surdensité accrète la matière dans ces régions et le phénomène s'intensifie au cours du temps, donnant naissance aux galaxies d'aujourd'hui.

Pour étudier le CMB, on réalise une projection multipolaire sur le ciel des fluctuations de température. Celles-ci sont décomposées en une combinaison linéaire d'harmoniques sphériques  $Y_{lm}$ , avec  $l$  le moment multipolaire qui est relié à l'angle de séparation dans le ciel  $\Delta\theta$  via  $l = \pi/\Delta\theta$ . La variance  $C_l$  de l'amplitude de chaque harmonique  $Y_{lm}$  permet de décrire la distribution des variations de température des photons du CMB. Cette décomposition en harmoniques sphériques est l'analogie, sur la sphère, au spectre de puissance du chapitre 2.

La figure 1.2 montre les variances  $D_l = l(l+1)C_l/2\pi$  en fonction du moment multipolaire  $l$ . Le modèle  $\Lambda$ CDM, en rouge, est ajusté sur les données sur une large gamme de multipôles. Cet ajustement remarquable justifie l'utilisation de  $\Lambda$ CDM comme modèle standard de la cosmologie. Le spectre de puissance peut être divisé en trois régions (pour plus de détails, voir le chapitre 7 de Dodelson 2003). La première région,  $l < 100$ , décrit l'univers lors de l'inflation et donne

---

<sup>1</sup>En cosmologie, on utilise improprement le terme de baryons pour désigner la matière ordinaire, par opposition à la matière noire.

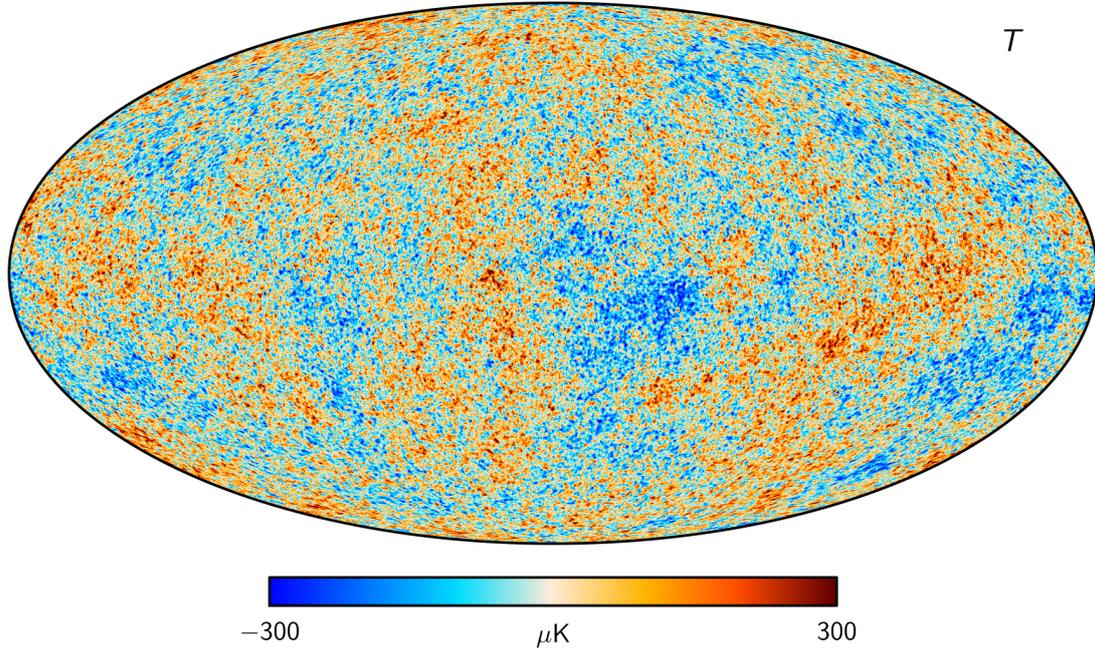


FIGURE 1.1: Carte du rayonnement fossile mesuré par le satellite Planck. Source : [https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/File:Cmb\\_inpaint\\_T\\_commander\\_v1.png](https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/File:Cmb_inpaint_T_commander_v1.png)

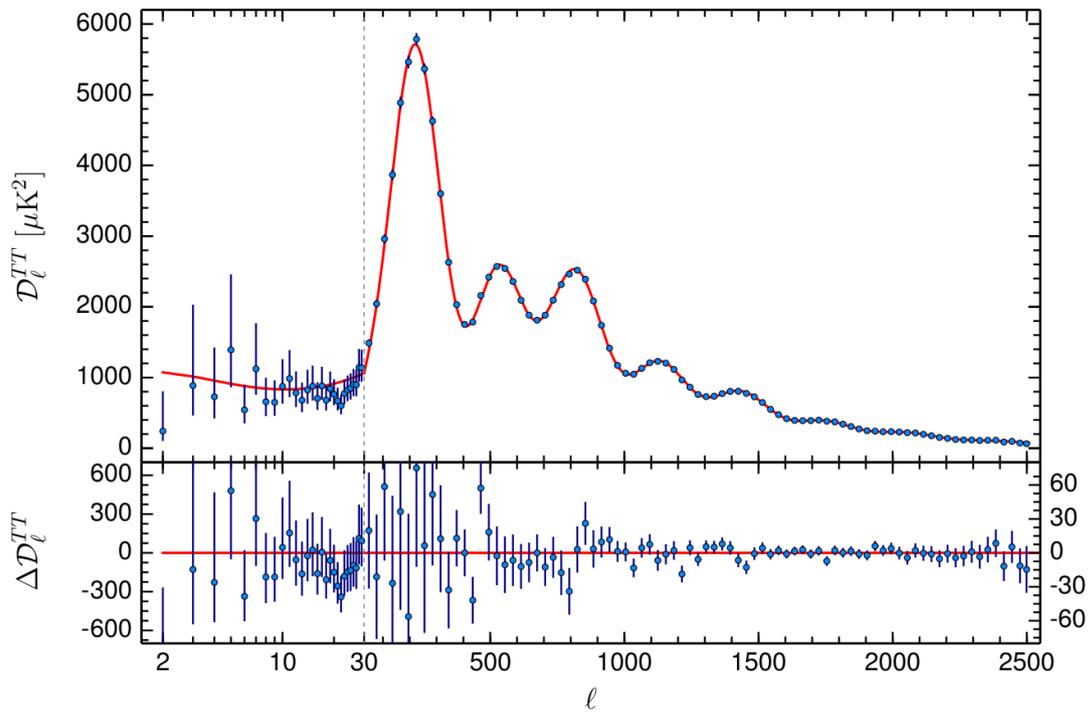


FIGURE 1.2: Évolution des variances du CMB  $D_l = l(l+1)C_l/2\pi$ , mesurées par le satellite Planck, en fonction du moment multipolaire  $l$ . Source : Planck Collaboration et al. (2016)

donc des informations sur les fluctuations primordiales qui ont été créées à cette époque. On remarque que la précision des mesures est plus faible pour les petits  $l$ , ce phénomène est dû à la variance cosmique qui augmente à grande échelle<sup>2</sup>. La deuxième région,  $100 < l < 1000$ , est marquée par une série d'oscillations, les oscillations acoustiques dont on parlera à la sous-section 1.4.2. Les pics sont créés par une série de compression-décompression dans le plasma primordial et leur position angulaire dépend donc de la composition énergétique de ce plasma. La troisième région,  $l > 1000$ , décrit l'univers à petite échelle. On remarque la progressive suppression des anisotropies à grand  $l$  qui est causée par l'augmentation du libre parcours moyen des photons.

### 1.1.3 La mesure des paramètres de $\Lambda$ CDM

La mesure du spectre des fluctuations de températures du CMB permet d'ajuster les six paramètres indépendants du modèle  $\Lambda$ CDM :

- **L'indice spectral des perturbations primordiales scalaires**  $n_s$ , qui est défini par :  $P_s(k) = A_s k^{n_s-1}$ , où  $P_s$  désigne le spectre de puissance des fluctuations primordiales scalaire<sup>3</sup>. Puisque le paramètre  $n_s$  est mesuré au moment de l'inflation, il est contraint par la région  $l < 100$  de la figure 1.2.
- **L'amplitude  $A_s$  des perturbations primordiales scalaires**, qui est également contrainte par la région  $l < 100$ .
- **L'angle apparent de la position du premier pic acoustique**  $\theta_s$ , qui mesure la position du premier pic acoustique de la figure 1.2.
- **Le paramètre de densité de matière noire**  $\Omega_c h^2$ , qui rend compte de la contribution actuelle de la matière noire froide à l'énergie totale de l'univers.  $\Omega_c$  est défini par l'équation (1.30), et  $h = H_0/100$ , où  $H_0$  est la constante de Hubble aujourd'hui<sup>4</sup>. Les caractéristiques physiques des oscillations acoustiques qui se propagent dans le plasma primordial dépendent de la densité énergétique des baryons et de la matière noire. C'est pourquoi  $\Omega_c h^2$  est contraint par la position angulaire des pics acoustiques du spectre de puissance du CMB.
- **Le paramètre de densité baryonique**  $\Omega_b h^2$ , qui rend compte de la contribution actuelle de la matière baryonique à l'énergie totale de l'univers. De même que  $\Omega_c h^2$ , le paramètre  $\Omega_b h^2$  est contraint par la position angulaire des pics acoustiques du spectre de puissance du CMB ainsi que par le rapport des amplitudes du deuxième et du premier pics. Mais également par le taux d'amortissement à petite échelle. En effet, la présence de baryons diminue le libre parcours moyen des photons et donc l'amortissement à grand  $l$ .

---

<sup>2</sup>La variance cosmique désigne l'incertitude sur les mesures cosmologiques liée au fait qu'on ne peut observer qu'un seul univers.

<sup>3</sup>Le mode se réfère à la nature de la perturbation sur la géométrie de l'univers. Un mode scalaire correspond à une perturbation scalaire. Par opposition, les modes tensoriels désignent des perturbations tensorielles de l'espace-temps, les ondes gravitationnelles, et sont contraints par la polarisation du CMB.

<sup>4</sup>L'indice 0 signifie qu'il s'agit d'une valeur mesurée aujourd'hui. Par exemple,  $t_0$  désigne le temps présent et  $a_0 = a(t_0)$  est la valeur actuelle du facteur d'échelle.

- **L'épaisseur optique de réionisation**  $\tau$ , qui quantifie le nombre moyen de diffusions Thomson subies par un photon émis lors de la recombinaison. Ce paramètre est contraint à grande échelle et donc petit  $l$ .

Dans la suite, on utilisera les paramètres ajustés sur les mesures de la collaboration Planck (Planck Collaboration et al. 2016), qui sont rappelés en annexe A.

## 1.2 L'Univers en équations

Pour simplifier les formules, j'adopte la convention :

$$c = k_B = \hbar = 1. \quad (1.1)$$

### 1.2.1 Les équations de champ d'Einstein

La théorie de la Relativité Générale d'Einstein part du postulat simple qu'une chute libre dans un champ gravitationnel est équivalente à une accélération. Autrement dit, un observateur, enfermé dans une capsule sans fenêtre, n'est pas capable de distinguer si l'accélération qu'il ressent est due à l'attraction par un objet massif, une planète par exemple, ou est induite par les moteurs de la capsule. L'idée d'Einstein consiste à considérer la gravitation non plus comme une force mais comme une propriété de l'espace-temps. Ainsi, la géométrie de l'espace-temps est déterminée par la distribution de masse :

$$G_{\mu\nu} \propto T_{\mu\nu}. \quad (1.2)$$

Dans l'équation 1.2,  $G_{\mu\nu}$  est le tenseur d'Einstein, défini par  $G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ .  $R$  et  $R_{\mu\nu}$  désignent, respectivement, le scalaire et de tenseur de Ricci qui sont entièrement déterminés par la métrique  $g_{\mu\nu}$ , qui décrit la géométrie de l'Univers (sous-section 1.2.2). Dans l'équation 1.2,  $T_{\mu\nu}$  est le tenseur énergie-impulsion. Il décrit la composition énergétique de l'Univers (sous-section 1.2.3). Pour trouver le facteur de proportionnalité, on se place à la limite de champ gravitationnels faibles pour lesquels on doit retrouver les lois de Newton. Ce qui donne :

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \quad (1.3)$$

Quand Einstein construit la théorie de la Relativité Générale, il pense que l'Univers est statique. Or, l'équation précédente n'est pas celle d'un Univers statique. Il introduit alors la *constante cosmologique*, notée  $\Lambda$ , qui serait une constante fondamentale de la gravitation, au même titre que la constante de Newton,  $G$ . Ainsi, l'équation 1.3 devient :

$$G_{\mu\nu} + g_{\mu\nu}\Lambda = 8\pi G T_{\mu\nu}. \quad (1.4)$$

La découverte de l'expansion de l'univers dans les années 30 ruine cette construction et Einstein abandonne la constante cosmologique, la qualifiant (ou pas) de "plus grosse erreur de sa vie" (O'Raifeartaigh & Mitton 2018). Mais suite à la découverte de l'accélération de l'expansion, elle réapparaît dans les années 90, et s'impose à travers le modèle  $\Lambda$ CDM.

Dans les sous sections suivantes, nous étudierons dans un premier temps, la forme de la métrique  $g_{\mu\nu}$  qui permet de décrire un univers homogène, isotrope et en expansion. Puis, on répertoriera les différentes composantes de l'univers pour en déduire la forme du tenseur  $T_{\mu\nu}$ .

### 1.2.2 Décrire l'espace

Pour décrire la géométrie de n'importe quel espace, il faut définir une métrique  $g_{\mu\nu}$  qui permet de calculer les distances entre les points de cet espace. Soient  $P(x^0, x^1, x^2, x^3)$  et  $P'(x^0 + dx^0, x^1 + dx^1, x^2 + dx^2, x^3 + dx^3)$  deux points de l'espace-temps, le carré de la distance entre  $P$  et  $P'$  s'écrit :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (1.5)$$

La métrique  $g_{\mu\nu}$  doit inclure les propriétés d'isotropie et d'homogénéité de l'Univers. Au milieu des années 30, Robertson et Walker ont montré que les seules solutions répondant à cette double exigence étaient de la forme (Robertson 1936, Walker 1937) :

$$ds^2 = dt^2 - \frac{dr^2}{1 - \kappa r^2} - r^2 d\theta^2 - r^2 \sin^2 \theta d\phi^2, \quad (1.6)$$

avec  $(r, \theta, \phi)$  le système de coordonnées sphériques et  $\kappa$ , le *paramètre de courbure*, décrivant la forme de l'Univers (figure 1.3) :

$$\begin{cases} \kappa > 0 & \text{Univers fermé} \\ \kappa = 0 & \text{Univers plat} \\ \kappa < 0 & \text{Univers ouvert} \end{cases} \quad (1.7)$$

Dans le cas  $\kappa > 0$ , l'univers est dit fermé et il est équivalent à la surface d'une sphère. En particulier, la somme des angles d'un triangle y est supérieure à  $180^\circ$ . Dans le cas  $\kappa = 0$ , l'Univers est plat et on retrouve la géométrie euclidienne : la somme des angles d'un triangle y est égale à  $180^\circ$ . Enfin, dans le cas  $\kappa < 0$ , l'Univers est ouvert et ressemble à une "selle de cheval". La somme des angles d'un triangle y est inférieure à  $180^\circ$ . La question de savoir si l'univers est plat, ouvert ou fermé est primordiale. En effet, de là résulte la bonne estimation des angles et des distances cosmologiques (voir section 1.3).

À cause de l'expansion de l'univers, la distance  $r$  va être modifiée au cours du temps. Pour rétablir l'indépendance entre les paramètres du système de coordonnées  $(t, r, \theta, \phi)$ , on fait le changement de variable :

$$r \longrightarrow \frac{r}{a(t)}, \quad (1.8)$$

avec  $a(t)$ , appelé *facteur d'échelle*, qui décrit entièrement l'évolution temporelle des distances créée par l'expansion. Par convention, la valeur actuelle du facteur d'échelle, notée  $a_0$ , est égale à 1. Ainsi, le facteur d'échelle  $a(t)$  est le rapport entre la distance physique mesurée au temps  $t$  et la distance physique actuelle. Par conséquent,  $a(t)$  doit être inférieur à 1 quand  $t$  est négatif et être une fonction croissante. Soit un objet de taille physique actuelle  $d_0$  (figure 1.4). Alors, les tailles  $d(t_1)$  et  $d(t_2)$  aux temps respectifs  $t_1$  et  $t_2$  sont liées par la relation :

$$\frac{d(t_1)}{a(t_1)} = \frac{d(t_2)}{a(t_2)} = d_0. \quad (1.9)$$

Le rapport entre la distance physique et le facteur d'échelle est constant et indépendant de l'expansion.

L'expansion de l'univers étant homogène et isotrope, le facteur d'échelle ne dépend pas des coordonnées spatiales mais seulement du temps. En introduisant  $a$  dans l'équation (1.6), elle

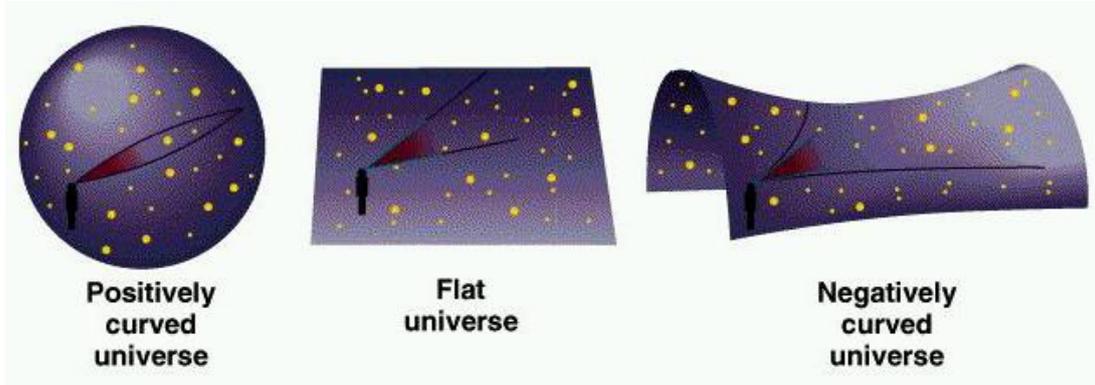


FIGURE 1.3: Le paramètre  $\kappa$  décrit la forme de l'univers. Si  $\kappa = 0$ , alors l'univers est plat et la somme des angles d'un triangle est égale à  $180^\circ$ . Si  $\kappa > 0$ , l'univers est fermé et ressemble à la surface d'une sphère : la somme des angles d'un triangle est supérieure à  $180^\circ$ . Si  $\kappa < 0$ , l'Univers ressemble à une "selle de cheval" : la somme des angles d'un triangle est inférieure à  $180^\circ$ . Source : [http://www.astro.umass.edu/~myun/teaching/a100\\_old/longlecture25.html](http://www.astro.umass.edu/~myun/teaching/a100_old/longlecture25.html)

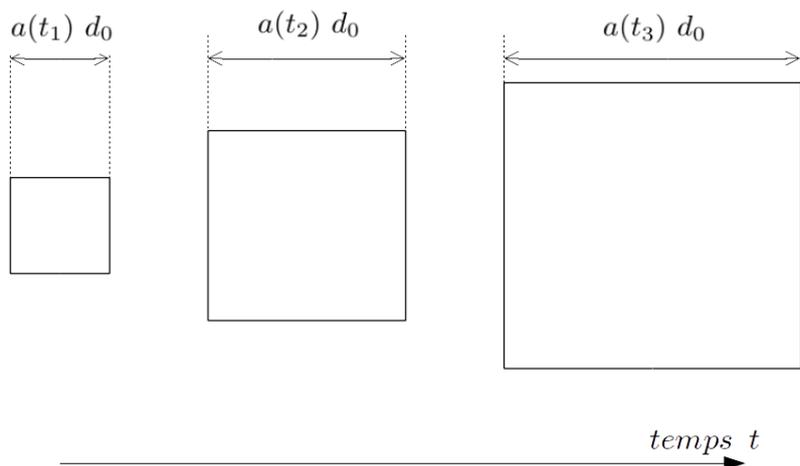


FIGURE 1.4: Le facteur d'échelle  $a(t)$  décrit l'effet de l'expansion au temps  $t$ . Un objet de taille  $d_0$  dans le référentiel comobile, mesure  $a(t_1) d_0$  à  $t_1$ , puis  $a(t_2) d_0$  à  $t_2$  ... Puisque l'univers est en expansion,  $a$  est une fonction croissance du temps.

devient :

$$ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right]. \quad (1.10)$$

Pour simplifier l'équation (1.10), on introduit la *distance comobile*  $\chi$  telle que :

$$ds^2 = c^2 dt^2 - a^2(t) \left[ d\chi^2 + \chi^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right]. \quad (1.11)$$

$\chi$  dépend de la courbure  $\kappa$  de l'espace :

$$\chi = \begin{cases} |\kappa|^{-1/2} \sinh^{-1} \sqrt{\kappa} r & \text{si } \kappa > 0 \\ r & \text{si } \kappa = 0 \\ |\kappa|^{-1/2} \sin^{-1} \sqrt{\kappa} r & \text{si } \kappa < 0 \end{cases} \quad (1.12)$$

La distance comobile correspond à la distance  $d_0$  de la figure 1.4, c'est la distance mesurée dans le référentiel qui évolue avec l'expansion.

Finalement, la métrique  $g_{\mu\nu}$  s'écrit :

$$g_{\mu\nu} = \begin{pmatrix} c^2 & 0 & 0 & 0 \\ 0 & -a^2(t) & 0 & 0 \\ 0 & 0 & -a^2(t)\chi^2 & 0 \\ 0 & 0 & 0 & -a^2(t)\chi^2 \sin^2 \theta \end{pmatrix} \quad (1.13)$$

Cette métrique est appelée *métrique de Friedmann-Lemaître-Robertson-Walker*, FLRW dans la suite du manuscrit (Weinberg 1972).

L'expression de  $G_{\mu\nu}$  de l'équation (1.4) fait intervenir le scalaire  $R$  et le tenseur  $R_{\mu\nu}$  de Ricci de la métrique FRLW. Le calcul est long mais ne présente pas de difficultés particulières. C'est pourquoi je ne l'exposerai pas dans cette thèse. Pour plus de détails, le lecteur intéressé peut se référer au chapitre 4 de Dodelson (2003).

### 1.2.3 La composition énergétique de l'Univers

Les constituants de l'Univers se manifestent par leur densité d'énergie et sont décrits par le tenseur énergie-impulsion,  $T_{\mu\nu}$ , de l'équation 1.4. On peut choisir un référentiel dans lequel  $T_{\mu\nu}$  est diagonal et s'écrit :

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \quad (1.14)$$

La densité d'énergie  $\rho$  et la pression  $p$  dépendent de la nature des composantes du milieu : photons, matière ou neutrinos.

En considérant que l'Univers est en équilibre thermique et dynamique<sup>5</sup>, la première loi de la thermodynamique appliquée à un volume d'univers s'écrit :

$$dU + pdV = 0. \quad (1.15)$$

Dans l'univers en expansion, le volume au temps  $t$  peut s'écrire :  $V(t) = a(t)^3 V_0$ , avec  $V_0$  le volume comobile indépendant de  $t$ . L'énergie interne  $U$  est définie par la somme des énergies des constituants dans le volume  $V$ , soit  $\rho V$ . Finalement, l'équation 1.15 devient :

$$\frac{d(\rho a^3)}{dt} + p \frac{d(a^3)}{dt} = 0, \quad (1.16)$$

Cette équation, appelée *équation de continuité*, peut encore s'écrire sous la forme :

$$\frac{d\rho}{dt} + 3(p + \rho)a^{-1} \frac{da}{dt} = 0. \quad (1.17)$$

### 1.2.3.1 Particules non-relativistes ou matière

Dans ce modèle, le terme "matière" désigne la matière ordinaire et la matière noire. Dans les deux cas, il s'agit de particules non relativistes<sup>6</sup> pour lesquelles la pression  $p_m$  est négligeable devant la densité d'énergie  $\rho_m$  (Kolb & Turner 1990). Dans ce cas, l'équation (1.17) devient :

$$\frac{1}{\rho_m} \frac{d\rho_m}{dt} + 3 \frac{1}{a} \frac{da}{dt} = 0. \quad (1.18)$$

Soit, en intégrant,

$$\rho_m(t) = \rho_{m,0} a^{-3}(t), \quad (1.19)$$

Cette équation traduit le fait que la masse totale de matière dans l'univers est constante, et que, du fait de l'expansion, la matière est diluée dans l'espace au cours du temps.

### 1.2.3.2 Particules relativistes ou rayonnement

Le terme "rayonnement" désigne les particules relativistes, c'est-à-dire qui se déplacent avec une vitesse proche de celle de la lumière. Dans ce modèle cosmologique, il s'agit des photons et des neutrinos. On peut montrer, en utilisant les expressions intégrales de la densité d'énergie et de la pression, que dans ce cas, l'équation d'état s'écrit (Kolb & Turner 1990, Dodelson 2003, chapitre 3) :

$$\rho_r = 3p_r. \quad (1.20)$$

Ainsi, l'équation (1.17) devient :

$$\frac{1}{\rho_r} \frac{d\rho_r}{dt} + 4 \frac{1}{a} \frac{da}{dt} = 0. \quad (1.21)$$

<sup>5</sup>L'univers n'est pas rigoureusement en équilibre thermique et dynamique, sinon il serait figé. Mais cette hypothèse est une bonne approximation. Pour décrire les évolutions thermique et dynamiques, on peut utiliser une approche perturbative du modèle en équilibre (voir chapitre 3 de Dodelson 2003).

<sup>6</sup>Dans d'autres modèles cosmologiques que  $\Lambda$ CDM, la matière noire est chaude et donc ses particules sont relativistes.

L'évolution temporelle de la densité d'énergie de rayonnement est donnée par :

$$\rho_r(t) = \rho_{r,0} a^{-4}(t). \quad (1.22)$$

De même que la matière, les photons et les neutrinos sont dilués par l'expansion de l'univers. Mais à cette dilution s'ajoute une perte de l'énergie propre à chaque particule, qui est à l'origine du phénomène de *redshift* (cf. sous-section 1.3.1).

Dans la première phase de l'Univers, *i.e.* quand  $a$  est très petit, l'énergie de rayonnement est dominante à cause de la forme de  $\rho_r$  (équation 1.22). L'ère de domination du rayonnement, en bleu sur la figure 1.5, s'achève quand la densité d'énergie de rayonnement égale la densité de matière. Elle est suivie par l'ère de domination de la matière, en rouge, puis de l'énergie noire, en noir.

## 1.2.4 Les équations de Friedmann-Lemaître

### 1.2.4.1 Première équation de Friedmann-Lemaître

La composante 00 de l'équation 1.4 donne la première équation de Friedmann-Lemaître :

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{\kappa c^2}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{3} \rho. \quad (1.23)$$

On introduit le *paramètre de Hubble*,  $H(t)$ , qui désigne la variation relative du facteur d'échelle avec le temps :

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} \quad \text{avec} \quad \dot{a}(t) = \frac{da(t)}{dt}. \quad (1.24)$$

L'équation 1.23 s'écrit alors :

$$H(t)^2 + \frac{\kappa}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{3} \rho. \quad (1.25)$$

En utilisant les expressions 1.19 et 1.22,

$$H^2(t) = \frac{8\pi G}{3} \left( \rho_{m,0} a^{-3}(t) + \rho_{r,0} a^{-4}(t) \right) - \frac{\kappa c^2}{a^2(t)} + \frac{\Lambda}{3} \quad (1.26)$$

Par analogie avec la matière et le rayonnement, on peut considérer la constante cosmologique sous la forme d'une nouvelle composante énergétique : l'*énergie noire*, pour laquelle la densité d'énergie  $\rho_\Lambda$  est constante et égale à :

$$\rho_\Lambda = \frac{\Lambda}{2\pi G}. \quad (1.27)$$

On a alors :

$$H^2(t) = \frac{8\pi G}{3} \left( \rho_{m,0} a^{-3}(t) + \rho_{r,0} a^{-4}(t) + \rho_{\Lambda,0} \right) - \kappa c^2 a^{-2}(t). \quad (1.28)$$

Pour simplifier l'équation, on introduit la *densité critique*  $\rho_c$  :

$$\rho_c = \frac{3H_0^2}{8\pi G}. \quad (1.29)$$

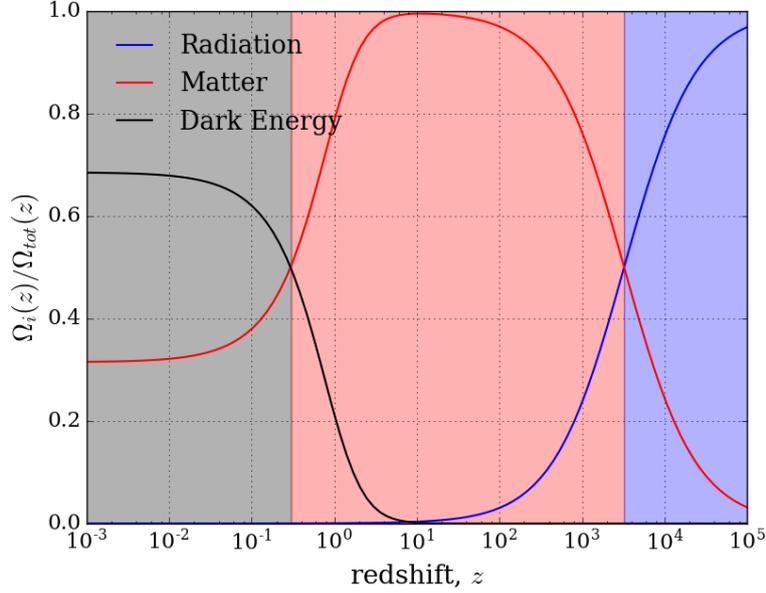


FIGURE 1.5: Mesure des paramètres de densité d'énergie de rayonnement (en bleu), de la matière (en rouge) et de l'énergie noire (en noir) sur la somme totale des paramètres de densité d'énergie. On remarque que l'histoire de l'univers peut se découper en trois grandes phases : l'ère de domination du rayonnement pour  $z \in [3246, 10^5]$ , puis l'ère de domination de la matière pour  $z \in [0.5, 3246]$  suivie de l'ère de domination de l'énergie noire.

La densité critique  $\rho_c$  correspond à la densité énergétique moyenne actuelle d'un univers plat, homogène, isotrope et en expansion.

On définit alors les paramètres de densité de matière  $\Omega_m$ , de rayonnement  $\Omega_r$  et d'énergie noire  $\Omega_\Lambda$  par :

$$\Omega_i = \frac{\rho_{i,0}}{\rho_c} \quad \text{avec} \quad i = m, r \text{ ou } \Lambda. \quad (1.30)$$

Par analogie, on définit aussi le paramètre de densité de courbure,  $\Omega_\kappa$  :

$$\Omega_\kappa = -\frac{3\kappa}{8\pi G} \quad (1.31)$$

Finalement, la première loi de Friedmann-Lemaître donne l'évolution temporelle du paramètre de Hubble en fonction des paramètres de densité d'aujourd'hui :

$$\frac{H^2(t)}{H_0^2} = \Omega_\Lambda + \Omega_\kappa a^{-2}(t) + \Omega_m a^{-3}(t) + \Omega_r a^{-4}(t). \quad (1.32)$$

Je représente sur la figure 1.5, l'évolution en *redshift* des paramètres de densité d'énergie de rayonnement (en bleu), de matière (en rouge) et d'énergie noire (en noir) sur la somme totale des paramètres des densités,  $\Omega_{tot} = \Omega_m + \Omega_r + \Omega_\Lambda = 1$  (pour  $\Omega_\kappa = 0$ ). Pour tracer ce graphique, j'ai utilisé les valeurs de  $\Omega_m$ ,  $\Omega_\Lambda$  et  $\Omega_r$  données en annexe A. On remarque que  $\Omega_{tot}$  est d'abord dominé par le rayonnement, jusqu'au *redshift*  $z_{r/m}$  tel quel que  $\Omega_m(z_{r/m}) = \Omega_r(z_{r/m})$ ,

soit  $z_{r/m} = 3246$ . Puis vient la phase de domination de la matière suivie par la domination de l'énergie noire, dans laquelle nous nous trouvons actuellement, à partir du *redshift*  $z_{m/\Lambda} = 0.5$ . Cette transition marque le passage de l'expansion décélérée à l'expansion accélérée.

### 1.2.4.2 Deuxième équation de Friedmann-Lemaître

La composante 11 (identique aux composantes 22 et 33) de l'équation 1.4 donne la deuxième équation de Friedmann-Lemaître :

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{\kappa c^2}{a^2} - \Lambda c^2 = -8\pi G p. \quad (1.33)$$

En introduisant  $H$  et sa dérivée par rapport au temps  $\dot{H}$ ,

$$2\dot{H}(t) + 3H(t) + \frac{\kappa c^2}{a^2} - \Lambda c^2 = -\frac{8\pi G}{c^2} p. \quad (1.34)$$

Cette équation donne l'évolution du paramètre de Hubble en fonction de la pression des constituants. De même que précédemment, on peut voir la constante cosmologique comme de l'énergie noire dont la pression  $p_\Lambda$  est égale à :  $-\frac{\Lambda}{2\pi G}$ , c'est-à-dire, en utilisant l'équation 1.27, que l'équation d'état de l'énergie noire est donnée par :

$$p_\Lambda = -\rho_\Lambda. \quad (1.35)$$

On peut étendre le modèle  $\Lambda$ CDM au modèle  $w$ CDM dans lequel la densité d'énergie noire n'est plus constante dans le temps. Son équation d'état est donnée par :

$$p_\Lambda = w\rho_\Lambda. \quad (1.36)$$

Le paramètre  $w$  peut être fixe ou varier au cours du temps. Dans ce cas, la densité d'énergie noire est donnée par :

$$\rho_\Lambda(t) = \rho_{\Lambda,0} a^{-(w+1)}(t). \quad (1.37)$$

Quand  $w = -1$ , on retrouve le modèle  $\Lambda$ CDM.

Finalement, la deuxième équation de Friedmann-Lemaître peut s'écrire :

$$2\dot{H}(t) + 3H(t) = -\frac{8\pi G}{c^2} (p_r + p_\Lambda) - \frac{\kappa c^2}{a^2} \quad (1.38)$$

## 1.3 Les distances en cosmologie

Dans un univers plat et statique, la notion de distance est intuitive : c'est à la fois le plus court rayon entre deux points et le chemin emprunté par un photon pour aller d'un point à l'autre. En revanche, dans un espace courbe, ces deux notions sont différentes. Laquelle retenir alors pour caractériser les distances ? Avec l'expansion, les choses se compliquent encore : la mesure d'une distance au temps  $t_1$  sera différente de la mesure à  $t_2 \neq t_1$ . Nous voyons ainsi la nécessité de clarifier la notion de "distance". Dans cette section, nous verrons ainsi les distances les plus couramment utilisées en cosmologie.

Dans cette section, j'abandonne la convention  $c = 1$ .

### 1.3.1 Le *redshift*

On a vu précédemment que l'expansion de l'univers faisait diminuer l'énergie propre des particules relativistes, en particulier des photons. Puisque la fréquence d'un photon est proportionnelle à son énergie, on va voir que l'expansion de l'univers crée un décalage spectral, appelé *redshift*.

Considérons un volume comobile d'univers, noté  $V_0$ , constitué de  $n_\gamma$  photons de fréquence  $\nu(t)$  à l'instant  $t$ . D'après la relation de Planck-Einstein, l'énergie des photons du système au temps  $t$  est donnée par :

$$E_\gamma(t) = n_\gamma h \nu(t). \quad (1.39)$$

En utilisant l'équation 1.22, on peut aussi écrire :

$$E_\gamma(t) = \rho_{\gamma,0} V_0 a^{-1}(t). \quad (1.40)$$

Ainsi, le rapport des fréquences aux temps  $t_1$  et  $t_2$  est donné par :

$$\frac{\nu(t_1)}{\nu(t_2)} = \frac{a(t_2)}{a(t_1)}. \quad (1.41)$$

Un photon émis à la longueur d'onde  $\lambda_e$  et au temps  $t_e$  est observé aujourd'hui à la longueur d'onde  $\lambda_{obs}$  définie par :

$$\lambda_{obs} = a(t_e) \lambda_e = (1 + z(t_e)) \lambda_e. \quad (1.42)$$

Le décalage spectral  $z$  est appelé *redshift* et défini par :

$$1 + z(t) \equiv \frac{1}{a(t)}. \quad (1.43)$$

Le *redshift* permet de quantifier les distances par rapport à l'observateur d'aujourd'hui : plus une source est éloignée, plus son *redshift* est grand. Le *redshift* a l'avantage d'être facilement mesurable sur les spectres et ne nécessite aucune hypothèse sur le modèle d'univers, à l'inverse des distances que l'on va décrire plus loin.

### 1.3.2 La distance comobile

La distance comobile est la distance mesurée à n'importe quelle époque dans le référentiel qui évolue avec l'expansion, elle a été introduite à la section 1.2.2.

Soit une galaxie au *redshift*  $z$  qui émet un photon en  $(t_e, \chi_e, \theta_e, \phi_e)$  et qui est observée aujourd'hui à  $(t_0, \chi = 0)$ . Les photons se propagent à  $(\theta, \phi)$  constants et vérifient  $ds^2 = 0$ . D'après l'équation 1.11, on a alors :

$$\frac{cdt}{a(t)} - d\chi = 0, \quad (1.44)$$

et la distance comobile  $\chi_e$  s'écrit :

$$\chi_e = \int_{t_e}^{t_0} \frac{cdt}{a(t)}. \quad (1.45)$$

Puisqu'on n'a pas directement accès à  $t_e$  lors de la mesure, on préfère utiliser le *redshift*. En utilisant l'équation 1.43, la distance comobile au *redshift*  $z$  est définie par :

$$\chi(z) \equiv \int_0^z \frac{cdz'}{H(z')}. \quad (1.46)$$

Le calcul de la distance comobile nécessite de connaître l'expression de  $H(z)$  et donc de présupposer un modèle d'univers.

### 1.3.3 La distance angulaire

Dans un univers statique, l'angle d'observation  $\Delta\theta$  d'un objet au *redshift*  $z$ , placé perpendiculaire à la ligne de visée et de taille physique  $l$ , est donné par (équation 1.10) :

$$\Delta\theta = \frac{l}{r}, \quad (1.47)$$

où  $r$  est le rayon entre l'observateur et l'objet au *redshift*  $z$ . En inversant les relations de l'équation 1.12, l'équation 1.47 peut s'écrire en fonction de la distance comobile  $\chi$  :

$$\Delta\theta = \frac{l}{S_\kappa(\chi)}, \quad (1.48)$$

avec  $S_\kappa$  qui permet de passer de  $\chi$  à  $r$ , et défini par :

$$r = S_\kappa(\chi) = \begin{cases} |\kappa|^{-1/2} \sinh \sqrt{\kappa} \chi & \text{si } \kappa > 0 \\ \chi & \text{si } \kappa = 0 \\ |\kappa|^{-1/2} \sin \sqrt{\kappa} \chi & \text{si } \kappa < 0 \end{cases}. \quad (1.49)$$

À cause de l'expansion de l'univers, la longueur physique  $l$  mesurée au *redshift*  $z = 0$  est transformée en  $l/(1+z)$  au *redshift*  $z$ . D'où,

$$\Delta\theta = \frac{l}{(1+z)S_\kappa(\chi)}. \quad (1.50)$$

On définit la *distance angulaire*,  $D_A$ , comme le rapport entre taille physique de l'objet  $l/(1+z)$  et l'angle de séparation  $\Delta\theta$ . Soit,

$$D_A(z) \equiv \frac{l}{(1+z)\Delta\theta}, \quad (1.51)$$

qui peut encore s'écrire :

$$D_A(z) = \frac{S_\kappa(\chi)}{1+z}. \quad (1.52)$$

Ainsi, en mesurant  $\Delta\theta$  et  $z$ , on peut calculer la distance angulaire avec l'équation 1.51. Puis en déduire des contraintes cosmologiques sur la courbure de l'Univers et l'intégrale du taux d'expansion via l'équation 1.52.

On peut également définir la *distance angulaire comobile*,  $D_M$ , qui sera utile dans les chapitres suivants :

$$D_M(z) \equiv (1+z)D_A(z) = \frac{l}{\Delta\theta} \quad (1.53)$$

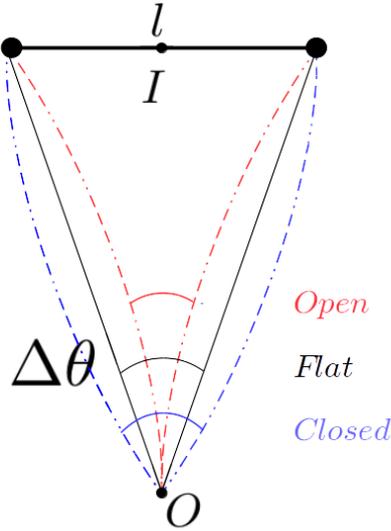


FIGURE 1.6: L'angle d'observation  $\Delta\theta$  est différent selon que l'univers est plat (en noir), ouvert (en rouge) ou fermé (en bleu). Soit un objet, de taille physique  $l$ , placé perpendiculairement à la ligne de visée. Connaissant  $l$ , la mesure de l'angle  $\Delta\theta$  donne une mesure indirecte du taux d'expansion au *redshift* de la mesure.

En connaissant  $l$  et en mesurant  $\Delta\theta$ , on peut déduire une valeur intégrée de la distance de Hubble et contraindre ainsi, de façon indirecte, le taux d'expansion au *redshift* moyen de la mesure.

### 1.3.4 La distance de Hubble

Considérons maintenant que l'objet de taille physique  $l$  est placé le long de la ligne de visée à grande distance de l'observateur<sup>7</sup> (figure 1.7). La taille physique  $l$ , mesurée au *redshift* moyen  $\bar{z} = \frac{z_2+z_1}{2}$ , est égale à la différence des longueurs comobiles :

$$l = \chi(z_2) - \chi(z_1) = \int_{z_1}^{z_2} \frac{cdz}{H(z)}. \quad (1.54)$$



FIGURE 1.7: Soit un objet, de taille physique  $l$ , placé à grande distance de l'observateur et le long de la ligne de visée. Connaissant  $l$ , la mesure du décalage spectral  $\delta z = z_1 - z_2$  donne une mesure directe du taux d'expansion en  $\bar{z} = \frac{z_1+z_2}{2}$ .

<sup>7</sup>La taille physique  $l$  de l'objet doit être négligeable devant la séparation entre l'observateur et l'objet.

Dans le cas où l'objet est très éloigné, *i.e.*  $l \ll \chi(z_1)$ , on peut négliger la variation du paramètre de Hubble entre  $z_1$  et  $z_2$  et écrire :  $H(z_1) \simeq H(z_2) \simeq H(\bar{z})$  avec  $\bar{z} = \frac{z_1+z_2}{2}$ . Finalement,

$$l = \frac{c}{H(\bar{z})} \Delta z \quad \text{avec} \quad \Delta z = z_2 - z_1. \quad (1.55)$$

On définit la *distance de Hubble*,  $D_H$  :

$$D_H(\bar{z}) \equiv \frac{l}{\Delta z}. \quad (1.56)$$

En connaissant  $l$  et en mesurant  $\Delta z$ , on peut en déduire la distance de Hubble et contraindre directement le taux d'expansion au *redshift* moyen de la mesure. Dans la sous-section 1.4.2, on utilisera les distances  $D_H$  et  $D_A$  pour mesurer l'échelle standard des oscillations acoustiques baryoniques.

### 1.3.5 La distance de luminosité

Considérons une source lumineuse au *redshift*  $z$ , qui émet  $N$  photons de fréquence  $\nu$  pendant l'intervalle de temps  $\Delta t$ . Le flux lumineux  $\phi$  reçu par l'observateur est égal au rapport de la luminosité reçue,  $L_r$ , sur la surface de la sphère centrée sur la source et passant par l'observateur. Soit,

$$\phi = \frac{L_r}{4\pi D_L^2}, \quad (1.57)$$

avec  $D_L$  la distance de séparation entre l'observateur et la source. La luminosité reçue est égale à variation temporelle d'énergie lumineuse reçue. D'après la loi de Planck-Einstein,

$$L_r = \nu_r \frac{dn}{dt}, \quad (1.58)$$

où  $\nu_r$  est la fréquence des photons reçus et  $n$  le nombre de photons reçus. Finalement, le flux lumineux s'écrit :

$$\phi = \frac{\nu_r}{4\pi D_L^2} \frac{dn}{dt}. \quad (1.59)$$

On va calculer le flux lumineux reçu dans le cas d'un univers plat et statique, puis courbe et en expansion.

Dans le cas de l'univers **statique**, la luminosité reçue est égale à la luminosité émise et le flux lumineux s'écrit :

$$\phi_s = \frac{\nu N}{4\pi D_L^2}. \quad (1.60)$$

Si l'univers est **plat** et **statique**, il n'y a pas d'ambiguïté sur la notion de distance et  $D_L$  désigne indifféremment  $r$  (équation 1.10) et  $\chi$  (équation 1.10).

Dans le cas d'un univers **en expansion**, il y a une perte de l'énergie propre des photons pendant leur déplacement. D'après l'équation 1.43, la fréquence  $\nu_r$  des photons reçus est  $(1+z)^{-1}\nu$ . À cause de l'expansion, les intervalles de temps varient également entre la source et l'observateur. Ainsi, si  $N\Delta t$  photons sont émis par la source pendant  $\Delta t$ , ils sont reçus par

l'observateur pendant  $\Delta t' = (1+z)\Delta t$  :

$$\frac{dn}{dt} = \frac{N\Delta t}{\Delta t'} = \frac{N}{1+z}. \quad (1.61)$$

Si l'univers est **courbe**, le rayon de la sphère qui relie l'observateur à la source est donné par la composante  $r$  des coordonnées sphériques, qui peut s'écrire également  $S_\kappa(\chi)$  (équation 1.49). Finalement, le flux lumineux d'un univers courbe en expansion est donné par :

$$\phi_e = \frac{\nu N}{4\pi(1+z)^2 S_\kappa(\chi)}. \quad (1.62)$$

Par analogie avec l'équation (1.60) on définit la distance de luminosité  $D_L$  dans un espace courbe et en expansion :

$$D_L = (1+z)S_\kappa(\chi). \quad (1.63)$$

La distance de luminosité est surtout utilisée dans les études de chandelles standard, dont on parlera à la sous-section 1.4.1.

## 1.4 Mesurer l'expansion de l'Univers

À la sous-section 1.1.2, j'ai décrit le CMB, qui donne une image très nette de l'état de l'univers au moment de la recombinaison ( $z \simeq 1100$ ). Pour détecter l'accélération et caractériser l'expansion de l'univers, il faut la mesurer à plus bas *redshift*. Dans cette section, je décris deux sondes cosmologiques : les supernovae de type Ia et les oscillations acoustiques baryoniques. Pour une description d'autres sondes (*e.g.*, les lentilles gravitationnelles, les amas de galaxies), voir par exemple, [Huterer & Shafer \(2018\)](#).

### 1.4.1 Les supernovae de type Ia

Une chandelle standard est un objet astrophysique idéalisé dont on suppose que la luminosité intrinsèque est reproductible. Ainsi, en mesurant la luminosité apparente de ces objets, on peut en déduire leur distance de luminosité  $D_L$  (sous-section 1.3.5). En pratique, on préfère utiliser une échelle logarithmique et considérer les magnitudes plutôt que les luminosités. La différence entre la magnitude apparente  $m$ , mesurée par l'observateur, et la magnitude absolue  $M$  (la magnitude apparente qu'elle aurait à une distance de 10 pc de la source) est appelée module de distance, noté  $\mu$ , et dépend des paramètres cosmologiques à travers  $D_L$  :

$$\mu \equiv m - M = 5 \log_{10} \left( \frac{D_L(\Omega_m, \Omega_\Lambda, H_0)}{10 \text{ pc}} \right). \quad (1.64)$$

Parmi les chandelles standard, on trouve les supernovae de type Ia (SNIa dans la suite du manuscrit). Une SNIa est créée par l'explosion thermonucléaire d'une naine blanche, qui a accréte suffisamment de matière d'une étoile compagnon (*e.g.*, une géante rouge) pour approcher la masse limite de Chandrasekhar. Le spectre d'une SNIa est identifiable, notamment parce qu'il contient des absorptions caractéristiques de Si II ([Astier 2012](#)).

Les SNIa permettent de contraindre le modèle  $\Lambda$ CDM lors de la période de domination de l'énergie noire et ont été la première sonde utilisée pour montrer l'accélération de l'expansion de l'univers ([Riess et al. 1998](#), [Perlmutter et al. 1999](#)). J'ai réalisé ma thèse au LPNHE où le

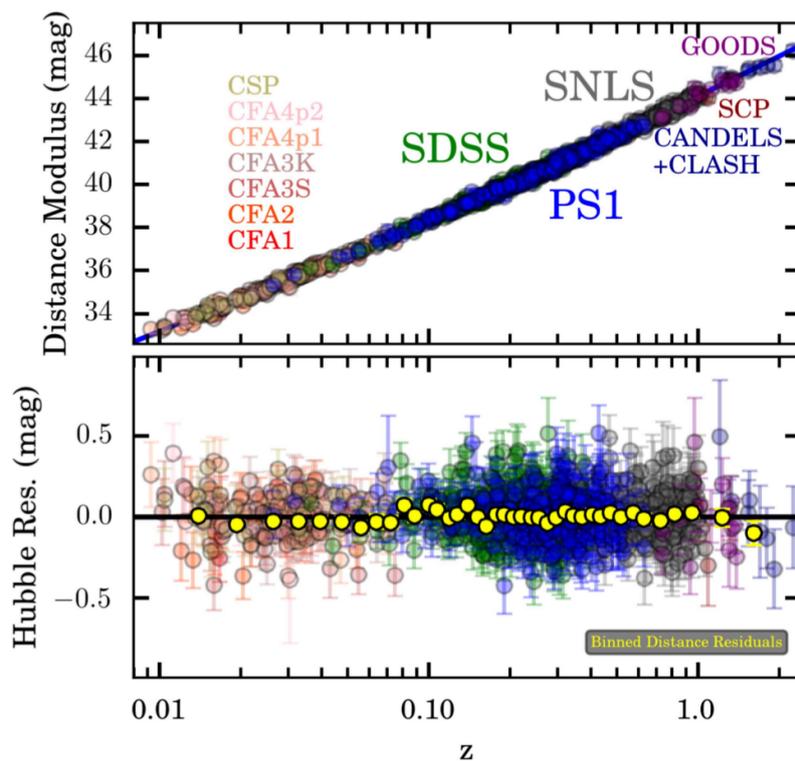


FIGURE 1.8: Modules de distance  $m - M$  des SNIa de l'échantillon *Pantheon*, en fonction de leur *redshift*. Source : Scolnic et al. (2018)

groupe de cosmologie est très impliqué dans la collaboration SNLS (SuperNova Legacy Survey) qui a détecté et mesuré, pendant 5 ans, plusieurs centaines de SNIa entre les *redshifts* 0.2 et 1.0 (Astier et al. 2006, Sullivan et al. 2011). Pour améliorer la calibration, Betoule et al. (2014) ont utilisé les données provenant d'expériences différentes (principalement SNLS et SDSS II) et réalisé la première inter-calibration entre ces jeux de données, ce qui a permis de réduire les erreurs systématiques liées à la calibration. Scolnic et al. (2018) proposent une compilation des lots de données SNIa actuels en inter-calibrant 1048 SNIa pour  $z \in [0.01, 2.3]$ . Ils obtiennent la meilleure contrainte, en utilisant des SNIa, sur la mesure des paramètres  $\Omega_m$  et  $\Omega_\Lambda$ .

La figure 1.8, appelée *diagramme de Hubble*, présente les mesures des modules de distance des SNIa de l'échantillon *Pantheon*, présenté par Scolnic et al. (2018), en fonction de leur *redshift*. Les paramètres cosmologiques  $\Omega_m$  et  $\Omega_\Lambda$  sont ajustés directement sur les données via l'équation 1.64. Dans le cas du modèle  $\Lambda$ CDM ouvert ( $\Omega_\kappa \neq 0$ ), le meilleur ajustement donne :

$$\Omega_m = 0.319 \pm 0.070 \quad \Omega_\Lambda = 0.733 \pm 0.113 \quad (1.65)$$

Ces valeurs sont parfaitement cohérentes avec les contraintes obtenues avec le CMB (voir annexe A). Dans le chapitre 7, on comparera ce résultat à celui donné par la sonde BAO, décrite dans la sous-section suivante.

La sonde SNIa est limitée par ses incertitudes systématiques, qui proviennent, la plupart du temps, de la calibration photométrique entre SNIa observées par des instruments différents.

Cette difficulté devrait être supprimée par les grands relevés photométriques à venir comme LSST et WFIRST (Huterer & Shafer 2018).

## 1.4.2 Les oscillations acoustiques baryoniques

### 1.4.2.1 La formation des oscillations acoustiques baryoniques

L'inflation cosmique, qui a lieu dans les premiers instants après le Big Bang, génère des fluctuations de la densité d'énergie de matière noire, de matière baryonique et des photons. Ces fluctuations sont à l'origine des structures à grande échelle visibles aujourd'hui et des oscillations acoustiques baryoniques (Eisenstein et al. 2007).

Sur la figure 1.9, tirée de Eisenstein et al. (2007), on suit l'évolution au cours du temps, dans le référentiel comobile, d'une perturbation de la densité d'énergie de matière noire (en noir), de matière baryonique (en bleu) et de photons (en rouge). Initialement (non représenté), les profils de masse des trois composantes sont ponctuelles et localisés près de l'origine. Ils sont identiques pour les trois composantes. À cause de la haute température du milieu, les photons et les baryons sont couplés par interaction Thomson et forment un plasma. Le plasma baryons-photons est soumis à deux forces contraires :

- La pression de radiation des photons, qui "pousse" la perturbation de la densité de plasma vers l'extérieur,
- L'interaction entre la matière noire et la matière ordinaire, qui "retient" la perturbation de la densité de plasma au centre.

La compétition entre gravitation et pression de radiation donne naissance à la propagation d'une onde acoustique de plasma (au centre). Ce phénomène est appelé *Oscillation Acoustique Baryonique* (*Baryonic Acoustic Oscillations* en anglais, noté BAO dans la suite du manuscrit). La vitesse du front d'onde, qui peut être calculée par une approche perturbative des équations de champ d'Einstein et des équations de Boltzmann (Dodelson 2003, chapitre 8), est donnée par :

$$c_s = \frac{c}{\sqrt{3\left(1 + \frac{3\rho_b}{4\rho_\gamma}\right)}}. \quad (1.66)$$

La vitesse acoustique  $c_s$  est ainsi entièrement déterminée par la densité d'énergie de matière baryonique  $\rho_b$  et la densité d'énergie des photons  $\rho_\gamma$ .

À la recombinaison, quand le milieu a suffisamment refroidi pour permettre la formation des premiers atomes d'hydrogène, les photons "s'échappent" du plasma. Les profils de masse de matière baryonique et des photons se découplent (figure 1.9, au milieu). Privées de la pression de radiation, les BAO sont figées dans le référentiel comobile. La position  $r_d$  de la perturbation de densité d'énergie de la matière ordinaire correspond au chemin parcouru par les BAO jusqu'à

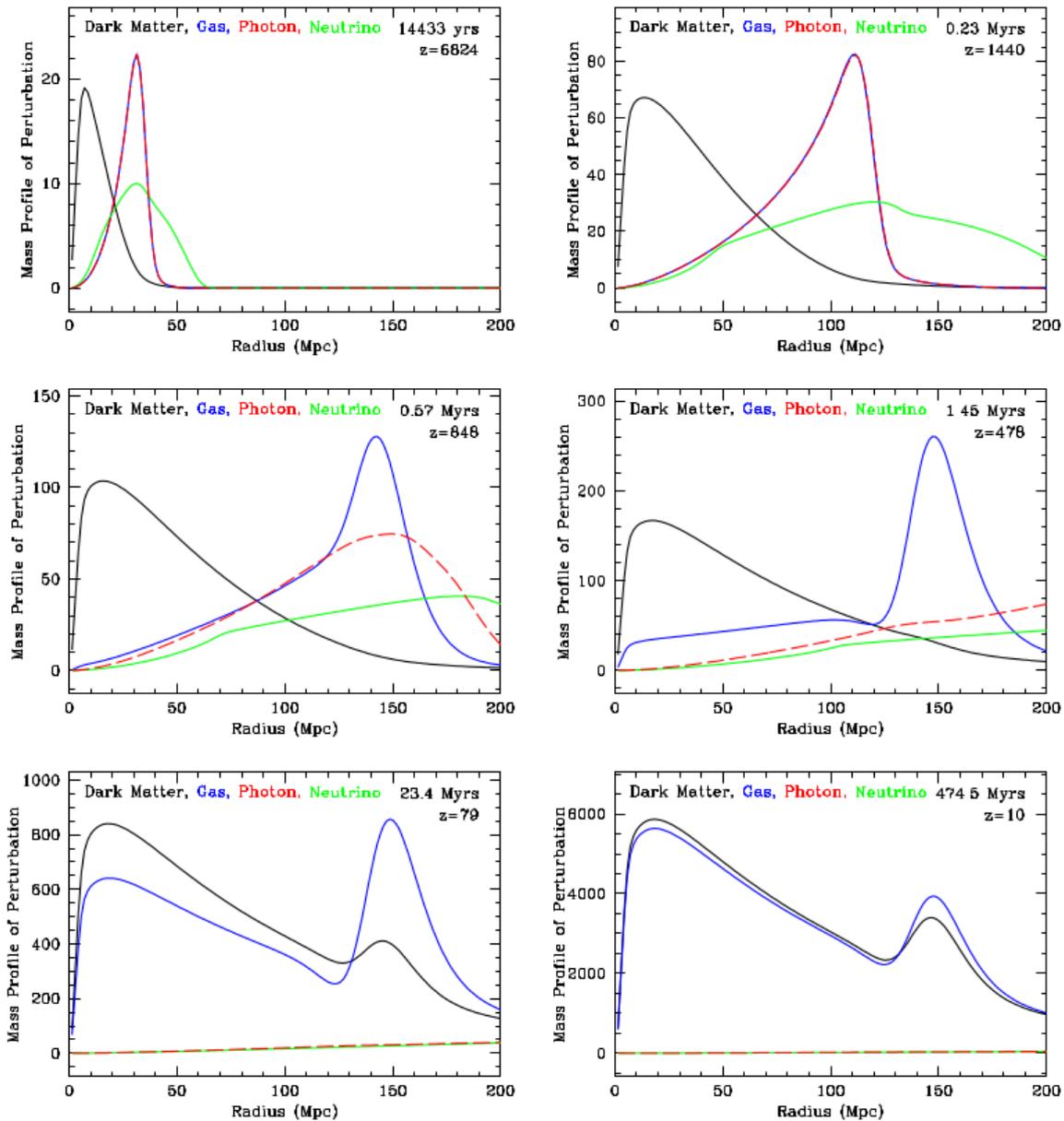


FIGURE 1.9: Évolution du profil de masse d'une sur-densité d'énergie constituée de matière noire (en noir), de matière ordinaire (en bleu), de photons (en rouge) et de neutrinos (en vert), en fonction de la distance comobile. Initialement (non représenté), toutes les espèces ont un profil de masse identique qui forme un pic proche de l'origine. Puis, à cause de la pression de radiation, le pic de sur-densité du plasma baryons-photons quitte le centre de la perturbation (en haut, à gauche). Le phénomène se poursuit et, à cause de l'interaction gravitationnelle entre la matière noire et la matière ordinaire, les pics des sur-densités de matière noire et de plasma sont élargis (en haut, à droite). Au moment de la recombinaison, la matière ordinaire et les photons sont découplés. Les photons se déplacent alors librement dans le milieu, tandis que la matière noire et la matière ordinaire continuent d'interagir (au milieu). Finalement, les photons ont complètement diffusé et les profils de masse de matière noire et de matière ordinaire sont très semblables (en bas). On remarque que la position comobile du pic BAO (à 150 Mpc) n'a pas évolué depuis la recombinaison, ce qui veut dire que cette position peut être mesurée à différentes périodes pour retrouver les taux d'expansion à ces mêmes périodes. Source : Eisenstein et al. (2007)

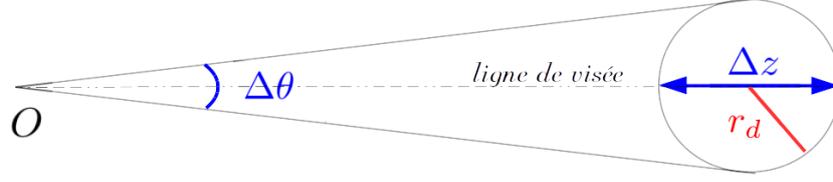


FIGURE 1.10: Au *redshift* moyen  $\bar{z}$ , la taille caractéristique des BAO le long de la ligne de visée est donnée par la distance de Hubble  $D_H(\bar{z})$ . Perpendiculairement à la ligne de visée, elle est donnée par la distance angulaire comobile  $D_M(\bar{z})$ . En mesurant  $(\Delta z, \Delta\theta)$  et en connaissant  $r_d$ , on peut accéder aux valeurs de  $(D_H(\bar{z}), D_M(\bar{z}))$  puis contraindre les paramètres cosmologiques.

la recombinaison<sup>8</sup>. Soit,

$$r_d = \int_{z_d}^{\infty} \frac{c_s(z) dz}{H(z)}, \quad (1.67)$$

avec  $z_d$ , le *redshift* au moment de la recombinaison. Le pic localisé à  $r_d$  dans le profil de masse de la matière baryonique est appelé *pic BAO*. Les BAO apparaissent sous la forme d'oscillations (d'où leur nom) dans le spectre de puissance du CMB (voir figure 1.2).

Au cours du temps, l'interaction entre la matière noire et la matière baryonique modifie les profils de masse des deux espèces. La matière noire attire la matière baryonique au centre, et la matière baryonique attire la matière noire au niveau du pic BAO (figure 1.9, en bas). À cause de la quantité supérieure de matière noire, le pic BAO est atténué à bas *redshift* mais sa position reste inchangée, dans le référentiel comobile, depuis la recombinaison.

On verra au chapitre suivant que le profil de masse de la perturbation de matière est donné par la fonction de corrélation à deux points de la matière. La première détection des BAO a été réalisée en utilisant des galaxies comme traceur de masse (Cole et al. 2005, Eisenstein et al. 2005).

#### 1.4.2.2 La mesure de l'échelle acoustique

La position du pic BAO est appelée *échelle acoustique*. Elle est fixe dans le référentiel comobile. Ce qui signifie que si on connaît l'échelle acoustique à deux époques quelconques, on peut retrouver le rapport des taux d'expansion entre ces deux époques.

L'échelle acoustique peut être mesurée parallèlement à la ligne de visée (voir figure 1.10). Dans ce cas, la différence de *redshift*  $\Delta z$  de part et d'autre de la perturbation s'écrit (sous-section 1.3.4) :

$$\Delta z = \frac{r_d}{D_H(z)}, \quad (1.68)$$

où  $z$  est le *redshift* moyen de la perturbation.

<sup>8</sup>En toute rigueur  $r_d$  désigne le chemin parcourru jusqu'au moment, appelé *drag epoch*, où les photons cessent de sentir les effets dynamiques des baryons, qui a lieu environ 20 000 ans après la recombinaison. L'horizon à la recombinaison, noté  $r^*$ , est mesuré dans le CMB. La différence entre les deux distances est inférieure à 3 Mpc (soit 2%).

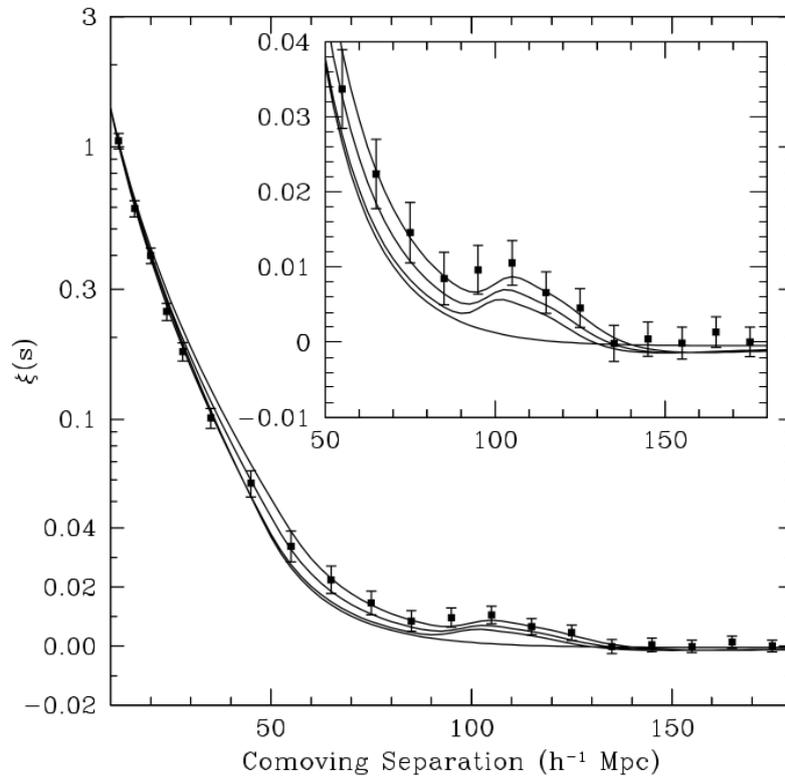


FIGURE 1.11: Première détection (à  $3.4\sigma$ ) du pic BAO dans la fonction de corrélation à deux points des galaxies, au *redshift* moyen 0.35. On remarque que l'amplitude du pic BAO est faible, c'est pourquoi il faut un grand nombre de données pour le détecter. Source : [Eisenstein et al. \(2005\)](#)

Perpendiculairement à la ligne de visée, la mesure angulaire  $\Delta\theta$  de la perturbation s'écrit :

$$\Delta\theta = \frac{r_d}{D_M(z)}. \quad (1.69)$$

Ainsi, la mesure de l'échelle acoustique dans les deux directions donne accès à deux mesures indépendantes de  $D_H$  et  $D_A$  au *redshift*  $z$  considéré, qui sont, respectivement, une mesure directe et une mesure intégrée du taux d'expansion à  $z$ . D'où l'on peut déduire des contraintes sur les paramètres cosmologiques  $\Omega_\Lambda$  et  $\Omega_m$ .

On verra dans le chapitre 2 que la longueur caractéristique (ou échelle acoustique) du BAO est imprimée sous la forme d'un pic dans les fonctions de corrélation à deux points des traceurs de la matière. La première détection (à  $3.4\sigma$ ) du pic BAO a été réalisée dans la fonction de corrélation des galaxies au *redshift* 0.35 (voir figure 1.11). On remarque que l'amplitude du pic BAO a diminué entre les figures 1.9 et 1.11. On verra à la section 6.1.1 que ce phénomène est dû à la formation des structures qui lisse progressivement le pic au cours du temps. Ainsi, la détection du pic BAO à bas *redshift* nécessite un volume de données très important.

Au chapitre 7, je présenterai les mesures les plus récentes des distances  $D_H(z)$  et  $D_A(z)$  de l'horizon acoustique à différents *redshifts*. Je déduirai de l'ensemble de ces données, des contraintes sur  $\Omega_\Lambda$  et  $\Omega_m$ .



## Chapitre 2

# Fonction de corrélation de la matière et forêts Ly $\alpha$

### 2.1 La fonction de corrélation à deux points de la matière

On a vu au chapitre précédent que l'échelle caractéristique des BAO était toujours imprimée dans la distribution de matière de l'univers, et que la mesure de cette échelle permettait de contraindre les paramètres cosmologiques  $\Omega_m$  et  $\Omega_\Lambda$ . Dans ce chapitre, on étudiera dans un premier temps la fonction de corrélation à deux points de la matière, qui permet de décrire simplement la répartition de celle-ci dans l'univers et ainsi de mesurer l'échelle acoustique. Puis dans un deuxième temps, on se concentrera sur un traceur particulier de la matière : les absorptions Ly $\alpha$  visibles sur les spectres des quasars à haut *redshift*. Chaque absorption correspond à la présence d'un nuage absorbant sur la ligne de visée du quasar. Ainsi, en calculant la fonction de corrélation à deux points des absorptions Ly $\alpha$ , on décrit la répartition des nuages d'hydrogène neutre, dans laquelle est inscrite la longueur caractéristique des BAO.

#### 2.1.1 Définition

Dans un univers où la matière est distribuée aléatoirement dans l'espace avec la densité uniforme  $\rho_0$ , la probabilité de trouver un objet 2 dans un volume  $dV_2$  à la distance  $\vec{r}$  d'un objet 1, sélectionné au hasard, dans un volume  $dV_1$  est donnée par :

$$dN = \rho_0^2 dV_1 dV_2. \quad (2.1)$$

Comme on l'a vu dans le chapitre précédant, la matière n'est pas disposée parfaitement au hasard. Pour décrire l'écart par rapport à la distribution moyenne uniforme, on introduit la *fonction de corrélation à deux points* de la matière. Notée  $\xi(\vec{r})$ , elle quantifie l'excès de probabilité (par rapport à la distribution moyenne uniforme) de trouver un objet 2 à une distance  $\vec{r}$  d'un objet 1, à la position  $\vec{x}$  sélectionnée au hasard. C'est-à-dire que, pour l'univers réel, l'équation 2.1 peut s'écrire (Peebles 1980) :

$$d^2N(\vec{r}) = \rho_0^2 (1 + \xi(\vec{r})) dV_1 dV_2, \quad (2.2)$$

avec  $\rho_0$  la densité moyenne :  $\rho_0 \equiv \langle \rho(\vec{x}) \rangle_{\vec{x}}$ . Le symbole  $\langle \rangle_{\vec{x}}$  désigne une moyenne sur toutes les positions  $\vec{x}$  de l'objet 1, *i.e.* sur tout l'espace.

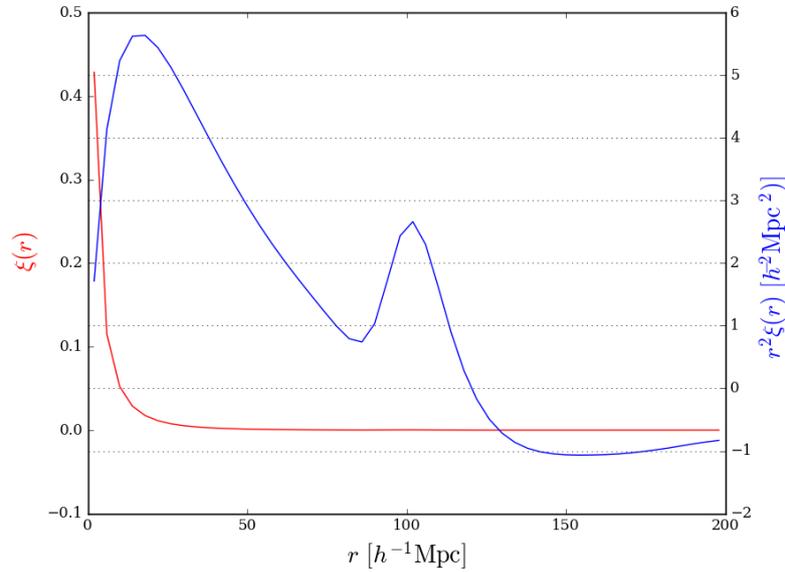


FIGURE 2.1: En rouge, fonction de corrélation à deux points de la matière, à  $z = 2.3$ , calculée avec l'outil CAMB où les paramètres du modèle  $\Lambda$ CDM sont donnés par l'annexe A. Puisque le pic BAO, autour de  $100 h^{-1}\text{Mpc}$ , représente un excès de probabilité de seulement 1%, il n'est pas visible sur  $\xi(r)$ . Pour le faire apparaître, on multiplie  $\xi(r)$  par  $r^2$  (en bleu).

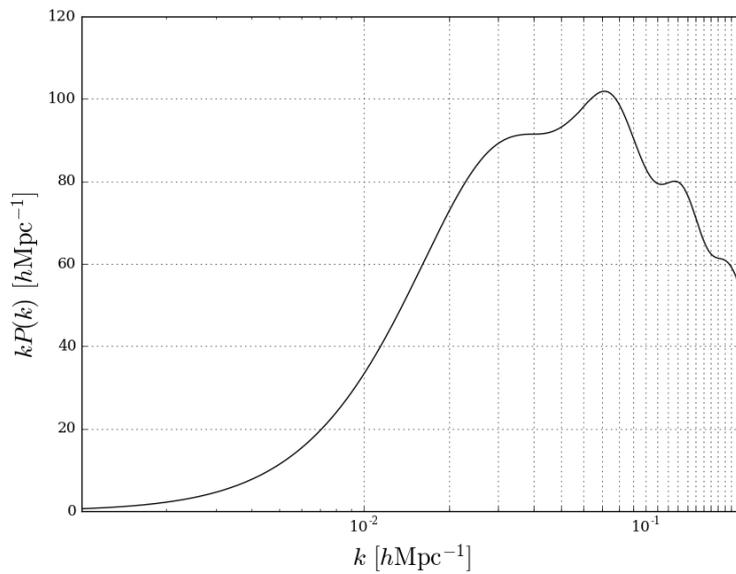


FIGURE 2.2: Spectre de puissance de la matière, à  $z = 2.3$ , calculé avec l'outil CAMB où les paramètres du modèle  $\Lambda$ CDM sont donnés par l'annexe A.

On peut également écrire l'équation 2.1 en notant  $\rho(\vec{x})$  et  $\rho(\vec{x} + \vec{r})$  les densités respectives des objets 1 et 2 :

$$d^2 N(\vec{x}, \vec{r}) = \rho(\vec{x})\rho(\vec{x} + \vec{r})dV_1 dV_2. \quad (2.3)$$

On définit le contraste de densité  $\delta(\vec{x})$  qui mesure l'écart à la position  $\vec{x}$  entre la densité mesurée  $\rho(\vec{x})$  et la densité moyenne  $\rho_0$  :

$$\delta(\vec{x}) \equiv \frac{\rho(\vec{x})}{\rho_0} - 1. \quad (2.4)$$

En remplaçant dans l'équation 2.3 et en moyennant sur l'ensemble des positions  $\vec{x}$ , on obtient :

$$d^2 N(\vec{r}) = \rho_0^2 \left(1 + \langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) \rangle_{\vec{x}}\right) dV_1 dV_2. \quad (2.5)$$

Par identification des équations 2.2 et 2.5, la fonction de corrélation à deux points s'écrit :

$$\xi(\vec{r}) = \langle \delta(\vec{x})\delta(\vec{x} + \vec{r}) \rangle_{\vec{x}}. \quad (2.6)$$

Pour le calcul de la fonction de corrélation de la matière, on utilise le module **CAMB**<sup>1</sup> (Lewis et al. 2000). Ce module calcule la fonction de corrélation à deux points et le spectre de puissance de la matière dans le cadre du modèle  $\Lambda$ CDM, en utilisant notamment une approche perturbative des équations de champ d'Einstein et des équations de Boltzmann. Je représente en rouge sur la figure 2.1, la fonction de corrélation calculée avec les paramètres cosmologiques de l'annexe A. On voit le premier pic de corrélation centré en zéro et dont la largeur, de l'ordre de  $10 h^{-1}\text{Mpc}$ , correspond à la taille moyenne des amas de galaxies. Le pic BAO n'est pas visible sur cette courbe puisqu'il représente un excès de probabilité de seulement 1%. Il apparaît autour de  $100 h^{-1}\text{Mpc}$  quand on multiplie par la distance au carré (en bleu).

On verra plus tard que, pour simplifier les calculs, il peut être utile de passer dans l'espace de Fourier et de transformer les distances  $\vec{x}$  en fréquences  $\vec{k}$ . La transformée de Fourier de la fonction de corrélation est appelée spectre de puissance et définie par :

$$P(\vec{k}) = \frac{1}{(2\pi)^3} \int \xi(\vec{r}) e^{i\vec{k}\vec{r}} d^3 \vec{k}. \quad (2.7)$$

Je représente figure 2.2 le spectre de puissance de la matière à  $z = 2.3$  calculé avec **CAMB**, sur lequel on voit apparaître les BAO sous la forme d'oscillations comme on l'a vu au chapitre 1 dans le spectre de puissance du CMB (voir figure 1.2).

### 2.1.2 Effets de l'expansion et de la formation des structures

À cause de l'expansion de l'univers et de la formation des structures, la répartition de matière dépend du *redshift*. Dans le cadre de l'approximation de Zel'dovitch, c'est-à-dire pour des fluctuations  $\delta$  inférieures ou de l'ordre de 1 (Dodelson 2003, chapitre 7), on peut découpler la position et le *redshift* dans l'expression du contraste de densité :

$$\delta(r, z) = D(z) \times \delta(r, z = 0). \quad (2.8)$$

<sup>1</sup><https://camb.info/>

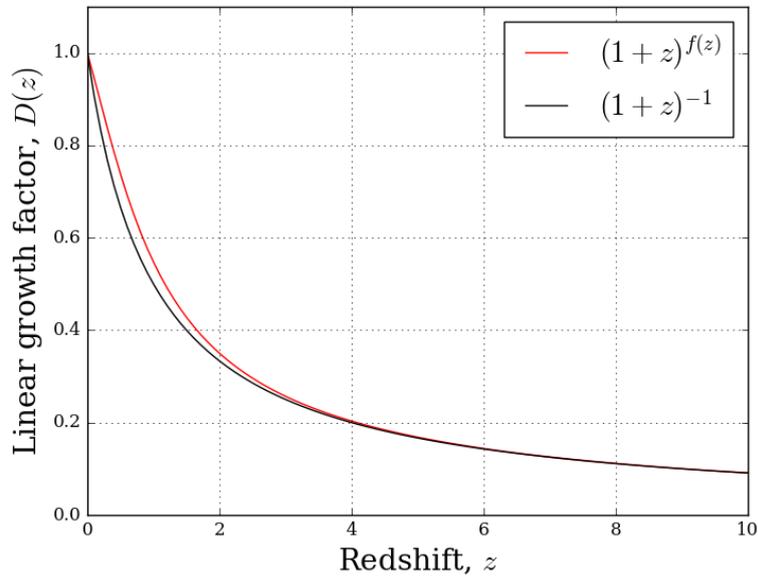


FIGURE 2.3: taux de croissance linéaire  $D(z)$  en fonction du *redshift*  $z$ . En rouge, on représente l'approximation de Lahav et al. (1991) de l'équation 2.11 valable dans les ères de dominations de la matière et de l'énergie noire. En noir, on approxime  $G(z)$  par  $1/(1+z)$ . La différence entre les deux approximations est inférieure à 5% pour  $z > 2$ .

$D$ , appelé taux de croissance linéaire, décrit la formation des structures à grande échelle en régime linéaire.

En remplaçant dans l'équation 2.6, la fonction de corrélation de la matière au *redshift*  $z$  s'écrit alors :

$$\xi(\vec{r}, z) = D(z)^2 \times \xi(\vec{r}, z = 0). \quad (2.9)$$

Usuellement, pour simplifier les calculs, on définit le *taux de croissance logarithmique*,  $f$  :

$$f(z) \equiv -\frac{d \log D(z)}{d \log(1+z)}. \quad (2.10)$$

Pour les ères de domination de la matière et de l'énergie noire (voir figure 1.5), on peut estimer  $f$  par (Lahav et al. 1991) :

$$f(z) = \left( \Omega_M (1+z)^3 \right)^{0.6} + \frac{\Omega_\Lambda}{70} \left( 1 + \frac{\Omega_M}{2} \right). \quad (2.11)$$

D'où on peut déduire l'expression de  $D$  via (2.10). Je représente sur la figure 2.3, le taux de croissance linéaire en fonction du *redshift* (en rouge), que je compare avec l'expression approchée  $1/(1+z)$  (en noir). On voit que l'écart entre les deux expressions est inférieur à 5% quand le *redshift* est supérieur à 2. On utilisera donc dans les chapitres suivants l'approximation  $D(z) \simeq (1+z)^{-1}$  pour simplifier les calculs.

### 2.1.3 Traceur de la matière et biais

Pour étudier la distribution de matière dans l'univers, on considère la distribution des objets appartenant à une même catégorie astrophysique (galaxies, quasars, nuages absorbants, ...). On dit qu'une catégorie est un bon traceur si ses variations de densité suivent celles de la matière. En réalité, un traceur ne trace jamais parfaitement la matière, et la densité de contraste du traceur  $t$  est reliée à la densité de contraste de la matière via le biais  $b_t$  (Kaiser 1984) :

$$\delta_t(\vec{x}) = b_t \times \delta(\vec{x}). \quad (2.12)$$

En pratique, on ne connaît pas le biais de chaque objet observé à la position  $\vec{x}$ , mais on fait l'approximation que  $b_t$  est identique pour tous les objets d'une même classe  $t$ . Puisque degré d'évolution des structures dépend du *redshift*, on peut inclure une dépendance en *redshift* du biais. La forme de  $b_t$  dépend du type de traceur, on étudiera à la section 2.2, le cas particulier des absorptions Ly $\alpha$ .

Ainsi, la fonction de corrélation à deux points mesurée avec le traceur  $t$  s'écrit :

$$\xi_t(\vec{r}) = b_t^2 \times \xi(\vec{r}). \quad (2.13)$$

Le biais du traceur impose l'amplitude de la fonction de corrélation, ce qui signifie que plus  $b_t$  est grand plus il est facile de détecter le pic BAO.

Expérimentalement, on choisit le type de traceur  $t$  à utiliser en fonction de sa densité d'observation au *redshift* auquel on souhaite mesurer la fonction de corrélation (voir sous-section 3.3.1). Puis, on établit un relevé des objets de type  $t$  observés, où chaque objet est repéré par deux coordonnées angulaires et un *redshift* (voir figure 1.10). On va voir dans la sous-section suivante que l'effet des vitesses particulières induit une mauvaise estimation des distances de séparation et distord la fonction de corrélation mesurée.

### 2.1.4 Espace réel et espace des *redshifts*

La différence de fréquence entre un photon émis par un objet et le photon reçu par l'observateur est l'effet combiné de l'expansion de l'univers et de la vitesse particulière de l'objet. Ainsi, le *redshift* apparent  $\tilde{z}$  s'écrit :

$$1 + \tilde{z} = (1 + z)\left(1 + \frac{v_{\parallel}}{c}\right), \quad (2.14)$$

avec  $z$  le *redshift* cosmologique, c'est-à-dire  $1 + z = a^{-1}$ , et  $v_{\parallel}$ , la projection le long de la ligne de visée de la vitesse particulière de l'objet.

La confusion entre *redshift* cosmologique et *redshift* apparent est à l'origine de distorsions entre la distribution réelle de matière et la distribution dans les relevés spectroscopiques. Ces distorsions sont appelées distorsions en *redshift* (*Redshift Space Distorsions* en anglais, RSD dans la suite du manuscrit). Je présente sur la figure 2.4, une représentation schématique des RSD, tirée de Hamilton (1998). Considérons une surdensité sphérique dans l'espace réel. La matière présente aux abords de la surdensité est attirée vers son centre. Dans le régime linéaire (en haut), quand les vitesses particulières sont faibles devant le rayon de la surdensité, celle-ci apparaît écrasée perpendiculairement à la ligne de visée. Plus le rayon diminue, plus l'écrasement apparent augmente jusqu'à ressembler à une ligne quand les vitesses particulières annulent le taux d'expansion (au milieu). Aux échelles encore plus petites (en bas), à cause des effets non-linéaires

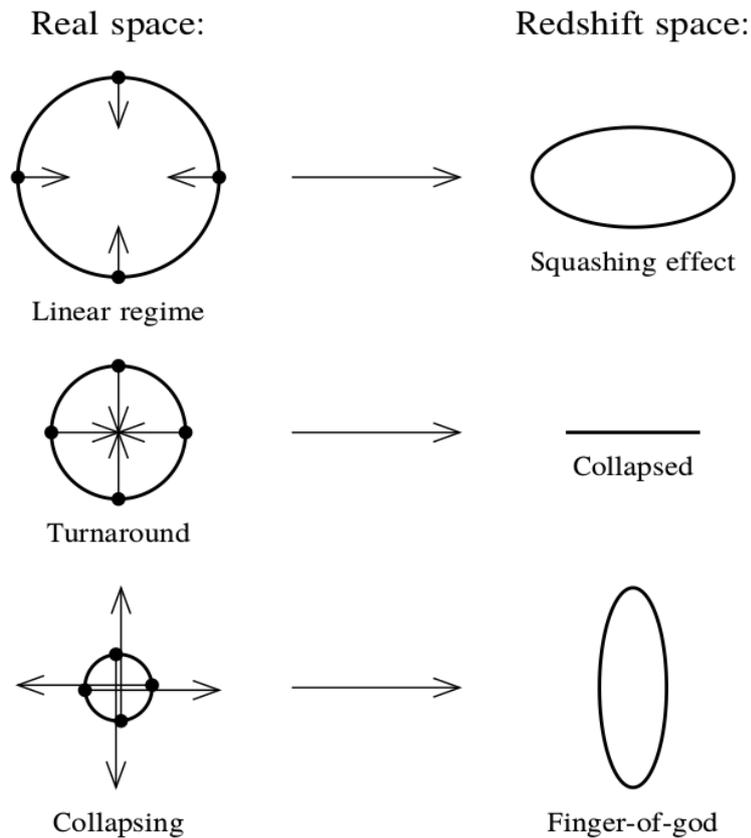


FIGURE 2.4: Les points représentent des objets célestes (par exemple, des galaxies) qui sont attirés par une zone de surdensité sphérique, et les flèches représentent leurs vitesses particulières. On va étudier la forme dans l'espace des *redshifts*, de la "coquille" de la surdensité (qui est sphérique dans l'espace réel). Aux grandes échelles (en haut), les vitesses particulières sont faibles comparées au rayon de la coquille, celle-ci apparaît écrasée perpendiculairement à la ligne de visée. À plus petite échelle, non seulement le rayon de la coquille est plus petit, mais les vitesses particulières sont plus grandes. Au point critique (au milieu), les vitesses particulières annulent le taux d'expansion et la coquille apparaît totalement effondrée. Pour les échelles encore plus petites (en bas), à cause des effets non-linéaires liés à l'effondrement gravitationnel, la coquille apparaît allongée le long de la ligne de visée, ce phénomène est appelé *finger-of-God*. Source : [Hamilton \(1998\)](#)

causés par l'effondrement gravitationnel, les RSD ont changé de sens et apparaissent allongées le long de la ligne de visée. Ces figures sont appelées *fingers-of-God* et sont caractéristiques des relevés spectroscopiques (Kaiser 1987). On présente dans cette sous-section une étude brève des RSD linéaires. Pour plus d'information, le lecteur peut se référer à l'étude de Hamilton (1998).

Puisqu'on n'a seulement accès au *redshift* apparent  $\tilde{z}$  et non pas au *redshift* cosmologique  $z$ , on peut calculer la distance comobile  $s = \chi(\tilde{z})$  dans l'espace des *redshifts* mais pas la vraie distance  $r = \chi(z)$ . Dans le cas où  $v_{\parallel} \ll c$ , on peut faire un développement limité des distances comobiles (équation 1.46). On obtient :

$$s \simeq r \left(1 + \frac{u_{\parallel}}{r}\right) \quad \text{avec} \quad u_{\parallel} = \frac{(1+z)v_{\parallel}}{H(z)}. \quad (2.15)$$

En écrivant l'équation de conservation de la quantité de matière dans les deux espaces, on obtient la densité de contraste  $\delta_s$  dans l'espace des *redshifts* :

$$\delta_s \simeq \delta - \frac{du_{\parallel}}{dr}. \quad (2.16)$$

Cette approximation n'est valable que dans le cas où  $u_{\parallel} \ll r$ , ce qui est vrai à haut *redshift*.

En passant dans l'espace de Fourier et en utilisant l'équation de continuité, on peut montrer que (Hamilton 1998) :

$$\tilde{\delta}_s(\vec{k}, z) = (1 + f\mu_k^2) \times \tilde{\delta}(\vec{k}, z), \quad (2.17)$$

où  $\tilde{\delta}$  désigne la transformée de Fourier du contraste de densité, et  $\mu_k$ , la projection du vecteur  $\vec{k}$  sur la ligne de visée.

L'effet des RSD sur la fonction de corrélation sera étudié dans la sous-section 6.1.2.

### 2.1.5 Formule de Kaiser

Si on suppose que le champ des vitesses particulières du traceur  $t$  est égal au champ des vitesses particulières de la matière (Hamilton 1998), alors la transformée de Fourier,  $\tilde{\delta}_{t,s}$  du contraste de densité du traceur  $t$ , mesuré dans l'espace des *redshifts*, s'écrit :

$$\tilde{\delta}_{t,s}(\vec{k}) = b_t(1 + \beta_t\mu_k^2)\tilde{\delta}(\vec{k}), \quad (2.18)$$

où  $\beta_t = f/b_t$  est appelé *paramètre RSD* du traceur  $t$ .

Finalement, on peut écrire le spectre de puissance du traceur  $t$  mesuré dans le cadre d'un relevé en *redshift* :

$$P_{t,s}(\vec{k}) = b_t^2(1 + \beta_t\mu_k^2)^2 P(\vec{k}). \quad (2.19)$$

Cette formule, appelée *formule de Kaiser*, a été introduite pour la première fois par Kaiser (1987). Elle rend compte des effets de biais et de vitesses particulières sur le spectre de puissance d'un traceur de la matière. La fonction de corrélation du traceur  $t$  est obtenue en faisant la transformée de Fourier de l'équation 2.19.

Dans la section suivante, on se concentre sur un traceur particulier : les absorptions Lyman- $\alpha$  visibles dans les spectres des quasars, qui sont l'objet d'étude central de cette thèse.

## 2.2 Les forêts Lyman- $\alpha$

### 2.2.1 Quasars

Dans les années 1950, l'avènement des radio télescopes permet aux astronomes de voir le rayonnement radio de sources lumineuses (galaxies, étoiles ...) précédemment observées avec des télescopes optiques. Ils découvrent alors un type d'objet très particulier : son émission radio, très étendue, montre qu'il s'agit d'une source très énergétique, mais la contrepartie optique correspond à des objets de très petite taille, à peine celle d'une étoile (Shields 1999). On leur donne alors le nom de *Quasi Stellar Objects*, plus tard raccourci en *quasars*. Une décennie plus tard, Schmidt (1963) observe pour la première fois le spectre optique d'un quasar. Grâce à l'étude des raies d'émission, il en déduit que cet objet est très lointain et donc extrêmement lumineux. À cette époque, les astronomes sont surpris par la gigantesque densité des quasars mais ne savent pas comment l'expliquer.

Aujourd'hui, on pense que les quasars sont créés par un trou noir super-massif au centre d'une galaxie (Shields 1999). Le trou noir accrète la matière environnante dont les particules sont accélérées par la chute dans le potentiel gravitationnel. Les particules relativistes sont ensuite canalisées par le champ magnétique pour former des jets extrêmement lumineux en direction des pôles du trou noir (voir figure 2.5).

Les quasars sont parmi les objets les plus lumineux de l'univers et permettent donc de tracer la matière à très haut *redshift*. Ils peuvent être ainsi utilisés de deux façons :

- directement, à la manière des galaxies, comme traceur discret de la matière,
- indirectement, en étudiant les raies d'absorption visibles sur leurs spectres.

C'est cette deuxième méthode que je mets en oeuvre dans cette thèse, à travers l'étude des régions d'absorption Ly $\alpha$  des spectres des quasars.

### 2.2.2 L'idée de Gunn-Peterson

Le flux émis par une source lumineuse à haut *redshift* est en partie absorbé par la matière présente sur sa ligne de visée. Ainsi, l'étude des séries d'absorption visibles sur les spectres des quasars donne des indications sur la physique et la composition chimique du milieu intergalactique. Cet outil est précieux car c'est la seule méthode connue pour tracer le gaz du milieu intergalactique, qui compose la grande majorité de la matière baryonique de l'univers (Murdin 2001).

À cause de l'expansion de l'univers, l'absorption de la longueur d'onde  $\lambda_{\text{abs}}$  par un nuage au *redshift*  $z_{\text{abs}}$ , est vue sur le spectre à la longueur d'onde  $(1 + z_{\text{abs}}) \times \lambda_{\text{abs}}$ . Puisque l'hydrogène est le composant principal du gaz du milieu intergalactique et que la transition Lyman- $\alpha$  (notée Ly $\alpha$  dans la suite du manuscrit) est la transition de l'hydrogène avec la plus grande section

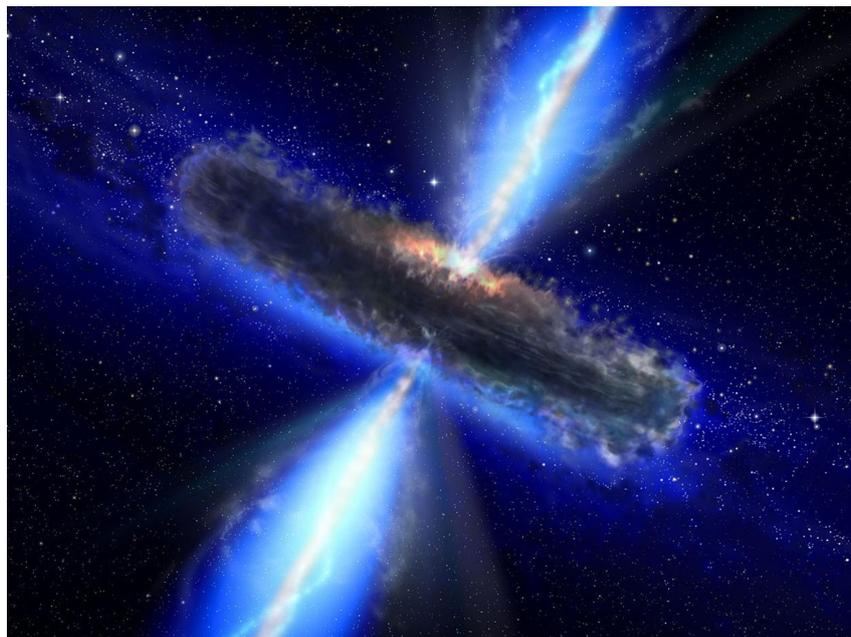


FIGURE 2.5: Vue d'artiste d'un quasar. Au centre, un trou noir accrète la matière environnante et une partie de l'énergie potentielle est convertie et libérée sous la forme de deux jets extrêmement lumineux dirigés vers les pôles magnétiques du trou noir. Source : <https://www.jpl.nasa.gov/spaceimages/details.php?id=PIA16114>

efficace, on s'attend à voir beaucoup plus d'absorption pour les longueurs d'onde inférieures à  $(1 + z_q) \times \lambda_{Ly\alpha}$ , où  $z_q$  est le *redshift* du quasar. Cet phénomène, connu sous le nom d'*effet Gunn-Peterson*, a été prédit puis observé pour la première fois par [Gunn & Peterson \(1965\)](#).

Cet effet est visible sur la figure 2.6, qui représente le spectre d'un quasar au *redshift* 2.77. On voit, autour de 460 nm, la raie d'émission Ly $\alpha$ , qui est créée par la désexcitation de l'hydrogène au voisinage du quasar. On remarque que l'amplitude et la fréquence des absorptions à gauche de la raie d'émission Ly $\alpha$  sont beaucoup plus grandes qu'à droite. La région du spectre à gauche de la raie d'émission Ly $\alpha$  (coloriée en gris) est appelée *forêt Ly $\alpha$*  à cause de l'aspect que lui donnent les séries d'absorption par le gaz du milieu intergalactique. Dans le référentiel au repos des nuages absorbants, les absorptions ont lieu dans l'UV lointain à 121.6 nm mais, à cause du *redshift*, elles sont décalées jusque dans le spectre visible.

Considérons le flux incident  $F$  qui traverse un nuage d'hydrogène de densité  $n_H$  et de largeur  $dl$ . Le flux  $dF$  absorbé en Ly $\alpha$  par le nuage s'écrit :

$$dF = -n_H \sigma_\alpha F dl, \quad (2.20)$$

où  $\sigma_\alpha$  désigne la section efficace de la transition Ly $\alpha$ . En intégrant,

$$F = \exp\left(-\int \tau\right) \quad \text{avec} \quad \tau = \int n_H \sigma_\alpha dl. \quad (2.21)$$

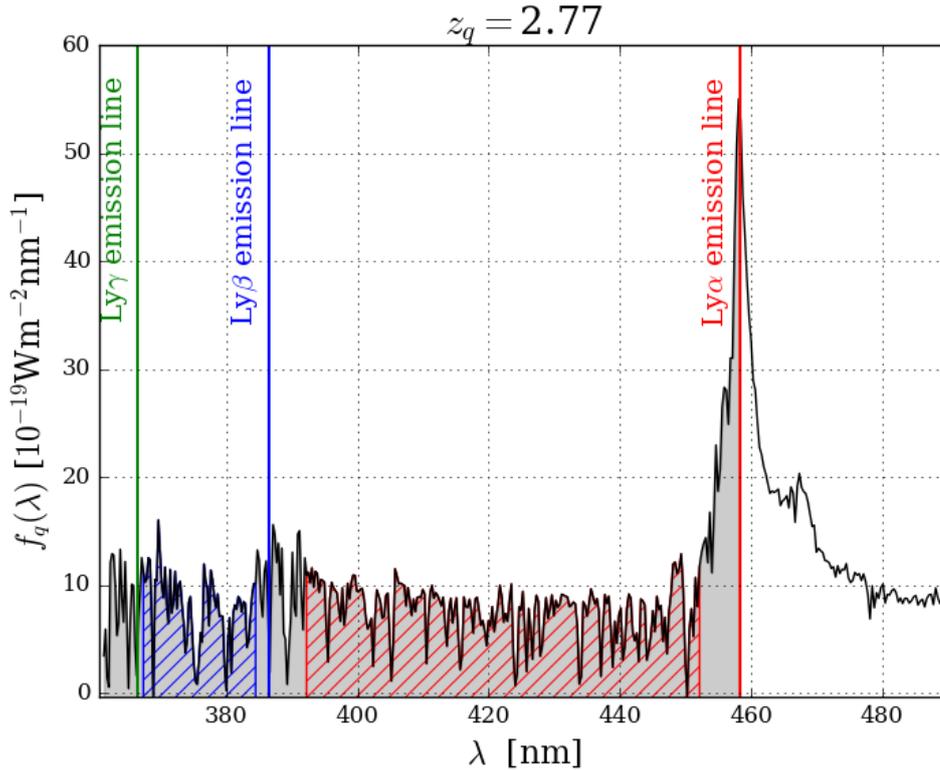


FIGURE 2.6: Spectre d'un quasar du relevé eBOSS-SDSS IV. Les raies d'émissions Ly $\alpha$  (en rouge), Ly $\beta$  (en bleu) et Ly $\gamma$  (en vert) sont représentées par les traits pleins. Les raies d'émission Ly $\beta$  et Ly $\gamma$  ne sont pas visibles sur le spectre car le nombre de photons correspondants émis au voisinage du quasar est trop faible. La région spectrale à gauche de la raie d'émission Ly $\alpha$  (coloriée en gris) est appelée *forêt Ly $\alpha$* . Sa forme particulière de "forêt" lui est donnée par les séries d'absorption Ly $\alpha$  par la matière présente sous forme de gaz le long de la ligne de visée du quasar. On peut définir des régions spécifiques de la forêt Ly $\alpha$  en fonction des transitions atomiques qui y ont lieu. Par exemple, la région Ly $\alpha$  (hachurée en rouge) ne contient que des absorptions Ly $\alpha$ , alors que la région Ly $\beta$  (hachurée en bleu) contient à la fois des absorptions Ly $\alpha$  et Ly $\beta$ .

en étudiant la physique du milieu intergalactique, on peut montrer<sup>2</sup> que la profondeur optique  $\tau$  peut s'écrire en fonction du contraste de densité de l'hydrogène neutre  $\delta_H$  :

$$\tau = a_1(1 + \delta_H)^{a_2}. \quad (2.22)$$

Typiquement,  $a_1 \in [0.15, 0.6]$  et  $a_2 \simeq 1.6$  (Seljak 2012).

Les équations 2.21 et 2.22 montrent que les variations de densité de l'hydrogène neutre sont amplifiées exponentiellement dans le flux.

Finalement, l'utilisation des absorptions Ly $\alpha$  comme traceur de la matière présente quatre avantages principaux :

- Les absorptions Ly $\alpha$  suivent la densité d'hydrogène neutre du gaz du milieu intergalactique, qui représente la majorité de la matière baryonique présente dans l'univers.
- Les variations de la densité d'hydrogène neutre sont facilement détectables car elle sont amplifiées exponentiellement dans le flux.
- Puisque les quasars sont parmi les objets les plus lumineux de l'univers, ils peuvent éclairer le milieu intergalactique à très haut *redshift*.
- Un spectre de quasar donne accès à plusieurs centaines d'absorbants.

### 2.2.3 Régions Lyman- $\alpha$ et Lyman- $\beta$

On a vu que toutes les absorptions de type Ly $\alpha$  étaient situées à gauche de la raie d'émission Ly $\alpha$  sur le spectre (en rouge sur la figure 2.6). De même, les absorptions Ly $\beta$  sont situées à gauche de la raie d'émission Ly $\beta$ <sup>3</sup> (en bleu). Ainsi, entre les raies d'émission Ly $\alpha$  et Ly $\beta$ , la seule transition de l'hydrogène neutre observée est la transition Ly $\alpha$ . Cette région spectrale (hachurée en rouge) est appelée *région Ly $\alpha$* . Le domaine de définition exact de la région Ly $\alpha$  sera explicité au chapitre 4.

On peut également définir la *région Ly $\beta$*  (hachurée en bleu), entre les raies d'émission Ly $\beta$  et Ly $\gamma$ . Dans cette région, il n'est pas possible de distinguer une absorption Ly $\alpha$  d'une absorption Ly $\beta$ . Néanmoins, puisque la section efficace de Ly $\beta$  est cinq fois plus faible que celle de Ly $\alpha$ , le nombre d'absorption Ly $\beta$  est cinq fois plus faible que le nombre d'absorption Ly $\alpha$  (Seljak 2012). On verra au chapitre 5 que la contamination par les absorptions Ly $\beta$  n'est pas problématique pour le calcul de la fonction de corrélation des absorptions Ly $\alpha$ . Pour la première fois, je calculerai la fonction de corrélation Ly $\alpha$  à deux points en utilisant les absorptions de la région Ly $\beta$ . La motivation principale à la prise en compte de cette région est l'augmentation du nombre de données. En effet, en ajoutant la région Ly $\beta$ , on accroît de 30% environ la taille du domaine spectral étudié.

<sup>2</sup><https://home.strw.leidenuniv.nl/~jarle/Teaching/GalaxyFormation/lecture14.html>

<sup>3</sup>La raie d'émission Ly $\beta$  n'est pas visible sur le spectre car le nombre de photons correspondants émis à proximité du quasar est trop faible.

## 2.3 Mesure de la position du pic BAO avec les forêts Lyman- $\alpha$

Pour mon analyse, je réduis désormais le champ d'intérêt au calcul de la fonction de corrélation en utilisant les absorptions Ly $\alpha$  comme traceur de masse, appelée *fonction de corrélation Ly $\alpha$*  dans la suite du manuscrit.

Pour la première fois, [McDonald \(2003\)](#) évoque l'idée d'utiliser la fonction de corrélation Ly $\alpha$  pour contraindre le paramètre cosmologique  $\Omega_\Lambda$ . Il réalise des simulations hydrodynamiques de l'univers puis calcule la fonction de corrélation Ly $\alpha$  pour en déduire la précision de mesure attendue sur  $\Omega_\Lambda$  en fonction des conditions expérimentales (nombre, signal sur bruit et résolution des spectres). À cette époque, le relevé SDSS II est en cours mais les quelques milliers de spectres disponibles ne suffisent pas encore pour mesurer la fonction de corrélation Ly $\alpha$  à deux points. [McDonald & Eisenstein \(2007\)](#) complètent le travail précédent en faisant une analyse statistique sur la précision de la mesure de la position du pic BAO en fonction de la densité de quasars, du volume du ciel observé, de la résolution et du signal sur bruit des spectres.

Pour la première fois, [Slosar et al. \(2011\)](#) calculent la fonction de corrélation Ly $\alpha$  en utilisant les régions Ly $\alpha$  des 14 000 spectres de quasars de la première année de l'expérience BOSS (voir chapitre 3). Le nombre de données ne permet pas encore de voir le pic BAO mais le premier pic de corrélation (distances de séparation inférieures à  $60 h^{-1}\text{Mpc}$ , voir figure 2.1) est détecté. Les auteurs en déduisent des contraintes sur le biais et le paramètre RSD des absorptions Ly $\alpha$ . Deux ans plus tard, l'ajout d'environ 40 000 nouveaux spectres permet de détecter pour la première fois les BAO dans la fonction de corrélation Ly $\alpha$  ([Busca et al. 2013](#), [Slosar et al. 2013](#)). Les deux articles proposent des analyses indépendantes qui conduisent à une détection du pic entre  $3$  et  $5\sigma$  selon les modèles utilisés, voir figure 2.7. Les rapports  $D_H(z=2.4)/r_d$  et  $D_M(z=2.4)/r_d$  sont mesurés avec les précisions respectives 3.4% et 7.2% ([Slosar et al. 2013](#)). En combinant leurs résultats avec des mesures du CMB, [Busca et al. \(2013\)](#) obtiennent  $H(z=2.3)$  avec une précision de 4% et présentent une analyse des contraintes sur  $\Omega_M$  et  $\Omega_\Lambda$ . Deux ans et 80 000 spectres plus tard, [Delubac et al. \(2015\)](#) mesurent  $D_H(z=2.4)/r_d$  et  $D_M(z=2.4)/r_d$  avec les précisions respectives 3.0% et 5.8%.

Dans tous les travaux précédents ([Slosar et al. 2011](#), [Busca et al. 2013](#), [Slosar et al. 2013](#), [Delubac et al. 2015](#)), la position du pic BAO est mesurée en ajustant un modèle de la fonction de corrélation Ly $\alpha$  sur les données. À cause des erreurs systématiques, la transformée de Fourier de l'équation 2.19 ne permet pas, avec seulement les deux paramètres  $b_{Ly\alpha}$  et  $\beta_{Ly\alpha}$ , d'expliquer les données. C'est pourquoi les erreurs systématiques sont ajoutées au modèle sous la forme d'une fonction purement mathématique, appelée *broadband* (voir section 6.4). [Bautista et al. \(2017\)](#) marquent une rupture avec les papiers précédents en proposant une nouvelle procédure d'analyse qui permet d'expliquer, sans l'ajout de la fonction *broadband*, la forme de la fonction de corrélation.

Dans cette thèse, je décris l'analyse de la fonction de corrélation Ly $\alpha$  mesurée avec les premières données d'eBOSS, présentée dans [de Sainte Agathe et al. \(2019\)](#) (reproduit en annexe D). Par rapport aux travaux précédents, je calcule pour la première fois la fonction de corrélation Ly $\alpha$  en utilisant la région Ly $\beta$  et je raffine le modèle d'analyse présenté dans [Bautista et al. \(2017\)](#). Je comparerai les deux résultats dans le chapitre 7.

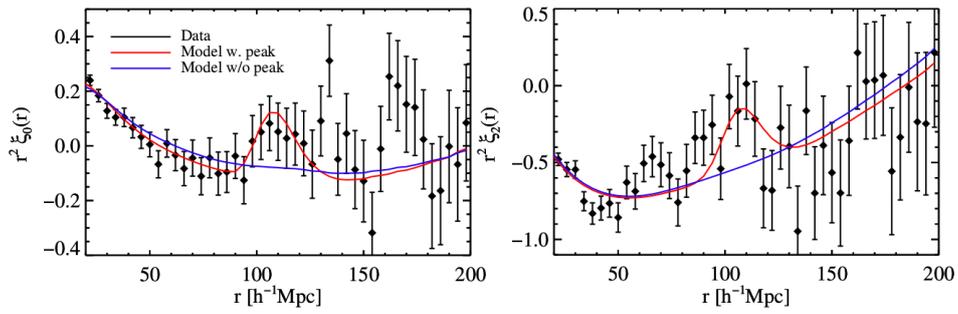


FIGURE 2.7: Première détection du pic BAO dans la fonction de corrélation Ly $\alpha$  à deux points par Busca et al. (2013). On représente le monopôle (à gauche) et le quadrupôle (à droite) de la décomposition de Legendre de la fonction de corrélation. On superpose aux données (en noir), le modèle avec BAO (en rouge) et sans BAO (en bleu). Le pic BAO est détecté à  $4.3\sigma$ . Source : Busca et al. (2013)



## Chapitre 3

# L'expérience eBOSS du relevé SDSS-IV

Les données utilisées dans cette thèse sont celles du relevé *extended Baryon Oscillation Spectroscopic Survey* (eBOSS dans la suite du manuscrit). Dans ce chapitre, je décris l'instrument du *Sloan Digital Sky Survey* (SDSS dans la suite du manuscrit) utilisé pour prendre les données eBOSS. Puis la phase de prise de mesure, suivie du traitement des données. À la fin du processus, la chaîne d'extraction de SDSS fournit, pour chaque objet observé, le flux calibré et sa variance sur le domaine spectral. Les spectres des quasars sont ensuite identifiés et analysés, et les résultats sont enregistrés dans le catalogue des quasars du 14e relevé SDSS, DR14Q (Pâris et al. 2018). Cette phase de préparation réalisée par les équipes de SDSS et d'eBOSS est cruciale pour l'analyse que j'ai réalisée, qui sera présentée dans les chapitres suivants.

### 3.1 Le relevé SDSS

Le relevé SDSS a débuté en 2000 (York et al. 2000) après une décennie de construction et de conception. Depuis cette date, les nouvelles données sont publiées chaque année, faisant du télescope de SDSS l'un des plus productifs au monde (Madrid & Macchetto 2009).

Ces dix-neuf dernières années, quatre phases du relevé se sont succédées donnant lieu à plus de 7 700 publications et 376 000 citations (<https://www.sdss.org/science/>). Pendant huit ans, SDSS-I et II ont mesuré les positions et spectres de 120 000 quasars et 930 000 galaxies, constituant la plus grande carte de l'Univers jamais réalisée. L'instrument se composait alors d'un télescope de 2.5m situé à Apache Point (voir sous-section 3.2.1), d'une caméra pour la photométrie et d'un système de 630 fibres optiques reliées à une paire de spectrographes pour la spectroscopie (York et al. 2000). Pour sa troisième phase (2008-2014, SDSS III), l'instrument a été amélioré avec l'ajout de 370 nouvelles fibres optiques et d'une nouvelle paire de spectrographes plus performants (Alam et al. 2015). Depuis juillet 2014 la quatrième phase de SDSS a commencé et utilise les mêmes instruments que SDSS-III.

Trois expériences participent au relevé SDSS IV :

- APOGEE-2, Apache Point Observatory Galactic Evolution Experiment 2, observe les spectres des étoiles de la Voie Lactée pour étudier la composition et les mouvements stellaires au sein de notre galaxie (<https://www.sdss.org/surveys/apogee-2/>).

- MANGA, Mapping Nearby Galaxies at APO, utilise les spectres des galaxies proches pour l'étude du "cycle de vie" des galaxies depuis leur formation jusqu'à leur mort (<https://www.sdss.org/surveys/manga/>).
- eBOSS, extended Baryon Oscillation Spectroscopic Survey, mesure l'empreinte des oscillations acoustiques baryoniques dans la distribution des galaxies, quasars et absorbants Ly $\alpha$  dans l'Univers (<https://www.sdss.org/surveys/eboss/>). Mon étude s'inscrit dans cette expérience.

Dans la section suivante, je décris les différentes parties de l'instrument tel qu'il est utilisé par eBOSS.

## 3.2 L'instrument

En mode spectroscopique, l'instrument se compose du télescope (sous-section 3.2.1), d'une paire de spectrographes et d'une cartouche (sous-section 3.2.3) dans laquelle les fibres optiques sont montées. L'ensemble est schématisé figure 3.1. En phase photométrique, la cartouche est remplacée par la caméra (sous-section 3.2.2). Pour passer de la phase photométrique à la phase spectrométrique, il suffit de remplacer la caméra par le système composé par la cartouche et la paire de spectrographes.

### 3.2.1 Le télescope

Cette sous-section décrit brièvement le télescope de SDSS (figure 3.2) et s'inspire largement de l'article dédié [Gunn et al. \(2006\)](#).

L'idée de réaliser un relevé du ciel à grande échelle est née au milieu des années 80 avec l'explosion des semi-conducteurs et des capacités de calculs informatiques. Afin de prendre simultanément un maximum d'images et de spectres d'objets très éloignés, le futur télescope devait répondre à quatre exigences :

- un plan focal très large,
- une faible distorsion sur le plan focal,
- un pointage précis des objets célestes,
- un passage rapide de l'imagerie à la spectroscopie et vice-versa.

Un télescope de type Richey-Chrétien ordinaire, c'est-à-dire avec deux miroirs hyperboliques, semblait approprié pour obtenir un champ large mais donnait un astigmatisme beaucoup trop important pour les objectifs du relevé. La solution a été apportée en ajoutant un ensemble de

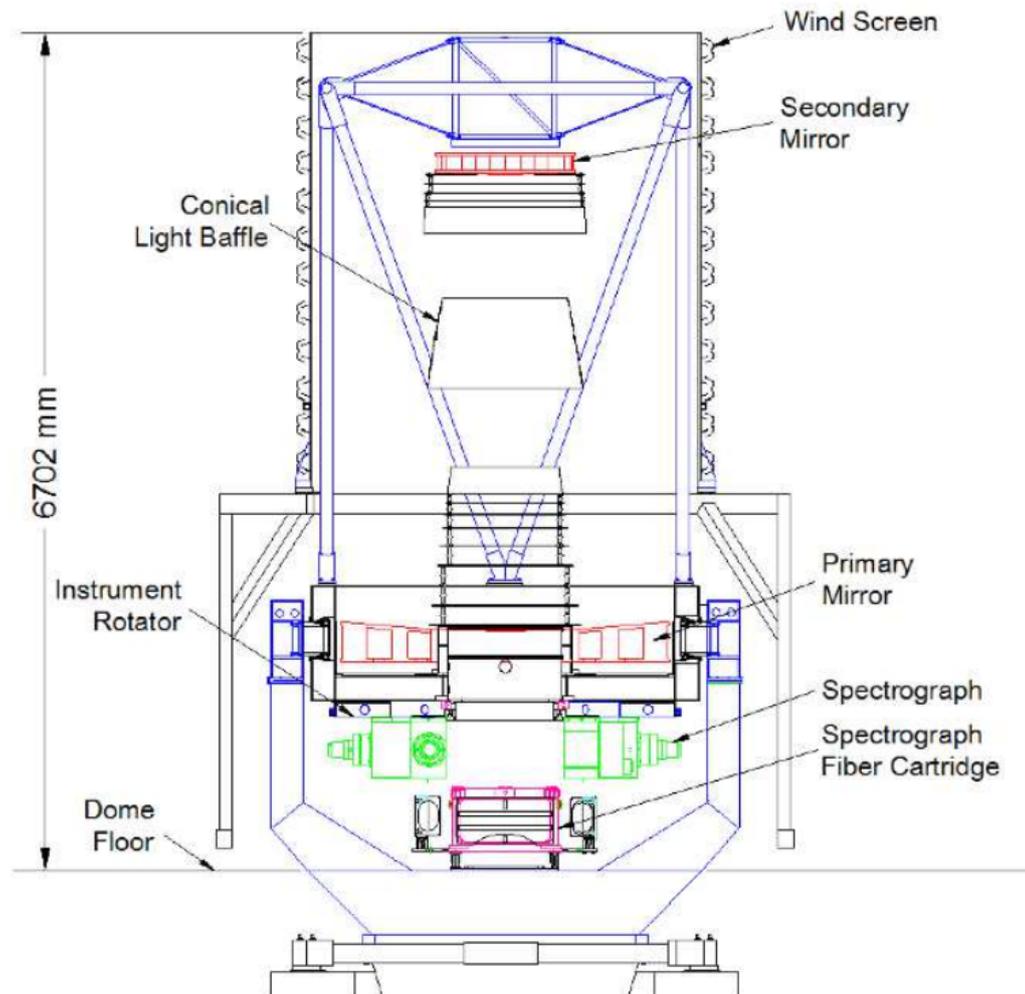


FIGURE 3.1: Schéma de l'instrument de SDSS. En rouge, sont représentés les miroirs primaire et secondaire. On voit en noir les barrières coniques (*conical light baffle* en anglais) qui éliminent une partie des lumières parasites. En rose, on peut voir une cartouche (*cartridge* en anglais) qui n'est pas encore installée. Lors de l'observation, elle est placée entre les deux spectrographes (en vert). Source : [Smee et al. \(2013\)](#)



FIGURE 3.2: Le télescope de 2.5m de type Richey-Chrétien modifié est localisé à Apache Point Observatory dans l'état du Nouveau-Mexique (États-Unis). Spécialement construit pour le relevé SDSS, il est opérationnel depuis 1998. Source : <https://www.sdss.org/education/sdss-telescope-small/>

deux correcteurs au système. Le premier correcteur se compose d'une lentille de 802 mm de diamètre et d'environ 12 mm d'épaisseur. Le second correcteur varie selon le mode d'observation. En phase photométrique, il est constitué d'une lentille mince d'épaisseur 45 mm au centre, avec une surface d'entrée asphérique<sup>1</sup> et une surface de sortie plane, qui fait partie intégrante de la caméra (voir sous-section 3.2.2). En phase spectroscopique, la lentille correctrice est encore plus mince (10 mm d'épaisseur au centre pour 727 mm de diamètre) et spécialement optimisée pour la spectroscopie.

Le miroir primaire a une ouverture géométrique de  $f/2.25$  et forme un anneau de diamètre intérieur de 1.17m et de diamètre extérieur de 2.5m. Une barrière conique est montée en aval du miroir pour éliminer les lumières parasites. Le miroir secondaire est placé à 3.6m du primaire et mesure 1.08m de diamètre pour un rayon de courbure de 7330mm environ. L'ensemble a une ouverture géométrique de  $f/5$ , réalisant ainsi la troisième condition avec un champ de vue de  $3\text{deg}^2$ . Le plan focal du système est à 0.76m sous le miroir primaire pour permettre un accès facile aux instruments de mesure.

Un système, composé d'un ascenseur hydraulique et de loquets pneumatiques, permet de remplacer la caméra par la cartouche, et vice-versa. La procédure est optimisée pour pouvoir être réalisée rapidement, de nuit et en toute sécurité.

La conception et la construction du télescope a duré une dizaine d'années. Situé à Apache Point Observatory (APO) dans l'état du Nouveau-Mexique (États-Unis), il observe le ciel depuis 1998.

---

<sup>1</sup>Une lentille asphérique permet d'avoir de meilleures performances optiques en périphérie par rapport à une lentille sphérique. Source : [https://en.wikipedia.org/wiki/Aspheric\\_lens](https://en.wikipedia.org/wiki/Aspheric_lens)

### 3.2.2 La caméra

Dans cette sous-section, nous décrivons rapidement la caméra utilisée pour le relevé photométrique. Pour plus de détails, le lecteur peut se référer à [Gunn et al. \(1998\)](#).

L'élément de base de la caméra est le dispositif à transfert de charge (*Charged Coupled Device* en anglais, CCD dans la suite du manuscrit). C'est un dispositif qui transforme les photons reçus en électrons qui sont collectés puis dénombrés. La quantité d'électrons en sortie du CCD est proportionnelle au nombre de photons reçus en entrée.

Pour avoir une image de bonne qualité, le temps d'exposition d'un CCD doit être suffisamment long et/ou l'objet bien éclairé. Dans le cas d'un relevé du ciel, on a le double inconvénient d'observer des sources lumineuses à la fois faibles et mouvantes. Pour pallier ce problème, on utilise la technique du décalage temporel et intégration (*time delay and integrate* en anglais, TDI dans la suite du manuscrit). Au lieu d'utiliser une seule rangée de CCD, on utilise une série de rangées de CCD. Les charges d'une source lumineuse entrant dans le champ sont d'abord collectées par la première rangée de CCD de la série puis par la deuxième,... jusqu'à la dernière. Les charges de chaque rangée CCD s'additionnent simultanément au mouvement de la source.

La caméra de SDSS est constituée de deux types de CCD :

- 30 CCD photométriques de  $2\,048 \times 2\,048$  *pixels*, répartis en 6 séries de 5 capteurs. Chaque CCD est monté avec un filtre optique particulier de façon à observer les cinq bandes photométriques ( $u, g, r, i, z$ ) dans une même série.
- 24 CCD de  $2\,048 \times 400$  *pixels* dont 2 sont dévolus à la calibration et 22 à l'astrométrie.

La figure 3.3 représente la disposition des CCD sur le plan focal de la caméra et la figure 3.4 est une image de la caméra. Les CCD photométriques sont situés au centre du plan et les CCD de calibration et d'astrométrie, à la périphérie. La lecture par TDI se fait le long de l'axe  $Y$  descendant : les photons d'un objet qui entre dans le plan focal rencontrent en premier lieu les CCD d'astrométrie et de calibration puis un CCD de la bande  $r$ , puis  $i$ , puis  $u$ , etc ... jusqu'à rencontrer de nouveau les CCD d'astrométrie et de calibration.

Les CCD photométriques ne sont pas disposés au hasard sur le plan focal mais de façon à maximiser la détection des photons entrants. La figure 3.5 montre l'efficacité quantique du système en fonction de la bande photométrique. Les courbes supérieures et inférieures prennent en compte les effets instrumentaux (transmission des filtres, efficacité quantique des CCD, pertes liées à l'optique). Les extinctions atmosphériques sont soustraites sur les courbes inférieures. L'efficacité quantique donne le rapport entre le nombre de charge collectées par le système sur le nombre de photons incidents. Ainsi, dans la bande  $r$ , environ 60% des photons incidents sont détectés. Les bandes  $z$  et  $u$  affichent les moins bonnes performances avec moins de 15% de détection. Dans le cas de  $u$ , l'extinction atmosphérique est responsable de la perte de sensibilité, pour  $z$ , c'est l'efficacité quantique des CCD qui diminue. Afin de maximiser la détection, on choisit de mettre les bandes avec une plus faible sensibilité au centre du plan focal, c'est-à-dire à l'emplacement le plus éclairé, et celle de plus forte sensibilité en périphérie. D'où la disposition des CCD visible sur la figure 3.3.

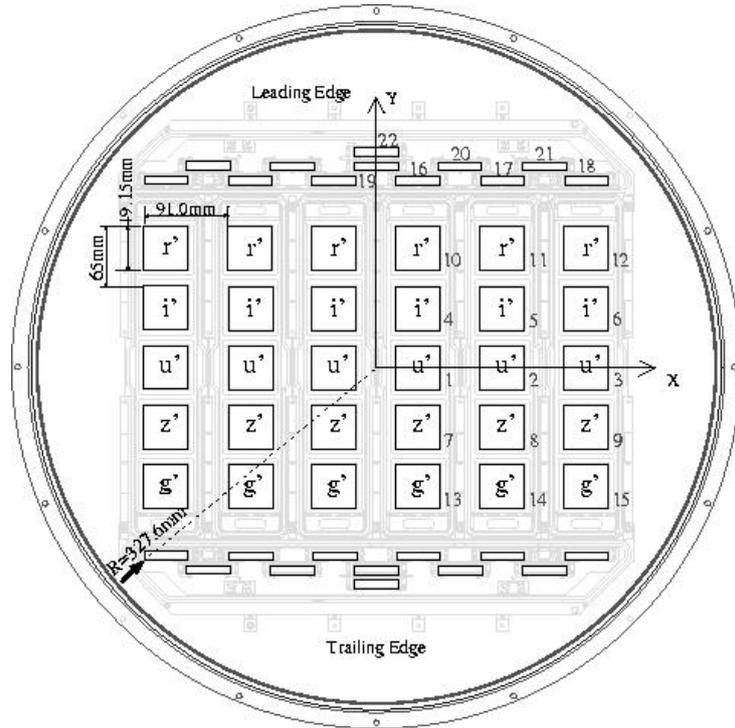


FIGURE 3.3: Disposition sur le plan focal des CCD de la caméra. Au centre, on trouve les 5×6 CCD dédiés à la photométrie répartis en bandes et à la périphérie, les 24 CCD dédiés à l'astrométrie. La lecture se fait du haut vers le bas. Source : [Gunn et al. \(1998\)](#)

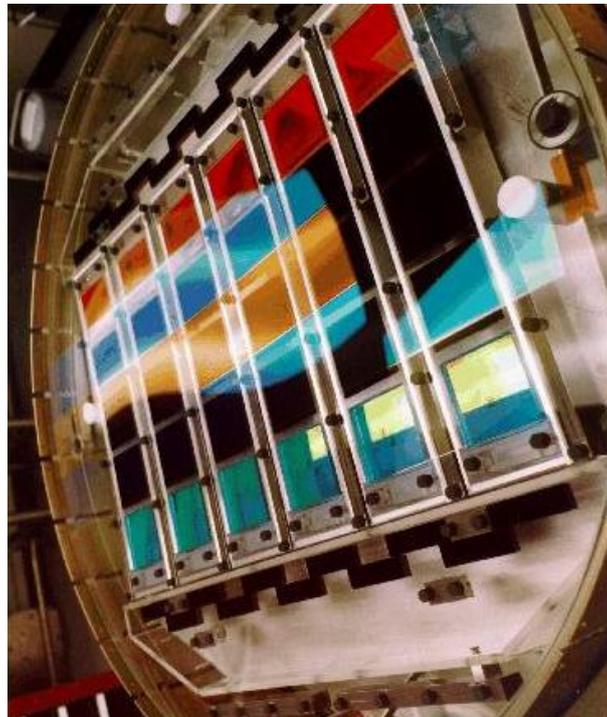


FIGURE 3.4: Photographie de la caméra de SDSS. Source : [Gunn et al. \(1998\)](#)

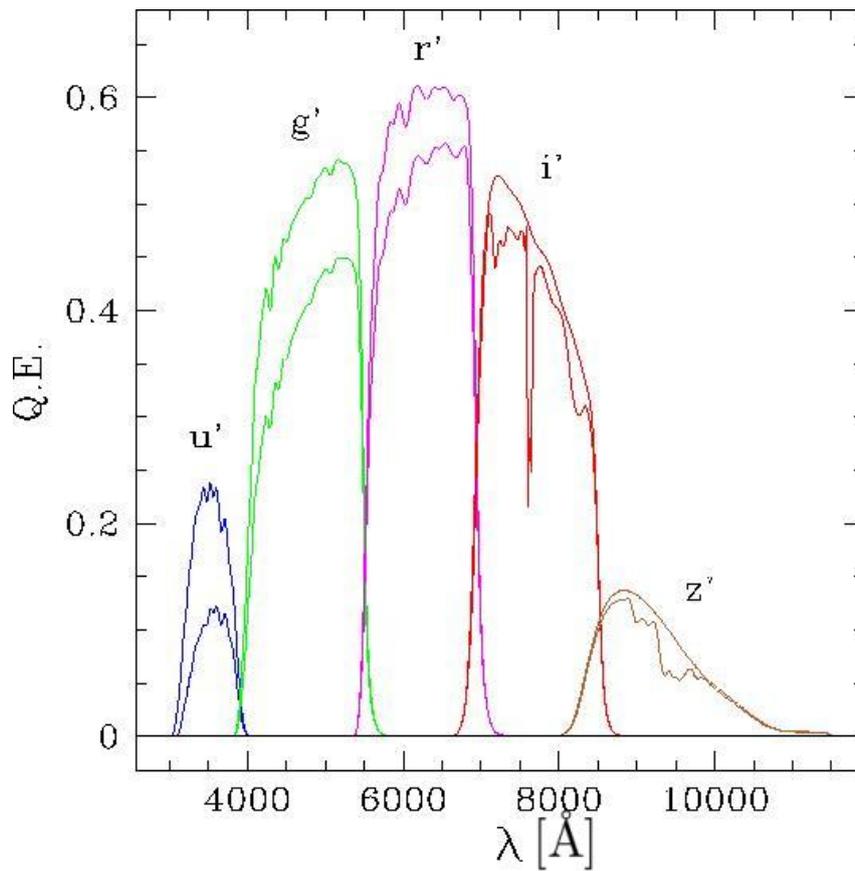


FIGURE 3.5: Efficacité quantique du système dans chaque bande photométrique. Les courbes supérieures et inférieures prennent en compte les systématiques de l'instrument (transmission des filtres, l'efficacité quantique des CCD, pertes liées à l'optique). Les courbes inférieures prennent également en compte l'extinction atmosphérique. Source : [Gunn et al. \(1998\)](#)

### 3.2.3 Le système de spectroscopie

Le système de spectroscopie de SDSS original a été mis en service en automne 1999, puis amélioré en 2009 pour les besoins de l'expérience BOSS (*Baryon Oscillation Spectroscopic Survey*). On donne dans cette sous-section un aperçu de son fonctionnement. Pour plus d'information, voir [Smee et al. \(2013\)](#).

L'instrument de spectroscopie de BOSS se compose de trois éléments principaux :

- Un ensemble de 1 000 fibres optiques qui collectent la lumière dans le plan focal du télescope.
- Une cartouche mobile, facilement positionnable, qui supporte et protège l'ensemble des fibres.
- Un jeu de deux spectrographes composés chacun d'un filtre dichroïque qui oriente les composantes rouge et bleue de la lumière vers deux caméras (bleue et rouge).

#### 3.2.3.1 Les fibres optiques

1 000 fibres optiques de diamètre  $120\mu\text{m}$  (soit  $2''$  sur le plan focal) sont placées sur le plan focal du télescope pour collecter la lumière incidente et l'acheminer vers les spectrographes.

#### 3.2.3.2 Les cartouches

La cartouche désigne l'ensemble du dispositif qui maintient les fibres. Elle est conçue pour protéger les fibres lors des phases d'installation et de stockage.

L'extrémité des fibres est maintenue sur le plan focal du télescope grâce à une plaque d'aluminium de 3.2mm d'épaisseur et 800mm de diamètre (voir figure 3.6). On a percé à l'avance des trous sur les positions des objets à observer et les fibres sont glissées à l'intérieur. L'ensemble des fibres est divisé en deux jeux de 500, un jeu par spectrographe. L'autre extrémité des fibres est maintenue dans une demi-plaque coudée qui est encastrée dans le spectrographe lors de l'observation (voir figure 3.7).

Les cartouches sont préparées pendant la journée et installées de nuit en quelques minutes. Un monte charge est utilisé par l'opérateur pour permettre la mise en place et le stockage de la cartouche, qui pèse 145kg.

#### 3.2.3.3 Les spectrographes

La figure 3.8 illustre le trajet de la lumière depuis les fibres jusqu'aux CCD de la caméra. Le faisceau lumineux rencontre un collimateur qui la réfléchit vers un filtre dichroïque. Les longueurs d'onde supérieures à 640nm sont orientées vers le canal rouge et les inférieures, vers le canal bleu. Le faisceau de chaque canal rencontre un grisme<sup>2</sup> qui disperse la lumière à l'entrée de

---

<sup>2</sup>Un grisme est un prisme avec une face sur laquelle est gravé un réseau de diffraction. Source : <https://fr.wikipedia.org/wiki/Grisme>

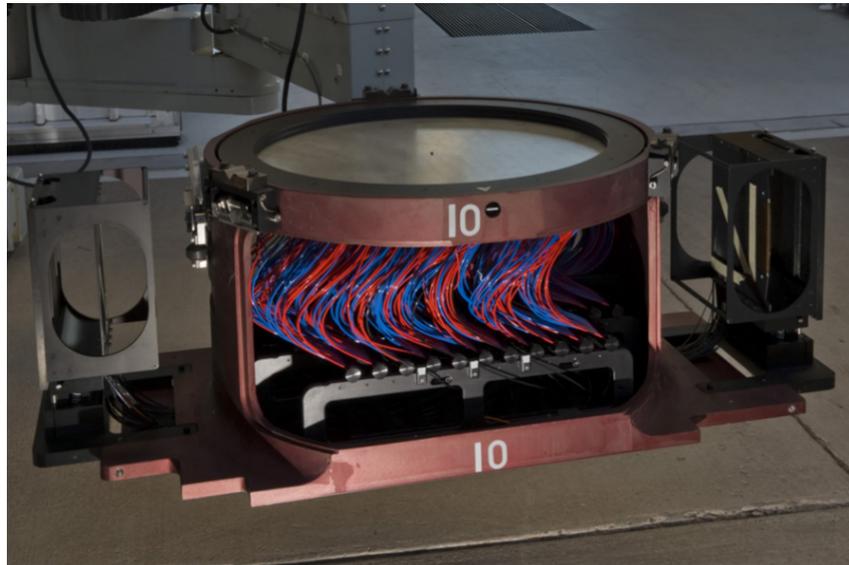


FIGURE 3.6: Photographie d'une cartouche prête à être installée sous le télescope. Les fibres optiques, visibles en bleu et rouge, sont positionnées sur le plan focal grâce à une plaque d'aluminium percée, puis elles sont séparées en deux groupes, dirigés vers les demi-plaques coudées de part et d'autre de la cartouche. Lors de l'installation, les demi-plaques sont encastrées dans les spectrographes. L'ensemble du dispositif permet de protéger les fibres lors des phases de manutention et de stockage. Source : [Smee et al. \(2013\)](#)

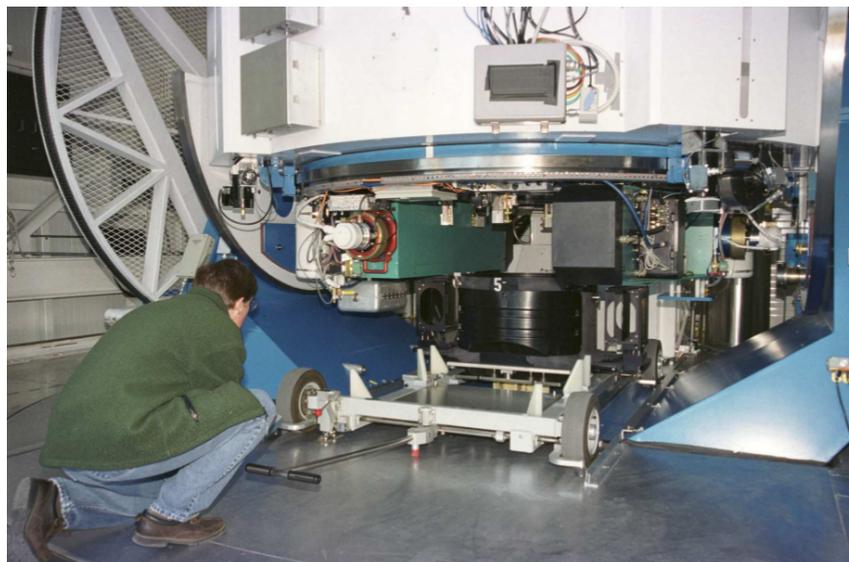


FIGURE 3.7: Installation d'une cartouche (en noir) sous le télescope. La procédure doit pouvoir être effectuée de nuit en quelques minutes. Une fois positionnée, chaque demi-plaquette est encastrée dans un spectrographe (en vert). Source : [Smee et al. \(2013\)](#)

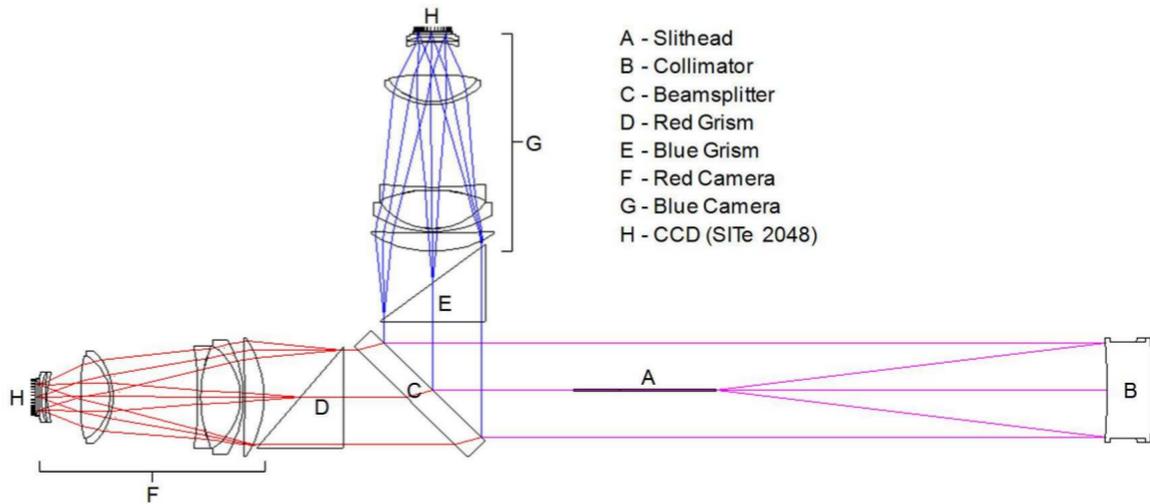


FIGURE 3.8: Le faisceau lumineux quitte les fibres en A et rencontre le collimateur en B. En passant à travers le filtre dichroïque en C, il est séparé en deux faisceaux. Le canal rouge est composé des longueurs d'onde supérieures à 640nm et le canal bleu, inférieures à 640nm. Dans chaque canal, la lumière est dispersée par un grisme (D et E) avant d'entrer dans la caméra. Source : [Smee et al. \(2013\)](#)

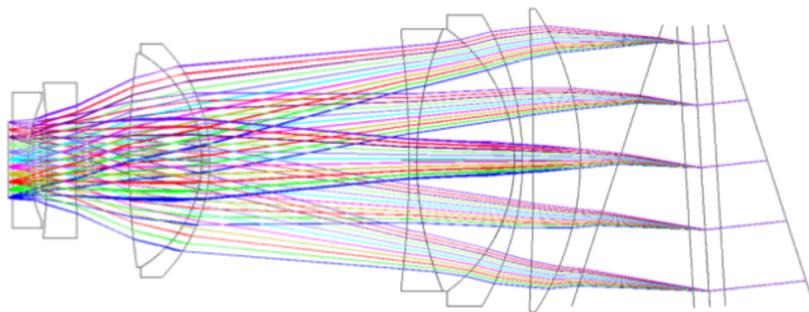


FIGURE 3.9: Parcours de la lumière dans la caméra de BOSS. Le grisme (à droite) diffracte le faisceau lumineux, qui est focalisé sur les CCD (à gauche) grâce à un ensemble de huit lentilles. Source : [Smee et al. \(2013\)](#)

la caméra. Le faisceau est focalisé par un ensemble de huit lentilles sur les CCD de  $4\,000 \times 4\,000$  *pixels* (figure 3.9) qui convertissent le signal lumineux reçu en signal électrique. Le traitement du signal de sortie sera décrit dans la sous-section 3.3.2.

Dans la section suivante, je décris les étapes de la prise de données par l'instrument.

## 3.3 La prise de données

### 3.3.1 Objectifs de eBOSS

L'objectif des expériences BOSS et eBOSS est de mesurer les distances cosmologiques  $D_M(z)$  et  $D_H(z)$  avec une précision de 1 à 2% pour des *redshifts* compris entre 0.6 et 2.5 (<https://www.sdss.org/surveys/eboss/>). Il s'agit du plus grand volume jamais observé par un relevé cosmologique (Dawson et al. 2013), qui donnera accès à la mesure du taux d'expansion au cours la transition entre l'ère de domination de la matière et l'ère de domination de l'énergie noire.

Le signal BAO est visible dans la fonction de corrélation à deux points de la matière sous la forme d'une sur-probabilité de 1% pour deux objets d'être séparés par la distance caractéristique des BAO, environ  $100 h^{-1}\text{Mpc}$  (chapitre 2). Ainsi, cette détection nécessite de tracer la matière sur des volumes très larges et de choisir un traceur avec une densité d'observation suffisante. À cause de la formation hiérarchique des structures de l'Univers et de la baisse de luminosité pour les objets lointains, la densité d'observation d'un traceur dépend du *redshift*.

Afin de mesurer les distances cosmologiques à différents *redshifts*, on sélectionne plusieurs traceurs de la matière (figure 3.10) :

- Les galaxies ont une répartition très dense dans l'Univers mais elles ne sont pas assez lumineuses pour être observées en assez grande quantité à haut *redshift*.
- Les quasars sont les objets les plus lumineux de l'Univers mais leur densité d'observation chute au delà du *redshift* 2.2 (Ata et al. 2017).
- Les absorptions Ly $\alpha$  sont présentes par centaines sur les spectres des quasars à haut *redshift*. Chaque absorption équivaut à la présence de matière sur la ligne de visée. Ainsi, même si la densité de quasars est faible, puisque chaque spectre nous donne accès à des centaines d'objets, on peut quand même détecter le pic BAO.

La figure 3.10 (Raichoor et al. 2017) montre la répartition des traceurs observés par SDSS en fonction du *redshift*. Les galaxies observées (échantillons LRG, LOWZ, CMASS et ELG) sont uniformément réparties pour  $z < 1.0$  et permettent d'accéder à l'ère de domination de l'énergie noire. Elles sont une sonde complémentaire aux mesures de Supernovae de type Ia (Dawson et al. 2013). Les quasars sont utilisés comme traceur pour  $z < 2.1$  et donnent accès à une mesure intermédiaire lors de la phase de décélération de l'expansion, à  $z \sim 1.5$  (Ata et al. 2018). Pour les quasars dont le *redshift* est supérieur à 2.0, les absorptions Ly $\alpha$  visibles sur les spectres sont

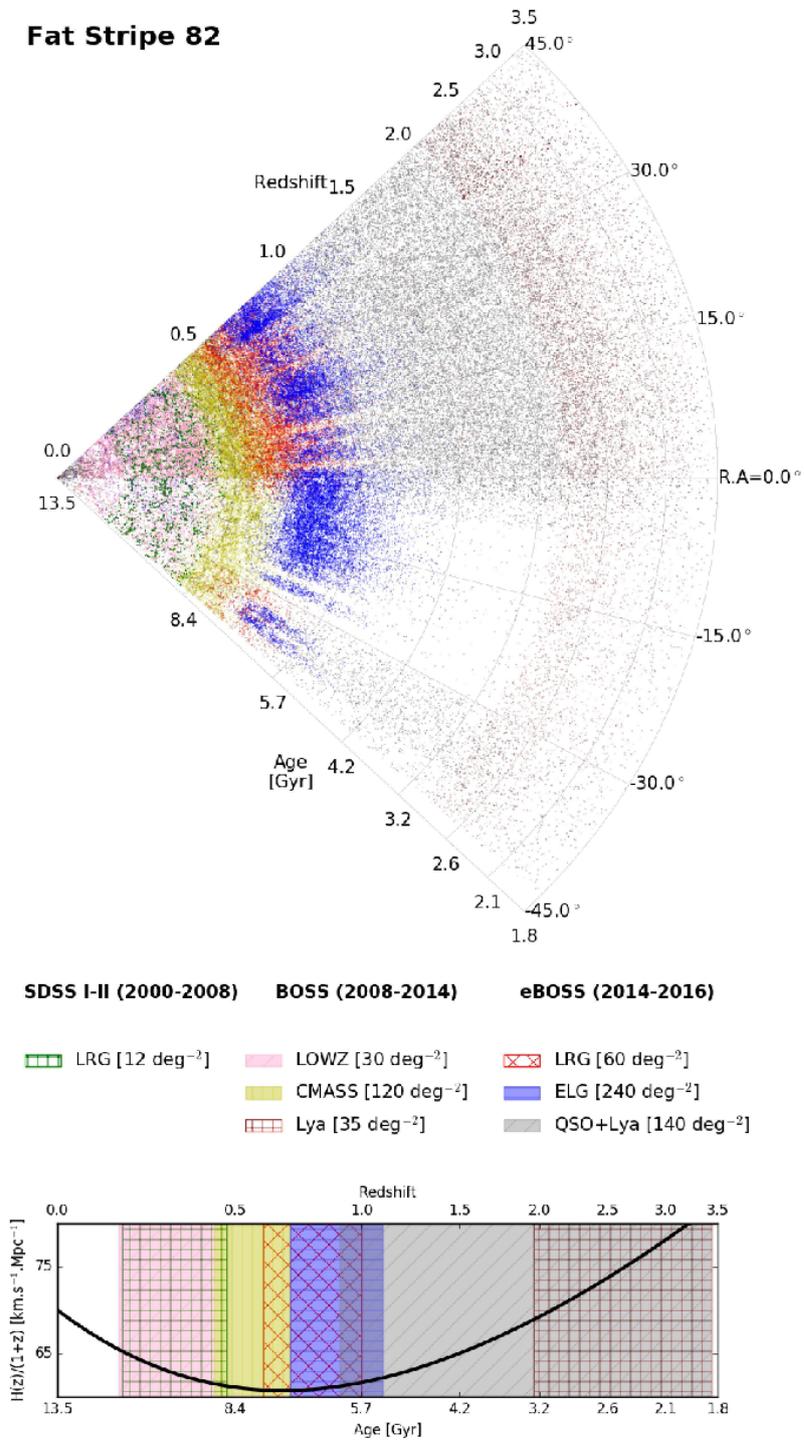


FIGURE 3.10: Répartition en *redshift* des traceurs observés par les relevés SDSS successifs.  
 Source : Raichoor et al., 2017

utilisées comme traceur continu de la matière pour  $z \sim 2.3$  (section 2.2).

La prise des données s'organise en trois étapes :

- La phase photométrique : on répertorie les différents objets célestes sur le champ de vue de l'instrument.
- La sélection des cibles : on établit une stratégie d'observation permettant de répondre aux objectifs de l'expérience.
- La phase spectroscopique : on mesure les spectres des objets sélectionnés lors de l'étape précédente.

### 3.3.2 Phase photométrique

La caméra de SDSS (voir section 3.2.2) a été mise en service dans la nuit du 19 septembre 1998 et a fonctionné jusqu'au 18 novembre 2009. Ces neuf années d'imagerie ont permis de cataloguer près d'un demi milliard d'objets répartis sur  $14\,055\text{ deg}^2$ , soit un tiers du ciel (Aihara et al. 2011).

La lumière de chaque objet est décomposée selon les cinq bandes spectrales *ugrid* (voir figure 3.5). Pour déterminer la nature de l'objet (galaxie, quasar, étoile ...), on mesure les magnitudes dans chaque bande  $u$ ,  $g$ ,  $r$ ,  $i$ , et  $d$ , respectivement  $m_u$ ,  $m_g$ ,  $m_r$ ,  $m_i$  et  $m_d$ . Chaque type de traceur peut être caractérisé par une série de coupures sur les paramètres  $m_u$ ,  $m_g$ ,  $m_r$ ,  $m_i$  et  $m_d$ . Selon le résultat obtenu, on peut donner avec plus ou moins de certitude la nature de l'objet. La pureté des catalogues, c'est-à-dire la proportion d'objets correctement identifiés, dépend de la nature de la source. Ainsi, les lignes d'émission des galaxies de type ELG rendent celles-ci facilement identifiables et le catalogue photométrique des ELG a une pureté d'environ 74 %. À l'inverse, le catalogue photométrique des quasars a une pureté de seulement 50% (Dawson et al. 2016).

La première identification des sources est réalisée à l'issue de la phase photométrique et permet de sélectionner les objets dont on mesurera les spectres. Une deuxième identification, plus précise, aura lieu à l'issue de la phase spectroscopique (voir sous-section 3.3.4).

Le relevé photométrique de SDSS (Aihara et al. 2011) n'est pas le seul relevé à être utilisé dans eBOSS. Selon le type de traceur étudié, des relevés photométriques annexes sont également pris en compte (Prakash & SDSS-IV/eBOSS 2016, Raichoor et al. 2017, Pâris et al. 2018).

### 3.3.3 Sélection des cibles

#### 3.3.3.1 Les *Luminous Red Galaxies*

Les *Luminous Red Galaxies*, LRG dans la suite du manuscrit, sont les galaxies les plus massives et les plus lumineuses à  $z < 1$  (Prakash & SDSS-IV/eBOSS 2016). Grâce à leur grande densité, c'est le premier traceur à avoir été utilisé pour la détection du pic BAO (Eisenstein et al. 2005).

TABLE 3.1: Volume du champ observé, densité requise, pureté du catalogue et densité observée pour les quatre types d'objets (LRG, ELG, "CORE quasars" et "Ly $\alpha$  quasars".)

Traceurs	<i>redshift</i>	Champ	Densité requise	Pureté	Densité observée
LRG	[0.6, 1.0]	7 500 deg <sup>2</sup>	40 deg <sup>-2</sup>	80%	50 deg <sup>-2</sup>
ELG	[0.6, 1.1]	1 500 deg <sup>2</sup>	126 deg <sup>-2</sup>	74%	220 deg <sup>-2</sup>
CORE quasars	[0.9, 2.2]	7 500 deg <sup>2</sup>	45 deg <sup>-2</sup>	50%	90 deg <sup>-2</sup>
Ly $\alpha$ quasars	> 2.1	7 500 deg <sup>2</sup>	-	-	26 deg <sup>-2</sup>

De 2009 à 2014, l'expérience BOSS a mesuré les spectres de 1.5 millions de LRG pour des *redshifts* inférieurs à 0.7 (LOWZ en rose sur la figure 3.10 et CMASS, en jaune). L'expérience eBOSS complète le relevé en ajoutant 300 000 LRG pour  $z \in [0.6, 1.0]$ , en bleu sur la figure 3.10. L'objectif est de parvenir à une précision de 1% sur la mesure des distances pour  $z < 1$  (Dawson et al. 2016). Pour cela, il faut 40 nouveaux spectres par deg<sup>2</sup> sur le champ de volume 7 500 deg<sup>2</sup>. En requérant une pureté des LRG du relevé photométrique de 80%, cela impose 50 LRG à observer par deg<sup>2</sup> (voir table 3.1).

Le processus de sélection des LRG à observer est décrit dans Prakash & SDSS-IV/eBOSS (2016).

### 3.3.3.2 Les *Emission Line Galaxies*

Initialement, seules les galaxies de type LRG devaient être observées par eBOSS. Mais il a été montré qu'une réduction de 16% du volume dédié aux LRG permettait d'observer suffisamment d'*Emission Lines Galaxies* (ELG dans la suite du manuscrit) pour obtenir une mesure supplémentaire des distances avec une précision de 2% pour un *redshift* supérieur à celui accessible avec les LRG et d'environ 0.85 (Dawson et al. 2016, Raichoor et al. 2017).

Pour pouvoir atteindre la précision de 2% avec les ELG, il faut que la densité d'objets observés soit grande. En même temps, il faut une mesure de 1% avec les LRG, c'est-à-dire que le nombre de fibres allouées aux ELG doit être faible. Finalement, la solution est d'observer un volume réduit du ciel (1 220 deg<sup>2</sup>) avec une densité requise d'objets élevée, d'environ 163 deg<sup>-2</sup> (Raichoor et al. 2017).

Les lignes d'émission des ELG sont corrélées avec les bandes photométriques, ce qui permet d'atteindre une pureté de 74% à l'issue de la phase photométrique (Dawson et al. 2016). Finalement, la densité d'ELG à observer est d'environ 220 deg<sup>-2</sup>.

Pour permettre un relevé uniforme, les fibres des ELG sont presque intégralement contenues sur 300 plaques dédiées. Pour ne pas perturber les résultats des LRG, ces plaques ont été observées seulement lors des deux dernières années d'eBOSS. Le processus de sélection des ELG est décrit dans Raichoor et al. (2017).

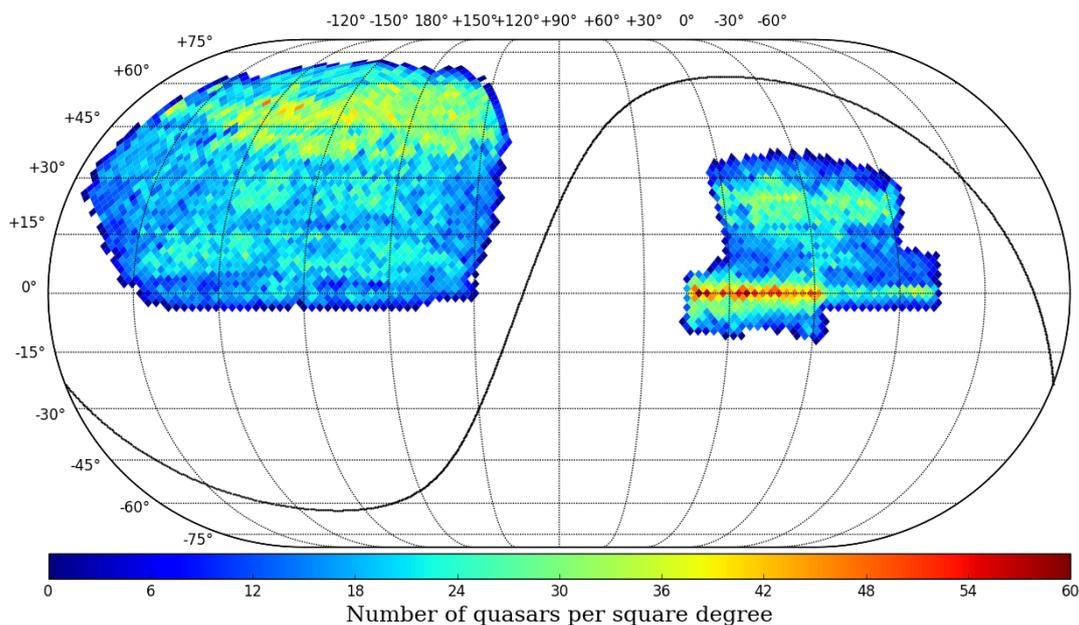


FIGURE 3.11: Empreinte sur le ciel des 216 162 " $\text{Ly}\alpha$  quasars" du catalogue DR14Q. La ligne noire donne le plan de la galaxie et sépare l'hémisphère sud de la galaxie, à gauche, de l'hémisphère nord de la galaxie, à droite.

### 3.3.3.3 Les quasars

On utilise les quasars de la gamme de *redshift*  $[0.9, 2.2]$  de façon analogue aux galaxies, c'est-à-dire comme traceur discret<sup>3</sup> de la matière. On les dénote "CORE quasars" pour les distinguer de ceux qui sont utilisés pour leurs absorptions  $\text{Ly}\alpha$ , notés " $\text{Ly}\alpha$  quasars". Le spectre d'un "CORE quasar" permet de confirmer qu'il est bien un quasar et de déterminer son *redshift*.

Pour atteindre une précision de 2% dans la mesure des distances cosmologiques à  $z \sim 1.5$ , il faut que la densité de "CORE quasars" soit de 58 objets par  $\text{deg}^2$  visible sur l'ensemble du champ eBOSS de  $7500 \text{ deg}^{-2}$  (Dawson et al. 2016) répartis uniformément dans le ciel. Sachant que BOSS avait déjà observé 13 "CORE quasars" par  $\text{deg}^2$ , la densité d'objets requise est de  $45 \text{ deg}^{-2}$ . On estime que 50% des objets définis comme quasars par le relevé photométrique sont bien des quasars. C'est pourquoi il faut observer le double de la densité requise pour satisfaire aux objectifs d'eBOSS. Cela donne 90 "CORE quasars" par  $\text{deg}^2$  (voir table 3.1), soit environ 340 000 objets supplémentaires à observer pendant les 6 ans d'eBOSS.

Les absorptions visibles sur les spectres des quasars  $\text{Ly}\alpha$  sont utilisées comme traceur continu de la matière. Puisque chaque spectre contient plusieurs centaines de longueurs d'onde en partie absorbées par le milieu intergalactique, il n'est pas nécessaire d'avoir une grande densité de " $\text{Ly}\alpha$  quasars" observés pour détecter le signal des BAO. En revanche, il faut des spectres de bonne qualité. Les estimations de McQuinn & White (2011) montrent que leur rapport signal sur bruit doit être supérieur à 3 pour améliorer de 50% les mesures de distances de BOSS. Les

<sup>3</sup>On parle de traceur "continu" de la matière quand on peut suivre continûment ses variations de densité. Les absorptions  $\text{Ly}\alpha$  sont un traceur continu car leur contraste de densité est une fonction continue de l'amplitude des absorptions, qui varie continûment. Au contraire, les galaxies, qui sont localisées en un point, sont des traceurs "discrets" car leur contraste de densité mesuré n'est pas continu.

quasars BOSS avec  $z > 2.1$  et dont le rapport signal sur bruit est inférieur à 3 sont ré-observés. On estime que chaque ré-observation améliore le signal sur bruit de 40% (Dawson et al. 2016). De plus, 60 000 nouveaux "Ly $\alpha$  quasars" ont été observés pendant les 6 ans de eBOSS. En utilisant les données de BOSS et eBOSS, on devrait obtenir, à la fin du relevé, 26 "Ly $\alpha$  quasars" par deg<sup>2</sup> avec un rapport signal sur bruit des spectres supérieur à 3 pour chaque longueur d'onde. Je représente sur la figure 3.11 une empreinte dans le ciel des "Ly $\alpha$  quasars" observés par BOSS et lors des deux premières années d'eBOSS.

Myers et al. (2015) détaille le processus de sélection des quasars à observer pour satisfaire aux exigences précédemment décrites.

### 3.3.4 Phase spectroscopique

#### 3.3.4.1 Prise de données

À l'issue de la phase de sélection des cibles, on a pour chaque traceur (LRG, ELG, "CORE quasar", "Ly $\alpha$  quasar") une liste de positions dans le ciel à observer. Chaque position dans le ciel est convertie en coordonnées dans le plan focal du télescope. Il faut maintenant établir une stratégie d'observation (Smee et al. 2013, Dawson et al. 2016) qui permette d'observer le maximum de cibles en utilisant un minimum de plaques d'observation tout en respectant les conditions suivantes :

- Chaque plaque doit contenir 1 000 fibres, 500 par spectrographe.
- La séparation minimale entre les fibres est de 260  $\mu\text{m}$ .
- 100 fibres par plaque doivent être assignées à la calibration.
- Environ 100 fibres par plaque sont dévolues à des programmes auxiliaires.

L'algorithme d'affectation de l'ensemble des fibres est décrit par Blanton et al. (2001).

Une fois que la stratégie d'observation est définie, les plaques sont percées plusieurs mois à l'avance puis acheminées sur site. L'opération de branchement de l'ensemble des fibres d'une cartouche dure entre 30 et 45 min (Smee et al. 2013) et est réalisée pendant la journée qui précède l'observation. Chaque plaque est exposée pendant 20 min, en incluant le temps d'acquisition des données et de lecture des CCD (Dawson et al. 2013). Les plaques sont observées plusieurs fois au cours de l'expérience et le nombre d'observation dépend de la qualité voulue des spectres.

#### 3.3.4.2 Traitement des données

La chaîne d'extraction de SDSS est appliquée en sortie des spectrographes et permet d'obtenir pour chaque objet observé, le flux et sa variance sur la bande spectrale de l'instrument, soit [360, 1000]nm. L'ensemble du processus est décrit dans Bolton et al. (2012).

Dans un premier temps, les signaux issus des différents CCD sont transformés en spectres à une dimension. Ces spectres sont ensuite calibrés en utilisant les mesures des étoiles standard. Le fond du ciel est estimé et soustrait par demi-plaque grâce à l'utilisation de fibres dédiées à l'observation du ciel. On estime ensuite la variance du flux en prenant en compte le nombre de photons détectés par les CCD, le bruit de mesure et les défauts connus de l'instrument.

Les spectres des expositions individuelles d'un même objet sont ensuite combinés pour former un spectre commun avec un rapport signal sur bruit plus grand. L'observation fractionnée d'un objet est privilégiée, car elle est moins sujette aux erreurs systématiques de mesure qu'une observation unique et plus longue. Typiquement, 4 à 6 expositions individuelles sont réalisées par objet.

Les mesures d'eBOSS se sont achevées en Février 2019 et seront rendues publiques en Décembre 2019. Dans l'étude que je présente dans cette thèse, j'utilise le relevé DR14 qui contient l'ensemble des données BOSS ainsi que les deux premières années d'eBOSS (Juillet 2014-Juillet 2016, [Abolfathi et al. 2018](#)).

## 3.4 Le catalogue des spectres des quasars

Dans cette section, je décris la constitution du catalogue de quasars uniquement ([Pâris et al. 2018](#)). Pour les autres traceurs, les méthodes sont analogues et décrites dans [Prakash & SDSS-IV/eBOSS \(2016\)](#) pour les LRG, et [Raichoor et al. \(2017\)](#) pour les ELG.

### 3.4.1 Identification des objets

Pour le relevé DR14, environ un million d'objets supposés être des quasars ont été observés et forment le *superset*. Or, on a vu précédemment que parmi toutes ces sources, seulement la moitié sont effectivement des quasars ([Dawson et al. 2016](#)). Dans cette sous-section, on décrit la phase d'identification des spectres qui aboutit à la création d'un catalogue de quasars avec une contamination de 1%, nécessaire pour satisfaire aux exigences de précision sur la mesure des distances ([Pâris et al. 2018](#)).

Premièrement, on applique sur l'ensemble du *superset* le processus d'identification SDSS décrit dans ([Bolton et al. 2012](#)). L'idée est d'ajuster, pour chaque spectre, des modèles standard de spectres d'étoile (STAR), de galaxies (GALAXY) et de quasars (QSO). La qualité des ajustements est mesurée par le paramètre  $\chi_{\min}^2$ , minimum de la fonction de vraisemblance. Plus  $\chi_{\min}^2$  est petit, plus l'ajustement est bon. À la fin de cette étape, on obtient pour chaque spectre une liste des cinq identifications les plus probables, c'est-à-dire celles correspondant aux cinq  $\chi_{\min}^2$  les plus petits, avec leur *redshift*  $z_{pip}$ . Pour les identifications QSO, un paramètre ZWARNING est affecté d'une valeur différente de zéro si la détermination du *redshift* est incertaine<sup>4</sup>.

La présence d'absorptions très importantes sur certains spectres de quasars rend leur identification difficile et entraîne une diminution de la pureté du catalogue de quasars par rapport aux galaxies ([Pâris et al. 2018](#)). Pour l'expérience BOSS, [Pâris et al. \(2012\)](#) ont vérifié visuellement

<sup>4</sup>On considère que la détermination du *redshift* est incertaine si la différence de vitesse entre le *redshift* de la première identification et le *redshift* de la seconde est supérieure à 1 000 km.s<sup>-1</sup>.

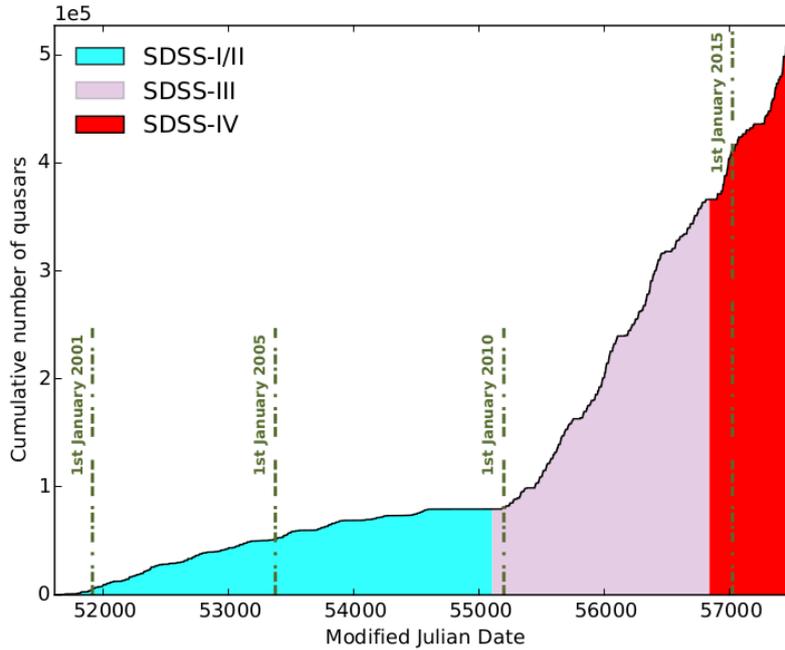


FIGURE 3.12: Nombre de quasars confirmés dans les relevés SDSS I, II, III et IV. Source : [Pâris et al. \(2018\)](#)

l'identification de chaque quasar en sortie du processus d'identification SDSS. Mais l'augmentation très importante du nombre de quasars (voir figure 3.12) dans l'expérience eBOSS a rendu impossible une inspection visuelle systématique. C'est pourquoi [Pâris et al. \(2018\)](#) ont développé un algorithme à appliquer en sortie du processus d'identification SDSS qui permet de sélectionner au préalable les spectres incertains requérant une inspection visuelle.

L'algorithme de tri de [Pâris et al. \(2018\)](#) est présenté figure 3.13. Il fait intervenir une série de conditions sur les cinq identifications les plus probables données par la chaîne d'extraction de SDSS. En sortie de cet algorithme, les objets sont répartis en quatre classes : STAR, GALAXY, QSO et UNKNOWN. Les objets de la classe UNKNOWN requièrent une inspection visuelle. Sur l'ensemble des 899 098 objets du *superset*, cela représente 42 729 spectres à visualiser. Sachant que 32 621 avaient déjà été observés lors des relevés SDSS I, II et III, il restait 10 108 spectres à classer visuellement pour eBOSS ([Pâris et al. 2018](#)).

Finalement, le catalogue de quasars, appelé DR14Q, est composé de 526 356 quasars dont 387 223 ont bénéficié d'une inspection visuelle lors des relevés précédents ou suite à la classification UNKNOWN par l'algorithme de tri. La pureté du catalogue DR14Q a été estimée sur les données du programme pilote SEQUELS<sup>5</sup> avec 36 489 objets. Elle est de 99.5%.

### 3.4.2 Détermination du *redshift*

Le spectre d'un quasar présente des raies d'émissions particulières créées par l'éjection de matière au voisinage du quasar. Parmi les raies les plus fortes (voir figure 3.14), on trouve les

<sup>5</sup>Sloan Extended Quasar, ELG, and LRG Survey (SEQUELS) est un programme auxiliaire de l'expérience BOSS, réalisé pour préparer eBOSS.

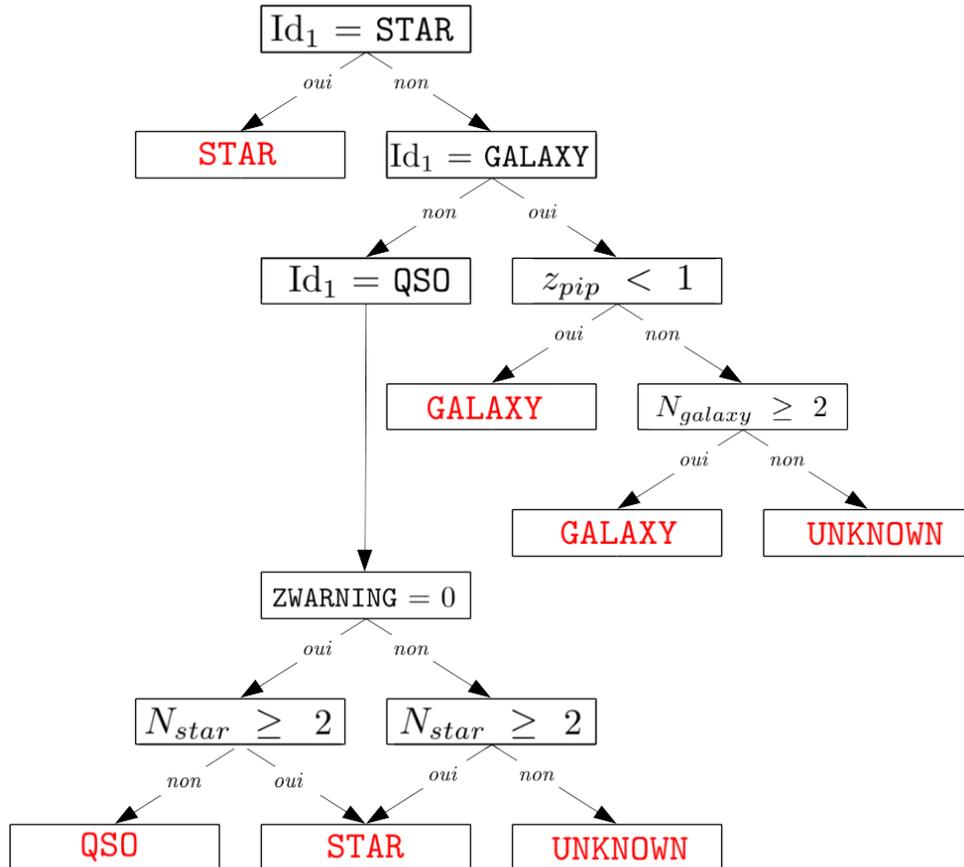


FIGURE 3.13: Schéma de l'algorithme décrit par Pâris et al. (2018)). Le processus d'identification SDSS (Bolton et al. 2012) donne une liste des cinq identifications les plus probables parmi STAR, GALAXY et QSO.  $Id_1$  désigne l'identification la plus probable et  $N_{galaxy}$  le nombre d'identifications GALAXY parmi les 4 suivantes. À la fin de l'algorithme, les objets sont répartis en quatre catégories : STAR, GALAXY, QSO et UNKNOWN. Les spectres UNKNOWN sont ensuite identifiés visuellement.

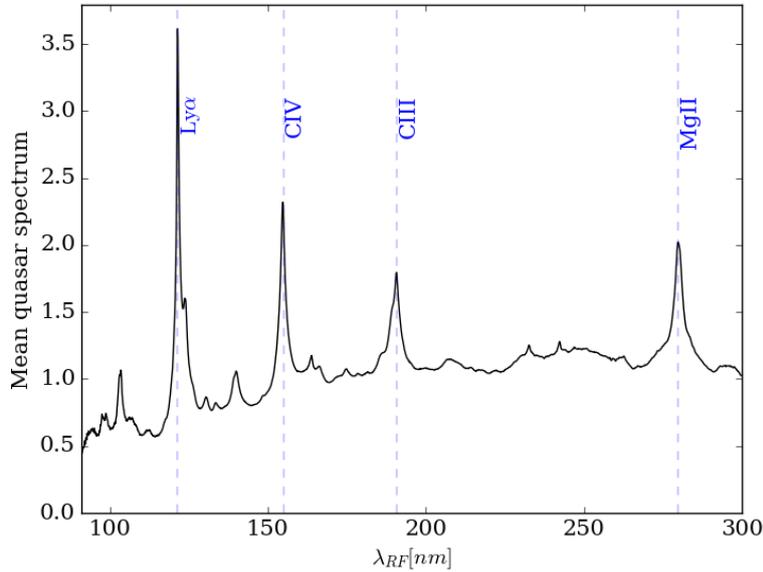


FIGURE 3.14: Spectre moyen des quasars de l'étude dans le référentiel au repos des quasars. On voit nettement apparaître les raies d'émission Ly $\alpha$ , MgII, CIV et CIII qui sont caractéristiques des spectres de quasar et permettent d'en identifier le *redshift*.

transitions Ly $\alpha$  à 1215.72 Å, CIV<sup>6</sup> à 1549.061 Å, CIII à 1908.734 Å et MgII à 2798.778 Å (Pâris et al. 2012). À cause de l'expansion de l'Univers et de la vitesse particulière du quasar, une raie émise à la longueur d'onde  $\lambda_e$  est observée à  $\lambda_e(1 + z_q)$ , où  $z_q$  est le *redshift* du quasar. Ainsi, en identifiant les raies d'émission d'un spectre, on est capable de déterminer son *redshift*.

Une analyse visuelle de tous les spectres de DR14Q serait trop longue, c'est pourquoi Pâris et al. (2012) ont développé une méthode automatisée de détermination du *redshift*. Dans un premier temps, on sélectionne un sous-échantillon d'environ 8 000 spectres avec un haut rapport signal sur bruit et représentatifs de la diversité des quasars de l'expérience<sup>7</sup>. On inspecte visuellement chaque spectre du sous-échantillon pour en déterminer le *redshift* et identifier les raies d'émission Ly $\alpha$ , MgII, CIV et CIII. Puis, grâce à une analyse en composantes principales (PCA dans la suite du manuscrit), chaque spectre de DR14Q est modélisé sous la forme d'une combinaison linéaire de 4 spectres du sous-échantillon. Le nombre de composantes a été choisi après plusieurs essais et assure que l'estimation du *redshift* soit robuste pour le maximum d'objets (Pâris et al. 2012). Les raies d'émission Ly $\alpha$ , MgII, CIV et CIII sont alors automatiquement identifiées et le *redshift* du quasar est déduit de leur position sur le spectre.

Dans les études spectroscopiques à venir (*e.g.*, DESI, 4MOST), le nombre de spectres observés sera trop important pour faire des vérifications visuelles. De nouvelles techniques d'identification, basées sur des réseaux de neurones, permettent d'analyser avec la même précision que Pâris et al.

<sup>6</sup>Dans la suite du manuscrit, j'adopte la convention astrophysique pour la notation des ions. Ainsi, je noterai CIV l'ion C<sup>3+</sup>.

<sup>7</sup>Pâris et al. (2012) imposent que le rapport signal sur bruit au carré soit supérieur à 5. Cette valeur faible permet de maximiser le nombre de quasars du sous-échantillon et ainsi, de s'assurer qu'il soit représentatif de la diversité spectrale de l'expérience. De plus, les spectres sélectionnés ne doivent pas présenter de BAL (voir sous-section 3.4.3).

(2018), de larges échantillons de données, sans aucune assistance humaine (Busca & Balland 2018).

### 3.4.3 Détection des *Broad Absorption Lines*

Certains quasars éjectent une grande quantité de matière qui absorbe une partie de leur flux lumineux au voisinage de ces derniers. Ce phénomène est visible sur les spectres sous la forme d'absorptions larges caractéristiques à gauche des pics d'émission, appelées *Broad Absorption Lines* et notées BAL dans la suite du manuscrit. La figure 3.15 représente, dans son référentiel au repos, le spectre d'un quasar avec des BAL. Ces absorptions sont identifiables à gauche des raies d'émission Ly $\alpha$  et CIV.

À cause de la composition chimique des gaz à proximité des quasars, les BAL sont surtout visibles pour la transition CIV. On les caractérise à l'aide de l'*indice de balnicité* (Weymann et al. 1991) qui mesure l'écart entre le spectre et un spectre moyen sans BAL. Plus l'indice de balnicité est élevé, plus l'absorption par les BAL est importante.

Nous verrons au chapitre 4, que les spectres utilisés dans notre étude ne doivent pas présenter une trop grande diversité. Or, les BAL dépendent du milieu environnant du quasar et sont donc très divers. C'est pourquoi, on va identifier les spectres avec BAL pour pouvoir ensuite les retirer de l'échantillon étudié.

Pour BOSS, tous les spectres ont été visualisés pour détecter les BAL. Mais cette opération n'était plus possible étant donné le nombre de spectres d'eBOSS. Pâris et al. (2018) ont mis en place une procédure automatisée, qui utilise l'analyse en composante principale décrite précédemment pour obtenir le spectre moyen sans BAL de chaque quasar et en déduire l'indice de balnicité. L'indice est ensuite enregistré dans le catalogue DR14Q sous l'entrée BI\_CIV. La condition BI\_CIV > 0 indique que le spectre présente des BAL.

#### Synthèse du chapitre

Dans ce chapitre, j'ai décrit le fonctionnement et la prise de données par l'instrument de eBOSS. Ces données sont ensuite traitées par la chaîne d'extraction de SDSS qui fournit, pour chaque objet observé, le flux calibré et sa variance sur l'ensemble du domaine spectral de eBOSS, soit [360,1000]nm. Une procédure automatisée permet ensuite de sélectionner les spectres de quasars et d'en identifier le *redshift*, ces informations sont ensuite stockées dans le catalogue de quasars DR14Q.

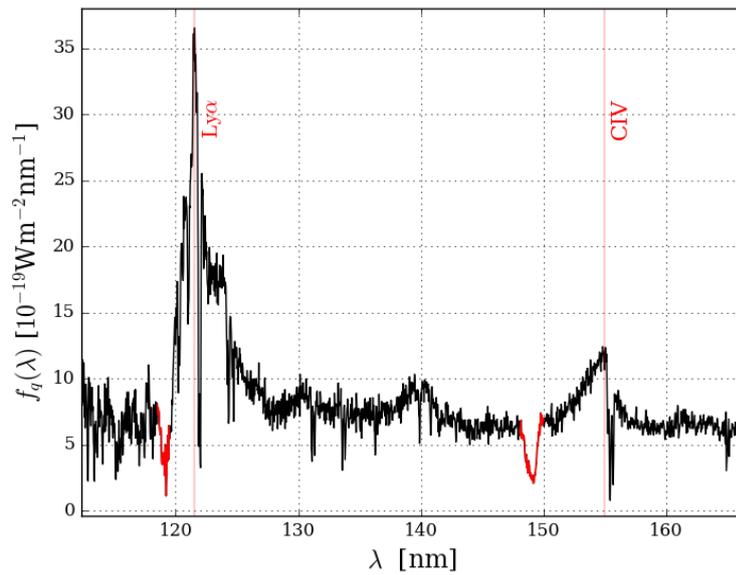


FIGURE 3.15: Spectre d'un quasar présentant des BAL (en rouge), dans son référentiel au repos. Ces absorptions caractéristiques sont clairement identifiables à gauche des raies d'émission  $\text{Ly}\alpha$  et  $\text{CIV}$ .

## Chapitre 4

# Du spectre au champ d'absorption

Dans le chapitre précédent, on a décrit la collecte et la calibration des spectres pour obtenir, pour chaque quasar  $q$  et longueur d'onde  $\lambda$ , le flux,  $f_q(\lambda)$ , et sa variance estimée par la chaîne d'extraction,  $\sigma_{f,q}^2(\lambda)$ . Dans le chapitre suivant, on calculera la fonction de corrélation Ly $\alpha$  comme une somme pondérée des absorptions par le milieu inter-galactique. Ce chapitre traite, pour sa part, du passage :

$$(f_q(\lambda), \sigma_{f,q}^2(\lambda)) \longrightarrow (\delta_q(\lambda), w_q(\lambda)), \quad (4.1)$$

où  $\delta_q(\lambda)$  est le champ d'absorption du couple  $(q, \lambda)$  et  $w_q(\lambda)$  son poids dans la fonction de corrélation (défini à la sous-section 4.1.3).

Cette section s'organise en cinq sous-sections :

- Dans la sous-section 4.1, je décris les modèles utilisés pour obtenir  $(f_q(\lambda), \sigma_{f,q}^2(\lambda))$ , et les raisons physiques et pratiques qui m'ont amenée à les choisir.
- Dans la sous-section 4.2, je détaille le processus de sélection des spectres et des régions spectrales à étudier.
- Dans la sous-section 4.3, je décris toutes les étapes qui permettent de calculer  $\delta_q(\lambda)$  et  $w_q(\lambda)$  à partir de  $f_q(\lambda)$  et  $\sigma_{f,q}^2(\lambda)$ .
- Le processus d'extraction des champs d'absorption corrèle les champs entre eux. Je décris dans la sous-section 4.4 comment contrôler ces effets sur la fonction de corrélation.
- Dans la sous-section 4.5, j'étudie la dépendance en *redshift* du champ d'absorption et j'introduis cette dépendance dans les poids  $w_q(\lambda)$ .

L'ensemble de ces étapes sont implémentées dans un module python, `picca`<sup>1</sup>, développé par le groupe Ly $\alpha$  d'eBOSS et présenté dans l'annexe C.

---

<sup>1</sup><https://github.com/igmhub/picca/tree/v3.DR14>

## 4.1 Description du modèle

Dans cette section, je décris la méthode à suivre pour obtenir  $(\delta_q(\lambda), w_q(\lambda))$  sans entrer dans les détails de la procédure qui seront exposés dans les sections suivantes.

### 4.1.1 Définition des régions Ly $\alpha$ et Ly $\beta$

Comme on l'a vu à la section 2.2, les absorptions Ly $\alpha$  sont visibles sur tout le domaine spectral aux longueurs d'onde plus grandes que la raie d'émission Ly $\alpha$  (voir figure 4.1). Les analyses antérieures à mon travail restreignent le domaine d'étude à la région Ly $\alpha$  localisée entre les pics d'émission Ly $\beta$  et Ly $\alpha$  pour éviter les contaminations par les autres transitions Lyman de l'hydrogène (Slosar et al. 2011, Busca et al. 2013, Delubac et al. 2015, Bautista et al. 2017). Dans la présente étude, j'inclus pour la première fois la région Ly $\beta$  entre les pics d'émission Ly $\gamma$  et Ly $\beta$ , ce qui permet d'augmenter d'environ 10% le nombre total des absorptions utilisées et de 30% le nombre de paires utilisées.

Dans ce chapitre, les deux régions sont traitées de façon complètement indépendante. Les outils développés pour l'extraction des champs d'absorption dans les régions Ly $\alpha$  et Ly $\beta$  sont les mêmes, seuls les domaines spectraux changent. On peut noter au passage que ces outils peuvent être également appliqués à d'autres forêts d'absorption comme celle du C IV (Blomqvist et al. 2018) ou du Mg II (du Mas des Bourboux et al. 2019).

### 4.1.2 Modéliser le champ d'absorption

La figure 4.2 représente les flux mesurés de deux quasars aux *redshifts*  $z = 2.80$  et  $z = 3.14$ . Dans les deux cas, on voit les raies d'émission Ly $\alpha$  et Ly $\beta$  qui ne sont pas observées aux mêmes longueurs d'onde puisque les *redshifts* des quasars sont différents. On remarque également les séries d'absorption Ly $\alpha$ , à gauche du pic d'émission Ly $\alpha$ , caractérisées par des variations rapides et de grande amplitude du flux lumineux.

Pour chaque spectre  $q$  à la longueur d'onde observée  $\lambda$ , le flux  $f_q$  d'une région (Ly $\alpha$  ou Ly $\beta$ ) peut s'écrire :

$$f_q(\lambda) = C_q(\lambda)(1 + \delta_q(\lambda)), \quad (4.2)$$

avec,

- une composante aux variations lentes : le flux moyen, noté  $C_q(\lambda)$  et appelé *continuum*,
- une composante aux variations rapides : le champ d'absorption, noté  $\delta_q(\lambda)$ , et appelé *champ de  $\delta$*  dans la suite.

Le champ de  $\delta$  est créé par les absorptions par la matière présente sur la ligne de visée. C'est la composante d'intérêt de cette étude.

Pour calculer  $\delta_q(\lambda)$ , il faut d'abord modéliser  $C_q(\lambda)$ . Physiquement, le flux moyen est le produit du flux émis par le quasar et de la transmission moyenne du milieu intergalactique dans la ligne de visée. On l'obtient en suivant deux étapes :

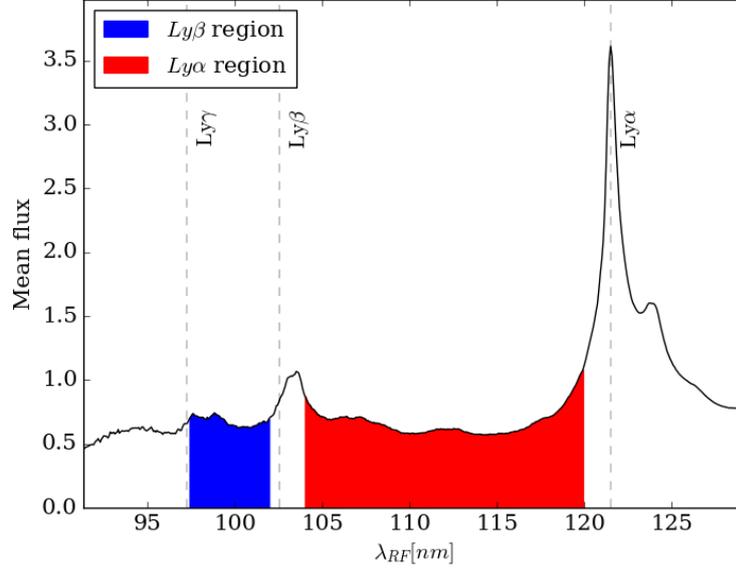


FIGURE 4.1: Moyenne, sur l'ensemble des quasars utilisés dans cette étude, des spectres dans le référentiel de chaque quasar. On représente les raies d'émission  $\text{Ly}\alpha$ ,  $\text{Ly}\beta$  et  $\text{Ly}\gamma$  qui correspondent à l'émission de ces transitions à proximité du quasar. La région  $\text{Ly}\alpha$  (en rouge) est localisée entre les pics  $\text{Ly}\beta$  et  $\text{Ly}\alpha$ , et la région  $\text{Ly}\beta$ , entre les pics  $\text{Ly}\gamma$  et  $\text{Ly}\beta$ .

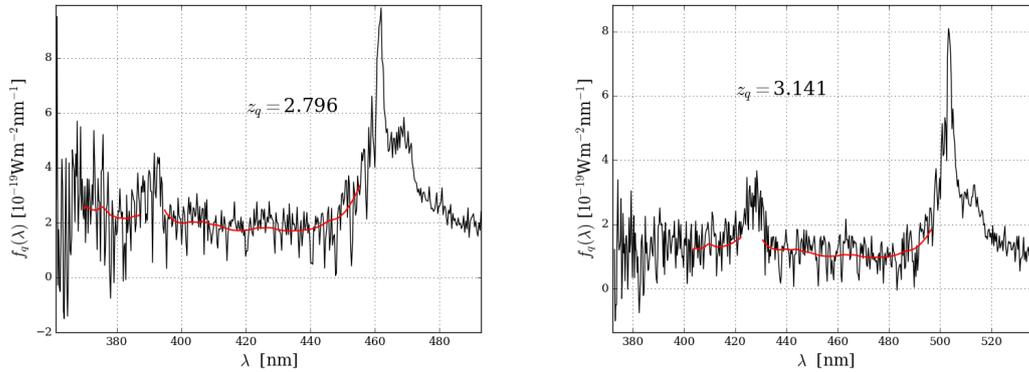


FIGURE 4.2: Flux mesurés de deux quasars aux *redshifts*  $z = 2.796$  et  $z = 3.141$ . Les continums modélisés des régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  sont représentés en rouge.

- **Étape 1** : On calcule le flux moyen d'un quasar  $\bar{f}$  en faisant la moyenne de tous les spectres dans le référentiel au repos.
- **Étape 2** : On ajuste sur chaque spectre un ou plusieurs paramètres pour rendre compte de sa spécificité.

#### 4.1.2.1 Étape 1

Dans un premier temps, on déplace chaque flux dans le référentiel au repos du quasar. Autrement dit, pour chaque spectre  $q$ , on réalise la transformation :

$$\lambda \longrightarrow \lambda_{\text{RF}} = \frac{\lambda}{1 + z_q}. \quad (4.3)$$

Le flux moyen  $\bar{f}$ , représenté figure 4.1, est la moyenne de tous les spectres de DR14Q dans leur référentiel au repos (voir équation 4.21).

#### 4.1.2.2 Étape 2

Dans un deuxième temps, on prend en compte la diversité spectrale des quasars. On reprend la forme proposée et validée par l'utilisation de simulations par [Bautista et al. \(2017\)](#) pour la modéliser :

$$C_q(\lambda) = (a_q + b_q \Lambda) \bar{f} \left( \frac{\lambda}{1 + z_q} \right) \quad \text{avec} \quad \Lambda = \log_{10} \lambda. \quad (4.4)$$

Le paramètre  $a_q$  décrit l'amplitude du flux dans la région étudiée ( $\text{Ly}\alpha$  ou  $\text{Ly}\beta$ ) et  $b_q$  traduit la variation de cette amplitude autour de la moyenne. On choisit de travailler avec une échelle logarithmique pour faciliter le passage du référentiel du quasar à celui de l'observateur qui se fait alors par une simple addition (*cf.* équation 4.3).

On représente sur la figure 4.2 les continus des régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  des deux spectres. Visuellement, le continuum correspond à la courbe qui passe par le milieu de toutes les absorptions.

Finalement, le champ de  $\delta$  s'exprime par :

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{C_q(\lambda)} - 1. \quad (4.5)$$

#### 4.1.3 L'estimation des poids

Le calcul de la fonction de corrélation nécessite de pondérer chaque élément  $\delta_q(\lambda)$ . L'estimateur des poids le plus adapté à la détection du pic BAO est celui qui maximise le signal sur bruit de la fonction de corrélation. Un tel estimateur est proposé par [McQuinn & White \(2011\)](#) et fait intervenir la variance des  $\delta_q(\lambda)$  ainsi qu'un terme dépendant de l'angle de visée de chaque quasar. Pour des raisons pratiques de temps de calcul, nous en utilisons une version approximée qui donne néanmoins une précision suffisante pour la mesure du pic BAO ([Busca et al. 2013](#)). On privilégie les  $\delta_q(\lambda)$  avec une variance faible en utilisant des poids donnés par l'inverse de la

variance des champs d'absorption :

$$w_q(\lambda) = 1/\sigma_q^2(\lambda). \quad (4.6)$$

La variance du champ d'absorption  $\sigma_q^2(\lambda)$  est la somme :

- d'une contribution instrumentale,  $\sigma_{q,inst}^2(\lambda)$  liée à l'incertitude de la mesure des flux.
- d'une contribution cosmologique,  $\sigma_{cosmo}^2(\lambda)$ , créée par la variation de densité des absorbants dans le milieu intergalactique. Puisque l'Univers est isotrope,  $\sigma_{cosmo}^2(\lambda)$  ne dépend pas de la direction d'observation mais seulement de la longueur d'onde.

En utilisant l'équation 4.5, la variance instrumentale du champ d'absorption s'écrit :

$$\sigma_{q,inst}^2(\lambda) = \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2}. \quad (4.7)$$

Le calcul de la variance du champ d'absorption hors des forêts Ly $\alpha$  montre que celle-ci est sous-estimée par la chaîne d'extraction SDSS pour les longueurs d'onde en deçà de 570 nm et surestimée au-delà (Palanque-Delabrouille et al. 2013, Lee et al. 2013). Bautista et al. (2017) modélisent cet effet en multipliant  $\sigma_{q,inst}^2$  par une fonction  $\eta(\lambda)$  dépendant de la longueur d'onde. L'équation 4.7 devient :

$$\sigma_{q,inst}^2(\lambda) = \eta(\lambda) \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2}. \quad (4.8)$$

Blomqvist et al. (2018) remarquent que pour les spectres à haut signal sur bruit, c'est-à-dire pour  $\sigma_{f,q}^2$  faible, la variance du champ de  $\delta$  est sous-estimée par l'expression 4.8. Ils ajoutent alors le terme,  $\epsilon(\lambda)$ , qui augmente la variance pour les  $\sigma_{f,q}^2$  faibles :

$$\sigma_{q,inst}^2(\lambda) = \eta(\lambda) \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2} + \epsilon(\lambda) \left( \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2} \right)^{-1}. \quad (4.9)$$

Finalement, la variance du champ d'absorption est modélisée par :

$$\sigma_q^2(\lambda) = \eta(\lambda) \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2} + \epsilon(\lambda) \left( \frac{\sigma_{f,q}^2(\lambda)}{C_q(\lambda)^2} \right)^{-1} + \sigma_{cosmo}^2(\lambda) \quad (4.10)$$

Les fonctions  $\eta$ ,  $\epsilon$  et  $\sigma_{cosmo}^2$  seront ajustées sur la variance des  $\delta$  à la section 4.3.

## 4.2 Préparation des spectres

Le catalogue de quasars DR14 décrit par Pâris et al. (2018) contient 526 356 objets. À l'issue de la phase de sélection exposée dans cette section, plus de la moitié des spectres est éliminée.

La phase de préparation des spectres s'effectue en trois temps :

TABLE 4.1: Liste des sélections opérées sur les spectres des quasars du catalogue DR14Q. Ne sont gardés que les objets clairement identifiés, dans la gamme de *redshift* correspondant à la région étudiée, sans BAL, d'une taille supérieure à 50 pixels et dont l'ajustement du continuum a fonctionné.

Conditions	Nombre de régions Ly $\alpha$	Nombre de régions Ly $\beta$
DR14	526 356	526 356
and thid>0	446 137	446 137
and z>zmin	225 366	95 449
and z<zmax	216 162	86 245
and BI_CIV<=0	201 286	80 443
and npixels > 50	182 529	57 105
and continuum > 0	179 965	56 154

TABLE 4.2: Définition des domaines spectraux des régions Ly $\alpha$  et Ly $\beta$  et gamme en *redshift* des quasars dont les spectres contiennent ces régions.

Region	$z_q$	$\lambda_{Obs}$ [nm]	$\lambda_{RF}$ [nm]
Ly $\beta$	[2.53,3.5]	[360,459]	[97.4,102]
Ly $\alpha$	[2.0,3.5]	[360,540]	[104,120]

TABLE 4.3: Liste des raies d'absorption et des raies du ciel fournie par la chaîne d'extraction SDSS (<http://classic.sdss.org/dr6/algorithms/linestable.html>). Les deux dernières colonnes donnent les intervalles de longueurs d'onde observées masquées sur tous les spectres de cette étude.

Raie	Type	$\lambda^{min}$ [nm]	$\lambda^{max}$ [nm]
Ca(K)	absorption	393.1	393.7
Ca(H)	absorption	396.6	397.2
	ciel	404.2	405.4
	ciel	435.5	436.5

- On extrait de chaque spectre la région voulue ( $\text{Ly}\alpha$  ou  $\text{Ly}\beta$ ) via une série de coupures sur les longueurs d'onde observées et dans le référentiel au repos du quasar.
- On sélectionne les spectres pour lesquels suffisamment de *pixels* sont disponibles dans la région étudiée ( $\text{Ly}\alpha$  ou  $\text{Ly}\beta$ ). Dans la suite de ce chapitre, on appelle *pixel* l'élément défini par  $(q, \lambda)$  avec  $q$  l'identification du quasar et  $\lambda$  la longueur d'onde observée sur le spectre. Typiquement, au moins 50 *pixels*.
- Pour chaque spectre sélectionné, on masque les raies du ciel et les absorptions saturées présentes dans la région étudiée.

## 4.2.1 Sélection des spectres

### 4.2.1.1 Limites des régions $\text{Ly}\alpha$ et $\text{Ly}\beta$

Dans chaque spectre, la région  $\text{Ly}\alpha$  est ceinte des pics d'émission  $\text{Ly}\beta$  et  $\text{Ly}\alpha$  (figure 4.1). L'amplitude et la largeur des pics d'émission dépendent de la distribution de gaz au voisinage du quasar et sont soumises à une grande variabilité qui se traduit par une augmentation brusque de la variance des  $\delta$  à proximité des pics d'émission, comme on peut le voir sur la figure 4.3. Pour s'émanciper de la forme particulière de chaque quasar, on supprime les régions spectrales pour lesquelles la variance du champ de  $\delta$ , calculée sur l'ensemble des spectres, évolue très rapidement, ce sont les zones en rouge sur la figure 4.3.

Finalement, la région  $\text{Ly}\alpha$  est définie pour chaque spectre  $q$  dans le référentiel propre du quasar par :

$$\text{Ly}\alpha : \lambda_{\text{RF}} \equiv \frac{\lambda}{1+z_q} \in [104.0 \text{ nm}, 120.0 \text{ nm}]. \quad (4.11)$$

De façon équivalente, on définit la région  $\text{Ly}\beta$ , ceinte des pics d'émission  $\text{Ly}\gamma$  et  $\text{Ly}\beta$ , par :

$$\text{Ly}\beta : \lambda_{\text{RF}} \equiv \frac{\lambda}{1+z_q} \in [97.4 \text{ nm}, 101.6 \text{ nm}]. \quad (4.12)$$

### 4.2.1.2 Tri des spectres

La première phase de tri consiste à supprimer les objets dont l'identification est incertaine. Le catalogue DR14Q les distingue par l'expression `thid == -1` (section 3.4). 80 119 spectres sont ainsi éliminés (tableau 4.1).

On supprime également 9 204 quasars dont la *redshift* est supérieur à 3.5 car la densité d'objets au delà chute fortement (Pâris et al. 2018) et n'est plus suffisante pour mesurer précisément la fonction de corrélation (Bautista et al. 2017). On pose :

$$z_{q,\text{max}} = 3.5, \quad (4.13)$$

ce qui induit la condition suivante sur la limite haute,  $\lambda_{\text{max}}$ , des longueurs d'onde observées :

$$\lambda_{\text{max}} = (1+z_{q,\text{max}})\lambda_{\text{RF},\text{min}} \quad (4.14)$$

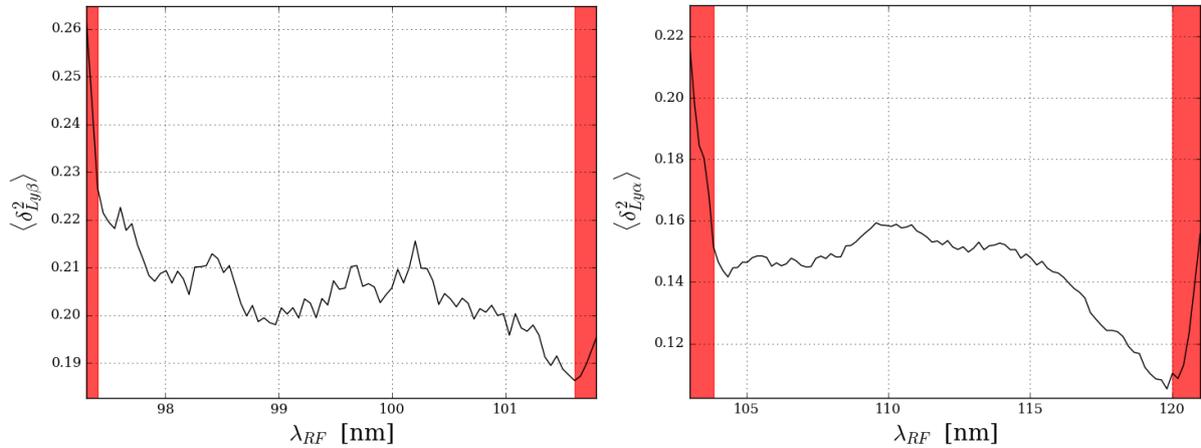


FIGURE 4.3: On sélectionne les domaines spectraux des régions  $\text{Ly}\beta$  (à gauche) et  $\text{Ly}\alpha$  (à droite) en vue de minimiser la variance des  $\delta$ . Cela implique de rester loin des pics d'émission dont la forme varie fortement d'un quasar à l'autre.

En deçà de 360 nm, l'efficacité des CCD devient faible (section 3.3.4) et l'atmosphère très opaque. On impose donc que la longueur d'onde observée  $\lambda$  soit supérieure à 360 nm :  $\lambda_{\min} = 360$  nm.

Pour pouvoir être visible sur les spectres au delà de 360 nm, la région  $\text{Ly}\alpha$  doit être suffisamment décalée vers le rouge. Cela impose une limite minimale sur le *redshift* des quasars sélectionnés :

$$z_{q,\min} + 1 = \frac{\lambda_{\min}}{\lambda_{RF,\max}} \quad (4.15)$$

Les valeurs des différentes limites obtenues pour les régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  sont regroupées dans le tableau 4.2.

Certains quasars éjectent une grande quantité de matière qui absorbe une partie du flux aux abords des lignes d'émission. Leurs spectres présentent de larges raies d'absorption, les BAL (sous-section 3.4.3), dont la largeur et l'amplitude dépendent du volume et de la vitesse de l'éjecta. La forme particulière et imprévisible de ces spectres leur vaut d'être éliminés de l'échantillon étudié.

Pour limiter le temps de calcul et le volume des données, on divise le nombre de longueurs d'onde observées en regroupant, pour chaque spectre, les pixels adjacents trois par trois (Bautista et al. 2017). La résolution spectrale est alors d'environ 1 Å par pixel.

Les spectres contenant moins de 50 pixels sont supprimés car ils n'offrent pas un bras de levier suffisant pour l'ajustement du continuum. On rejette également les spectres pour lesquels le continuum ajusté est négatif.

La table 4.1 donne, pour chaque phase de tri, le nombre de spectres éliminés dans les  $\text{Ly}\alpha$  et  $\text{Ly}\beta$ .

## 4.2.2 Sélection des longueurs d'onde

### 4.2.2.1 Masquage des lignes du ciel

Les éléments chimiques présents dans l'atmosphère créent des raies d'émission et d'absorption visibles sur les spectres. Elles sont identifiées grâce à l'utilisation de fibres dédiées (section 3.2.3) et répertoriées dans le tableau 4.3. Ces régions spectrales sont masquées dans les spectres, c'est-à-dire que leurs poids sont mis à zéro.

### 4.2.2.2 Traitement des *Damped Lyman- $\alpha$ systems*

Certains nuages sont si denses en hydrogène neutre qu'ils absorbent une grande partie du flux lumineux émis par le quasar sur des régions spectrales très étendues, parfois une dizaine de nanomètres. À la différence des BAL, ils sont localisés dans le milieu intergalactique et sont le lieu de formation des futures galaxies. Ces objets sont appelés *Damped Ly $\alpha$  systems* (DLA dans la suite du manuscrit) car leurs absorptions présentent des bords à l'aspect amorti (*damped* en anglais). Physiquement, leur densité de colonne d'hydrogène neutre est supérieure à  $10^{21}\text{cm}^{-2}$  (Noterdaeme et al. 2012).

Les DLA visibles dans les spectres des quasars de DR14Q sont identifiés visuellement puis classés automatiquement suivant la procédure décrite par Noterdaeme et al. (2012). Pour chacun des 34 541 objets du catalogue de DLA, on répertorie son *redshift* et sa densité de colonne puis on en déduit le flux absorbé en utilisant un profil de Voigt. Les régions dans lesquelles le flux transmis est inférieur à 80% du flux moyen sont ensuite masquées, comme on peut le voir sur la figure 4.4. De façon équivalente, certaines structures peuvent créer des absorptions Ly $\beta$  aux profils amortis qui sont prises en compte dans le traitement des régions Ly $\beta$  et traitées d'une manière analogue aux DLA.

## 4.3 Calcul du champ d'absorption

Dans cette section, on détaille le processus itératif utilisé pour ajuster le continuum (équation 4.4) de la région Ly $\alpha$  de chaque spectre ainsi que les différents termes décrits dans l'équation 4.10 (voir la routine `do_deltas.py` dans l'annexe C).

On commence par initialiser les valeurs des paramètres libres puis on fait des ajustements successifs jusqu'à obtention d'une solution stable (en pratique, après cinq itérations).

### 4.3.1 Initialisation

On initialise les coefficients de diversité spectrale ( $a_q, b_q$ ) de chacun des 182 529 spectres sélectionnés de la façon suivante :

$$a_q = \frac{\sum_{\lambda} w_q(\lambda) f_q(\lambda)}{\sum_{\lambda} w_q(\lambda)} \quad \text{et} \quad b_q = 0. \quad (4.16)$$

Les valeurs d'initialisation pour les paramètres  $\eta$ ,  $\sigma_{cosmo}^2$  et  $\epsilon$  de l'équation 4.10 sont choisies selon l'importance de leur contribution (voir figure 4.5). Pour chacune des longueurs d'onde observées dans la région Ly $\alpha$ ,

$$\eta(\lambda) = 1, \quad \sigma_{cosmo}^2(\lambda) = 0.2 \quad \text{et} \quad \epsilon(\lambda) = 0. \quad (4.17)$$

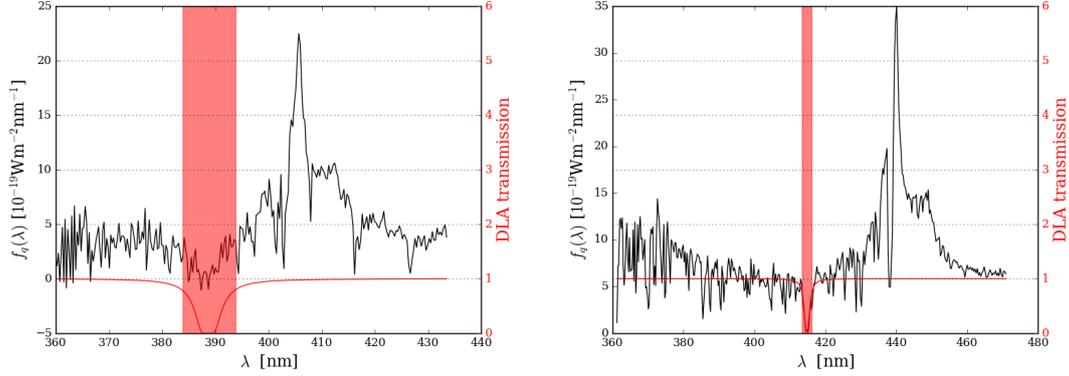


FIGURE 4.4: Flux mesurés de deux quasars présentant des DLA (en noir). En rouge, on représente le flux transmis calculé grâce à un profil de Voigt. Les régions pour lesquelles le flux transmis est inférieur à 80% du flux moyen sont masquées (bandes rouges sur les graphes).

Pour chacune des longueurs d'onde dans le référentiel propre du quasar, on prend :

$$\bar{f}(\lambda_{RF}) = 1. \quad (4.18)$$

### 4.3.2 Relation de récurrence

Le processus d'ajustement des variables se fait en trois étapes :

- pour chaque quasar  $q$ , un ajustement sur l'ensemble de la région spectrale de  $(a_q, b_q)$  est réalisé,
- puis on calcule le flux moyen  $\bar{f}(\lambda_{RF})$  dans le référentiel de chaque quasar,
- Enfin, pour chaque longueur d'onde observée  $\lambda$ , un ajustement sur l'ensemble des quasars de  $\eta(\lambda)$ ,  $\sigma_{cosmo}^2(\lambda)$  et  $\epsilon(\lambda)$  est réalisé.

Le nombre d'itérations correspond à l'indice  $l$  minimal pour lequel  $|\langle g_{l+1} - g_l \rangle| < 10^{-7}$  où  $g_l$  désigne, à l'itération  $l$ , tour à tour chacune des fonctions à ajuster. On constate que cinq itérations suffisent pour obtenir la convergence avec ce critère.

#### 4.3.2.1 Ajustement de $(a_q, b_q)$ sur le flux $f_q$

Le modèle du continuum du quasar  $q$ ,  $C_q(\lambda)$  (équation 4.4), doit être ajusté sur le flux  $f_q(\lambda)$  dont chaque élément est pondéré par l'inverse de la variance du flux :

$$w_q^f(\lambda) = 1/\sigma_{f,q}^2(\lambda). \quad (4.19)$$

Cela revient à minimiser par rapport au couple  $(a_q, b_q)$  la fonction de vraisemblance :

$$\chi_q^2(a_q, b_q) = \frac{\sum_{\lambda} w_q^f(\lambda) (f_q(\lambda) - (a_q + b_q \Lambda) \bar{f}(\lambda_{RF}))^2}{\sum_{\lambda} w_q^f(\lambda)} \quad (4.20)$$

Nous utilisons un modèle linéaire pour réaliser cette minimisation.

#### 4.3.2.2 Calcul du flux moyen $\bar{f}$

Le flux moyen  $\bar{f}$  calculé sur l'ensemble des spectres (voir figure 4.1) est défini par :

$$\bar{f}(\lambda_{\text{RF}}) = \frac{\sum_q f_q(\lambda_{\text{RF}})/\sigma_{f,q}^2(\lambda_{\text{RF}})}{\sum_q 1/\sigma_{f,q}^2(\lambda_{\text{RF}})}. \quad (4.21)$$

Pour réaliser ce calcul, on interpole chaque spectre dans le référentiel au repos du quasar puis on fait la moyenne de tous les spectres.

#### 4.3.2.3 Ajustement de $\eta$ , $\sigma_{\text{cosmo}}^2$ et $\epsilon$ sur la variance des $\delta$

On subdivise les variances du flux en 100 intervalles entre les valeurs  $10^{-5}$  et 2. Puis, on partitionne les spectres en autant de sous-groupes  $S_\lambda(\sigma_j^2)$ . À  $\lambda$  fixé, le spectre  $q$  appartient au sous-groupe  $S_\lambda(\sigma_j^2)$  si et seulement si la variance du flux  $\sigma_{f,q}^2(\lambda)$  est comprise entre  $\sigma_j^2$  et  $\sigma_{j+1}^2$ . Dans chaque sous groupe  $S_\lambda(\sigma_j^2)$ , on calcule les moyennes de  $\delta$  (donné par l'équation 4.5),  $\delta^2$  et  $\delta^4$  suivant la formule :

$$\langle \delta^n(\lambda) \rangle_{\sigma_j^2} = \frac{1}{N_{S_\lambda(\sigma_j^2)}} \sum_{q \in S_\lambda(\sigma_j^2)} \delta_q^n(\lambda), \quad (4.22)$$

avec  $N_{S_\lambda(\sigma_j^2)}$  le nombre d'éléments dans  $S_\lambda(\sigma_j^2)$ .

On peut ainsi calculer la variance des  $\delta$  du sous-groupe  $S_\lambda(\sigma_j^2)$  :

$$\mu_{2,\delta}(\lambda, \sigma_j^2) = \langle \delta^2(\lambda) \rangle_{\sigma_j^2} - \langle \delta(\lambda) \rangle_{\sigma_j^2}^2, \quad (4.23)$$

puis son moment d'ordre 4 :

$$\mu_{4,\delta}(\lambda, \sigma_j^2) = \langle \delta^4(\lambda) \rangle_{\sigma_j^2} - \mu_{2,\delta}(\lambda, \sigma_j^2)^2. \quad (4.24)$$

À  $\lambda$  fixé, la somme  $\eta(\lambda)\sigma^2 + \sigma_{\text{cosmo}}^2(\lambda) + \frac{\epsilon(\lambda)}{\sigma^2}$  doit reproduire au mieux la variance des  $\delta$  définie par l'équation 4.23. Cela revient à minimiser pour chaque triplet  $(\eta(\lambda), \sigma_{\text{cosmo}}^2(\lambda), \epsilon(\lambda))$  la fonction de vraisemblance suivante :

$$\chi_\lambda^2(\eta(\lambda), \sigma_{\text{cosmo}}^2(\lambda), \epsilon(\lambda)) = \sum_{\sigma^2} \left[ \left( \mu_{2,\delta}(\lambda, \sigma^2) - \left( \eta(\lambda)\sigma^2 + \sigma_{\text{cosmo}}^2(\lambda) + \frac{\epsilon(\lambda)}{\sigma^2} \right) \right)^2 / \mu_{4,\delta}(\lambda, \sigma^2) \right] \quad (4.25)$$

#### 4.3.3 Extraction du champ de $\delta$

À la fin du processus itératif, on estime la variance du champ d'absorption via l'équation 4.23. On représente en couleur sur la figure 4.5 les trois contributions de l'estimateur de la variance des  $\delta$  pour différentes valeurs de  $\sigma_{f,q}^2(\lambda)$ . En noir, on trace la variance des  $\delta$  définie par l'équation 4.23 à  $\sigma_{f,q}^2(\lambda)$  fixé. La contribution cosmologique (en bleu) est identique dans les trois cas puisqu'elle ne dépend que de la longueur d'onde. À bas signal sur bruit, la modélisation de [Bautista et al. \(2017\)](#) (le premier et le dernier terme de l'équation 4.10, en cyan sur la figure) est suffisante. Mais pas à haut signal sur bruit, où la différence entre les courbes noire et cyan est importante et justifie l'ajout du troisième terme (en rouge).

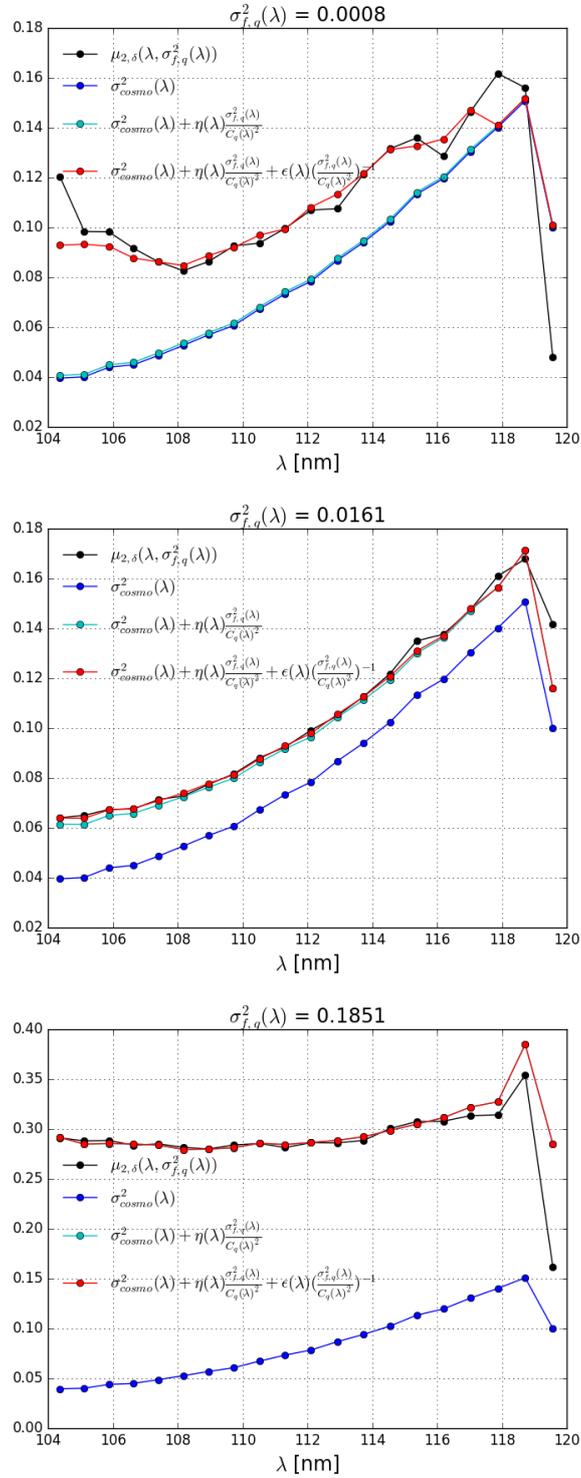


FIGURE 4.5: Variance mesurée du champ de  $\delta$  (en noir) en fonction de la longueur d'onde observée pour différentes valeurs de la variance du flux  $\sigma^2_{f,q} = \sigma^2$ . On compare la variance du champ de  $\delta$  avec les différents termes du modèle donné par l'équation 4.10 (en couleur).

À la fin des 5 itérations de la relation de récurrence, on obtient les valeurs ajustées des paramètres  $a_q$ ,  $b_q$  pour chaque quasar. On en déduit le champ d'absorption  $\delta_q$  via l'équation 4.5. Nous verrons dans le chapitre 5, que l'ajustement des paramètres  $a_q$  et  $b_q$ , qui est réalisée sur l'ensemble des longueurs d'onde d'un même spectre, a pour effet de corrélérer les pixels voisins du champ de  $\delta$ . Dans la section 4.4, nous allons voir comment le champ de  $\delta$  peut être transformé de façon à modéliser ces corrélations.

## 4.4 Prise en compte des effets liés à l'ajustement du continuum

L'ajustement du continuum sur l'ensemble des spectres et sa soustraction corréntent les éléments des champs de  $\delta$  et impactent la fonction de corrélation. Dans les premières analyses (Busca et al. 2013, Delubac et al. 2015) des simulations des champs d'absorptions Ly $\alpha$  ont montré que cet effet était visible sur l'ensemble de la fonction de corrélation et pouvait être pris en compte par l'ajout d'un terme purement mathématique dans le modèle. De façon à s'assurer de l'indépendance de la position du pic BAO vis-à-vis du modèle de continuum, plusieurs formes ont été testées (fonctions linéaire ou exponentielle) ne montrant aucune différence significative dans les résultats.

Bautista et al. (2017) introduisent une nouvelle façon de traiter l'effet de l'ajustement du continuum qui se déroule en deux étapes :

- On cherche un projecteur  $\mathbb{P}$  à appliquer sur le champ de  $\delta$  mesuré de façon à ce que la projection du champ mesuré égale la projection du champ vrai.
- On modélise l'effet de la projection sur la fonction de corrélation par une matrice, *la matrice de distorsion*, qui est appliquée au modèle.

Dans cette section, on exposera la première étape. La seconde étape sera discutée dans le chapitre 5.

Dans un premier temps, on montre que la différence entre le champ mesuré et le champ vrai peut s'écrire sous la forme d'une fonction affine de  $\lambda$ . On en déduira la forme du projecteur  $\mathbb{P}$  qui sera ensuite appliqué à l'ensemble des champs  $\delta_q(\lambda)$ .

### 4.4.1 Champ vrai et champ mesuré

Dans cette section, on note  $\delta^m$  le champ mesuré précédemment en utilisant l'ajustement du continuum, et  $\delta^v$  le champ vrai, *i.e.* le champ d'absorption réellement présent dans le milieu inter-galactique.

Pour simplifier les calculs, les paramètres spectraux  $(a_q, b_q)$  de l'équation 4.4 peuvent être redéfinis de façon à avoir :

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{f(\lambda_{RF})(a_q + b_q(\Lambda - \Lambda_q))} - 1, \quad (4.26)$$

avec  $\overline{\Lambda}_q = \frac{\sum_{\lambda} w_q(\lambda)\Lambda}{\sum_{\lambda} w_q(\lambda)}$ . Dans la suite, les barres désignent la moyenne dans une région d'absorption (Ly $\alpha$  ou Ly $\beta$ ), c'est-à-dire pour un quasar donné, la moyenne sur les longueurs d'onde.

L'étude de simulations du champ d'absorption démontre que l'ajustement réalisé section 4.3 est proche de la réalité physique (Bautista et al. 2017), c'est-à-dire que les  $(a_q, b_q)$  mesurés sont proches des valeurs vraies  $(a_q^v, b_q^v)$ . Cela nous permet de modéliser  $a_q$  et  $b_q$  comme des petites perturbations autour des valeurs vraies :

$$\begin{cases} a_q = a_q^v + \delta a_q \\ b_q = b_q^v + \delta b_q \end{cases} \quad (4.27)$$

avec  $(\delta a_q, \delta b_q)$  les perturbations d'ordre 1 des paramètres  $(a_q, b_q)$ .

En introduisant ces expressions dans l'équation 4.26, puis en faisant un développement limité au premier ordre en  $(\delta a_q, \delta b_q)$ , on obtient :

$$\delta_q^m(\lambda) = \delta_q^v(\lambda) - A_q - B_q(\Lambda - \overline{\Lambda}_q), \quad (4.28)$$

avec  $A_q$  et  $B_q$  des coefficients propres à chaque quasar.

Soit  $n$  le nombre de longueurs d'onde dans la région étudiée (Ly $\alpha$  ou Ly $\beta$ ), l'équation 4.28 peut s'écrire sous forme vectorielle :

$$\underbrace{\begin{pmatrix} \delta_q^m(\lambda_1) \\ \vdots \\ \delta_q^m(\lambda_n) \end{pmatrix}}_{\delta_q^m} = \underbrace{\begin{pmatrix} \delta_q^v(\lambda_1) \\ \vdots \\ \delta_q^v(\lambda_n) \end{pmatrix}}_{\delta_q^v} - A_q \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}}_{v_1} - B_q \underbrace{\begin{pmatrix} \Lambda_1 - \overline{\Lambda}_q \\ \vdots \\ \Lambda_n - \overline{\Lambda}_q \end{pmatrix}}_{v_2} \quad (4.29)$$

$$\Rightarrow \boxed{\delta_q^m = \delta_q^v - A_q v_1 - B_q v_2.} \quad (4.30)$$

En définissant le produit scalaire  $u \cdot v$  par :

$$u \cdot v = u^T v = \frac{\sum_k w_q(\lambda_k) u_k v_k}{\sum_k w_q(\lambda_k)}, \quad (4.31)$$

les normes des vecteurs  $v_1$  et  $v_2$  sont données par :

$$\|v_1\| = 1 \quad \text{et} \quad \|v_2\| = \frac{\sum_k w_q(\lambda_k) (\Lambda_k - \overline{\Lambda}_q)^2}{\sum_k w_q(\lambda_k)}.$$

Dans la sous-section suivante, les vecteurs  $v_1$  et  $v_2$  vont permettre de définir le projecteur tel que la projection du champ mesuré égale la projection du champ vrai.

#### 4.4.2 Projecteur $\mathbb{P}$

On cherche le projecteur  $\mathbb{P}$  tel que :

$$\mathbb{P} \delta_q^m = \mathbb{P} \delta_q^v, \quad (4.32)$$

En utilisant 4.30, cela revient à écrire :

$$\mathbb{P}(A_q v_1 + B_q v_2) = 0. \quad (4.33)$$

Or, on a :

$$v_1 \cdot v_2 = \frac{\sum_k w_q(\lambda_k)(\Lambda_i - \bar{\Lambda}_q)}{\sum_k w_q(\lambda_k)} = \frac{\sum_k w_q(\lambda_k)\Lambda_i}{\sum_k w_q(\lambda_k)} - \bar{\Lambda}_q = 0,$$

Puisque  $v_1$  et  $v_2$  sont perpendiculaires, le projecteur  $\mathbb{P}$  qui transforme un vecteur quelconque  $u$  en  $\tilde{u}$  est donné par :

$$\tilde{u}_i = [\mathbb{P}u]_i = u_i - \frac{v_{1,i}(v_1^T u)}{\|v_1\|^2} - \frac{v_{2,i}(v_2^T u)}{\|v_2\|^2}, \quad (4.34)$$

d'où,

$$\tilde{u}_i = u_i - \frac{\sum_k w_q(\lambda_k)u_k}{\sum_k w_q(\lambda_k)} - \frac{(\Lambda_i - \bar{\Lambda}_q)\sum_k w_q(\lambda_k)(\Lambda_k - \bar{\Lambda}_q)u_k}{\sum_k w_q(\lambda_k)(\Lambda_k - \bar{\Lambda}_q)^2}. \quad (4.35)$$

Cela peut être réécrit :

$$\tilde{u}_i = [\mathbb{P}u]_i = \sum_j P_{ij}u_j, \quad (4.36)$$

avec

$$P_{ij} = \delta_{ij}^K - \frac{w_q(\lambda_j)}{\sum_k w_q(\lambda_k)} - \frac{(\Lambda_i - \bar{\Lambda}_q)w_q(\lambda_j)(\Lambda_j - \bar{\Lambda}_q)}{\sum_k w_q(\lambda_k)(\Lambda_k - \bar{\Lambda}_q)^2}, \quad (4.37)$$

où  $\delta_{ij}^K$  désigne le symbole de Kronecker.

Nous avons obtenu le projecteur  $\mathbb{P}$  recherché. Dans la sous-section suivante, on l'applique au champ de  $\delta$  mesuré.

#### 4.4.3 Projection des $\delta_q(\lambda)$

On calcule l'estimateur  $\hat{\delta}_q(\lambda_i)$  de  $\delta_q(\lambda_i)$  en lui appliquant la projection définie par l'équation 4.37 :

$$\hat{\delta}_q(\lambda_i) = \sum_j P_{ij}\delta_q(\lambda_j) = \delta_q(\lambda_i) - \frac{\sum_j w_q(\lambda_j)\delta_q(\lambda_j)}{\sum_j w_q(\lambda_j)} - (\Lambda_i - \bar{\Lambda}_q) \frac{\sum_j w_q(\lambda_j)(\Lambda_j - \bar{\Lambda}_q)\delta_q(\lambda_j)}{\sum_j w_q(\lambda_j)(\Lambda_j - \bar{\Lambda}_q)^2}. \quad (4.38)$$

Finalement,

$$\hat{\delta}_q(\lambda_i) = \delta_q(\lambda_i) - \bar{\delta}_q - (\Lambda_i - \bar{\Lambda}_q) \frac{(\Lambda - \bar{\Lambda}_q)\delta_q(\lambda)}{(\Lambda - \bar{\Lambda}_q)^2}. \quad (4.39)$$

En résumé, cette projection des champs de  $\delta$  permet de transformer les  $\delta_q(\lambda)$  mesurés en  $\hat{\delta}(\lambda)$  tels que  $\hat{\delta}(\lambda) = \hat{\delta}^v(\lambda)$ . Cette projection induit une distorsion de la fonction de corrélation mesurée, qui sera ensuite calculée sous la forme d'une matrice définie à partir de ce projecteur (cf. section 6.3) puis appliquée au modèle.

## 4.5 Évolution du champ de $\delta$ en fonction du *redshift*

Puisque la répartition des nuages d'hydrogène de l'univers dépend de leur *redshift*, alors il doit en être de même du champ d'absorption. Dans ce chapitre, je vais étudier la relation entre

l'élément  $\delta_q(\lambda)$  et le *redshift*  $z$  du nuage, présent sur la ligne de visée du quasar  $q$ , qui absorbe la longueur d'onde  $\lambda$ , soit  $z = \frac{\lambda}{\lambda_{Ly\alpha}} - 1$ . J'introduirai cette dépendance dans les poids définis à la section 4.1.3, de façon à maximiser le signal sur bruit de la fonction de corrélation.

#### 4.5.1 Le spectre de puissance à une dimension

Du fait de l'expansion de l'Univers et de l'ionisation progressive de l'hydrogène du milieu intergalactique, la densité d'hydrogène neutre diminue dans le temps, entraînant une décroissance de l'amplitude de la fonction de corrélation et du spectre de puissance Ly $\alpha$ . Cela signifie que l'amplitude de la fonction de corrélation augmente quand le *redshift* augmente. Dans cette sous-section, on va étudier cet effet à travers l'évolution en *redshift* du spectre de puissance à une dimension,  $P_{1d}$ , dans les régions Ly $\alpha$ .

Je représente sur la figure 4.6, les spectres de puissance Ly $\alpha$  à une dimension mesurés à plusieurs *redshifts* par McDonald et al. (2006) et Chabanier et al. (2018). Dans les deux cas, on observe bien l'augmentation de l'amplitude des spectres de puissance quand  $z$  augmente. De plus, on remarque que malgré la différence d'amplitude, la forme générale des spectres de puissance est similaire quel que soit le *redshift*.

Suivant McDonald et al. (2006), je suppose que la dépendance en *redshift* du spectre de puissance à une dimension, au *redshift*  $z_i$  est donnée par :

$$P_{1d}(k, z_i) = \left( \frac{1 + z_i}{1 + z_0} \right)^{\alpha_{k, z_i}} P_{1d}(k, z_0), \quad (4.40)$$

où  $z_0$  désigne un *redshift* de référence.

Les coefficients  $\alpha_{k, z_i}$  peuvent s'écrire :

$$\alpha_{k, z_i} = \left. \frac{d \ln[P_{1d}(k, z)]}{d \ln(1 + z)} \right|_{z_i} \simeq \frac{\ln(P_{1d}(k, z_{i+1})/P_{1d}(k, z_{i-1}))}{\ln[(1 + z_{i+1})/(1 + z_{i-1})]}. \quad (4.41)$$

On représente sur la figure 4.7, les coefficients  $\alpha_{k, z_i}$  mesurés avec les résultats de Chabanier et al. (2018), en fonction du *redshift*  $z$  (en haut) et du mode  $k$  (en bas). On ne voit pas apparaître de claire dépendance entre ces coefficients et les valeurs de  $k$  ou  $z$ . On peut alors supposer que les coefficients  $\alpha_{k, z}$  sont indépendants du *redshift* et du mode. L'équation 4.40 devient alors :

$$P_{1d}(k, z_i) = \left( \frac{1 + z_i}{1 + z_0} \right)^{\alpha} P_{1d}(k, z_0), \quad (4.42)$$

avec  $\alpha$  la moyenne des  $\alpha_{k, i}$ .

En utilisant les résultats de McDonald et al. (2006) puis Chabanier et al. (2018) sans sélection de *redshift*, on trouve  $\alpha = 3.8$ . Je représente en tirets sur la figure 4.6 les quantités  $\left( \frac{1+z}{1+z_0} \right)^{3.8} P_k(k, z_0)$ , en choisissant pour *redshift* de référence  $z_0 = 2.2$ . On voit que l'approximation 4.42 n'est pas parfaite mais qu'elle donne globalement le bon comportement du spectre de puissance en fonction du *redshift*. En toute rigueur, il faudrait faire le calcul de  $\alpha$  en ne gardant que les spectres de puissance pour lesquels le *redshift* est inférieur à 3.5 (*redshift* maximal des quasars utilisés dans cette étude, voir table 4.2). Cela donne  $\alpha = 3.5$ . Je choisis d'utiliser  $\alpha = 3.8$  dans la suite car c'est la valeur implémentée dans les simulations des forêts Ly $\alpha$  de Bautista

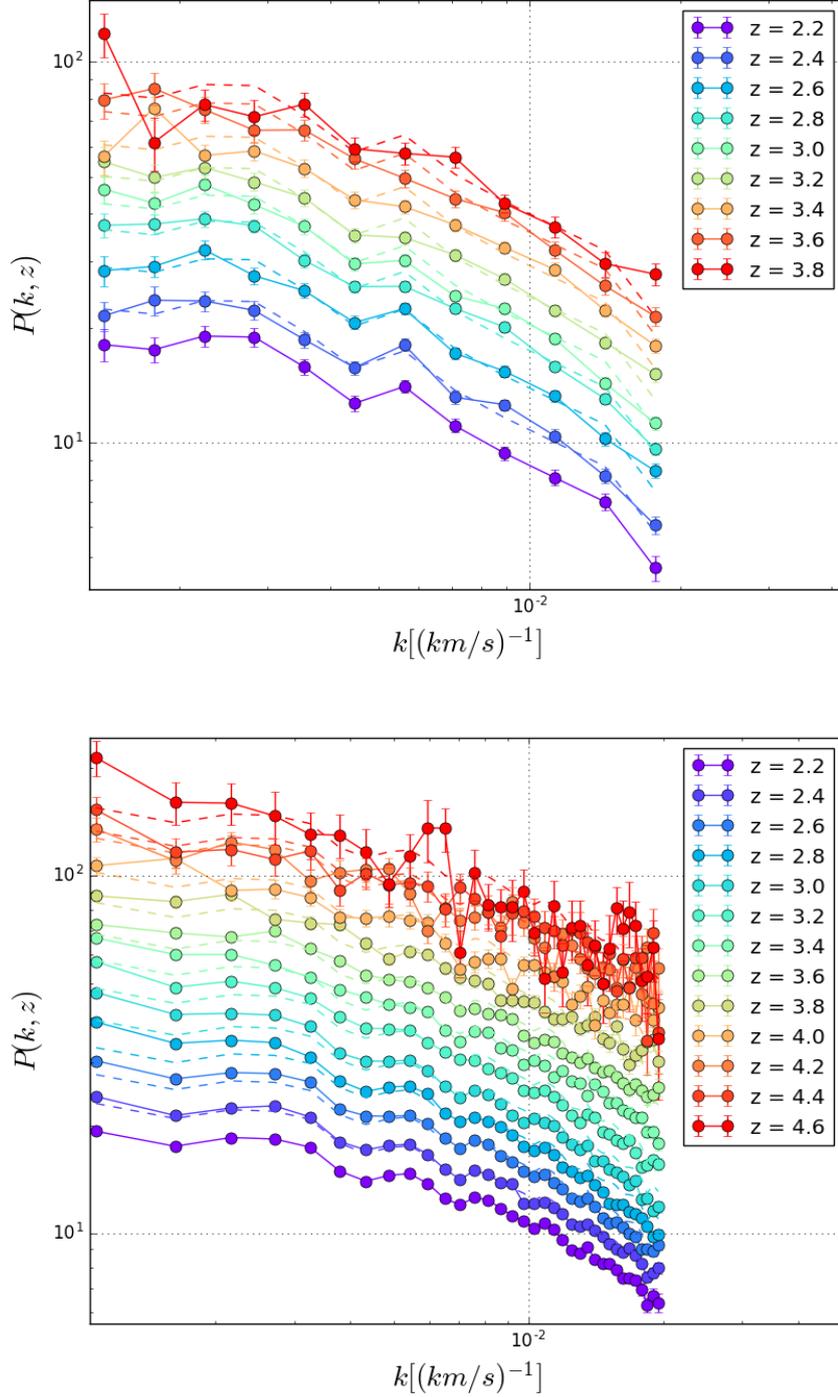


FIGURE 4.6: Mesure du spectre de puissance Ly $\alpha$  à une dimension en utilisant les données SDSS du deuxième relevé (McDonald et al. 2006, en haut) et du quatorzième relevé (Chabanier et al. 2018, en bas). On voit clairement apparaître l'augmentation de l'amplitude du spectre de puissance avec la *redshift*. Les lignes en pointillés représentent pour chaque *redshift*  $z$ , la quantité  $\left(\frac{1+z}{1+z_0}\right)^\alpha P(k, z_0)$  avec  $z_0 = 2.2$  et  $\alpha$ , la moyenne des  $\alpha_{k,i}$  (voir texte). Dans les deux cas on trouve  $\alpha = 3.8$ .

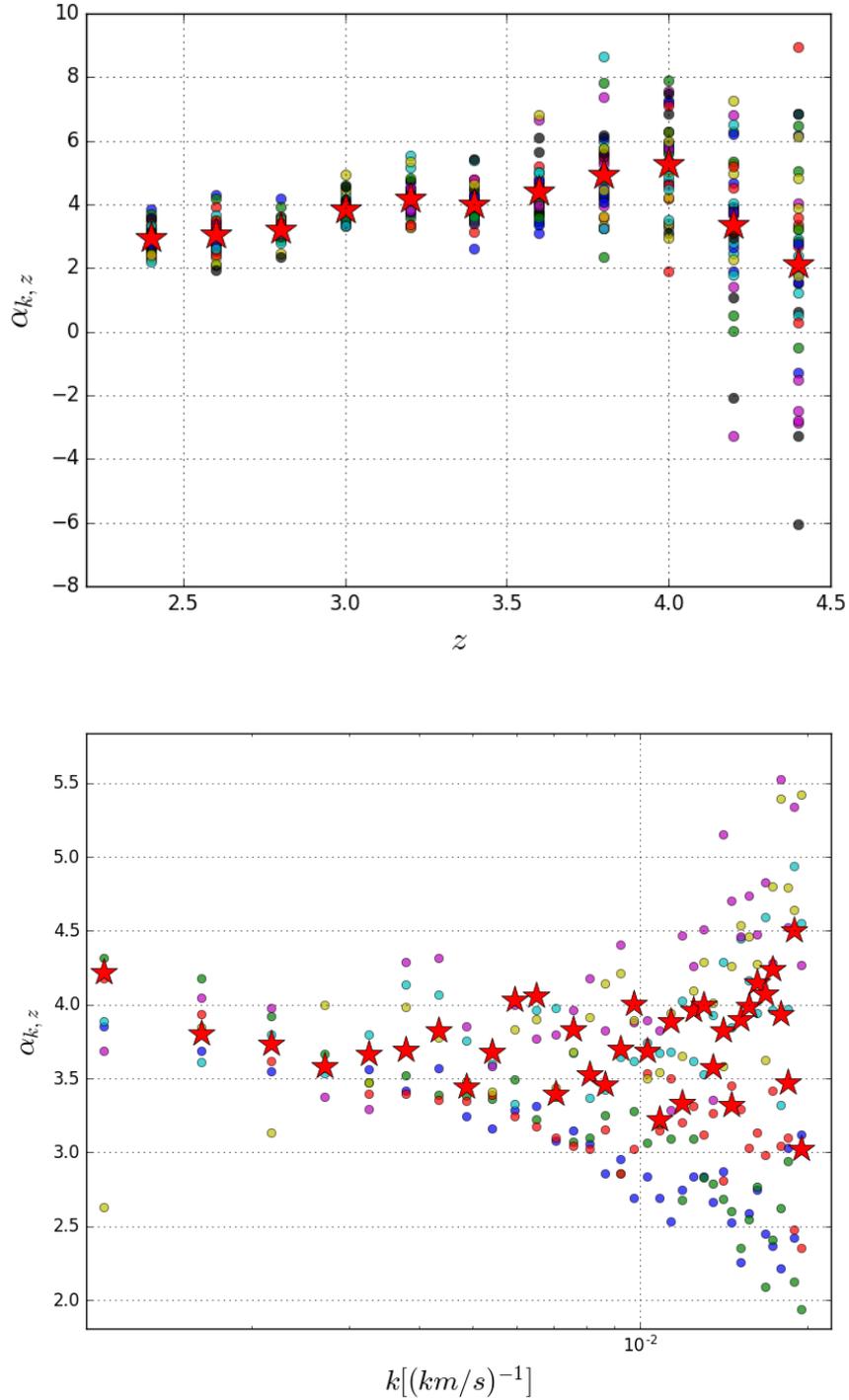


FIGURE 4.7: Coefficients  $\alpha_{k,z}$  d'évolution en *redshift* du spectre de puissance Ly $\alpha$  à une dimension, calculés avec les résultats de Chabanier et al. (2018). La figure du haut montre l'évolution en *redshift* et celle du bas, l'évolution en fonction du mode. Les étoiles rouges désignent la moyennes des  $\alpha_{k,z}$  à  $z$  fixé (en haut) et  $k$  fixé (en bas). On ne voit pas apparaître de dépendance en  $z$  ou en  $k$  de ces coefficients sur le domaine étudié.

et al. (2015), grâce auxquelles on a vérifié l'ensemble des modèles de cette étude.

Cette modélisation pourrait être améliorée en incluant une dépendance en longueur d'onde du paramètre  $\alpha$ . Cette meilleure prise en compte de la dépendance en *redshift* de l'absorbant n'aurait néanmoins aucun effet sur la position du pic BAO mais seulement sur le rapport signal sur bruit de la fonction de corrélation.

D'après la sous-section 2.1.2, on a :

$$P_{1d}(k, z) \propto b_{Ly\alpha}(z)^2 D(z)^2 P_{m,1d}(k, z_0), \quad (4.43)$$

avec  $P_{m,1d}$ , le spectre de puissance à une dimension de la matière,  $b_{Ly\alpha}$ , le biais des absorptions  $Ly\alpha$  et  $D(z)$  le taux de croissance linéaire des structures qui évolue en  $(1+z)^{-1}$  pour  $z > 2$  (voir sous-section 2.1.2).

On déduit de l'équation 4.42,

$$b_{Ly\alpha}(z) \propto (1+z)^\gamma \quad \text{avec} \quad \gamma = \alpha/2 + 1. \quad (4.44)$$

En utilisant,  $\alpha = 3.8$ ,  $\gamma = 2.9$ . Si  $\alpha = 3.5$ ,  $\gamma = 2.8$ . Le paramètre  $\gamma$  n'influe pas sur la position du pic BAO dans la fonction de corrélation mais seulement sur la précision de la mesure. Puisque l'écart entre les deux valeurs de  $\gamma$  est faible, la diminution du signal sur bruit de la fonction de corrélation est très petite.

Le biais doit reproduire la dépendance en *redshift* des champs de  $\delta$  du traceur (sous-section 2.1.3), ici les absorptions  $Ly\alpha$ . On a donc :

$$\delta_q(\lambda) \propto \left(\frac{\lambda}{\lambda_{Ly\alpha}}\right)^{\gamma-1}. \quad (4.45)$$

Cela signifie que les absorptions sont plus marquées à haut *redshift*, c'est-à-dire quand l'Univers était plus jeune. Ce qui est cohérent avec le processus de formation des structures : les gaz du milieu intergalactique se condensent pour former les galaxies. Ainsi, à plus haut *redshift*, il y a plus de gaz et moins de galaxies.

Dans la suite, on notera  $\gamma_{Ly\alpha}$  le paramètre d'évolution en *redshift* du biais  $Ly\alpha$  pour pouvoir faire la distinction avec d'autres types d'absorbants (dont les métaux, voir section 5.2). J'adopte la valeur  $\gamma_{Ly\alpha} = 2.9$ .

## 4.5.2 Ré-étalonnage des poids

D'après l'équation 4.45, les absorptions à haut *redshift* sont plus marquées. Elles doivent donc être privilégiées dans la fonction de corrélation. J'introduis dans les poids la même dépendance que dans le champ de  $\delta$  :

$$w_q(\lambda) \propto \left(\frac{\lambda}{\lambda_{Ly\alpha}}\right)^{\gamma_{Ly\alpha}-1}, \quad (4.46)$$

Je redéfinit donc les poids introduits à la section 4.1.3 par :

$$w_q(\lambda) = \frac{1}{\sigma_q(\lambda)^2} \frac{(\lambda/\lambda_{Ly\alpha})^{\gamma_{Ly\alpha}-1}}{(1+z_{ref})^{\gamma_{Ly\alpha}-1}}, \quad (4.47)$$

où  $z_{ref}$ , le *redshift* de référence, doit être proche du *redshift* effectif de la fonction de corrélation. Je choisis  $z_{ref} = 2.3$ . La division de tous les poids par  $(1 + z_{ref})^{\gamma_{Ly\alpha} - 1}$  est un artifice de calcul pour éviter que les sommes qui interviennent dans le calcul de la fonction de corrélation ne soient trop grandes.

### Synthèse du chapitre

Dans ce chapitre, j'ai extrait des spectres des quasars, les champs d'absorption Ly $\alpha$  dans les régions spectrales Ly $\alpha$  et Ly $\beta$ . J'ai également défini les poids des absorptions à utiliser pour maximiser le signal sur bruit de la fonction de corrélation et donc améliorer la détection du pic BAO.

## Chapitre 5

# Mesure de la fonction de corrélation $\text{Ly}\alpha$

L'objet de ce chapitre est le calcul de la fonction de corrélation des absorptions  $\text{Ly}\alpha$ , à partir des champs de  $\delta$  calculés dans le chapitre précédent.

Soient  $\delta_1$  et  $\delta_2$  les contrastes de densité de deux traceurs quelconques. D'après l'équation 2.6, la fonction de corrélation théorique (à deux points) entre les traceurs 1 et 2 est définie par :

$$\xi(\vec{r}) = \langle \delta_1(\vec{x})\delta_2(\vec{x} + \vec{r}) \rangle_{\vec{x}}, \quad (5.1)$$

avec  $\vec{x}$  et  $\vec{x} + \vec{r}$  les positions respectives des objets 1 et 2. Si les traceurs 1 et 2 sont identiques, on parle d'auto-corrélation, s'il sont différents, de corrélation croisée. En pratique,  $\xi(\vec{r})$  désigne l'excès de probabilité par rapport à la répartition uniforme, de trouver un objet de type 2 à la distance  $\vec{r}$  d'un objet de type 1.

Dans cette étude, on considère les absorptions  $\text{Ly}\alpha$  des régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  des spectres des quasars comme traceur de masse. Le contraste de densité des absorptions  $\text{Ly}\alpha$  est donné par les champs de  $\delta$  calculés au chapitre précédent.

Ce chapitre s'organise en trois parties :

- Tout d'abord, je calcule la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  à trois dimensions qui permettra de mesurer la position du pic BAO.
- Puis, je calcule la fonction de corrélation à une dimension, c'est-à-dire quand les champs de  $\delta$  sont sur la même ligne de visée. Cette étape permettra de détecter les métaux contaminant la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  à trois dimensions.
- Enfin, je calculerai la covariance de la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , qui permet d'estimer la précision de la mesure.

## 5.1 La fonction de corrélation à trois dimensions

Afin de simplifier le texte, la notation suivante :

$$\text{absorption 1(région 1)} \times \text{absorption 2(région 2)}, \quad (5.2)$$

désignera la fonction de corrélation des absorptions 1 de la région 1 avec les absorptions 2 de la région 2.

Par exemple,  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ , désigne la fonction de corrélation des absorptions  $\text{Ly}\alpha$  de la région  $\text{Ly}\alpha$  entre elles et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , la fonction de corrélation entre les absorptions  $\text{Ly}\alpha$  de la région  $\text{Ly}\alpha$  et les absorptions  $\text{Ly}\alpha$  de la région  $\text{Ly}\beta$ .

### 5.1.1 Les types d'absorption dans les régions $\text{Ly}\alpha$ et $\text{Ly}\beta$

Dans cette étude, les contrastes de densité  $\delta_1$  et  $\delta_2$  de l'équation 5.1 peuvent désigner le champ de  $\delta$  de la région  $\text{Ly}\alpha$ , noté  $\delta^{\text{Ly}\alpha}$  dans cette section, ou celui de la région  $\text{Ly}\beta$ , noté  $\delta^{\text{Ly}\beta}$ .

#### 5.1.1.1 Cas $\delta_1 = \delta_2 = \delta^{\text{Ly}\alpha}$

Étant donnée la grande abondance d'hydrogène neutre dans le milieu intergalactique (section 2.2), on suppose en première approximation que toutes les absorptions visibles dans les champs de  $\delta$  sont dues à des transitions de type Lyman de l'hydrogène. Les absorptions  $\text{Ly}\beta$  par la matière sur la ligne de visée d'un quasar se produisent à des longueurs d'onde observées inférieures à la longueur d'onde l'émission  $\text{Ly}\beta$  du quasar. Elles ne peuvent donc pas contaminer les absorptions observées dans la région  $\text{Ly}\alpha$ . Celles-ci sont donc uniquement des absorptions  $\text{Ly}\alpha$ . Ce cas correspond à  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ .

#### 5.1.1.2 Cas $\delta_1 = \delta^{\text{Ly}\alpha}$ et $\delta_2 = \delta^{\text{Ly}\beta}$

La région  $\text{Ly}\beta$  est située à gauche des pics d'émission  $\text{Ly}\alpha$  et  $\text{Ly}\beta$ , on y observe les types d'absorption dues à ces deux transitions. Ainsi en corrélant  $\delta^{\text{Ly}\alpha}$  et  $\delta^{\text{Ly}\beta}$ , on peut calculer les deux fonctions de corrélation :  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$ .

La section efficace de l'absorption  $\text{Ly}\beta$  étant cinq fois plus faible que pour  $\text{Ly}\alpha$  (voir sous-section 2.2.3), l'amplitude de la corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$  est cinq fois plus faible que celle de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et nécessite donc un volume de données cinq fois plus grand pour pouvoir y détecter le signal BAO. J'ai calculé la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$  mais les données eBOSS ne sont pas suffisantes pour y détecter le signal BAO. C'est pourquoi je ne la prends pas en compte dans le reste de ma thèse.

#### 5.1.1.3 Cas $\delta_2 = \delta_1 = \delta^{\text{Ly}\beta}$

Ce cas correspond aux fonctions  $\text{Ly}\alpha(\text{Ly}\beta) \times \text{Ly}\alpha(\text{Ly}\beta)$ ,  $\text{Ly}\alpha(\text{Ly}\beta) \times \text{Ly}\beta(\text{Ly}\beta)$  et  $\text{Ly}\beta(\text{Ly}\beta) \times \text{Ly}\beta(\text{Ly}\beta)$ . À cause de la taille réduite de la région  $\text{Ly}\beta$ , le nombre de paires dans ces trois fonctions de corrélation est trop faible pour y détecter le signal du BAO. Je les ai donc ignorées dans ce travail.

Finalement, je ne considérerai dans cette thèse que les deux fonctions de corrélation avec le plus grand nombre de paires d'absorptions disponibles, *i.e.*  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ .

### 5.1.2 Les coordonnées

Dans cette sous-section, je vais expliciter les distances  $\vec{r}$  et  $\vec{x}$  de l'équation 5.1.

Comme on l'a vu au chapitre précédent, chaque absorption (ou *pixel*)  $i$  du champ de  $\delta$  est définie par le couple  $(q_i, \lambda_i)$  :  $q_i$  désigne le quasar, c'est-à-dire la direction de l'absorption, et  $\lambda_i$ , la longueur d'onde observée qui peut être convertie en *redshift* en supposant connu le type d'absorption.

La position de  $i$  sur la voûte céleste s'exprime avec les coordonnées équatoriales  $(\text{RA}_i, \text{Dec}_i)$ . On représente sur la figure 5.1 le système avec l'absorption  $i$  en I, le point T désignant la Terre et le plan équatorial étant représenté en bleu.

- $\text{RA}_i$ , l'ascension droite (*Right Ascension* en anglais), désigne l'angle entre TV et TH avec H la projection du point I sur l'équateur céleste et V le point vernal<sup>1</sup>.
- $\text{Dec}_i$ , la déclinaison, désigne l'angle entre TI et TH.

On peut ainsi passer des coordonnées sur le spectre à une position dans le ciel :

$$(q_i, \lambda_i) \longrightarrow (\text{RA}_i, \text{Dec}_i, z_i = \frac{\lambda_i}{\lambda_{abs}} - 1), \quad (5.3)$$

où  $\lambda_{abs}$  désigne la longueur d'onde de la transition considérée ( $\text{Ly}\alpha$  ou  $\text{Ly}\beta$ ).

Soient deux absorptions  $i$  et  $j$ . On note  $\theta_{ij}$  l'angle de séparation entre les lignes de visée des absorbants correspondants (figure 5.2). Soit  $\rho$  le rayon de la sphère céleste, les points I et J désignent respectivement les intersections des lignes de visée de  $i$  et  $j$  avec la sphère céleste. I et J peuvent s'écrire dans le système  $(T, x, y, z)$  (figure 5.1) :

$$I \begin{pmatrix} \rho \cos \text{Dec}_i \cos \text{RA}_i \\ \rho \cos \text{Dec}_i \sin \text{RA}_i \\ \rho \sin \text{Dec}_i \end{pmatrix} \quad J \begin{pmatrix} \rho \cos \text{Dec}_j \cos \text{RA}_j \\ \rho \cos \text{Dec}_j \sin \text{RA}_j \\ \rho \sin \text{Dec}_j \end{pmatrix} \quad (5.4)$$

Le produit scalaire des vecteurs  $\vec{TI}$  et  $\vec{TJ}$  donne accès à  $\cos \theta_{ij}$  :

$$\cos \theta_{ij} = \cos \text{Dec}_i \cos \text{Dec}_j \cos(\text{RA}_i - \text{RA}_j) + \sin \text{Dec}_i \sin \text{Dec}_j. \quad (5.5)$$

Le vecteur de séparation  $\vec{r}_{ij}$  entre les absorbants de  $i$  et  $j$  dépend des distances comobiles de  $i$  et  $j$  considérés,  $\chi(z_i)$  et  $\chi(z_j)$ , et se décompose en composantes parallèle et perpendiculaire à

<sup>1</sup>Le point vernal est intersection entre l'équateur céleste et l'écliptique.

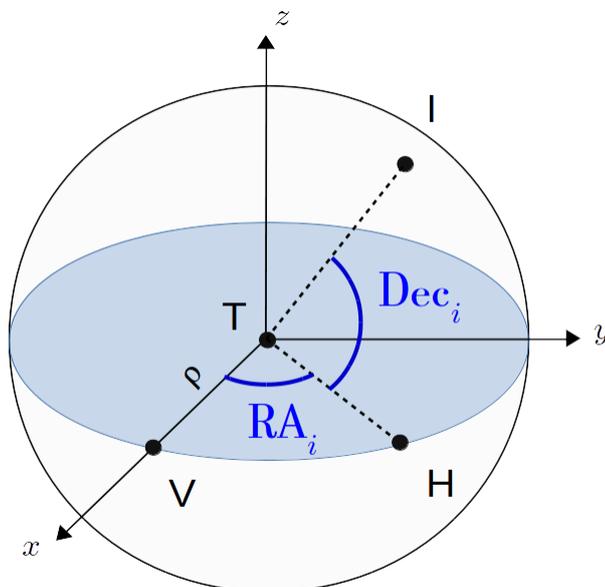


FIGURE 5.1: Schéma des coordonnées équatoriales de l'absorbant  $i$  au point I sur la voûte céleste. On représente en bleu le plan équatorial céleste. Le point T désigne la Terre et V, le point vernal (intersection entre l'équateur céleste et l'écliptique). L'ascension droite,  $RA_i$ , est l'angle entre TV et TH, où H est la projection de I sur l'équateur céleste. La déclinaison,  $Dec_i$ , est l'angle entre TH et TI.

la ligne de visée moyenne (figure 5.2) :

$$\begin{cases} r_{\parallel,ij} = |(\chi(z_i) - \chi(z_j)) \cos \frac{\theta_{ij}}{2}| \\ r_{\perp,ij} = (\chi(z_i) + \chi(z_j)) \sin \frac{\theta_{ij}}{2} \end{cases} \quad (5.6)$$

Le module de  $\vec{r}_{ij}$  est alors donné par :

$$|\vec{r}_{ij}| = r_{ij} = \sqrt{\chi(z_i)^2 + \chi(z_j)^2 - 2\chi(z_i)\chi(z_j) \cos \theta_{ij}}. \quad (5.7)$$

On peut également définir  $\mu_{ij}$ , le cosinus de l'angle de séparation entre la ligne de visée moyenne et le vecteur  $\vec{r}_{ij}$  :

$$\boxed{\mu_{ij} = \frac{r_{\parallel,ij}}{r_{ij}}}. \quad (5.8)$$

La séparation entre les absorbants peut s'exprimer indifféremment dans les systèmes  $(r_{\parallel}, r_{\perp})$  et  $(r, \mu)$ . On privilégiera le premier pour la représentation de la fonction de corrélation en deux dimensions et le second, pour la représentation en une dimension.

### 5.1.3 Séparation entre les absorptions Ly $\alpha$ et Ly $\beta$ dans la région Ly $\beta$

Nous avons vu précédemment que le passage des coordonnées spectrales aux coordonnées physiques (équation 5.3) nécessitait de connaître le type de l'absorption considérée. Or la région Ly $\beta$  est façonnée par des absorptions à la fois de type Ly $\alpha$  et Ly $\beta$ . Puisqu'il n'est pas possible de

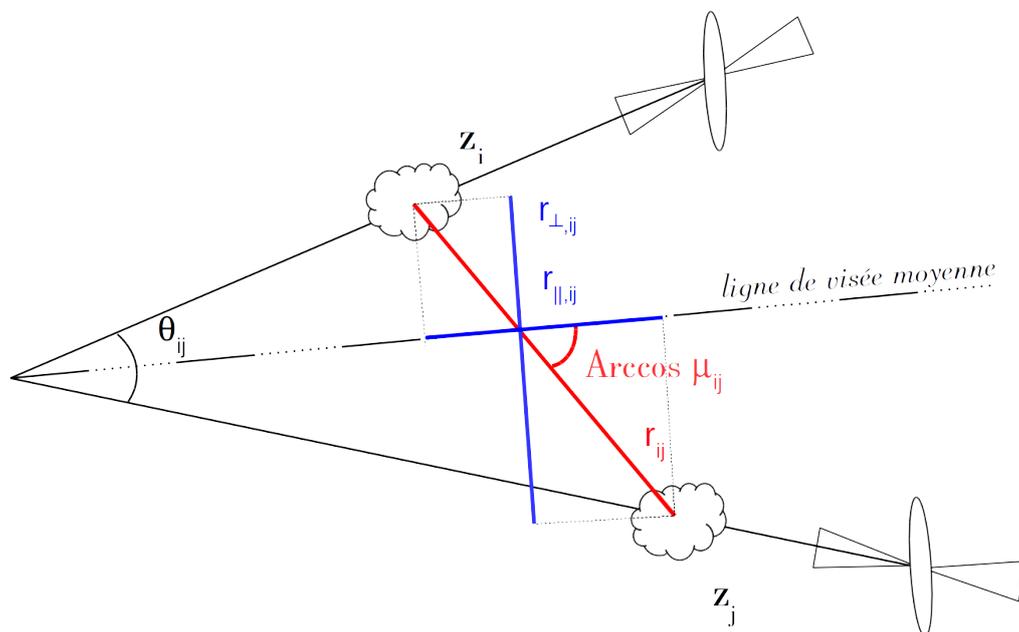


FIGURE 5.2: Schéma représentant la configuration physique des absorbers  $i$  et  $j$  séparés par l'angle  $\theta_{ij}$  et le vecteur  $\vec{r}_{ij}$ . On peut projeter  $\vec{r}_{ij}$  parallèlement et perpendiculairement à la ligne de visée moyenne et on obtient les coordonnées  $(r_{\parallel,ij}, r_{\perp,ij})$ . On peut alternativement utiliser le système  $(r_{ij}, \mu_{ij})$  avec  $r_{ij} = |\vec{r}_{ij}|$  et  $\mu_{ij} = \frac{r_{\parallel,ij}}{r_{ij}}$ .

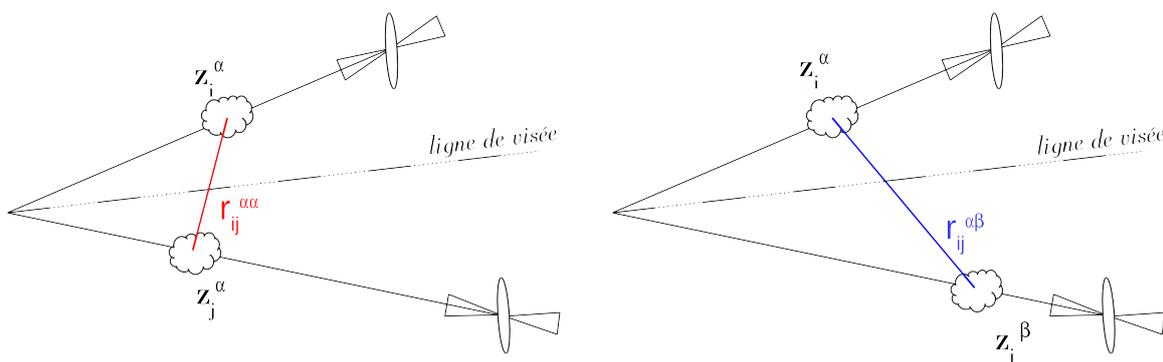


FIGURE 5.3: On considère une absorption  $i$  dans région Ly $\alpha$  d'un quasar à  $z_i$  et une absorption  $j$  dans la région Ly $\beta$  d'un quasar à  $z_j$ . Ce couple peut correspondre aux deux situations physiques exposées par ce schéma. À gauche, l'absorption  $j$  est de type Ly $\alpha$  à  $z_i = \frac{\lambda_i}{\lambda_\alpha} - 1$ , à droite, elle est de type Ly $\beta$  à beaucoup plus haut *redshift*,  $z_j = \frac{\lambda_j}{\lambda_\beta} - 1$ .

les distinguer, on fait l'hypothèse (incorrecte) que toutes les absorptions visibles dans la région Ly $\beta$  sont de type Ly $\alpha$ . J'étudie maintenant les conséquences d'une telle hypothèse.

Considérons les absorptions  $i$  dans la région Ly $\alpha$  d'un quasar  $q_i$  et  $j$  dans la région Ly $\beta$  d'un quasar  $q_j$ . Si  $j$  est de type Ly $\alpha$ , la distance de séparation s'écrit  $r_{ij}^{\alpha\alpha}$  (schéma de gauche sur la figure 5.3) et si  $j$  est de type Ly $\beta$ , elle s'écrit  $r_{ij}^{\alpha\beta}$  (schéma de droite sur la figure 5.3). Ainsi, en supposant incorrectement que  $j$  est de type Ly $\alpha$ , la composante parallèle de la distance de séparation est faussement estimée par  $r_{\parallel,ij}^F = r_{\parallel,ij}^{\alpha\alpha}$ , au lieu de la vraie mesure  $r_{\parallel,ij}^V = r_{\parallel,ij}^{\alpha\beta}$ . Cela crée un décalage en distance  $\Delta r_{\parallel,ij}$  parallèlement à la ligne de visée :

$$\Delta r_{\parallel,ij} = r_{\parallel,ij}^F - r_{\parallel,ij}^V = r_{\parallel,ij}^{\alpha\alpha} - r_{\parallel,ij}^{\alpha\beta} \quad (5.9)$$

Ce décalage en distance induit l'apparition d'un pic artificiel dans la fonction de corrélation autour de  $(\Delta r_{\parallel,ij}, \Delta r_{\perp,ij})$ , où  $\Delta r_{\perp,ij}$  est défini de manière analogue à  $\Delta r_{\parallel,ij}$ . Le pic artificiel correspondrait au décalage du premier pic de corrélation (visible pour les séparations  $r < 40 h^{-1}\text{Mpc}$  sur la figure 5.10).

Heureusement, sur la fenêtre d'étude choisie  $[0, 200] h^{-1}\text{Mpc}$  (*cf.* section 5.1.4), il n'existe aucun couple de longueurs d'onde  $(\lambda_i, \lambda_j)$  tel que  $\Delta r_{\parallel,ij} < 200 h^{-1}\text{Mpc}$  et  $r_{\parallel,ij}^{\alpha\beta} < 200 h^{-1}\text{Mpc}$ . Ce qui veut dire que les éventuels pics artificiels apparaissent pour des séparations supérieures à  $200 h^{-1}\text{Mpc}$ , et n'influent donc pas sur la fonction de corrélation mesurée pour  $r_{\parallel}$  et  $r_{\perp}$  dans  $[0, 200] h^{-1}\text{Mpc}$ . Ainsi, on peut librement calculer Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) en supposant que toutes les absorptions visibles dans la région Ly $\beta$  sont de type Ly $\alpha$ . Cette hypothèse aura un effet négligeable sur la fonction de corrélation dans le domaine de séparation étudié.

### 5.1.4 Estimateur de la fonction de corrélation

L'équation 5.1 peut être réécrite en utilisant les champs de  $\delta$  projetés calculés à la section 4.4 :

$$\hat{\xi}(\vec{r}) = \langle \hat{\delta}_i \hat{\delta}_j \rangle \quad \text{avec} \quad \vec{r}_{ij} = \vec{r} \quad (5.10)$$

où  $\langle \rangle$  désigne une moyenne d'ensemble sur tous les couples  $(i, j)$  tels que  $\vec{r}_{ij} = \vec{r}$ .

L'estimation de  $\hat{\xi}(\vec{r})$  va se faire en deux temps :

- Puisqu'on n'a pas accès à une mesure continue de  $\hat{\xi}(\vec{r})$ , je vais échantillonner  $\hat{\xi}$ , c'est-à-dire la moyenner sur des valeurs discrètes de  $\vec{r}$ .
- La moyenne d'ensemble  $\langle \hat{\delta}_i \hat{\delta}_j \rangle$  va être approximée en répartissant les paires  $(i, j)$  en sous-groupes et en calculant la moyenne des fonctions de corrélation de chaque sous-groupe.

#### 5.1.4.1 Échantillonnage

La fonction de corrélation est mesurée pour  $r_{\parallel}$  et  $r_{\perp}$  dans  $[0, 200] h^{-1}\text{Mpc}$  de façon à être centrée sur le pic BAO (autour de  $100 h^{-1}\text{Mpc}$ ).

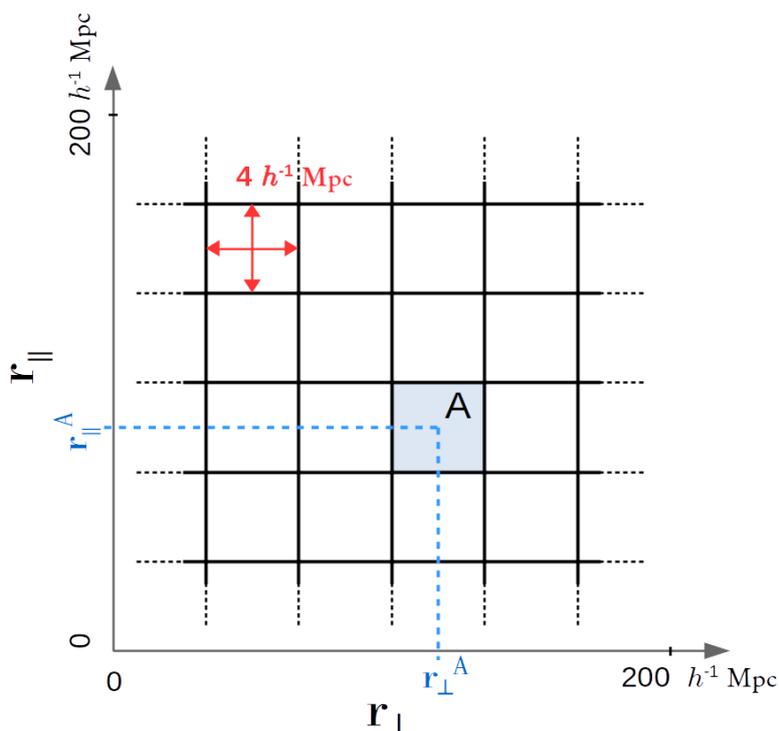


FIGURE 5.4: La grille en distance de séparation pave l'espace défini par  $r_{\parallel} \in [0, 200] h^{-1} \text{Mpc}$  et  $r_{\perp} \in [0, 200] h^{-1} \text{Mpc}$ . Un élément de la grille, appelé *bin*, couvre une surface de  $4 \times 4 h^{-2} \text{Mpc}^2$ . Le domaine d'étude regroupe  $50 \times 50 = 2500$  *bins*.

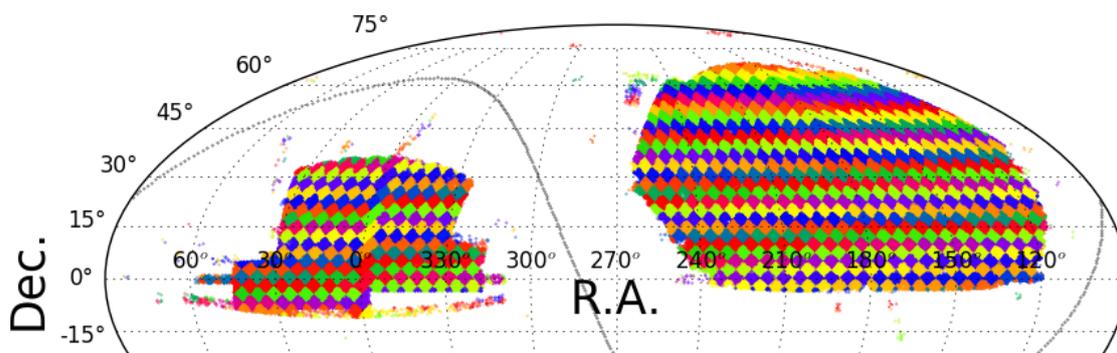


FIGURE 5.5: Les HEALPix (Górski et al. 2005) permettent de paver la voûte céleste en sous-ensembles de surfaces identiques. Chaque spectre de l'étude appartient à un seul HEALPix. Les surfaces colorées correspondent à l'empreinte dans le ciel des quasars eBOSS avec  $z_q > 2$ .

Je définis une grille en distance de séparation  $(r_{\parallel}, r_{\perp})$  (figure 5.4) et j'appelle *bins* les éléments de cette grille. Pour simplifier la modélisation (voir chapitre 6), j'utilise une grille carrée. Les largeurs d'un *bin* parallèlement et perpendiculairement à la ligne de visée,  $R_{\parallel}$  et  $R_{\perp}$ , sont choisies de manière à ce que les barres d'erreurs soient faibles lors de la mesure de la fonction de corrélation, tout en conservant une résolution assez fine pour voir le pic BAO étalé sur plusieurs *bins*. Ici, je reprends  $R_{\parallel} = R_{\perp} = 4 h^{-1}\text{Mpc}$ , valeurs précédemment utilisées dans Delubac et al. (2015), Bautista et al. (2017).

La grille doit paver tout l'espace pour  $r_{\parallel}$  et  $r_{\perp}$  dans  $[0, 200] h^{-1}\text{Mpc}$ . Elle se compose donc de  $50 \times 50 = 2500$  *bins* carrés de surface  $4 h^{-1}\text{Mpc} \times 4 h^{-1}\text{Mpc}$ .

Considérons le *bin*  $A$  dont la valeur centrale est  $(r_{\parallel}^A, r_{\perp}^A)$  (figure 5.4). On a :

$$(i, j) \in A \quad \Leftrightarrow \quad r_{\parallel, ij} \in \left[ r_{\parallel}^A - \frac{R_{\parallel}}{2}, r_{\parallel}^A + \frac{R_{\parallel}}{2} \right] \quad \text{et} \quad r_{\perp, ij} \in \left[ r_{\perp}^A - \frac{R_{\perp}}{2}, r_{\perp}^A + \frac{R_{\perp}}{2} \right]. \quad (5.11)$$

La fonction de corrélation mesurée au *bin*  $A$  est estimée par la moyenne des produits des  $\delta$  des paires appartenant à  $A$ , soit :

$$\hat{\xi}_A = \langle \hat{\delta}_i \hat{\delta}_j \rangle \quad \text{avec} \quad (i, j) \in A. \quad (5.12)$$

#### 5.1.4.2 Estimation de la moyenne d'ensemble

Je ré-écris l'équation 5.12 en pondérant les champs de  $\delta$  par les poids obtenus à la section 4.1.3 :

$$\hat{\xi}_A = \langle \hat{\delta}_i \hat{\delta}_j \rangle = \frac{\sum_{(i,j) \in A} w_i w_j \langle \hat{\delta}_i \hat{\delta}_j \rangle}{\sum_{(i,j) \in A} w_i w_j}. \quad (5.13)$$

La moyenne d'ensemble  $\langle \hat{\delta}_i \hat{\delta}_j \rangle$  désigne la moyenne du produit  $\hat{\delta}_i \hat{\delta}_j$  pour plusieurs réalisations de l'expérience. Je vais estimer cette moyenne en partitionnant aléatoirement les couples  $(i, j)$  et en calculant la fonction de corrélation dans chacun des sous-groupes. L'estimateur de la fonction de corrélation totale sera alors donné par la moyenne des fonctions de corrélations des sous-groupes. La séparation de la fonction de corrélation en sous-ensembles a l'avantage de favoriser la parallélisation du code et donc de réduire le temps de calcul. De plus, elle sera également utilisée à la section 5.3.2 lors du calcul de la matrice de covariance par la méthode de *subsampling*.

La première étape consiste à trouver une répartition des paires qui respecte les trois conditions suivantes :

- toutes les paires doivent être prises en compte,
- une paire ne peut appartenir qu'à un seul sous-ensemble,
- la répartition doit être aléatoire.

Il existe plusieurs façons de partitionner  $(i, j)$ . Delubac et al. (2015) et Bautista et al. (2017) utilisent, par exemple, une répartition par plaque d'observation (voir section 3.2.3). Dans cette

thèse, j'utilise une répartition par `HEALPix` (Górski et al. 2005). Les `HEALPix`s permettent de paver la voûte céleste en sous-ensembles de même surface (figure 5.5). Chaque position dans le ciel appartient alors à un `HealPix`, identifiable par un nombre.

Considérons le couple de *pixels*  $(i, j)$  tels que  $\text{RA}_i < \text{RA}_j$ . Le couple  $(i, j)$  appartient au sous-groupe défini par l'identifiant du `HEALPix` de  $i$ . On choisit le nombre de sous-ensembles pour s'assurer des deux conditions suivantes :

- Il faut qu'il y ait suffisamment de paires par sous-ensemble pour détecter un signal dans la fonction de corrélation associée.
- Il faut qu'il y ait assez de sous-ensembles pour que la moyenne soit stable.

En pratique, on détermine le nombre de sous-ensembles à utiliser en testant plusieurs valeurs sur des simulations de l'expérience. On vérifie que les variations, d'une simulation à l'autre, de la moyenne des fonctions de corrélation sont inférieures à la précision de la mesure. Je reprends le nombre de 1000 sous-ensembles utilisé dans les études précédentes (Delubac et al. 2015, Bautista et al. 2017).

Soit  $h$  un sous-ensemble, la fonction de corrélation  $\hat{\xi}_A^h$  mesurée dans  $h$  au *bin*  $A$  est la moyenne pondérée des paires  $\hat{\delta}_i \hat{\delta}_j$  appartenant à la fois à  $h$  et  $A$  (noté  $A_h$ ) :

$$\hat{\xi}_A^h = \langle \hat{\delta}_i \hat{\delta}_j \rangle_h \simeq \frac{\sum_{(i,j) \in A_h} w_i w_j \hat{\delta}_i \hat{\delta}_j}{\sum_{(i,j) \in A_h} w_i w_j}. \quad (5.14)$$

Finalement, l'estimateur de la fonction de corrélation mesurée au *bin*  $A$  est défini par :

$$\langle \hat{\xi}_A \rangle \equiv \frac{\sum_h W_A^h \hat{\xi}_A^h}{\sum_h W_A^h}, \quad (5.15)$$

avec  $W_A^h = \sum_{(i,j) \in A_h} w_i w_j$ .

Ce qui peut également s'écrire :

$$\langle \hat{\xi}_A \rangle = \frac{\sum_{(i,j) \in A} w_i w_j \hat{\delta}_i \hat{\delta}_j}{\sum_{(i,j) \in A} w_i w_j}. \quad (5.16)$$

### 5.1.5 Calcul de la fonction de corrélation

Dans cette sous-section, j'affine la définition 5.15 en supprimant certaines paires qui sont sources d'erreurs systématiques. Puis je réduis le temps de calcul en faisant une pré-sélection sur les couples de quasars. Dans la dernière partie, j'explique les valeurs moyennes de  $r_{\parallel}$ ,  $r_{\perp}$  et  $z$  dans un *bin* quelconque  $A$ .

#### 5.1.5.1 Sélection des paires

Les produits  $\hat{\delta}_i \hat{\delta}_j$  ne sont pas calculés pour tous les couples  $(i, j)$ . Comme expliqué à la section 3.3.4, on soustrait à chaque spectre le fond du ciel qui est estimé par la moyenne de tous

les spectres d'une même demi-plaque (plaque et spectrographe identiques). Ce processus crée, à longueur d'onde fixée, une corrélation positive entre les spectres d'une même demi-plaque. Pour ne pas contaminer par des erreurs systématiques la fonction de corrélation mesurée, je ne considère pas les produits  $\hat{\delta}_i \hat{\delta}_j$  lorsque  $i$  et  $j$  appartiennent à une même demi-plaque et que  $\lambda_i = \lambda_j$ . Cette sélection supprime environ 2 paires sur 10000 donc son effet ne dégrade pas la précision de la mesure.

### 5.1.5.2 Pré-sélection des voisins

Le calcul de la fonction de corrélation implique quatre sommes imbriquées : deux sur les  $n_\lambda$  longueurs d'onde et deux sur les  $n_q$  spectres. Soit  $n_q^2 \times n_\lambda^2$  opérations si on fait le calcul naïvement. Dans le cas de l'auto-corrélation des régions Ly $\alpha$ ,  $n_\lambda = 208$  et  $n_q = 179\,965$ , qui font plus de  $10^{15}$  opérations !

La paire d'absorbants  $(i, j)$  est prise en compte dans l'étude si :

$$r_{\perp,ij} \leq 200 \text{ h}^{-1}\text{Mpc} \quad (5.17)$$

En utilisant l'équation 5.6, cette condition peut également s'écrire en fonction de  $\theta_{ij}$ , la séparation angulaire entre les quasars  $q_i$  et  $q_j$  :

$$\sin \frac{\theta_{ij}}{2} \leq \frac{200 \text{ h}^{-1}\text{Mpc}}{\chi(z_i) + \chi(z_j)} \leq \frac{200 \text{ h}^{-1}\text{Mpc}}{2\chi(z_{\min})}, \quad (5.18)$$

avec  $z_{\min} = \frac{\lambda_{\min}}{\lambda_{\text{Ly}\alpha}} - 1$ .

Ainsi, l'équation 5.6 se transforme en une condition sur la séparation angulaire entre deux quasars :

$$\theta_{ij} \leq \theta_{\max} = 2 \arcsin \left( \frac{200 \text{ h}^{-1}\text{Mpc}}{2\chi(z_{\min})} \right). \quad (5.19)$$

En utilisant  $\lambda_{\min} = 360 \text{ nm}$  (table 4.2), je trouve  $\theta_{\max} = 3.33^\circ$ .

Lors du calcul de la fonction de corrélation, je ne corrèle avec l'absorption  $i$  que les absorptions  $j$  telles que l'angle de séparation entre les quasars  $q_i$  et  $q_j$  respecte la condition 5.19. Puisque chaque quasar  $q_i$  possède au plus quelques centaines de voisins  $q_j$  tels que  $\theta_{ij} \leq \theta_{\max}$ , cette étape préparatoire permet de diviser par trois ordres de grandeur le nombre d'opérations calculé précédemment.

### 5.1.5.3 Les trois dimensions de la fonction de corrélation

La fonction de corrélation est dite "à trois dimensions" car elle est fonction de  $r_{\parallel}$ ,  $r_{\perp}$  et  $z$ . On définit le *redshift* moyen du *bin* A,  $z^A$  par (Bautista et al. 2017) :

$$z^A = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \frac{z_i + z_j}{2}, \quad (5.20)$$

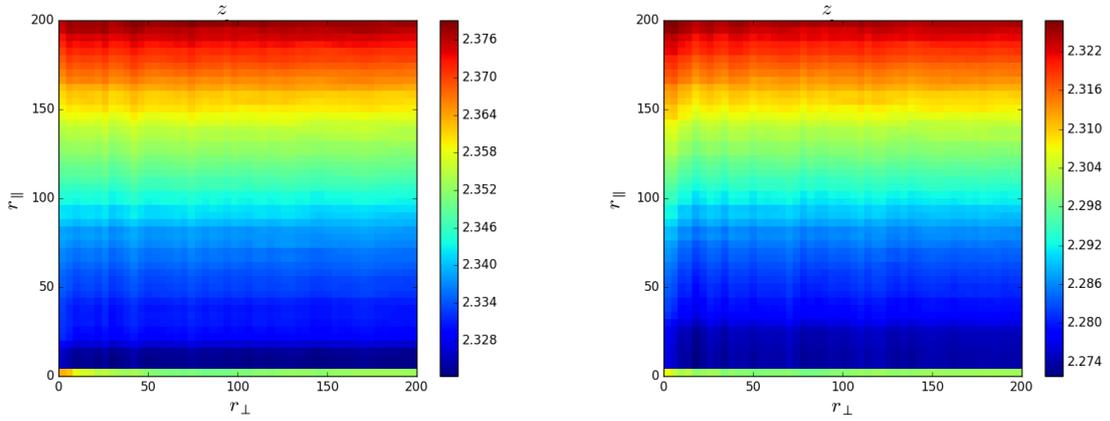


FIGURE 5.6: *Redshift* moyen des paires par *bin* dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (à gauche) et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (à droite). Dans les deux cas, l'amplitude de variation d'un *bin* à l'autre est faible (inférieure à 3%).

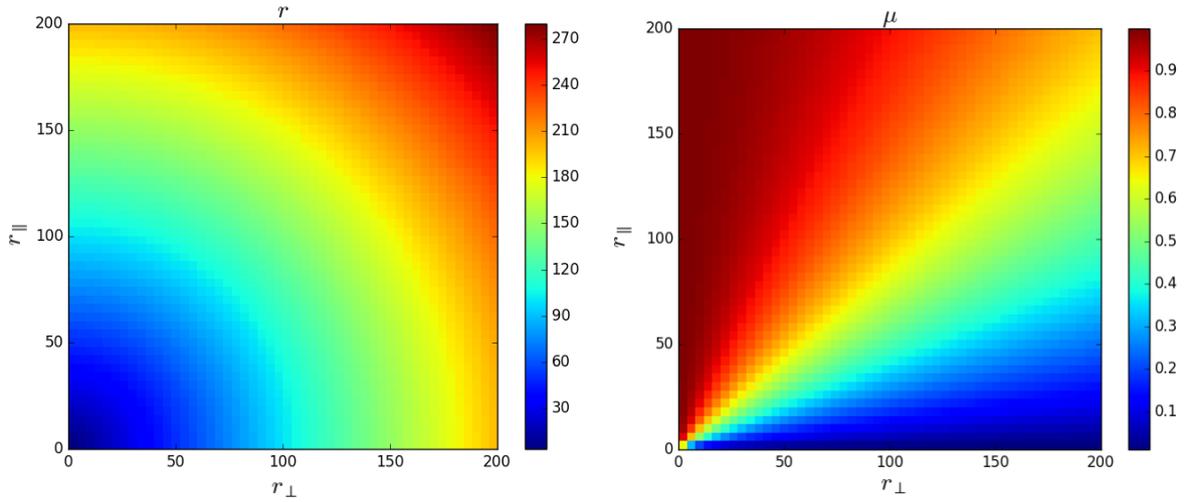


FIGURE 5.7: Grille des *bins* en  $(r, \mu)$ .

avec  $W_A = \sum_{(i,j) \in A} w_i w_j$ , le poids du *bin* A.

En pratique,  $z^A$  varie peu dans cette étude (voir figure 5.6) et la corrélation Ly $\alpha \times$  Ly $\alpha$  est mesurée à *redshift* (quasi) constant en fonction de  $r_{\parallel}$  et  $r_{\perp}$ . Le *redshift* effectif de la mesure est estimé par le *redshift* moyen  $\bar{z} = \sum_A W_A z^A / \sum_A W_A$ .

On redéfinit  $r_{\parallel}^A$  et  $r_{\perp}^A$  comme les moyennes pondérées des  $r_{\parallel,ij}$  et  $r_{\perp,ij}$  avec  $(i,j) \in A$ , respectivement :

$$r_{\parallel}^A = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \frac{r_{\parallel,i} + r_{\parallel,j}}{2}, \quad (5.21)$$

$$r_{\perp}^A = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \frac{r_{\perp,i} + r_{\perp,j}}{2}. \quad (5.22)$$

Ces trois quantités seront utilisées à la section 6.5 lors de l'ajustement de la fonction de corrélation.

On peut également mesurer la fonction de corrélation dans le système  $(r, \mu)$  en utilisant :

$$r^A = \sqrt{(r_{\parallel}^A)^2 + (r_{\perp}^A)^2}, \quad (5.23)$$

$$\mu^A = \frac{r_{\parallel}^A}{r^A}. \quad (5.24)$$

Je représente figure 5.7 les valeurs moyennées de  $(r, \mu)$  sur la grille  $(r_{\parallel}, r_{\perp})$ .

### 5.1.6 Résultats

Je représente figure 5.8 la fonction de corrélation estimée par l'équation 5.15 dans les cas Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) (à gauche) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) (à droite) en fonction de  $(r_{\parallel}, r_{\perp})$ . Les deux images sont très similaires : on ne voit aucune corrélation, excepté pour les petites séparations.

Sur la figure 5.9, je montre la fonction de corrélation multipliée par la distance de séparation au carré dans les cas Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) (à gauche) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) (à droite). On voit que le signal Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) est plus bruité que Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) à cause du nombre inférieur de paires. En revanche le pic BAO semble toujours invisible.

Pour pouvoir voir ce pic, il faut sommer la fonction de corrélation sur des portions de la grille pour en accroître le signal. Pour cela, je passe en coordonnées  $(r, \mu)$  et je somme sur des gammes de  $\mu$  en laissant  $r$  libre. J'obtiens la figure 5.10 où Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) est présentée en rouge et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) en bleu. Les erreurs sont définies comme la racine carré de l'inverse de la variance des *pixels*, voir section 5.3.

Dans le cas de Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ), on voit nettement le pic BAO à  $100 h^{-1}$ Mpc quand  $\mu$  est dans  $[0.80, 0.95]$  et  $[0.95, 1.00]$ . On remarque également un excès de corrélation pour  $r \in [40, 70] h^{-1}$ Mpc et  $\mu \in [0.95, 1.0]$ . On verra dans le chapitre suivant qu'il s'agit d'un artefact créé par les absorptions par les métaux du milieu intergalactique. En supposant que toutes les absorptions du champ de  $\delta$  étaient de type Ly $\alpha$ , j'ai mal estimé les distances de séparation de l'équation 5.6

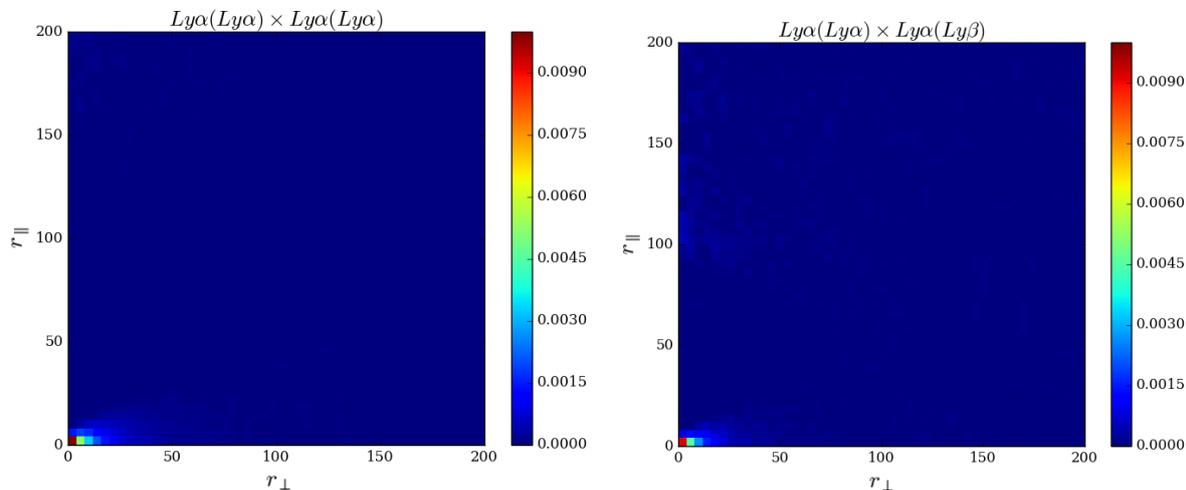


FIGURE 5.8: Représentation à deux dimensions selon les coordonnées  $(r_{\parallel}, r_{\perp})$  des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . À cause de sa faible amplitude, le pic BAO n'est pas visible.

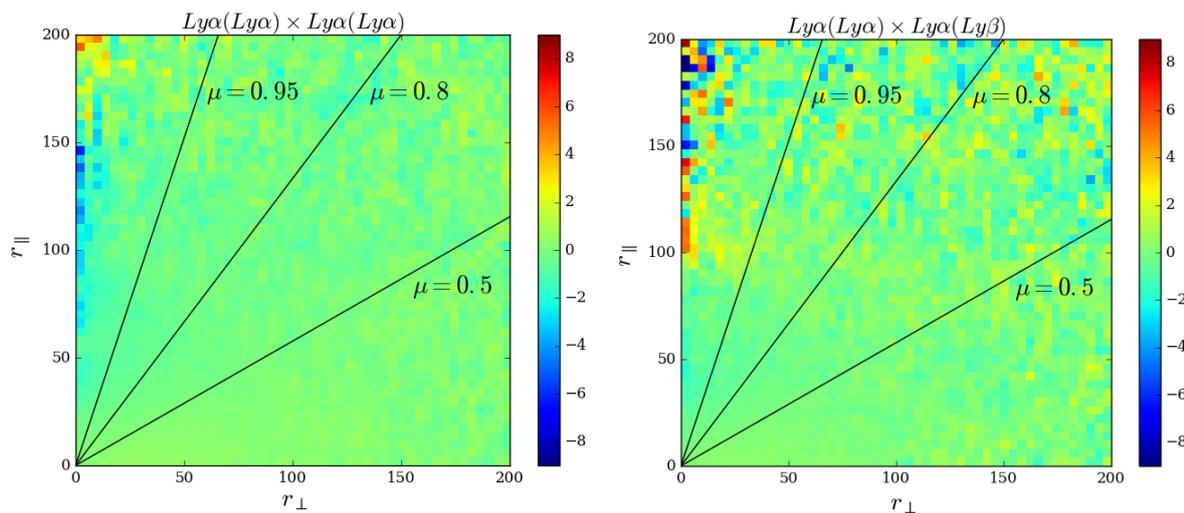


FIGURE 5.9: Représentation à deux dimensions selon les coordonnées  $(r_{\parallel}, r_{\perp})$  des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , multipliées par la distance de séparation au carré. Cette multiplication permet d'augmenter l'amplitude de la fonction aux grandes séparations. Cependant, le pic BAO n'est toujours pas visible.

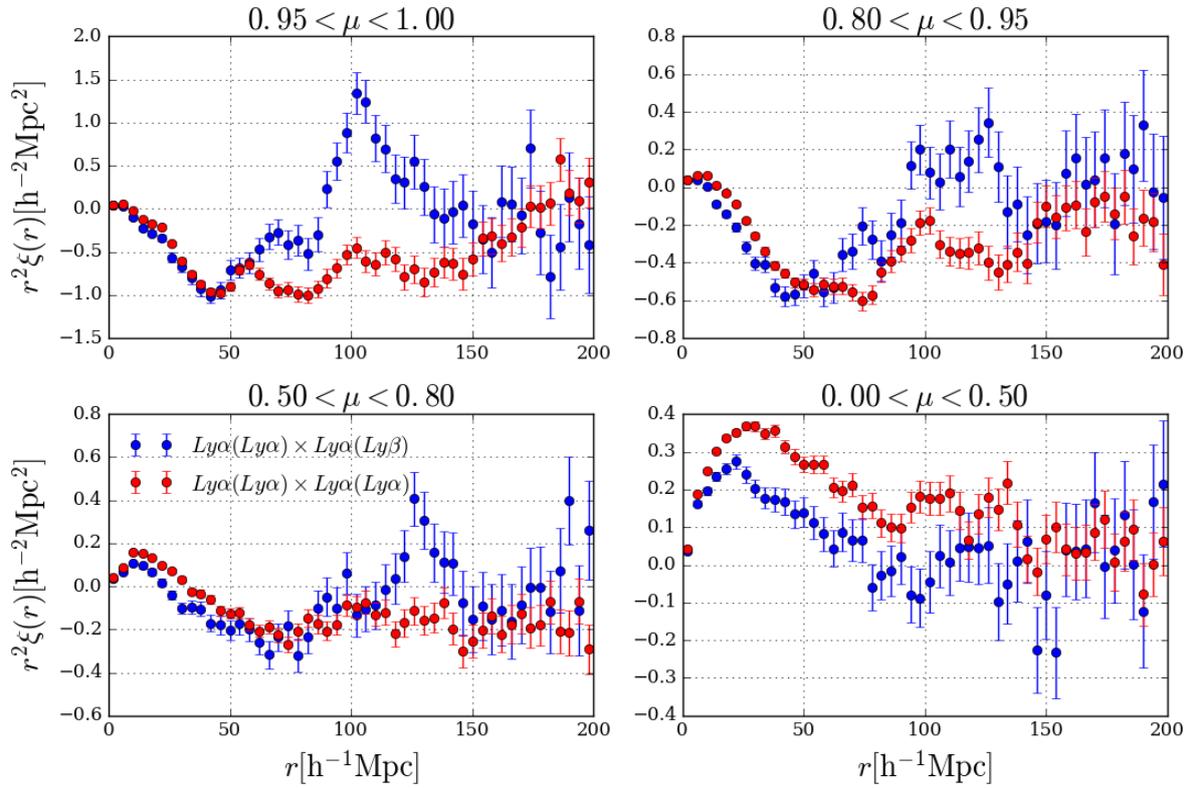


FIGURE 5.10: Mesure des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (en rouge) et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (en bleu) en fonction de la séparation  $r$ . Chacun des quadrants est obtenu en sommant la fonction de corrélation à 2D dans une gamme de  $\mu$  donnée, ce qui permet de voir le pic BAO autour de  $100$  h $^{-1}$  Mpc. Les erreurs sont estimées à partir de la matrice de covariance calculée par la méthode de *subsampling* (voir section 5.3.2).

et une partie de l'excès visible à bas  $r$  (avant  $40 h^{-1}\text{Mpc}$ ) a été translatée à plus grand  $r$  (voir section 6.2).

Dans le cas de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , le pic BAO n'est pas clairement visible. Le large excès de corrélation mesuré pour  $\mu \in [0.95, 1.00]$  est dû à la distorsion créée par la projection des  $\delta$  de la section 4.4. On verra que cet effet n'a pas d'influence sur la position du pic BAO (section 6.3).

En comparant  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  on remarque que :

- les barres d'erreurs de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  sont plus grandes que celles de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  à cause du nombre inférieur de paires dans le premier échantillon,
- l'amplitude de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  est plus grande que celle de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ , laissant penser que le biais des absorptions  $\text{Ly}\alpha$  n'est pas le même dans la région  $\text{Ly}\alpha$  que dans la région  $\text{Ly}\beta$ . Cet effet n'est pas physique, mais dû à projection des  $\delta$  qui n'a pas le même effet sur les deux fonctions de corrélation. On montrera dans le chapitre suivant que les biais sont effectivement identiques pour les deux jeux de données.

Dans la section suivante, j'étudie la fonction de corrélation à une dimension afin d'identifier les métaux du milieu intergalactique contaminants en utilisant la fonction de corrélation.

## 5.2 Détection des métaux

On sait que certaines absorptions du milieu intergalactique ne sont pas de type  $\text{Ly}\alpha$  et qu'elle peuvent donner lieu à des pics artificiels dans la fonction de corrélation (voir figure 5.10). Il est donc essentiel pour pouvoir modéliser la fonction de corrélation mesurée de connaître la nature des transitions contaminantes.

Pour cela, l'identification des contaminants se fait en deux étapes :

- Dans un premier temps, je détecte les métaux présents dans les DLA,
- puis j'identifie les paires de transitions qui contaminent la fonction de corrélation à une dimension.

J'introduis une notation spécifique pour les transitions métalliques : par exemple,  $\text{Si II}(1190)$  désigne la transition de type Si II autour de  $1190\text{\AA}$  (*cf.* annexe B).

### 5.2.1 Empilement des DLA

Les éléments métalliques sont principalement contenus dans les absorbants les plus denses (Pieri et al. 2010) et spécialement dans les DLA. Dans cette sous-section, on va répertorier les

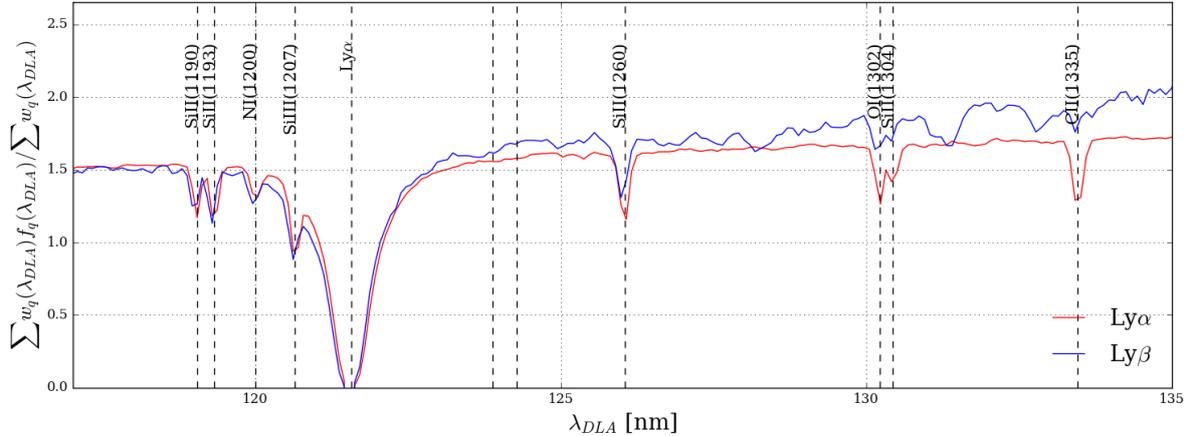


FIGURE 5.11: Empilement des spectres contenant un ou plusieurs DLA dans le référentiel au repos de chaque DLA. On voit apparaître des raies d'absorption dont la largeur est proportionnelle à la taille des nuages absorbants et la profondeur, à la densité d'absorbant. Afin de pouvoir visualiser les autres transitions atomiques, la raie Ly $\alpha$  a été coupée. On remarque que celle-ci est beaucoup plus profonde que les autres, ce qui est cohérent avec le fait que l'absorption Ly $\alpha$  est dominante dans le milieu intergalactique. On voit que la région Ly $\beta$  (en bleu) contient toutes les transitions déjà présentes dans Ly $\alpha$  (en rouge).

contaminants majeurs des DLA.

Le catalogue de DLA (sous-section 4.2.2.2) donne le *redshift*  $z_{\text{DLA}}$  de chaque DLA identifié. Ainsi, la longueur d'onde  $\lambda$  observée (depuis la Terre) est absorbée par le milieu à  $\lambda/(1+z_{\text{DLA}})$ . Pour faire apparaître les transitions absorbées dans ces systèmes, je déplace le flux  $f$  de chaque spectre contenant un DLA du référentiel de l'observateur terrestre à celui de l'absorbant :

$$f(\lambda) \longrightarrow f(\lambda_{\text{DLA}} = \frac{\lambda}{1+z_{\text{DLA}}}) \quad (5.25)$$

Je somme ensuite sur l'ensemble des DLA identifiés. J'obtiens alors la figure 5.11 où chaque transition majeure est visible sous la forme d'un puits dont la profondeur augmente avec la fréquence de l'absorption et la largeur dépend de la longueur moyenne des nuages absorbants. Comme attendu, la transition Ly $\alpha$  est très majoritaire dans les régions Ly $\alpha$  et Ly $\beta$  : alors que la profondeur des absorptions métalliques ne dépasse pas  $0.5 \cdot 10^{-19} \text{Wm}^{-2} \text{nm}^{-1}$ , elle est de plusieurs centaines voir milliers de fois supérieure dans le cas de Ly $\alpha$ .

On remarque que les transitions visibles dans la région Ly $\alpha$ , le sont aussi dans la région Ly $\beta$ . Les principaux métaux détectés à proximité de Ly $\alpha$  sont : SiII(1190), SiII(1193), NiI(200), SiIII(1207) et SiII(1260). Les autres sont trop éloignés en longueur d'onde pour contaminer la fonction de corrélation pour  $r < 200 h^{-1} \text{Mpc}$ .

Afin de limiter le nombre de contaminants à étudier, je fais une deuxième sélection en utilisant la fonction de corrélation à une dimension.

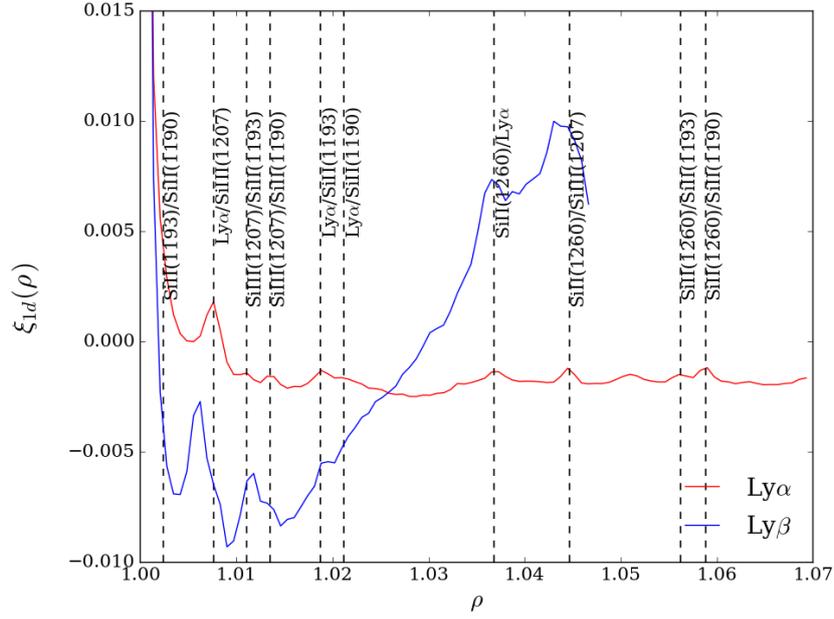


FIGURE 5.12: Corrélation entre les paires  $\hat{\delta}_i \hat{\delta}_j$  appartenant au même spectre ( $q_i = q_j$ ) en fonction du rapport de leurs longueurs d'onde  $\rho = \lambda_i / \lambda_j$ . Si les absorbants  $a$  et  $b$  sont présents dans les régions étudiées, des pics de corrélation apparaissent à  $\rho = \lambda_a / \lambda_b$  et  $\rho = \lambda_b / \lambda_a$ . On ne représente ici que les cas où  $\rho > 1$  puisque la fonction est symétrique par rapport à 1. La forme de la fonction de corrélation à 1D dans la région Ly $\beta$  est causée par la projection des  $\delta$  de la section 4.4. On verra dans la suite que, plus la région est petite, plus l'effet de la distorsion est grand sur les fonction de corrélation. Le pic visible dans la région Ly $\beta$  à  $\rho \simeq 1.06$  est créé par les corrélations O VI(1032)/Ly $\beta$  et O VI(1038)/O VI(1032). À cause de la séparation en longueur d'onde entre ces transitions et la transition Ly $\alpha$ , elles ne sont pas incluses parmi les contaminants métalliques de cette étude.

### 5.2.2 La fonction de corrélation à une dimension

Dans cette sous section, j'utilise la notation  $\hat{\delta}_{q_i}(\lambda_i)$  pour désigner le champ d'absorption au *pixel*  $i$ . Je rappelle que  $\hat{\delta}_{q_i}(\lambda_i) \equiv \hat{\delta}_i$ .

De façon analogue à l'équation 5.15, on définit la fonction de corrélation moyenne  $\xi^{1d}$  pour les paires d'absorbants appartenant au même spectre :

$$\xi^{1d}(\rho) = \frac{\sum_{q,\lambda} w_q(\lambda) w_q(\rho\lambda) \hat{\delta}_q(\lambda) \hat{\delta}_q(\rho\lambda)}{\sum_{q,\lambda} w_q(\lambda) w_q(\rho\lambda)}, \quad (5.26)$$

où  $\rho$  est le rapport des longueurs d'onde des paires.

Considérons un nuage absorbant les transitions  $a$  et  $b$  au *redshift*  $z_n$ . Sur le spectre, les longueurs d'onde observées  $(1+z_n)\lambda_a$  et  $(1+z_n)\lambda_b$  sont absorbées et il existe donc une corrélation positive entre  $\delta_q((1+z_n)\lambda_a)$  et  $\delta_q((1+z_n)\lambda_b)$ , qui se traduit par un excès de probabilité de  $\xi_{1d}(\lambda_b/\lambda_a)$  et  $\xi_{1d}(\lambda_a/\lambda_b)$ .

Je trace figure 5.12, la fonction de corrélation à une dimension  $\xi^{1d}(\rho)$  en rouge pour la région Ly $\alpha$ , et en bleu pour Ly $\beta$ . On voit apparaître des pics de corrélations correspondants aux paires d'absorbants présents dans ces régions. J'utilise les résultats de la sous-section précédente pour calculer les rapports des longueurs d'onde et identifier les pics.

Finalement, les métaux considérés dans cette étude sont : Si II(1190), Si II(1193), Si III(1207) et Si II(1260). La transition N V(1200) visible dans l'empilement des DLA n'est pas détectée dans la fonction de corrélation à une dimension, c'est pourquoi je ne le considérerai pas dans la suite.

La détection des métaux sera utilisée à la section 6.2 lors de la modélisation de la fonction de corrélation mesurée.

Dans la dernière section de ce chapitre, je vais m'intéresser à l'estimation de la matrice de covariance de la fonction de corrélation. La matrice de covariance sera ensuite utilisée pour estimer les erreurs et modéliser la fonction de corrélation.

## 5.3 La matrice de covariance

La matrice de covariance de la fonction de corrélation mesurée est définie par :

$$C_{AB} = \langle \hat{\xi}_A \hat{\xi}_B \rangle - \langle \hat{\xi}_A \rangle \langle \hat{\xi}_B \rangle, \quad (5.27)$$

où  $\langle \rangle$  désigne la moyenne sur les réalisations de l'expérience (différente de l'estimateur de l'équation 5.15).  $C_{AB}$  peut être calculée de deux manières :

- en reprenant la définition exacte de  $\hat{\xi}_A$  donnée par l'équation 5.13 et on développant les termes  $\langle \hat{\delta}_i \hat{\delta}_j \rangle$ ,
- en utilisant l'estimateur de  $\hat{\xi}_A$  de l'équation 5.15.

La première méthode est plus précise mais demande de grandes ressources de calcul. [Delubac et al. \(2015\)](#) y a recours pour justifier la validité de la seconde méthode. Je la présente dans ce manuscrit à titre informatif et j'adopte la seconde méthode pour mon étude.

### 5.3.1 Développement de Wick

En utilisant l'équation 5.13,  $C_{AB}$  devient :

$$C_{AB} = W_A^{-1} W_B^{-1} \sum_{(i,j) \in A} \sum_{(k,l) \in B} w_i w_j w_k w_l (\langle \hat{\delta}_i \hat{\delta}_j \hat{\delta}_k \hat{\delta}_l \rangle - \langle \hat{\delta}_i \hat{\delta}_j \rangle \langle \hat{\delta}_k \hat{\delta}_l \rangle). \quad (5.28)$$

Les moyennes d'ensemble désignent des moyennes sur plusieurs réalisations de l'expérience. En supposant que les  $\hat{\delta}$  sont gaussiens, on peut utiliser le théorème d'Isserlis<sup>2</sup> (ou théorème de Wick en physique des particules) qui transforme  $\langle \hat{\delta}_i \hat{\delta}_j \hat{\delta}_k \hat{\delta}_l \rangle$  en une somme de produits de moments d'ordre deux :

$$C_{AB} = W_A^{-1} W_B^{-1} \sum_{(i,j) \in A} \sum_{(k,l) \in B} w_i w_j w_k w_l (\langle \hat{\delta}_i \hat{\delta}_k \rangle \langle \hat{\delta}_j \hat{\delta}_l \rangle + \langle \hat{\delta}_i \hat{\delta}_l \rangle \langle \hat{\delta}_j \hat{\delta}_k \rangle). \quad (5.29)$$

La figure 5.13 représente pour deux couples de *pixels*  $(i, j)$  et  $(k, l)$ , les six configurations physiques possibles. Je décris chacun des types par ordre de contribution dans la covariance :

- T1 : Les couples sont identiques :  $(i, j) = (k, l)$ .
- T2 :  $i$  et  $k$  sont identiques.  $l$  et  $j$  sont sur le même spectre à différentes longueurs d'onde
- T3 :  $i$  et  $k$  sont sur le même spectre à différentes longueurs d'onde. Idem pour  $l$  et  $j$ .
- T4 :  $i$  et  $k$  sont identiques.  $l$  et  $j$  appartiennent à deux spectres différents.
- T5 :  $i$  et  $k$  sont sur le même spectre à différentes longueurs d'onde.  $l$  et  $j$  appartiennent à deux spectres différents.
- T6 :  $i$  et  $k$  appartiennent à deux spectres différents. Idem pour  $l$  et  $j$ .

La matrice de corrélation  $C_{AB}$  est ensuite découpée selon ces six contributions :

$$C_{AB} = C_{AB}^{T1} + C_{AB}^{T2} + C_{AB}^{T3} + C_{AB}^{T4} + C_{AB}^{T5} + C_{AB}^{T6}. \quad (5.30)$$

La plus grande contribution à la variance (60%, [Delubac et al. 2015](#)) provient du cas T1 où l'on corrèle deux fois la même paire. On peut négliger les termes  $\langle \delta_i \delta_j \rangle$  devant les  $\langle \delta_i \delta_i \rangle$ .  $C_{AA}^{T1}$  (équation 5.29) devient :

$$C_{AA}^{T1} \simeq W_A^{-2} \sum_{(i,j) \in A} w_i^2 w_j^2 \langle \hat{\delta}_i^2 \rangle \langle \hat{\delta}_j^2 \rangle. \quad (5.31)$$

<sup>2</sup>[https://en.wikipedia.org/wiki/Isserlis%27\\_theorem](https://en.wikipedia.org/wiki/Isserlis%27_theorem)

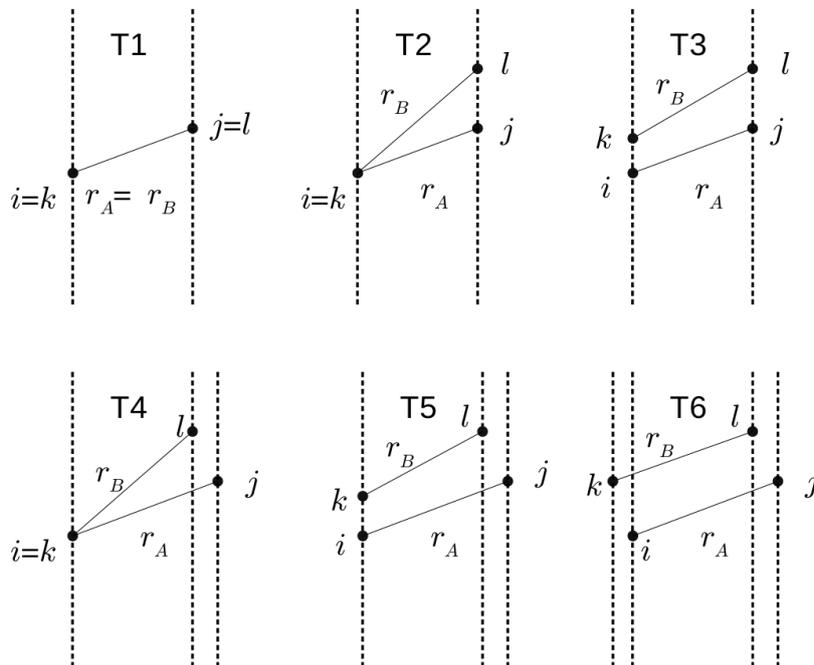


FIGURE 5.13: Configurations physiques des paires de *bins* possibles (Delubac et al. 2015). Les lignes en pointillés désignent les lignes de visée des quasars. La très grande majorité de la variance (97%) provient des cas où les paires de spectres sont identiques (configurations T1, T2 et T3). Parmi les configurations T2 et T4, on inclue également le cas symétrique  $l = j$  et  $i \neq k$ .

La deuxième contribution (25% de la variance totale) concerne les couples de paires avec une absorption ( $i = k$ ) et une forêt ( $q_l = q_j$ ) en commun. On a :

$$C_{AB}^{T2} \simeq W_A^{-1} W_B^{-1} \sum_{(i,j) \in A} \sum_{(i,l) \in B} w_i^2 w_j w_l \langle \hat{\delta}_i^2 \rangle \langle \hat{\delta}_j \hat{\delta}_l \rangle. \quad (5.32)$$

La troisième contribution (15% de la variance totale) considère les mêmes couples de forêts ( $q_i = q_k$  et  $q_l = q_j$ ) à des longueurs d'onde différentes. Puisque les termes  $C_{AB}^{T1}$ ,  $C_{AB}^{T2}$  et  $C_{AB}^{T3}$  font intervenir des paires appartenant au même spectre, on peut les calculer en utilisant les résultats de la section 5.2.2.

Les trois premiers termes sont suffisants pour le calcul de la matrice de covariance. En effet, les autres contributions sont inférieures à la précision statistique de la mesure (Delubac et al. 2015, Bautista et al. 2017).

### 5.3.2 Estimation par la méthode de *subsampling*

Introduit pour la première fois dans Delubac et al. (2015), le calcul de la covariance par la méthode de *subsampling* est beaucoup plus rapide que la méthode précédente et donne des résultats quasiment identiques (Delubac et al. 2015, Bautista et al. 2017). Ici, j'utilise l'approximation de la moyenne d'ensemble donnée par l'équation 5.15 dans l'équation 5.27.

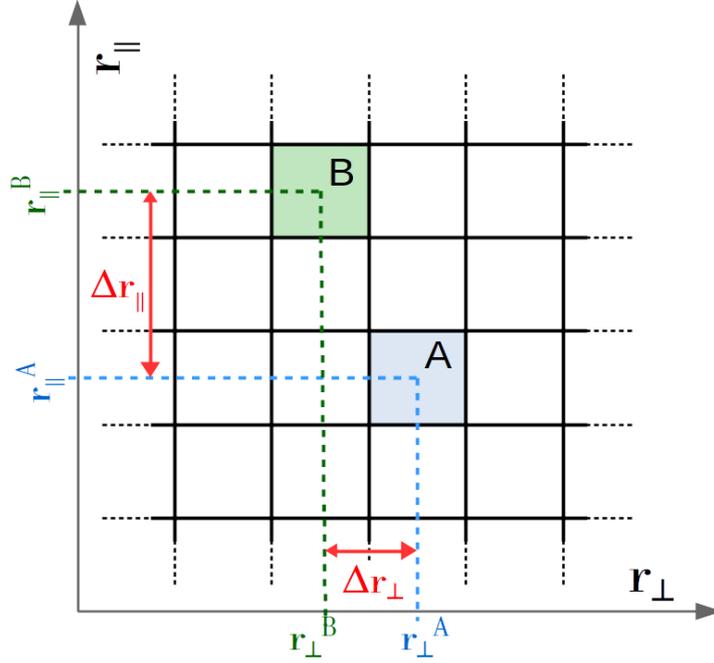


FIGURE 5.14: On étudie la covariance  $C_{AB}$  entre la fonction de corrélation calculée aux *bins*  $A$  et  $B$ . Delubac et al. (2015) montrent que la matrice de corrélation  $Corr_{AB} = C_{AB}/\sqrt{C_{AA}C_{BB}}$  ne dépend pas des positions de  $A$  et  $B$ , mais seulement des séparations  $\Delta r_{\parallel}$  et  $\Delta r_{\perp}$ .

### 5.3.2.1 Calcul de la matrice de covariance

Si les paires sont réparties de manière aléatoire dans des sous-ensembles  $h$  homogènes, alors l'équation 5.27 devient :

$$C_{AB} \simeq W_A^{-1} W_B^{-1} \sum_h W_A^h W_B^h (\hat{\xi}_A^h \hat{\xi}_B^h - \langle \hat{\xi}_A \rangle \langle \hat{\xi}_B \rangle). \quad (5.33)$$

Plus le nombre de sous-échantillons est grand, plus la méthode par *subsampling* donne un résultat proche de la vraie matrice de covariance. En pratique, on a accès à un nombre limité de sous-ensembles. Cela crée des corrélations entre les *bins* qui sont visibles dans la matrice de corrélation :

$$Corr_{AB} = \frac{C_{AB}}{\sqrt{C_{AA}C_{BB}}}. \quad (5.34)$$

La covariance est maximale dans le cas  $A = B$ , donc  $C_{AB} \leq \sqrt{C_{AA}C_{BB}}$  quels que soient  $A$  et  $B$ . Ainsi, la matrice de corrélation  $Corr_{AB}$  se compose de 1 sur la diagonale et de termes de valeur absolue strictement inférieure à 1, ailleurs.

Soient  $A$  et  $B$ , deux *bins* avec  $|r_{\parallel}^A - r_{\parallel}^B| = \Delta r_{\parallel}$  et  $|r_{\perp}^A - r_{\perp}^B| = \Delta r_{\perp}$  (voir le schéma de la figure 5.14). En utilisant le développement de Wick (sous-section 5.3.1), Delubac et al. (2015) puis Bautista et al. (2017) montrent que  $Corr_{AB}$  dépend de  $\Delta r_{\parallel}$  et  $\Delta r_{\perp}$  mais pas des positions respectives de  $A$  et  $B$  sur la grille. De plus, seules les contributions avec  $\Delta r_{\perp} = 0 \text{ h}^{-1}\text{Mpc}$  et

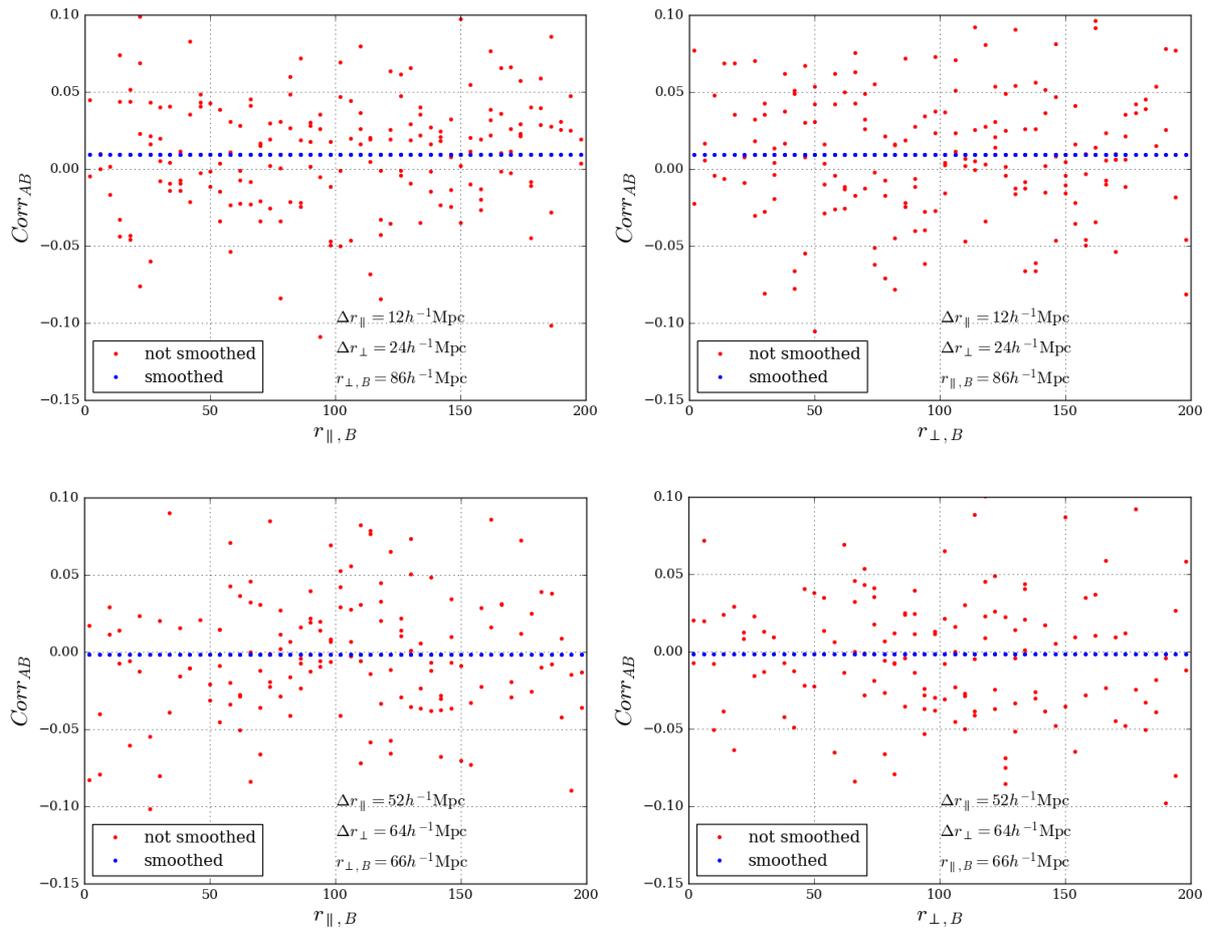


FIGURE 5.15:  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  : représentation des éléments de  $Corr_{AB}$  pour  $\Delta r_{\parallel}, \Delta r_{\perp}$  en fonction de  $r_{\parallel}^B$  (à gauche) et de  $r_{\perp}^B$  (à droite) pour les matrices lissées (bleu) et non lissée (rouge) dans deux exemples. Dans le cas de la matrice lissée, les éléments  $Corr_{AB}$  sont indépendants de  $r_{\parallel}^B$  et  $r_{\perp}^B$  alors qu'on voit une dispersion quand la matrice n'est pas lissée.

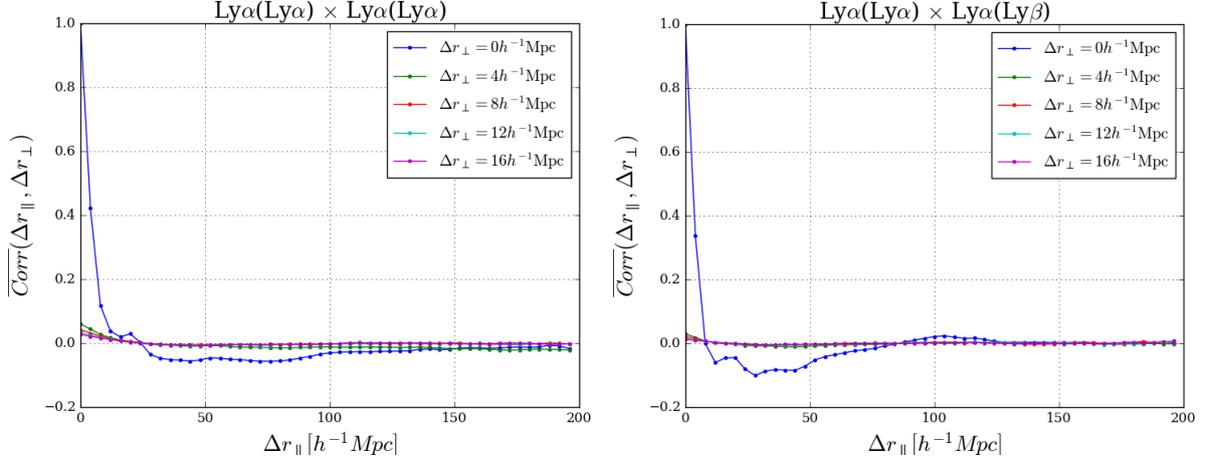


FIGURE 5.16: Représentation des termes de  $Corr_{AB}$  moyennés sur  $\Delta r_{\parallel}, \Delta r_{\perp}$  selon  $\Delta r_{\parallel}$  et pour différentes valeurs de  $\Delta r_{\perp}$  dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (à gauche) et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (à droite).

$\Delta r_{\parallel} \leq 20 h^{-1} \text{Mpc}$  sont plus grandes que la précision statistique du *subsampling*<sup>3</sup>.

Je représente figure 5.15 les éléments  $Corr_{AB}$  de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (en rouge) en fixant  $\Delta r_{\parallel}, \Delta r_{\perp}$  puis  $r_{\parallel}^B$  ou  $r_{\perp}^B$  dans deux exemples différents. Puisque  $Corr_{AB}$  ne doit pas dépendre de  $r_{\parallel}^B$  et  $r_{\perp}^B$ , on ne devrait voir aucune variation. Or on observe une dispersion due au nombre insuffisant de sous-échantillons. Il est important de noter que la répartition des points est aléatoire et donc d'origine statistique.

Je décris dans le paragraphe suivant le processus qui permet de lisser ces fluctuations statistiques.

### 5.3.2.2 Lissage

Je supprime la dispersion observée figure 5.15 en moyennant  $Corr_{AB}$  sur les *bins* de même séparation. La matrice de corrélation lissée  $Corr_{AB}^s$  est définie par :

$$Corr_{AB}^s = \frac{1}{N_{AB}} \sum_{A', B'} Corr_{A'B'}, \quad (5.35)$$

où  $A'$  et  $B'$  sont tels que  $|r_{\parallel}^{A'} - r_{\parallel}^{B'}| = |r_{\parallel}^A - r_{\parallel}^B|$  et  $|r_{\perp}^{A'} - r_{\perp}^{B'}| = |r_{\perp}^A - r_{\perp}^B|$ .  $N_{AB}$  désigne le nombre de couples  $A', B'$  satisfaisant à ces deux conditions.

Les éléments  $Corr_{AB}^s$  ne dépendent plus que de  $\Delta r_{\parallel}$  et  $\Delta r_{\perp}$ , et sont désignés par les points bleus sur la figure 5.15. Je représente figure 5.16 pour  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , les éléments moyennés de  $Corr_{AB}^s$  en fonction de  $\Delta r_{\parallel}$  pour différentes valeurs de  $\Delta r_{\perp}$ . On voit que la corrélation diminue rapidement quand  $\Delta r_{\perp}$  et  $\Delta r_{\parallel}$  augmentent et qu'elle est largement dominée par la variance.

<sup>3</sup>La précision statistique du *subsampling* est donnée par  $1/\sqrt{N_s}$ , avec  $N_s$ , le nombre de sous-échantillons

Les erreurs sur la mesure de  $\langle \hat{\xi}_A \rangle$  sont estimées par  $C_{AA}^{-1/2}$ .  $C_{AB}$  sera utilisé au chapitre 6 lors de l'ajustement du modèle théorique.

### 5.3.3 Covariance entre deux fonctions de corrélation

Dans le chapitre 7, je ferai un ajustement combiné des paramètres sur les deux fonctions de corrélation Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ). Je vérifie dans cette sous-section que les deux mesures sont indépendantes.

Pour cela je calcule la matrice de covariance croisée par *subsampling* entre les fonctions de corrélation Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ). Je représente figure 5.17 les éléments moyennés sur  $(\Delta r_{\parallel}, \Delta r_{\perp})$  de la matrice de corrélation croisée. On voit qu'ils sont négligeables devant 1. Ce qui revient à dire que la covariance croisée est négligeable devant la variance de chacune des mesures. On peut donc considérer que les deux mesures sont indépendantes.

#### Synthèse du chapitre

Dans ce chapitre, j'ai utilisé les champs d'absorption Ly $\alpha$  pour calculer la fonction de corrélation Ly $\alpha$  à deux points grâce à l'équation 5.15.

La fonction de corrélation à une dimension a permis de détecter les transitions contaminantes SiII(1190), SiII(1193), SiIII(1207) et SiII(1260). Leur contribution sera prise en compte dans le modèle, détaillé au chapitre suivant.

Pour estimer la précision de la mesure, j'ai calculé la matrice de covariance de la fonction de corrélation par la méthode de *subsampling*.

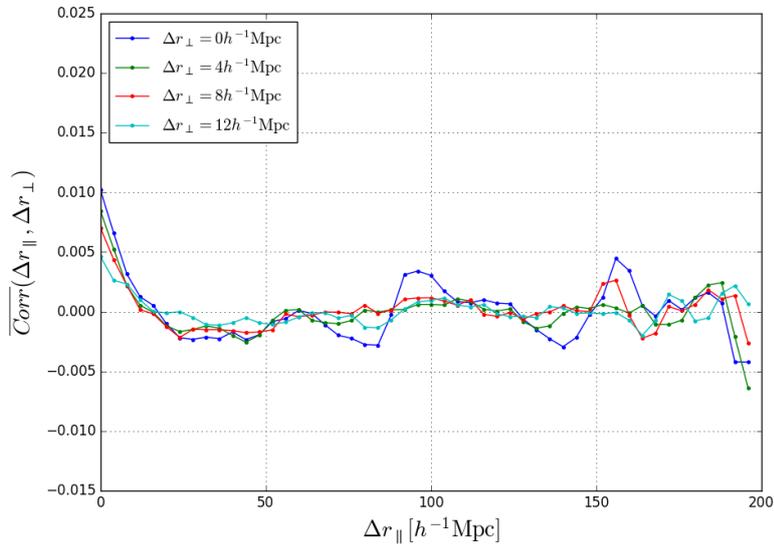


FIGURE 5.17: Éléments moyennés de la matrice de corrélation croisée entre  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . La covariance croisée est négligeable devant les variances des fonctions individuelles.



## Chapitre 6

# Modélisation et ajustement de la fonction de corrélation mesurée

L'objet de ce chapitre est d'établir et d'ajuster un modèle théorique des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  mesurées.

[Bautista et al. \(2017\)](#) proposent d'écrire le modèle sous la forme de 4 contributions :

$$\xi_A^{mod} = \sum_B D_{AB} \left( \xi_B^{Ly\alpha-Ly\alpha} + \xi_B^{metals} + \xi_B^{bb} \right) \quad (6.1)$$

- Le terme  $\xi^{Ly\alpha-Ly\alpha}$  représente la part de la fonction de corrélation totale qui est effectivement due aux couples d'absorptions  $\text{Ly}\alpha$ . Elle est dominante et nous permet de mesurer la position du pic BAO.
- Le terme  $\xi^{metals}$  modélise la contamination par les métaux.
- La contribution  $\xi^{bb}$ , appelée *broadband*, est purement mathématique. Optionnelle, elle peut être ajoutée pour inclure les erreurs systématiques non étudiées dans les parties précédentes et vérifier qu'elles n'ont pas d'influence sur la position du pic BAO.
- La matrice de distorsion  $D_{AB}$  prend en compte les effets de la projection des champs de  $\delta$  (section 4.4) sur la fonction de corrélation.

Dans les sections qui suivent, je vais expliciter chacun des quatre éléments de l'équation 6.1.

### 6.1 La fonction de corrélation $\text{Ly}\alpha$

Dans un premier temps, je détermine le spectre de puissance  $\text{Ly}\alpha \times \text{Ly}\alpha$ ,  $P^{Ly\alpha-Ly\alpha}$ , en partant du spectre de puissance linéaire de la matière auquel j'applique quatre transformations :

- extraction des oscillations acoustiques baryoniques,
- modélisation des distorsions dues au *redshift*,
- prise en compte des effets non-linéaires à grand  $k$ ,
- effet de l'échantillonnage.

Puis j'obtiens  $\xi^{Ly\alpha-Ly\alpha}$  par une transformation de Fourier de  $P^{Ly\alpha-Ly\alpha}$ .

### 6.1.1 Extraction du pic BAO du spectre de puissance linéaire

Kirkby et al. (2013) proposent de séparer le spectre de puissance en deux composantes : une composante lisse,  $P^{lisse}$ , et une composante qui décrit les oscillations baryoniques,  $P^{BAO}$ . De cette façon, la position du pic BAO est indépendante du modèle physique utilisé pour ajuster la composante lisse.

Pour cela, on calcule le spectre de puissance linéaire de la matière,  $P_{lin}$ , au *redshift* de référence  $z_{ref} = 2.3$  en utilisant l'outil CAMB (Lewis et al. 2000) appliqué à la cosmologie fiducielle donnée en annexe A.  $P^{lisse}$  est obtenu en 3 étapes :

- on calcule  $\xi_{lin}$ , la transformée de Fourier de  $P_{lin}$ ,
- on modélise la partie lisse de  $\xi_{lin}$  en interpolant  $\xi_{lin}$  dans la région  $[50,86] \cup [150,190]h^{-1}\text{Mpc}$  en utilisant une décomposition de Legendre<sup>1</sup>,
- on réalise une nouvelle transformée de Fourier pour obtenir  $P^{lisse}$ .

On calcule ensuite la composante  $P^{BAO}$  via la relation :

$$P_{lin}(\vec{k}, z_{ref}) = P^{lisse}(\vec{k}, z_{ref}) + P^{BAO}(\vec{k}, z_{ref}). \quad (6.2)$$

Je représente figures 6.1, le spectre de puissance de la matière (en noir), la composante lisse (c'est-à-dire sans oscillations acoustiques baryoniques, en rouge) et la composante BAO qui est la différence des deux précédents (en bleu). Les paramètres cosmologiques sont donnés dans l'annexe A. Le terme  $P^{BAO}$  peut être négatif car il n'a pas de réalité physique, il figure seulement l'écart entre deux modèles d'univers : avec et sans BAO. De façon analogue, je représente figure 6.2, la fonction de corrélation de la matière (en noir) et ses composantes lisse (en bleu) et BAO (en rouge).

À bas *redshift*, les effets non-linéaires de formation des structures influent sur la répartition de la matière dans l'univers. Les positions initiales de la matière (imprimées dans le CMB) sont décalées de quelques  $h^{-1}\text{Mpc}$ . Ce qui a pour effet d'éloigner de quelques  $h^{-1}\text{Mpc}$  les paires du

---

<sup>1</sup>La décomposition de Legendre permet d'inclure les effets non isotropes des RSD, cf. Hamilton (1998)

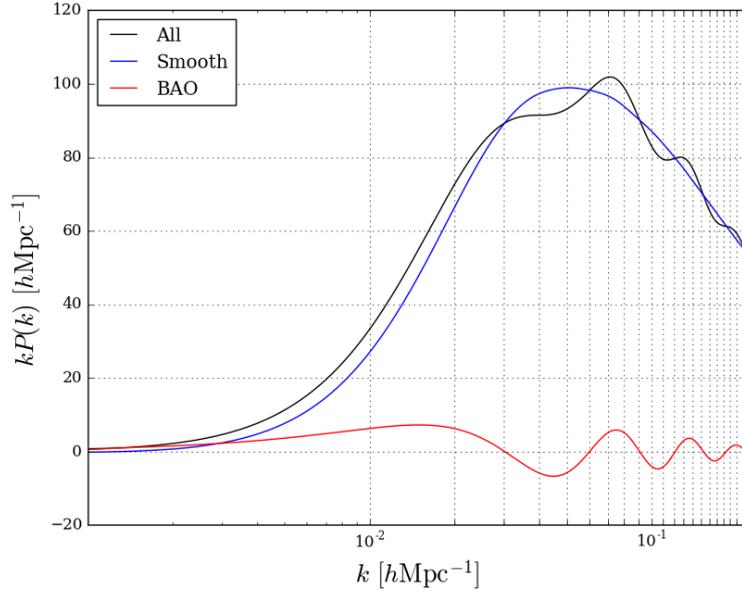


FIGURE 6.1: Le spectre de puissance linéaire  $P_{lin}$  calculé avec CAMB à  $z = 2.3$  (en noir) est la somme d'une composante lisse  $P^{lisse}$  (en bleu) et d'une composante oscillante  $P^{BAO}$  (en rouge). Le spectre de puissance lisse représente un modèle sans BAO.

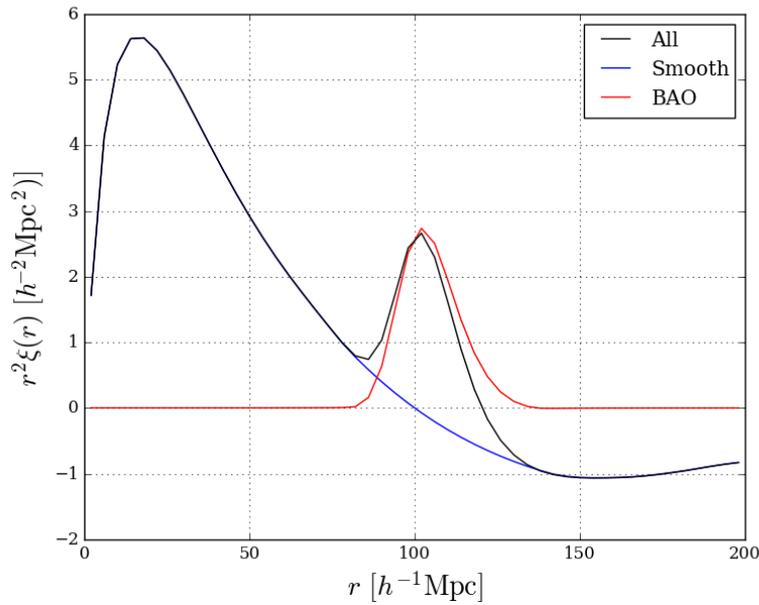


FIGURE 6.2: La fonction de corrélation de la matière calculé avec CAMB à  $z = 2.3$  (en noir) se décompose en partie lisse (en bleu) et BAO (en rouge).

pic initialement séparées de  $100 h^{-1}\text{Mpc}$ . Comme ce phénomène est moyenné sur l'ensemble des paires, la position du pic est inchangée et on observe seulement un lissage du pic. Ces effets sont étudiés en détail dans Eisenstein et al. (2007), qui traitent le phénomène comme un amortissement du spectre de puissance, modélisé par une gaussienne :

$$P_{QL}(\vec{k}, z_{ref}) = P^{\text{lis}}(\vec{k}, z_{ref}) + \exp\left(-\frac{\vec{k}^2 \vec{\Sigma}^2}{2}\right) P^{\text{BAO}}(\vec{k}, z_{ref}), \quad (6.3)$$

où  $\vec{\Sigma} = (\Sigma_{\parallel}, \Sigma_{\perp})$  désigne la moyenne quadratique des déplacements de matière parallèlement et perpendiculairement à la ligne de visée. En faisant une description analytique de ces déplacements, Eisenstein et al. (2007) calculent les valeurs de  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$  à différents *redshifts*. Kirkby et al. (2013) interpolent ces résultats à  $z = 2.4$  et obtiennent  $\Sigma_{\parallel} = 6.41 h^{-1}\text{Mpc}$  et  $\Sigma_{\perp} = 3.26 h^{-1}\text{Mpc}$ . En toute rigueur, il faudrait inclure une dépendance en *redshift* de  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$ . Cependant, puisque l'étalement du pic (quelques  $h^{-1}\text{Mpc}$ ) est faible devant la largeur du pic ( $\simeq 50 h^{-1}\text{Mpc}$ , cf. figure 6.2), ils proposent d'ignorer cette dépendance et de fixer  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$ . J'ai néanmoins vérifié lors de la phase d'ajustement, que laisser  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$  libres n'avait aucun effet significatif sur la valeur des autres paramètres.

## 6.1.2 Le modèle de Kaiser

Soient  $\delta_m$ , le contraste de densité de la matière et  $\delta_{Ly\alpha}$ , le champ d'absorption Ly $\alpha$ . Les paramètres de Kaiser  $b_{Ly\alpha}$  et  $\beta_{Ly\alpha}$  (sous-section 2.1.5) transforment  $\delta_m$  en  $\delta_{Ly\alpha}$  via la relation :

$$\delta_{Ly\alpha} = b_{Ly\alpha}(1 + \beta_{Ly\alpha}\mu^2)\delta_m. \quad (6.4)$$

### 6.1.2.1 Le biais

Le biais  $b_{Ly\alpha}$  contrôle l'amplitude de la fonction de corrélation Ly $\alpha \times$  Ly $\alpha$ . Une surdensité de matière correspond à  $\delta_m > 0$  et une absorption Ly $\alpha$  à  $\delta_{Ly\alpha} < 0$ . Par conséquent,  $b_{Ly\alpha}$  doit être négatif.

Puisque l'abondance des absorptions Ly $\alpha$  dépend du *redshift*, il doit en être de même du biais, et  $b_{Ly\alpha}$  doit suivre les évolutions en *redshift* du spectre de puissance à 1d (section 4.5). Nous adoptons la dépendance suivante :  $b(z) \propto (1+z)^{\gamma_{Ly\alpha}-1}$  avec  $\gamma_{Ly\alpha} = 2.9$ .

### 6.1.2.2 Les distorsions en *redshift*

Comme expliqué dans la sous-section 2.1.4, à cause des vitesses particulières, le *redshift* n'est pas un estimateur parfait des distances et l'écart entre la réalité et le relevé en *redshift* est maximal le long de la ligne de visée ( $\mu = 1$ ) et minimal perpendiculairement ( $\mu = 0$ ) à celle-ci. On peut observer l'effet des RSD sur la fonction de corrélation de la matière.

Sur la figure 6.3, je trace, à partir de l'équation 2.19 avec  $b_{Ly\alpha} = 1$  et  $\beta_{Ly\alpha} = 1.4$ , pour plusieurs gammes de  $\mu = r_{\parallel}/r$ , la fonction de corrélation en fonction de la distance de séparation  $r$  entre deux objets avec (en bleu) et sans (en noir) distorsions en *redshift*. On remarque que plus  $\mu$  est grand, plus le signal est amplifié.

Je représente figure 6.4 la même fonction de corrélation de la matière en fonction des projections de la distance de séparation entre deux objets parallèlement et perpendiculairement à

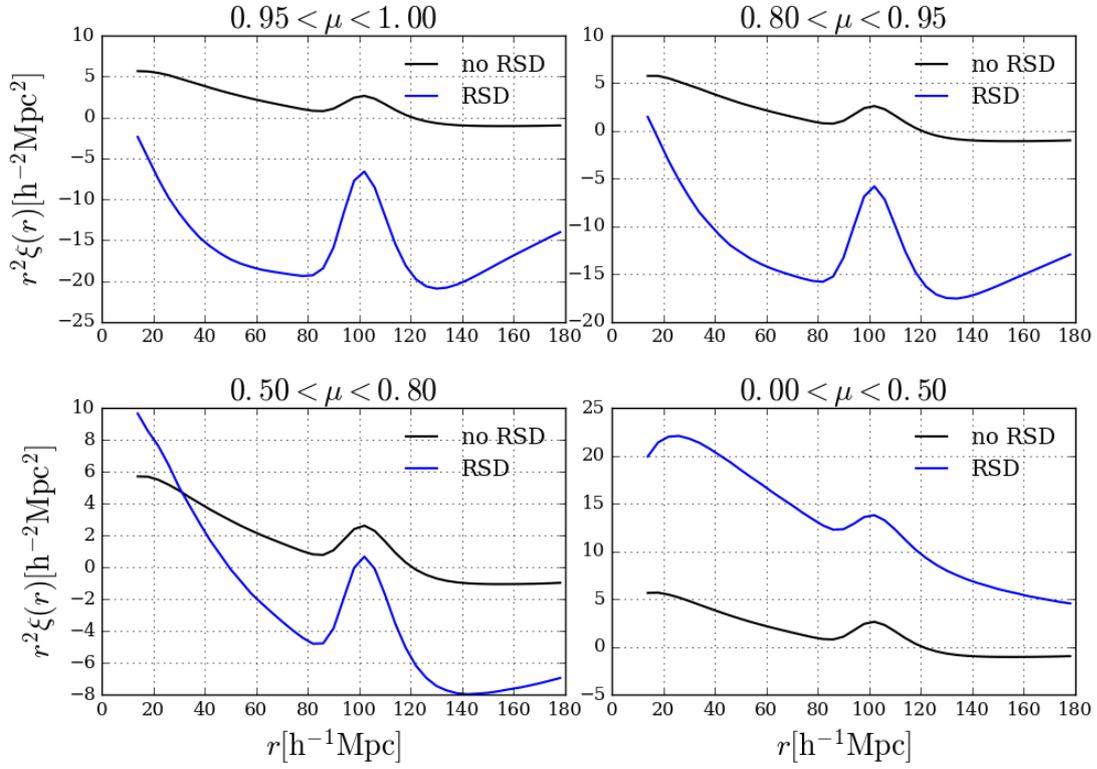


FIGURE 6.3: L'amplitude de la fonction de corrélation sans RSD (en noir) est constante. Les RSD amplifient le signal à grands  $\mu$  (en bleu).

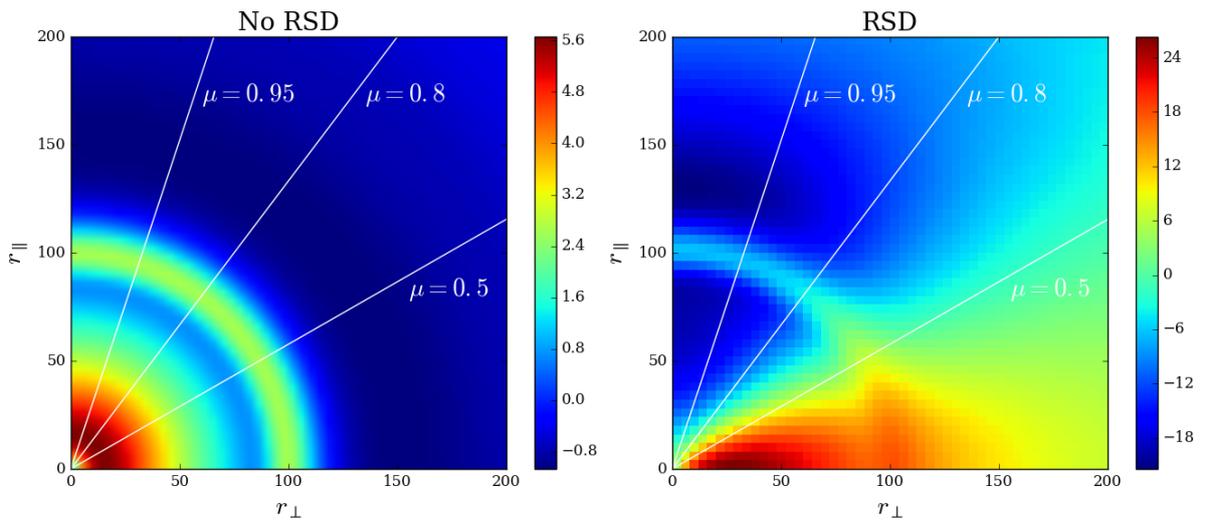


FIGURE 6.4: La fonction de corrélation sans RSD est isotrope (à gauche). Les RSD brisent cette isotropie en amplifiant le signal selon la valeur de  $\mu$  (à droite).

la ligne de visée. On remarque que l’isotropie du signal a disparu avec les distorsions en *redshift*. L’amplitude du signal varie avec  $\mu$  et elle est maximale pour  $\mu = 1$ .

### 6.1.2.3 L’effet des nuages denses

Les HCD (*High Column Density* en anglais) sont des absorbants dont la densité de colonne d’hydrogène neutre,  $N_{\text{HI}}$ , est supérieure à  $1.6 \times 10^{17}$  atomes par  $\text{cm}^2$  (Rogers et al. 2018). À cause de cette densité très grande, les profils des flux absorbés par des HCD sont élargis et présentent des bords amortis. Ces bords sont modélisables par une fonction de Voigt, convolution d’une Gaussienne, qui modélise l’élargissement Doppler<sup>2</sup>, et d’une Lorentzienne, qui modélise les élargissements naturels<sup>3</sup> et collisionnel<sup>4</sup>.

Les DLA, dont on a discuté à la sous-section 4.2.2.2, sont des HCD particulièrement denses ( $N_{\text{HI}} > 10^{21}$  atomes par  $\text{cm}^2$ ). Quand les DLA sont identifiées, le flux correspondant est traité en deux étapes :

- On masque la région où le flux relatif est inférieur à 80% du flux moyen.
- Les bords sont modélisés par un profil de Voigt qui est une fonction de la densité de colonne de l’absorbant.

Si tous les HCD pouvaient être identifiés, ils pourraient être masqués à la manière des DLA et leur impact sur la fonction de corrélation serait négligeable (Rogers et al. 2018). Malheureusement, il n’existe aucun catalogue exhaustif des HCD et particulièrement dans le cas des HCD qui ne sont pas des DLA ( $1.6 \times 10^{17} < N_{\text{HI}} < 10^{21}$  atomes par  $\text{cm}^2$ ). Cette contribution importante à la fonction d’auto-corrélation  $\text{Ly}\alpha$  a été modélisée pour la première fois dans Bautista et al. (2017). L’idée générale est qu’il faut traiter les HCD différemment des absorptions  $\text{Ly}\alpha$  moyennes. On modélise le spectre de puissance des HCD,  $P_{\text{HCD}}$ , par :

$$P_{\text{HCD}}(k_{\parallel}, k_{\perp}) = b_{\text{HCD}}^2 F_{\text{HCD}}^2(k_{\parallel}) (1 + \beta_{\text{HCD}} \mu^2)^2 P_{\text{lin}}(k_{\parallel}, k_{\perp}). \quad (6.5)$$

Les paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  sont les facteurs de Kaiser propres aux HCD. On s’attend en particulier à ce que  $b_{\text{HCD}}$  soit plus grand que  $b_{\text{Ly}\alpha}$  puisque le biais des absorptions croît avec la densité moyenne (Seljak 2012) et que celle-ci est plus grande dans les HCD que dans les autres absorbants  $\text{Ly}\alpha$ . À cause de la forme particulière des absorptions par les HCD, l’amplitude de  $P_{\text{HCD}}$  dépend de l’orientation par rapport à la ligne de visée (Font-Ribera & Miralda-Escudé

<sup>2</sup>Élargissement Doppler : l’effet Doppler provoque un décalage vers le rouge ou vers le bleu du rayonnement selon que la source s’éloigne ou se rapproche de l’observateur. Dans un gaz, toutes les particules sont en mouvement dans toutes les directions, ce qui provoque un élargissement des raies spectrales. Comme la vitesse des particules dépend de leur température : plus la température du gaz est élevée, plus les différences de vitesse sont grandes, et plus les raies sont larges. Cet effet est typiquement 100 fois plus intense que l’élargissement naturel (parfois environ 1 GHz). Source : [https://fr.wikipedia.org/wiki/Raie\\_spectrale](https://fr.wikipedia.org/wiki/Raie_spectrale)

<sup>3</sup>Élargissement naturel : le principe d’incertitude relie la durée de vie  $\Delta t$  d’un état excité et la précision de son niveau énergétique  $\Delta E$ , ainsi le même niveau excité a des énergies légèrement différentes dans différents atomes. Cet effet est assez faible (typiquement quelques MHz). Source : [https://fr.wikipedia.org/wiki/Raie\\_spectrale](https://fr.wikipedia.org/wiki/Raie_spectrale)

<sup>4</sup>Élargissement collisionnel : la collision entre particules (atomes ou molécules) modifie légèrement leurs niveaux énergétiques, d’où l’élargissement des raies. La grandeur de cet effet dépend de la densité du gaz. Source : [https://fr.wikipedia.org/wiki/Raie\\_spectrale](https://fr.wikipedia.org/wiki/Raie_spectrale)

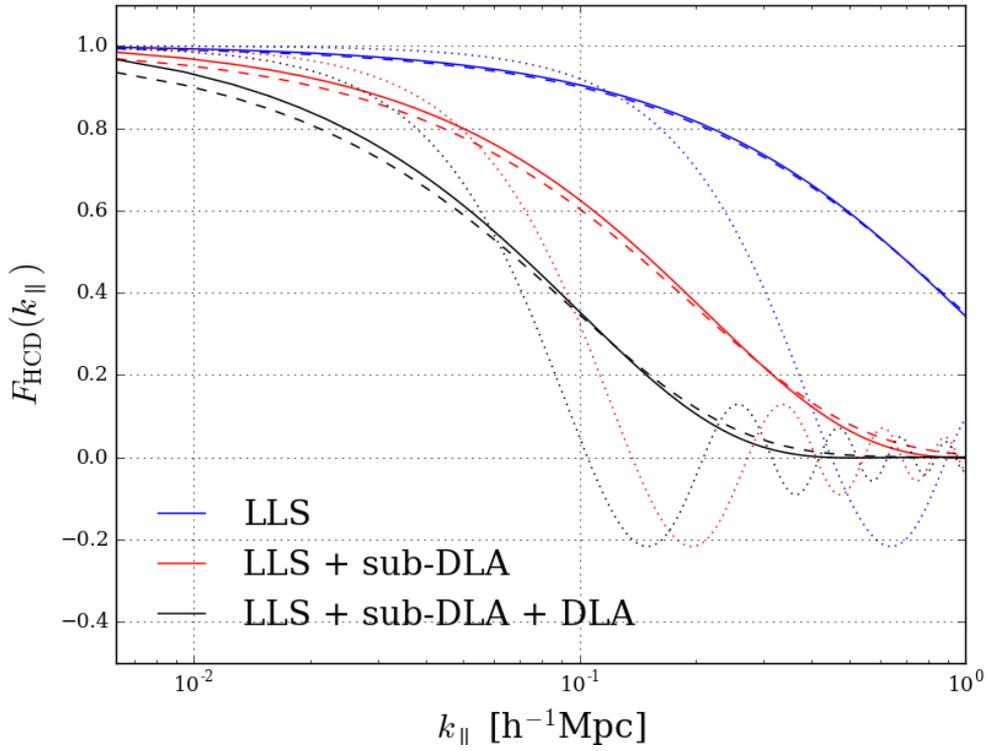


FIGURE 6.5:  $F_{\text{HCD}}$  en fonction de  $k_{\parallel}$  dans les cas  $1.6 \times 10^{17} < N_{\text{HI}} < 10^{19}$  (en bleu),  $10^{19} < N_{\text{HI}} < 2 \times 10^{20}$  (en rouge) et  $N_{\text{HI}} > 2 \times 10^{20}$  (en noir). En traits pleins, je représente les résultats du modèle complet de l'équation 6.6, avec les tirets, le modèle de Rogers et al. (2018) de l'équation 6.12 et en pointillés, le modèle de Bautista et al. (2017) en sinus cardinal.

2012). Cet effet est modélisé par la fonction  $F_{HCD}(k_{\parallel})$  qui est égale à la transformée de Fourier du profil des absorptions par les HCD.

Bautista et al. (2017) proposent d’approximer le profil amorti des HCD par une fonction rectangle. Ainsi,  $F_{HCD}(k_{\parallel})$  devient la transformée de Fourier d’une fonction rectangle, c’est-à-dire  $F_{HCD}(k_{\parallel}) = \text{sinc}(L_{HCD}k_{\parallel})$ , avec  $L_{HCD}$  un paramètre libre. Cette modélisation a été validée grâce à l’utilisation de 100 simulations (décrites dans Bautista et al. 2015) qui reproduisent les caractéristiques physiques et instrumentales de l’ensemble des spectres utilisés dans l’analyse Ly $\alpha$  de l’expérience BOSS.

Rogers et al. (2018) utilisent des simulations hydrodynamiques pour valider une expression plus précise de  $F_{HCD}$  :

$$F_{HCD}(k_{\parallel}) = \int_{\min(N_{HI})}^{\max(N_{HI})} dN_{HI} f(N_{HI})V(k_{\parallel}, N_{HI}), \quad (6.6)$$

avec  $f(N_{HI})$  la fonction de répartition des densités de colonne  $N_{HI}$  dans le milieu intergalactique et  $V(k_{\parallel}, N_{HI})$ , la transformée de Fourier du profil de Voigt des absorptions par les HCD. Notons que la forme proposée par Bautista et al. (2017) est une approximation de l’équation 6.6 avec un profil des absorptions par les HCD rectangulaire et une répartition uniforme de la densité de colonne. Les simulations étudiées par Rogers et al. (2018) ont permis de montrer que cette approximation était justifiée pour les distances de l’ordre du BAO mais pas à plus petite échelle.

Je représente sur la figure 6.5,  $F_{HCD}$  en fonction de  $k_{\parallel}$  dans les cas  $1.6 \times 10^{17} < N_{HI} < 10^{19}$  (en bleu),  $10^{19} < N_{HI} < 2 \times 10^{20}$  (en rouge) et  $N_{HI} > 2 \times 10^{20}$  (en noir). On remarque que plus la densité de colonne augmente, plus les profils sont amortis. Les lignes pleines désignent les résultats de Rogers et al. (2018) du modèle complet de l’équation 6.6. Les tirets représentent l’ajustement de  $F_{HCD}$  par une fonction exponentielle :  $F_{HCD}(k_{\parallel}) \simeq \exp(-L_{HCD}k_{\parallel})$ , où les valeurs de  $L_{HCD}$  sont ajustées sur les courbes en traits pleins. Les pointillés désignent le modèle en sinus cardinal proposé par Bautista et al. (2017). On remarque que ce dernier s’éloigne du modèle complet pour les  $k_{\parallel}$  intermédiaires et quand la densité de colonne des HCD augmente. De plus, on voit apparaître des oscillations qui ne sont pas physiques et qui créent des dégénérescence avec les autres paramètres<sup>5</sup>. L’utilisation du modèle de Bautista et al. (2017) est justifiée pour  $\mu$  petits mais s’éloigne de la réalité quand le terme  $F_{HCD}$  devient dominant, c’est-à-dire quand on se rapproche de la ligne de visée.

Finalement, le champ de  $\delta$  du chapitre 4, peut s’écrire sous la forme d’une contribution des HCD et d’une contribution des autres absorptions Ly $\alpha$  :

$$\delta = \delta_{HCD} + \delta_{Ly\alpha}. \quad (6.7)$$

Ainsi, la contamination par les HCD transforme le spectre de puissance :

$$P_{Ly\alpha-Ly\alpha} \longrightarrow P_{Ly\alpha-Ly\alpha} + 2P_{Ly\alpha-HCD} + P_{HCD-HCD}. \quad (6.8)$$

<sup>5</sup>Grâce à l’analyse présentée dans la sous-section 6.5.2, j’ai pu observer que le biais du Si III(1207) était très dégénéré avec  $L_{HCD}$  dans le cas du modèle de Bautista et al. (2017).

Soit,

$$P_{Ly\alpha-Ly\alpha}(\vec{k}) = \left( b'_{Ly\alpha} (1 + \beta'_{Ly\alpha} \mu^2) \right)^2 P_m(\vec{k}), \quad (6.9)$$

avec :

$$b'_{Ly\alpha} = b_{Ly\alpha} + b_{HCD} F_{HCD}(k_{\parallel}), \quad (6.10)$$

$$\beta'_{Ly\alpha} = \frac{b_{Ly\alpha} \beta_{Ly\alpha} + b_{HCD} \beta_{HCD} F_{HCD}(k_{\parallel})}{b_{Ly\alpha} + b_{HCD} F_{HCD}(k_{\parallel})}. \quad (6.11)$$

Dans un premier temps, j'ai repris le modèle de  $F_{HCD}$  proposé par [Bautista et al. \(2017\)](#) mais je me suis aperçue que, s'il semblait approprié à la fonction Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ), il n'était pas adapté à Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) (probabilité de  $\chi^2$ , cf. chapitre 7, de quelques pourcents). C'est pourquoi, j'ai utilisé l'approximation du modèle de [Rogers et al. \(2018\)](#) donnée par :

$$F_{HCD}(k_{\parallel}) = \exp(-L_{HCD} k_{\parallel}). \quad (6.12)$$

La valeur de  $L_{HCD}$  doit être fixée à cause des dégénérescences avec d'autres paramètres du modèle (principalement  $\beta_{Ly\alpha}$  et  $b_{SiIII}$ ). On impose  $L_{HCD} = 10 h^{-1} \text{Mpc}$ , correspondant à  $N_{HI} < 10^{21}$  atomes par  $\text{cm}^2$  (conversation privée avec Keir Rogers et Andreu Font-Ribera). Puisque la valeur de  $L_{HCD}$  dépend de  $N_{HI}$ , elle dépend également de la proportion de DLA détectées dans les catalogues (sous-section 4.2.2.2). Puisque que l'on n'a pas accès à cette information, j'ai ajusté le modèle en utilisant plusieurs valeurs pour  $L_{HCD}$  comprises entre 7 et 13  $h^{-1} \text{Mpc}$  pour vérifier l'absence d'effet significatif sur les paramètres de position du pic BAO  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  (définis plus loin, dans l'équation 6.23). Je trouve les écarts maximum  $\Delta\alpha_{\parallel} = 0.001$  et  $\Delta\alpha_{\perp} = 0.003$ , largement négligeables devant les incertitudes de mesure  $\sigma_{\alpha_{\parallel}} = 0.034$  et  $\sigma_{\alpha_{\perp}} = 0.048$  (cf. équation 7.4).

### 6.1.3 Les effets non-linéaires

Dans l'équation 6.3, on a vu que la formation des structures était à l'origine de petits déplacements autour de  $r_d$  qui entraînait un amortissement des oscillations acoustiques, modélisé par la multiplication par une exponentielle décroissante de  $P_{BAO}$ . Dans cette section, on va étudier les effets non-linéaires qui s'appliquent cette fois à l'ensemble du spectre de puissance.

À cause des effets conjoints de la formation des structures, des vitesses particulières et des effets de pression dans le milieu inter-galactique, le spectre de puissance Ly $\alpha$  s'éloigne de la théorie linéaire aux petites échelles. [McDonald \(2003\)](#) introduit la fonction  $D_{NL}(\vec{k})$  pour décrire l'impact des effets non-linéaires sur le spectre de puissance Ly $\alpha$  à trois dimensions :

$$P_{Ly\alpha-Ly\alpha}(\vec{k}, z_{ref}) \longrightarrow P_{Ly\alpha-Ly\alpha}(\vec{k}, z_{ref}) D_{NL}(\vec{k}). \quad (6.13)$$

À cause de la complexité des phénomènes impliqués, une approche purement théorique est à proscrire dans cette étude. En s'inspirant des modèles proposés pour la matière noire froide, [McDonald \(2003\)](#) infère une expression de  $D_{NL}$  dans le cas des absorptions Ly $\alpha$  en supposant que ces effets doivent disparaître aux grandes échelles. Cette expression est testée et validée grâce à l'utilisation de simulations hydrodynamiques. Il propose :

$$D_{NL}^{\text{McDonald}}(k, \mu) = \exp \left[ \left( \frac{k}{k_{NL}} \right)^{\alpha_{NL}} - \left( \frac{k}{k_P} \right)^{\alpha_P} - \left( \frac{\mu k}{k_{V0}(1 + k/k_{V'})} \right)^{\alpha_V} \right]. \quad (6.14)$$

Les différents termes de cette équation font apparaître les trois principaux effets non-linéaires :

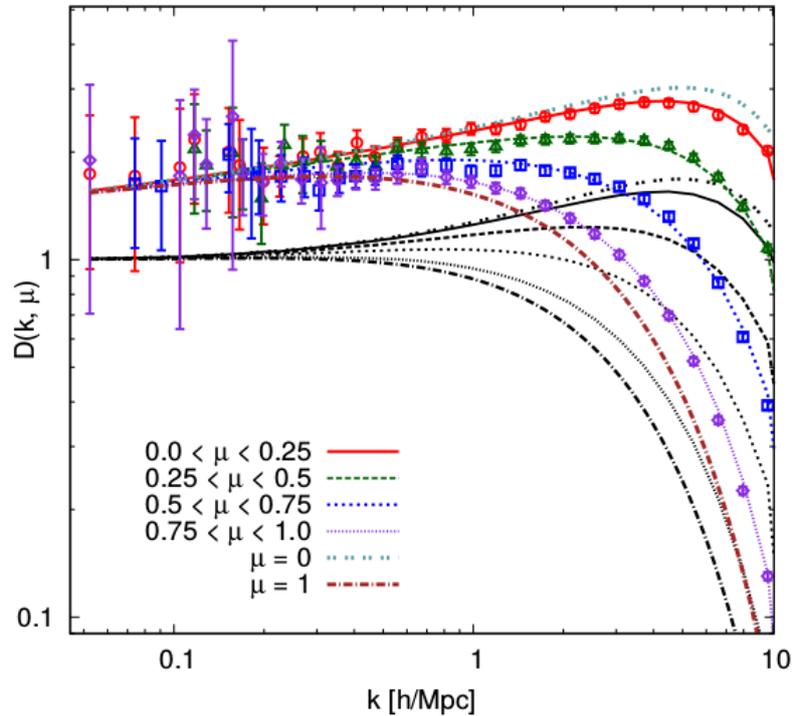


FIGURE 6.6: Les simulations hydrodynamiques réalisées par [Arinyo-i-Prats et al. \(2015\)](#) permettent d'étudier le comportement non-linéaire du spectre de puissance  $\text{Ly}\alpha$  aux petits  $k$ . Les valeurs de  $D_{NL}$  obtenues grâce aux simulations sont donnés par les symboles. Les courbes en couleurs représentent le modèle proposé par [Arinyo et al. \(2015\)](#), équation 6.16, et celles en noires, le modèle de [McDonald \(2003\)](#), équation 6.14. Les résultats sont donnés dans plusieurs gammes de  $\mu = k_{\parallel}/k$ . On remarque que dans tous les cas, le modèle d'[Arinyo-i-Prats et al. \(2015\)](#) est plus proche des simulations que [McDonald \(2003\)](#). Source : [Arinyo-i-Prats et al. \(2015\)](#)

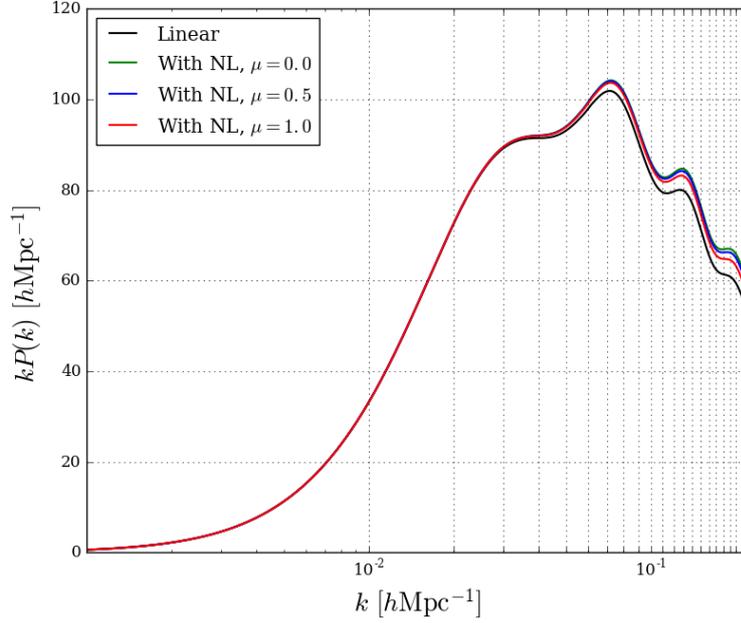


FIGURE 6.7: On représente en noir le spectre de puissance linéaire (calculé avec CAMB) puis en couleurs les spectres de puissances incluant des non-linéarités. L'effet est visible à grand  $k$  et maximal quand  $\mu = 0$ .

- La croissance des structures amplifie le spectre de puissance  $\text{Ly}\alpha$  (premier terme).
- Les effets de pression dans le milieu inter-galactique atténuent le spectre de puissance (deuxième terme).
- Les vitesses particulières atténuent également le spectre mais cet effet n'est pas isotrope et une dépendance en  $\mu$  apparaît (troisième terme).

Les huit paramètres  $k_{NL}$ ,  $\alpha_{NL}$ ,  $k_P$ ,  $\alpha_P$ ,  $k_{V0}$ ,  $k_{V'}$ ,  $\alpha_{V'}$  et  $\alpha_V$  sont ajustés grâce aux simulations hydrodynamiques. Dans [Bautista et al. \(2017\)](#), la déviation au modèle linéaire est estimée par la formule 6.14 et les valeurs des paramètres sont données par la table 1 de [McDonald \(2003\)](#).

[Arinyo-i-Prats et al. \(2015\)](#) continuent le travail entrepris par [McDonald \(2003\)](#) en améliorant les simulations hydrodynamiques dont la taille passe de  $40 h^{-1}\text{Mpc}$  à  $60 h^{-1}\text{Mpc}$ , et le nombre de particules, de  $256^3$  à  $512^3$ . Ils proposent également une nouvelle expression de  $D_{NL}$  qui part du postulat de départ que les corrections non-linéaires doivent obéir à  $D_{NL}(\vec{k}) - 1 \propto k^3 P_{lin}(\vec{k})$  quand  $k$  est faible. En effet, la théorie des perturbations prédit que le premier ordre du spectre perturbé doit dépendre de l'intégrale des produits de 4 (une pour chaque dimension) perturbations linéaires. En posant :

$$\Delta^2(\vec{k}) = \frac{1}{2\pi^2} k^3 P_{lin}(\vec{k}), \quad (6.15)$$

la déviation du spectre de puissance par rapport au régime linéaire est donnée par :

$$D_{NL}^{\text{Arinyo}}(k, \mu) = \exp \left[ q_1 \Delta^2(k) \left( 1 - \left( \frac{k}{k_v} \right)^{a_v} \mu^{b_v} \right) - \left( \frac{k}{k_p} \right)^2 \right]. \quad (6.16)$$

De même que précédemment, on peut identifier trois termes qui correspondent aux trois principaux effets non-linéaires :

- La croissance des structures est modélisée par le terme  $q_1 \Delta^2(k)$ , qui n'inclut pas l'amortissement des oscillations décrites dans 6.3 et assure qu'il n'y a pas de double comptage de certains effets non-linéaires.
- Les vitesses particulières sont prises en compte à travers  $\left( \frac{k}{k_v} \right)^{a_v} \mu^{b_v}$ , d'où la dépendance en  $\mu$ . Le spectre de puissance est lissé le long de la ligne de visée.
- $\left( \frac{k}{k_p} \right)^2$  modélise l'effet de la pression du gaz au delà de la longueur de Jeans<sup>6</sup>.

Arinyo-i-Prats et al. (2015) comparent les deux modèles aux simulations hydrodynamiques, voir figure 6.6. Pour différentes gammes de  $\mu$ , on représente les valeurs de  $D_{NL}$  des simulations (symboles), du modèle d'Arinyo-i-Prats et al. (2015) (lignes colorées) et du modèle de McDonald (2003) (lignes noires). On remarque que les nouvelles simulations semblent privilégier une correction linéaire non unitaire pour les  $k$  faibles. Le modèle d'Arinyo-i-Prats et al. (2015) paraît le plus adapté, c'est pourquoi je l'ai utilisé dans l'analyse présentée dans cette thèse. Les valeurs des paramètres sont données par la table 7 d'Arinyo-i-Prats et al. (2015). Pour vérifier que le modèle choisi pour  $D_{NL}$  n'avait pas d'influence sur la position du pic BAO, j'ai réalisé l'ajustement avec les deux modèles. Les écarts sur les paramètres de position du pic (définis plus loin, dans l'équation 6.23) sont  $\Delta\alpha_{\parallel} = 0.0001$  et  $\Delta\alpha_{\perp} = 0.0012$ , qui sont largement négligeables devant les incertitudes de mesure  $\sigma_{\alpha_{\parallel}} = 0.034$  et  $\sigma_{\alpha_{\perp}} = 0.048$  (cf. équation 7.4).

On représente figure 6.7 les spectres de puissance avec et sans effets non-linéaires. L'effet est faible et visible aux petites échelles.

#### 6.1.4 Modéliser l'échantillonnage

On a vu à la sous-section 5.1.4, que la fonction de corrélation était échantillonnée sur des *bins* carrés de taille  $R_{\parallel} \times R_{\perp}$  (avec  $R_{\parallel} = R_{\perp}$ ), c'est-à-dire que l'on mesure la fonction de corrélation pour des valeurs discrètes de  $(r_{\parallel}, r_{\perp})$ .

Soit  $A$  un *bin* de la grille en distance de séparation (figure 5.4). La fonction de corrélation échantillonnée au *bin*  $A$ ,  $\xi_A$ , peut s'écrire comme le produit de convolution de la fonction de

---

<sup>6</sup>Quand le rayon d'un système est inférieur à la longueur de Jeans, il est gravitationnellement stable. Au delà, sa pression interne ne compense pas les forces gravitationnelles, il s'effondre. Source : [www.simondriver.org/Teaching/AS2001/ce06.pdf](http://www.simondriver.org/Teaching/AS2001/ce06.pdf)

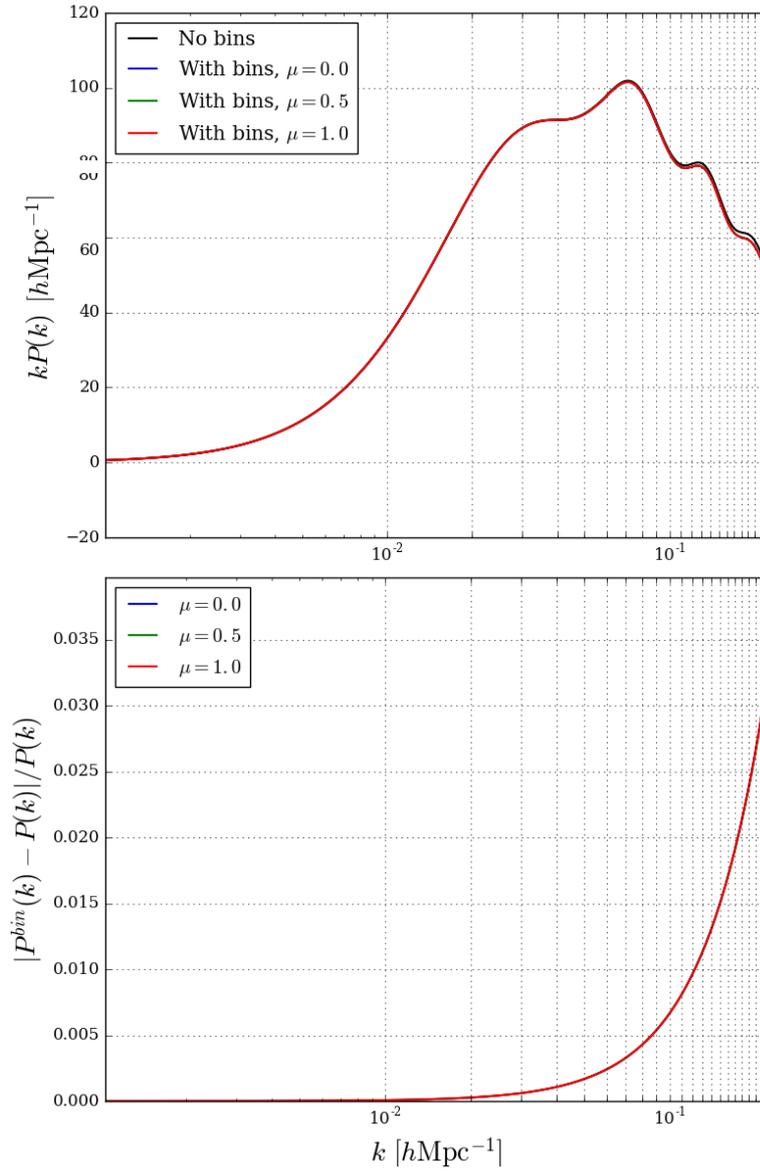


FIGURE 6.8: Comparaison des spectres de puissance avec (en couleur) et sans (en noir) échantillonnage de la fonction de corrélation pour différentes valeurs de  $\mu$ . Le graphique du bas montre que les écarts relatifs sont surtout visibles à grand  $k$  mais restent néanmoins inférieurs à 3.5%.

corrélation et des fonctions de répartition de  $r_{\parallel}$  et  $r_{\perp}$  sur les *bins*,  $f_{\parallel}$  et  $f_{\perp}$  :

$$\xi^{bin}(r_{\parallel}^A, r_{\perp}^A) = \int_{-\infty}^{\infty} dr_{\parallel} dr_{\perp} \xi(r_{\parallel}, r_{\perp}) f_{\parallel}(r_{\parallel}^A - r_{\parallel}) f_{\perp}(r_{\perp}^A - r_{\perp}). \quad (6.17)$$

Puisque la transformée de Fourier d'un produit de convolution est égale au produit des transformées de Fourier, le spectre de puissance incluant l'échantillonnage peut s'écrire comme le produit du spectre de puissance non-échantillonné et de la fonction  $G(\vec{k})$  définie par :

$$G(\vec{k}) = \tilde{f}_{\parallel}(k_{\parallel}) \tilde{f}_{\perp}(k_{\perp}), \quad (6.18)$$

avec  $\tilde{f}_{\parallel}$  et  $\tilde{f}_{\perp}$  les transformées de Fourier de  $f_{\parallel}$  et  $f_{\perp}$  respectivement.

Bautista et al. (2017) remarquent que  $f_{\parallel}$  est quasiment uniforme et donc modélisable par la fonction rectangle :

$$\text{rect}(r_{\parallel}) = \begin{cases} 1 & \text{si } r_{\parallel} \in \left[-\frac{R_{\parallel}}{2}, \frac{R_{\parallel}}{2}\right] \\ 0 & \text{sinon} \end{cases}$$

D'où,  $\tilde{f}_{\parallel} = \text{sinc}\left(\frac{k_{\parallel} R_{\parallel}}{2}\right)$ . En revanche, ils observent que  $f_{\perp}$  est proportionnelle à  $r_{\perp}$ . Ceci implique de recalculer, à chaque itération de l'ajustement, la transformée de Fourier de l'équation 6.17. Ainsi, la prise en compte de la non-uniformité de  $f_{\perp}$  demande un temps de calcul prohibitif. Je suppose donc que  $f_{\perp}$  est uniforme. En utilisant 100 simulations de l'ensemble des spectres BOSS utilisés pour l'analyse Ly $\alpha$ , Bautista et al. (2017) montrent que cette simplification donne un modèle suffisamment précis compte tenu des barres d'erreurs de la fonction de corrélation mesurée.

Finalement, je modélise le spectre de puissance échantillonné en multipliant le spectre de puissance non-échantillonné par :

$$G(\vec{k}) = \text{sinc}\left(\frac{k_{\parallel} R_{\parallel}}{2}\right) \text{sinc}\left(\frac{k_{\perp} R_{\perp}}{2}\right). \quad (6.19)$$

Je représente figure 6.8 les écarts entre les spectres avec et sans échantillonnage. On voit que cette correction est faible (inférieure à 3.5%) et visible à grand  $k$ .

### 6.1.5 Du spectre de puissance à la fonction de corrélation

Le spectre de puissance Ly $\alpha$ -Ly $\alpha$  est donné par le spectre de puissance quasi-linéaire (sous-section 6.1.1) incluant les effets liés à la nature du traceur et aux RSD (sous-section 6.1.2), des effets non-linéaires (sous-section 6.1.3) et de l'échantillonnage (sous-section 6.1.4). Ainsi, le spectre de puissance Ly $\alpha$ -Ly $\alpha$  au *redshift* de référence  $z_{ref} = 2.3$  s'écrit :

$$P_{Ly\alpha-Ly\alpha}(\vec{k}, z_{ref}) = \left(b'_{Ly\alpha}(z_{ref})(1 + \beta'_{Ly\alpha}\mu^2)\right)^2 D_{NL}(\vec{k}) G(\vec{k}) \times \left(P^{lisse}(\vec{k}, z_{ref}) + \exp\left(-\frac{\vec{k}\vec{\Sigma}}{2}\right) P^{BAO}(\vec{k}, z_{ref})\right). \quad (6.20)$$

Pourvu que l'on reste au voisinage de  $z_{ref}$ , on peut supposer que la dépendance en *redshift* du spectre de puissance Ly $\alpha$ -Ly $\alpha$  est contenue dans le biais Ly $\alpha$  :

$$P_{Ly\alpha-Ly\alpha}(\vec{k}, \bar{z}) = \left( \frac{b'_{Ly\alpha}(\bar{z})}{b'_{Ly\alpha}(z_{ref})} \right)^2 P_{Ly\alpha-Ly\alpha}(\vec{k}, z_{ref}) \quad (6.21)$$

Grâce à l'étude du spectre de puissance Ly $\alpha$  à une dimension (McDonald et al. 2006, Chabanier et al. 2018), on a vu que la dépendance en *redshift* du biais peut s'écrire comme une loi de puissance :

$$b'_{Ly\alpha}(\bar{z}) = \left( \frac{1 + \bar{z}}{1 + z_{ref}} \right)^{\gamma_{Ly\alpha} - 1} b'_{Ly\alpha}(z_{ref}) \quad \text{avec} \quad \gamma_{Ly\alpha} = 2.9 \quad (6.22)$$

Finalement, le modèle de la fonction de corrélation Ly $\alpha$   $\times$  Ly $\alpha$  au *bin* A (figure 5.4) est donné par :

$$\xi_A^{Ly\alpha-Ly\alpha} = \xi^{Ly\alpha-Ly\alpha}(r_{\parallel}^A, r_{\perp}^A) = \xi^{lisse}(r_{\parallel}^A, r_{\perp}^A) + A_{BAO} \xi^{BAO}(\alpha_{\parallel} r_{\parallel}^A, \alpha_{\perp} r_{\perp}^A), \quad (6.23)$$

avec  $\xi^{lisse}$  et  $\xi^{BAO}$  les transformées de Fourier respectives des parties lisse et oscillante de  $P_{Ly\alpha-Ly\alpha}$ ,  $A_{BAO}$  l'amplitude du pic BAO et  $(\alpha_{\parallel}, \alpha_{\perp})$  les paramètres de position du pic BAO.

D'après la sous-section 1.4.2, les quantités  $\Delta z$  et  $\Delta\theta$  (figure 1.10), mesurées au *redshift* moyen  $\bar{z}$ , sont reliées à la position du pic BAO  $(r_{\parallel}^{BAO}, r_{\perp}^{BAO})$  via :

$$\begin{cases} \Delta z = r_{\parallel}^{BAO} / D_H(\bar{z}) \\ \Delta\theta = r_{\perp}^{BAO} / D_M(\bar{z}) \end{cases} \quad (6.24)$$

Dans l'équation 6.23, on a introduit les  $(\alpha_{\parallel}, \alpha_{\perp})$  pour mesurer l'écart entre la cosmologie fiducielle (donnée en annexe A) et la mesure. Ainsi, si le pic BAO est positionné à  $(r_{\parallel}^{BAO}, r_{\perp}^{BAO})$  pour la cosmologie fiducielle, il est mesuré à  $(\alpha_{\parallel} r_{\parallel}^{BAO}, \alpha_{\perp} r_{\perp}^{BAO})$  par cette étude. On a vu que la position du pic BAO était donnée par l'échelle acoustique :  $r_{\parallel}^{BAO} = r_{\perp}^{BAO} = r_d$ . Finalement, en utilisant 6.24, les paramètres de position du pic BAO sont donnés par :

$$\begin{cases} \alpha_{\parallel} = \frac{[D_H(\bar{z})/r_d]_m}{[D_H(\bar{z})/r_d]_{fid}} \\ \alpha_{\perp} = \frac{[D_M(\bar{z})/r_d]_m}{[D_M(\bar{z})/r_d]_{fid}} \end{cases}, \quad (6.25)$$

avec  $m$  qui désigne la mesure et *fid* la cosmologie fiducielle.

L'objectif de cette étude est la mesure des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ , d'où l'on déduira de nouvelles contraintes cosmologiques au chapitre 7. Tous les autres paramètres sont introduits pour vérifier que la position du pic est indépendante du modèle.

## 6.2 Les métaux

On a vu dans la section 5.2, que le milieu intergalactique n'était pas seulement composé d'hydrogène neutre absorbant la transition Ly $\alpha$  mais de nombreux autres éléments. Grâce à la fonction de corrélation Ly $\alpha$  à une dimension, nous avons pu identifier les sources de contami-

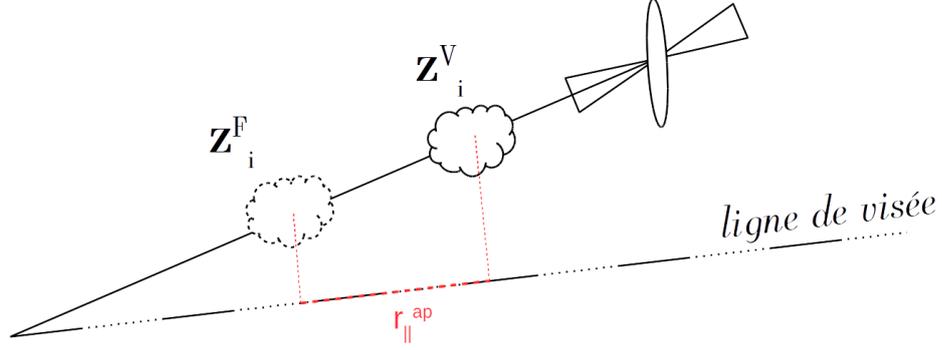


FIGURE 6.9: Supposer incorrectement qu’une absorption de type Si II au *redshift*  $z_i^V = \lambda_i/\lambda_{\text{SiII}} - 1$  est de type Ly $\alpha$ , revient à “créer” un absorbant Ly $\alpha$  à  $z_i^F = \lambda_i/\lambda_{\text{Ly}\alpha} - 1$ . La moyenne de la distance de séparation parallèle entre ces deux absorbants est donnée par  $r_{\parallel} = (1 + \bar{z})D_H(\bar{z})(\lambda_{\text{Ly}\alpha} - \lambda_{\text{SiII}})/\lambda_{\text{Ly}\alpha}$  où  $\bar{z}$  est le *redshift* moyen de la mesure.

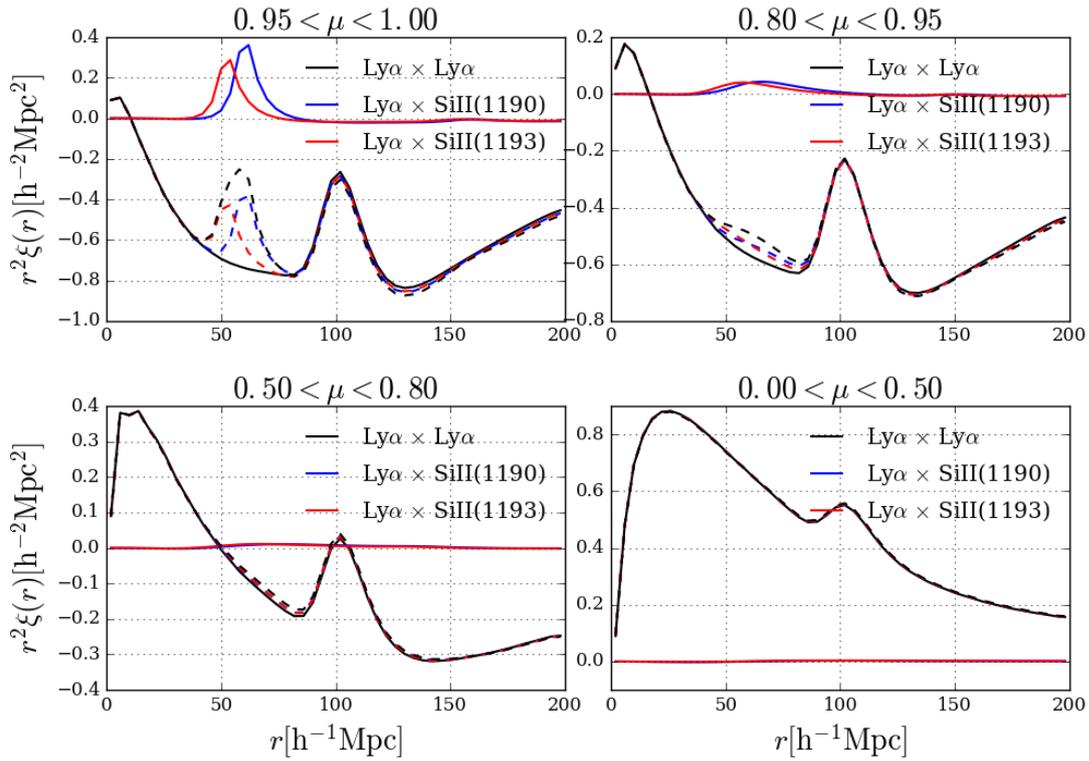


FIGURE 6.10: La mauvaise estimation du type d’absorption crée un décalage des fonctions de corrélations des métaux (en bleu : Ly $\alpha$   $\times$  SiII(1190), en rouge : Ly $\alpha$   $\times$  SiII(1193)). Plus l’écart spectral du métal avec Ly $\alpha$  est grand, plus le déplacement est important. La fonction de corrélation Ly $\alpha$   $\times$  Ly $\alpha$  (en noir) est contaminée par les corrélations métalliques et on voit apparaître à grand  $\mu$ , des pics fictifs dans la fonction de corrélation totale (tirets). Les tirets bleus désignent,  $\xi^{\text{Ly}\alpha-\text{Ly}\alpha} + \xi^{\text{Ly}\alpha-\text{SiII}(1190)}$ , la somme de la fonction de corrélation Ly $\alpha$   $\times$  Ly $\alpha$  et de la contribution Ly $\alpha$   $\times$  SiII(1190). Les tirets rouges désignent  $\xi^{\text{Ly}\alpha-\text{Ly}\alpha} + \xi^{\text{Ly}\alpha-\text{SiII}(1193)}$  et les noirs,  $\xi^{\text{Ly}\alpha-\text{Ly}\alpha} + \xi^{\text{Ly}\alpha-\text{SiII}(1190)} + \xi^{\text{Ly}\alpha-\text{SiII}(1193)}$ .

TABLE 6.1: Paires d’absorbants contaminant la fonction de corrélation  $\text{Ly}\alpha$ - $\text{Ly}\alpha$ . Le couple  $\text{CIV}(\text{eff})/\text{CIV}(\text{eff})$  n’est visible que dans les simulations alors que les autres sont détectés dans la fonction de corrélation à une dimension (figure 5.12). On définit la séparation parallèle apparente,  $r_{\parallel}^{\text{ap m-n}} = (1 + \bar{z})D_H(\bar{z})(\lambda_m - \lambda_n)/\lambda_{\text{Ly}\alpha}$  où  $\bar{z}$  est le *redshift* moyen de la mesure. Les valeurs des transitions atomiques sont données en annexe B.

Transitions	$r_{\parallel}^{\text{ap}} [h^{-1}\text{Mpc}]$
CIV(eff)/CIV(eff)	0
SiII(1193)/SiII(1190)	7
Ly $\alpha$ (1216)/SiIII(1207)	21
SiIII(1207)/SiII(1193)	31
SiIII(1207)/SiII(1190)	38
Ly $\alpha$ (1216)/SiII(1193)	52
Ly $\alpha$ (1216)/SiII(1190)	59
SiII(1260)/Ly $\alpha$ (1216)	105
SiII(1260)/SiIII(1207)	126
SiII(1260)/SiII(1193)	157
SiII(1260)/SiII(1190)	164

nation. Dans cette section, nous étudierons les effets des métaux sur la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  et leur prise en compte dans le modèle.

### 6.2.1 Le champ d’absorption des métaux

Pour simplifier, je considère dans un premier temps que le milieu intergalactique absorbe uniquement les transitions  $\lambda_{\text{Ly}\alpha}$  et  $\lambda_{\text{SiII}}$ . C’est-à-dire que le champ de  $\delta$  se compose d’une contribution des absorptions  $\text{Ly}\alpha$  et d’une contribution des absorptions  $\text{SiII}$  :

$$\delta = \delta^{\text{Ly}\alpha} + \delta^{\text{SiII}}. \quad (6.26)$$

La fonction de corrélation totale des  $\delta$  s’écrit somme la somme des contributions  $\text{Ly}\alpha \times \text{Ly}\alpha$ ,  $\text{Ly}\alpha \times \text{SiII}$  et  $\text{SiII} \times \text{SiII}$  :

$$\xi(r_{\parallel}, r_{\perp}) = \xi^{\text{Ly}\alpha - \text{Ly}\alpha}(r_{\parallel}, r_{\perp}) + 2\xi^{\text{SiII} - \text{Ly}\alpha}(r_{\parallel}, r_{\perp}) + \xi^{\text{SiII} - \text{SiII}}(r_{\parallel}, r_{\perp}). \quad (6.27)$$

L’importance des contributions dépend de la valeur des biais et plus exactement de la concentration atomique dans le milieu inter-galactique (section 5.2). Ainsi, du fait de la très grande abondance de l’hydrogène neutre dans l’Univers, la contribution principale est  $\xi^{\text{Ly}\alpha - \text{Ly}\alpha}$ , suivie de  $\xi^{\text{SiII} - \text{Ly}\alpha}$  puis  $\xi^{\text{SiII} - \text{SiII}}$ . Quand la concentration d’un élément métallique devient importante, sa corrélation croisée avec  $\text{Ly}\alpha$  ou un autre métal et son auto-corrélation perturbent le signal de l’auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ .

Puisque la détermination des distances de séparation  $(r_{\parallel}, r_{\perp})$  (équation 5.6) dépend du type d’absorption, le calcul de 6.27 nécessite de connaître *a priori* la décomposition du champ de  $\delta$  de l’équation 6.26. Dans la pratique ce n’est pas possible, c’est pourquoi on fait l’hypothèse (incorrecte) que toutes les absorptions sont de type  $\text{Ly}\alpha$ .

Considérons une absorption au *pixel*  $i$  de type  $\text{SiII}$  et faussement attribuée à  $\text{Ly}\alpha$  :

- Le *redshift* vrai de l'absorbant est donné par  $z_i^V = \lambda_i/\lambda_{\text{SiII}} - 1$ .
- Le *redshift* incorrectement estimé est  $z_i^F = \lambda_i/\lambda_{\text{Ly}\alpha} - 1$ .

La mauvaise estimation du type d'absorbant "créé" un faux nuage au *redshift*  $z_i^F$ , voir figure 6.9. Cela revient à déplacer le champ d'absorption SiII :

$$\delta_i^{\text{SiII}} \longrightarrow \delta_k^{\text{SiII}} \quad \text{avec} \quad \frac{\lambda_i}{\lambda_{\text{Ly}\alpha}} = \frac{\lambda_k}{\lambda_{\text{SiII}}}. \quad (6.28)$$

En introduisant  $\eta_{ik}^{\text{Ly}\alpha/\text{SiII}}$ , la matrice qui transforme  $i$  en  $k$ , le champ d'absorption SiII déplacé  $\tilde{\delta}^{\text{SiII}}$  au *pixel*  $i$  peut s'écrire :

$$\tilde{\delta}_i^{\text{SiII}} = \sum_k \eta_{ik}^{\text{Ly}\alpha/\text{SiII}} \delta_k^{\text{SiII}}, \quad (6.29)$$

avec

$$\eta_{ik}^{\text{Ly}\alpha/\text{SiII}} = \begin{cases} 1 & \text{si } \frac{\lambda_k}{\lambda_{\text{SiII}}} = \frac{\lambda_i}{\lambda_{\text{Ly}\alpha}} \\ 0 & \text{sinon} \end{cases}$$

L'équation 6.26 devient alors :

$$\delta \longrightarrow \delta^{\text{Ly}\alpha} + \tilde{\delta}^{\text{SiII}}. \quad (6.30)$$

La fonction de corrélation totale 6.27 est également impactée à travers le déplacement de  $\xi^{\text{SiII-Ly}\alpha}$  et  $\xi^{\text{SiII-SiII}}$  :

$$\xi(r_{\parallel}, r_{\perp}) \longrightarrow \xi^{\text{Ly}\alpha-\text{Ly}\alpha}(r_{\parallel}, r_{\perp}) + 2\tilde{\xi}^{\text{SiII-Ly}\alpha}(r_{\parallel}, r_{\perp}) + \tilde{\xi}^{\text{SiII-SiII}}(r_{\parallel}, r_{\perp}). \quad (6.31)$$

Je représente sur la figure 6.10 en tirets noirs, la fonction de corrélation totale 6.31 contaminée par SiII(1190) et SiII(1193). La contribution Ly $\alpha$   $\times$  Ly $\alpha$  est donnée par les traits pleins noirs et les contributions SiII(1190) et SiII(1193) par les traits pleins bleu et rouge (respectivement). Le premier pic de corrélation des contributions métalliques est décalé d'une distance qui dépend de la longueur d'onde. Le déplacement parallèle moyen correspond à la distance de séparation moyenne entre les nuages réel et fictif de la figure 6.9 :

$$r_{\parallel}^{\text{ap SiII-Ly}\alpha} = \int (\chi(\frac{\lambda}{\lambda_{\text{SiII}}} - 1) - \chi(\frac{\lambda}{\lambda_{\text{Ly}\alpha} - 1})) d\lambda \simeq (1 + \bar{z}) D_H(\bar{z}) \frac{\lambda_{\text{Ly}\alpha} - \lambda_{\text{SiII}}}{\lambda_{\text{Ly}\alpha}}. \quad (6.32)$$

L'approximation est valide dans le cas où  $\lambda_{\text{Ly}\alpha}$  et  $\lambda_{\text{SiII}}$  sont proches l'un de l'autre.  $\bar{z}$  désigne le *redshift* moyen des absorbants de la mesure.

À cause de la petitesse des angles de séparation entre les quasars (sous-section 5.1.5) et de la présence du sinus dans l'expression de  $r_{\perp}$  (équation 5.6), le déplacement perpendiculaire est quasi nul. Ceci explique que la contamination métallique ne soit visible sur la figure 6.10 que pour  $\mu > 0.95$ .

Dans la sous-section suivante, je vais introduire la *matrice des métaux* qui permet le passage de la fonction de corrélation vraie à la fonction de corrélation déplacée et permet d'écrire l'équation 6.30 sous forme matricielle.

### 6.2.2 La matrice des métaux et la fonction de corrélation métallique

En reprenant l'équation 5.16, l'estimateur de la corrélation croisée  $\text{Ly}\alpha \times \text{Si II}$  au *bin*  $A$  est donné par :

$$\langle \xi_A^{Ly\alpha-SiII} \rangle = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \delta_i^{SiII} \delta_j^{Ly\alpha}. \quad (6.33)$$

On en déduit l'estimateur de la corrélation  $\text{Ly}\alpha \times \text{Si II}$  déplacée en remplaçant  $\delta^{SiII}$  par l'expression de  $\tilde{\delta}^{SiII}$  (voir équation 6.29) :

$$\langle \xi_A^{Ly\alpha-SiII} \rangle = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \sum_k \eta_{ik}^{SiII/Ly\alpha} \delta_k^{SiII} \delta_j^{Ly\alpha}. \quad (6.34)$$

Afin d'augmenter le signal sur bruit de la mesure, les poids ont été ré-étalonnés à la sous-section 4.5.2 :

$$w_i \propto \left( \frac{\lambda_i}{\lambda_{Ly\alpha}} \right)^{\gamma_{Ly\alpha}-1}. \quad (6.35)$$

Dans l'équation 6.34, je supprime cette dépendance pour les poids des  $\delta_k^{SiII}$  et j'ajoute une nouvelle dépendance analogue pour les absorptions de type  $\text{Si II}$  :

$$\langle \tilde{\xi}_A^{Ly\alpha-SiII} \rangle = W_A^{-1} \sum_{(i,j) \in A} w_i w_j \left( \frac{\lambda_i}{\lambda_{Ly\alpha}} \right)^{1-\gamma_{Ly\alpha}} \sum_k \left( \frac{\lambda_k}{\lambda_{SiII}} \right)^{\gamma_{SiII}-1} \eta_{ik}^{SiII/Ly\alpha} \delta_k^{SiII} \delta_j^{Ly\alpha}. \quad (6.36)$$

En développant les sommes dans l'équation 6.36, on peut montrer que l'équation 6.36 peut s'écrire sous la forme :

$$\langle \tilde{\xi}_A^{Ly\alpha-SiII} \rangle = \sum_B M_{AB}^{Ly\alpha-SiII} \xi_B^{Ly\alpha-SiII}, \quad (6.37)$$

avec  $M_{AB}^{Ly\alpha-SiII}$  la *matrice des métaux* dans le cas de la corrélation croisée  $\text{Ly}\alpha \times \text{Si II}$ . Elle est définie par :

$$M_{AB}^{Ly\alpha-SiII} \equiv W_A^{-1} \sum_{(i,j) \in A} w_i w_j \left( \frac{\lambda_i}{\lambda_{Ly\alpha}} \right)^{1-\gamma_{Ly\alpha}} \left( \frac{\lambda_j}{\lambda_{Ly\alpha}} \right)^{1-\gamma_{Ly\alpha}} \sum_{(k,l) \in B} \left( \frac{\lambda_k}{\lambda_{SiII}} \right)^{\gamma_{SiII}-1} \left( \frac{\lambda_l}{\lambda_{Ly\alpha}} \right)^{\gamma_{Ly\alpha}-1} \eta_{ik}^{Ly\alpha/SiII} \eta_{jl}^{Ly\alpha/Ly\alpha}. \quad (6.38)$$

La formule 6.37 peut être généralisée au cas où deux types d'absorptions  $m$  et  $n$  sont incorrectement assimilées à  $\text{Ly}\alpha$  :

$$\langle \tilde{\xi}_A^{m-n} \rangle = \sum_B M_{AB}^{m-n} \xi_B^{m-n}. \quad (6.39)$$

Le terme  $\xi_B^{m-n}$  peut être modélisé de façon analogue à  $\text{Ly}\alpha \times \text{Ly}\alpha$  dans la sous-section 6.1.5. En revanche, étant donné que les contributions métalliques sont seulement des perturbations, il est inutile de considérer la décomposition quasi-linéaire du spectre (voir sous-section 6.1.1). De plus, la quantité de métaux du milieu intergalactique est faible comparée à l'hydrogène neutre. Ainsi, les absorptions métalliques saturées sont extrêmement rares. C'est pourquoi on ne considérera pas non plus la contamination par les HCD dans le cas des métaux. Finalement, le spectre de puissance  $m-n$  s'écrit :

$$P^{m-n}(\vec{k}, z_{ref}) = b_m(z_{ref})(1 + \beta_m \mu^2) b_n(z_{ref})(1 + \beta_n \mu^2) G(\vec{k}) P_{lin}(\vec{k}, z_{ref}). \quad (6.40)$$

Le terme  $\xi^{m-n}$  est alors obtenu par la transformée de Fourier de  $P^{m-n}$ . Les paramètres de biais  $b_m$  et  $b_n$ , qui seront ajustés sur la fonction de corrélation mesurée, permettent d'évaluer les fréquences respectives des transitions  $m$  et  $n$ . Plus elles sont fréquentes, plus la contamination de  $\xi^{mod}$  par  $\xi^{m-n}$  est grande.

Pour chaque couple  $m-n$  identifié (*cf.* table 6.1), je calcule la matrice  $M^{m-n}$  à partir des données. Pour des raisons de temps de calcul, la prise en compte de toutes les paires  $(i,j)$  et  $(i,k)$  de l'équation 6.38 est à proscrire. On a vérifié que l'utilisation de seulement 1% des paires sélectionnées aléatoirement donnait une précision suffisante.

### 6.2.3 Contribution totale des métaux

La contribution totale des métaux à la fonction de corrélation mesurée est estimée par :

$$\xi_A^{metals} = \sum_{m,n} M_{AB}^{m-n} \xi_B^{m-n}, \quad (6.41)$$

où  $(m,n)$  désignent les paires de métaux de la table 6.1. Les valeurs des transitions atomiques sont données en annexe B.

La corrélation CIV(eff)-CIV(eff)<sup>7</sup> est particulière, car n'étant pas décalée par rapport à  $Ly\alpha \times Ly\alpha$ , elle n'est pas détectée dans la fonction de corrélation à 1d (sous-section 5.2.2). En effet, la fonction de corrélation à 1d dépend de la distance apparente de séparation entre les deux types d'absorptions vues au même *redshift*. Ainsi, puisque dans les deux cas ( $Ly\alpha \times Ly\alpha$  et CIV(eff) $\times$ CIV(eff)) les deux types d'absorption sont identiques, alors la distance apparente de séparation est nulle dans les deux cas. La corrélation CIV(eff)-CIV(eff) ne crée donc pas de nouveau pic dans la fonction de corrélation mesurée, mais elle augmente son amplitude (et donc son biais) car elle vient se superposer au signal de  $\xi^{Ly\alpha-Ly\alpha}$ . Cette contribution a été détectée par [Bautista et al. \(2017\)](#) grâce aux simulations des champs d'absorption dans les régions  $Ly\alpha$ .

Hormis CIV(eff)/CIV(eff), l'ensemble des couples de la table 6.1 a été mis en évidence grâce à la fonction de corrélation à une dimension.

## 6.3 La matrice de distorsion

### 6.3.1 Effet de l'ajustement du continuum sur la fonction de corrélation

L'instrument eBOSS mesure les flux émis par des quasars lointains et transformés par le milieu inter-galactique traversé. Pour le calcul de la fonction de corrélation  $Ly\alpha$ , il est nécessaire d'extraire le champ d'absorption en éliminant au maximum la dépendance à la physique des quasars. La procédure d'extraction, décrite au chapitre 4, consiste à ajuster sur chaque spectre un continuum modélisant le flux émis par le quasar. Mais cet ajustement, réalisé sur toute une région, corréle les *pixels* voisins entre eux et distord la fonction de corrélation  $Ly\alpha \times Ly\alpha$ .

Pour étudier cet effet, [Bautista et al. \(2017\)](#) calculent les fonctions de corrélation  $Ly\alpha \times Ly\alpha$  à partir de 100 simulations du champ d'absorption dans les régions  $Ly\alpha$  (sans métaux). Le résultat est représenté par les lignes fines noires dans la figure 6.11 et la moyenne, par les points noirs.

---

<sup>7</sup>CIV(eff) désigne la raie d'absorption effective du doublet CIV(1548)-CIV(1551). Ces deux absorptions sont trop proches pour pouvoir être discriminées dans cette étude.

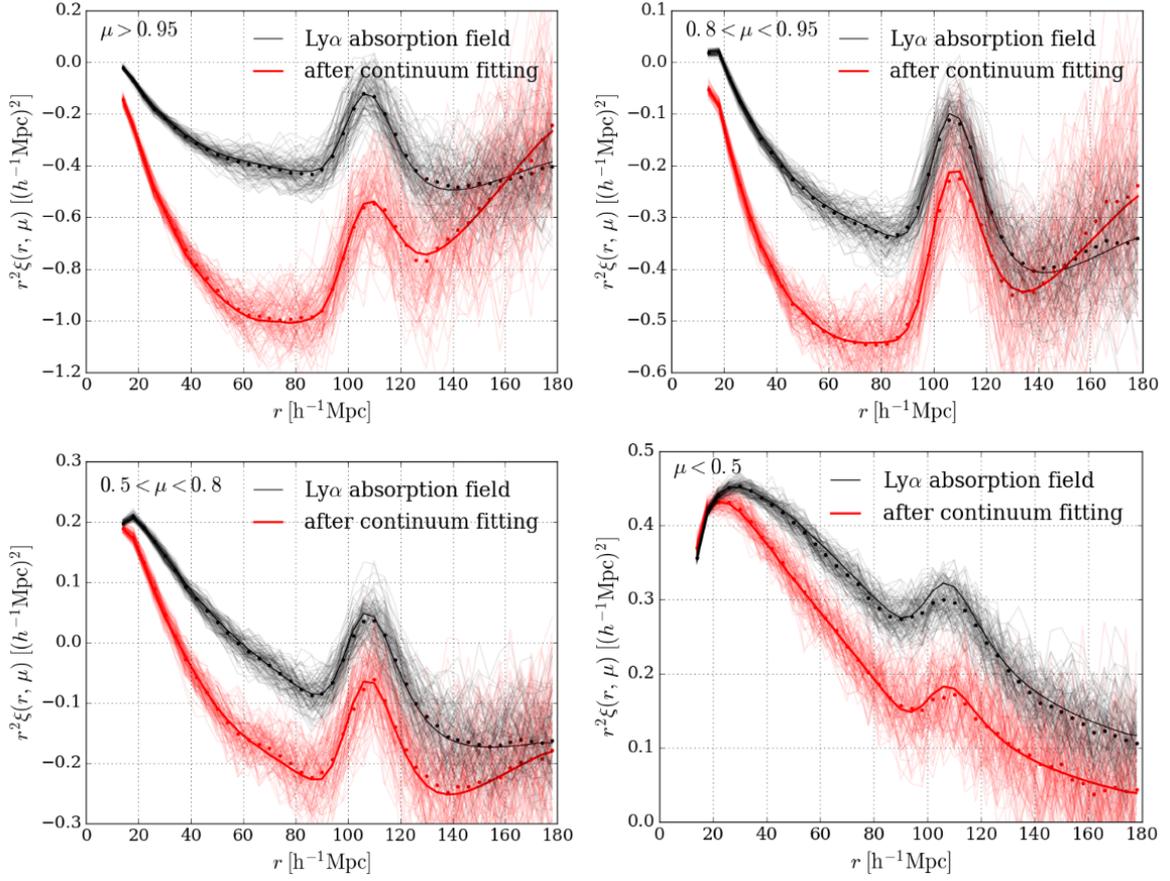


FIGURE 6.11: Fonctions de corrélations mesurées dans les 100 simulations de forêts Ly $\alpha$  (sans métaux) réalisées par [Bautista et al. \(2015\)](#) dans quatre gammes de  $\mu = r_{\parallel}/r$ . Pour chaque simulation, une ligne noire fine désigne la fonction de corrélation calculée sans l'addition du flux du quasar et une ligne rouge fine, avec addition du continuum. Les points désignent les moyennes de l'ensemble des 100 simulations. La ligne épaisse noire montre le modèle proposé par [Bautista et al. \(2017\)](#) pour la fonction de corrélation Ly $\alpha \times$  Ly $\alpha$ . En rouge, ce modèle est multiplié par la matrice de distorsion. Source : [Bautista et al. \(2017\)](#)

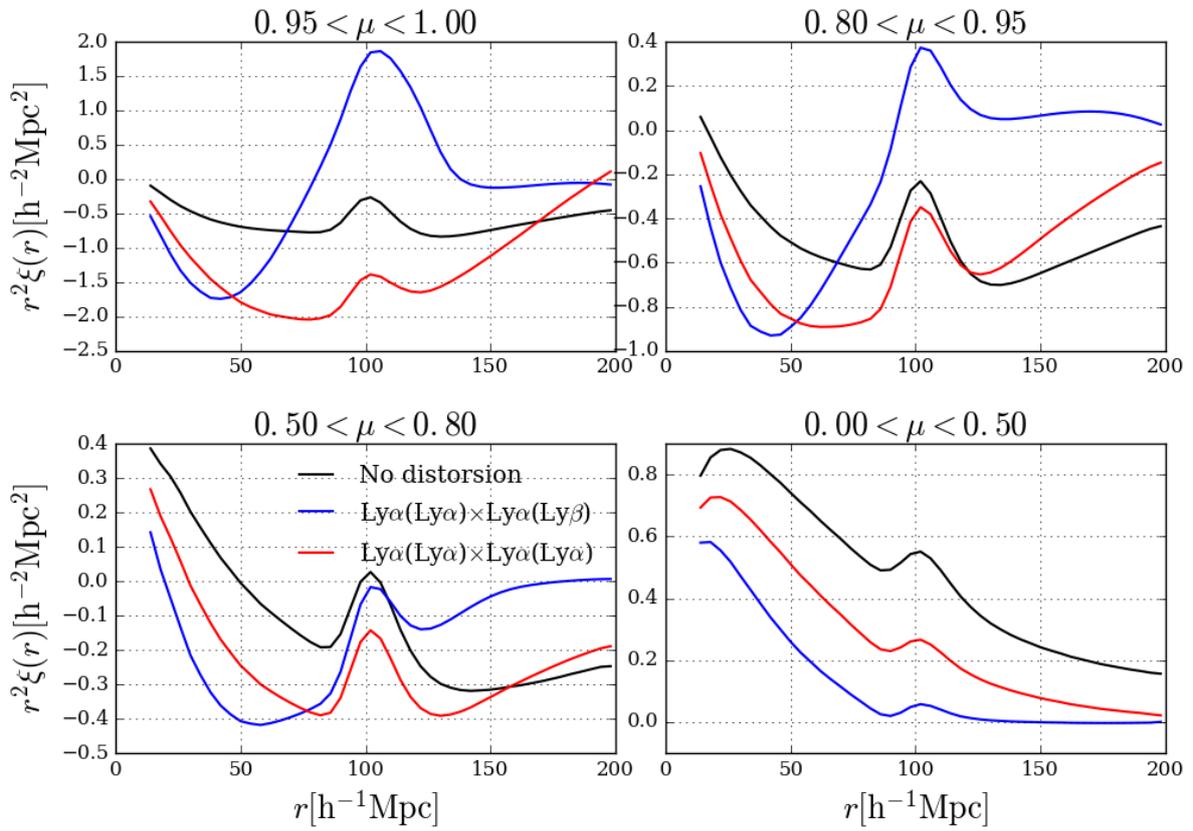


FIGURE 6.12: Effets des matrices de distorsion  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (en rouge) et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (en bleu) sur la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ .

La ligne noire épaisse désigne le modèle proposé par [Bautista et al. \(2017\)](#), qui est très proche de celui présenté à la sous-section 6.1.5.

Dans un deuxième temps, les auteurs modélisent les spectres des données BOSS en ajoutant un continuum à chaque simulation du champ d'absorption. Ces données sont traitées de la même façon que les données réelles : extraction et projection du champ d'absorption, puis calcul de la fonction de corrélation. La différence entre les courbes rouges et noires montre l'importance de l'effet de la projection des  $\delta$  sur la fonction de corrélation.

Dans les études initiales ([Busca et al. 2013](#), [Delubac et al. 2015](#)), il n'y avait pas de projection des champ d'absorption et l'effet de l'ajustement du continuum était pris en compte dans le modèle par l'ajout d'une composante purement mathématique, appelée terme de *broadband* (voir section 6.4). Plusieurs types de continuum étaient testés de façon à garantir que les résultats soient indépendants du modèle.

[Bautista et al. \(2017\)](#) introduisent une nouvelle approche qui permet de modéliser les effets de l'ajustement du continuum à travers une matrice, la *matrice de distorsion*. L'intérêt de cette méthode est qu'elle permet d'obtenir un modèle qui explique les données sans avoir à rajouter de terme non-physique (le terme de *broadband*). Dans mon étude, je reprends cette approche.

### 6.3.2 La projection des $\delta$ et ses effets sur la fonction de corrélation

Le traitement des effets de l'ajustement du continuum se fait en deux étapes :

- On applique un projecteur  $\mathbb{P}$  sur le champ de  $\delta$  de façon à ce que le champ de  $\delta$  projeté égale le champ d'absorptions vrai projeté (section 4.4).
- L'effet de la projection est modélisé par la matrice de distorsion, qui est appliquée au modèle.

En remplaçant les termes  $\hat{\delta}$  de l'estimateur 5.16 par leur expression (équation 4.39), on trouve :

$$\langle \hat{\xi}_A \rangle = \sum_B \left\{ W_A^{-1} \sum_{(i,j) \in A} w_i w_j \left( \sum_{(i',j') \in B} P_{ii'} P_{jj'} \right) \delta_{i'} \delta_{j'} \right\}. \quad (6.42)$$

Où les coordonnées  $P_{ij}$  du projecteur sont données par l'équation 4.37. De même que lors du calcul de la matrice des métaux (sous-section 6.2.2), on remplace  $\langle \delta_{i'} \delta_{j'} \rangle$  par  $\xi_B$  :

$$\langle \hat{\xi}_A \rangle = \sum_B \underbrace{W_A^{-1} \sum_{ij \in A} w_i w_j \left( \sum_{i'j' \in B} P_{ii'} P_{jj'} \right)}_{D_{AB}} \xi_B. \quad (6.43)$$

Finalement,

$$\langle \hat{\xi}_A \rangle = \sum_B D_{AB} \xi_B, \quad (6.44)$$

où l'on introduit la *matrice de distorsion*  $D$  définie par :

$$D_{AB} = W_A^{-1} \sum_{ij \in A} w_i w_j \sum_{i'j' \in B} P_{ii'} P_{jj'}. \quad (6.45)$$

Pour des raisons de temps de calcul, on ne prend en compte que 5% des paires (sélectionnées aléatoirement) de l'équation 6.45. Cette simplification a été validée par [Bautista et al. \(2017\)](#) grâce à l'utilisation des simulations.

La ligne épaisse rouge de la figure 6.11 montre le modèle transformé par la matrice de distorsion et valide l'utilisation de cette méthode pour la prise en compte des corrélations liées à l'ajustement du continuum.

Le modèle de la fonction de corrélation présenté à la sous-section 6.1.5 est valable à la fois pour  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et pour  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . En revanche, puisque la matrice de distorsion est calculée sur les données, les résultats seront différents pour les deux fonctions. On montre en noir sur la figure 6.12, la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  sans projection ni distorsion. En rouge, le modèle est transformé par la matrice de distorsion de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et en bleu, par  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . On remarque que dans les deux cas, l'effet de la projection des  $\delta$  est très important et dépend de l'angle  $\mu$ . Dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , il peut sembler inquiétant de voir que la distorsion est importante au niveau du pic BAO pour  $\mu$  dans  $[0.95, 1]$ . Nous allons voir dans la sous-section suivante que cet effet n'impacte pas la mesure de la position du pic BAO et quels sont les moyens qui peuvent être envisagés pour le supprimer.

### 6.3.3 Le cas $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$

Puisque les paramètres de positions du pic BAO  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  vont être utilisés pour mesurer les contraintes cosmologiques, il est essentiel qu'ils soient indépendants du modèle et plus particulièrement de la matrice de distorsion.

Au cours de mon étude, j'ai réalisé les deux tests suivants :

- L'ajustement de la fonction de corrélation en excluant les valeurs de  $\mu$  dans  $[0.95, 1]$  de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . On observe un décalage du pic de  $\Delta\alpha_{\parallel} = 0.009$  et  $\Delta\alpha_{\perp} = 0.004$  par rapport à l'ajustement standard, négligeable devant les incertitudes de mesure  $\sigma_{\alpha_{\parallel}} = 0.034$  et  $\sigma_{\alpha_{\perp}} = 0.048$  (*cf.* équation 7.4).
- J'ai utilisé les simulations en préparation pour l'analyse finale d'eBOSS pour vérifier que  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  étaient compatibles dans les cas avec et sans ajout du continuum.

L'observation de cet effet dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et pas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ , est dû à la taille réduite de la forêt  $\text{Ly}\beta$ . L'extension des bornes de la région  $\text{Ly}\beta$  permet de l'éliminer. Quand les régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$  sont de taille similaire, les matrices de distorsion sont très proches.

## 6.4 Modéliser les erreurs systématiques inconnues

Le modèle standard de l'ajustement est décrit par les termes  $\xi^{Ly\alpha Ly\alpha}$ ,  $\xi^{metals}$  et  $D$  de l'équation 6.1 et sera utilisé pour contraindre les paramètres cosmologiques. Avant [Bautista et al. \(2017\)](#), la contribution du terme  $\xi^{bb}$  de l'équation 6.1 jouait un rôle central puisqu'elle permettait de décrire les effets de l'ajustement du continuum sur la fonction de corrélation. Dans mon étude, le terme de *broadband*  $\xi^{bb}$  est optionnel et permet de décrire l'écart entre la fonction de corrélation mesurée et le modèle, ce qui revient à inclure les erreurs systématiques inconnues. Je vérifie ainsi que les paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont stables même si le modèle est incomplet.

On identifie deux types d'erreurs systématiques en fonction de leurs effets sur la fonction de corrélation :

- effet à petite échelle : l'effet est localisé (*e.g.*, la contamination métallique qui est principalement visible à  $\mu = 0$ ),
- effet à grande échelle : l'effet est visible sur l'ensemble de la fonction de corrélation.

Les effets à petite échelle agissent de façon ciblée sur quelques *bins* de la fonction de corrélation ainsi, les paramètres qui les décrivent sont ajustés sur des régions très spécifiques de la fonction de corrélation et sont donc faciles à contraindre (*cf.* section 6.5.2). *A contrario*, les effets à grande échelle sont plus difficiles à prendre en compte car ils agissent sensiblement de la même façon sur l'ensemble des *bins* de la fonction de corrélation.

Pour modéliser les erreurs systématiques des effets à grande échelle, le terme  $\xi^{bb}$  doit être une fonction de  $r$  et  $\mu$  lentement variable. Le domaine de définition de cette fonction est réduit aux régions où la fonction de corrélation varie faiblement. Ainsi, le modèle avec terme de *broadband* est ajusté pour  $r$  dans  $[40, 180] h^{-1}\text{Mpc}$ .

On reprend la forme proposée par [Bautista et al. \(2017\)](#) :

$$\xi_A^{bb} = \sum_{j=0}^{j_{max}} \sum_{i=i_{min}}^{i_{max}} a_{ij} \frac{L_j(\mu^A)}{(r^A)^i} \quad \text{avec } j \text{ pair.} \quad (6.46)$$

$L_j$  désigne le polynôme de Legendre d'ordre  $j$ .

La condition " $j$  pair" vient du fait que puisque la fonction de corrélation est paire, les termes impairs de sa décomposition de Legendre doivent être nuls. Pour choisir les valeurs appropriées de  $i_{min}$ ,  $i_{max}$  et  $j_{max}$ , on remarque que  $r^2\xi(r)$  varie peu pour  $r$  dans  $[40, 180] h^{-1}\text{Mpc}$  (figure 5.10). On choisit  $(i_{min}, i_{max}) = (0, 2)$  qui correspond à l'équation d'une parabole et permet de laisser libre  $r^2\xi^{lisse}(r)$  (voir équation 6.23) et de ses deux premières dérivées sur  $r$  dans  $[40, 180] h^{-1}\text{Mpc}$ . On choisit  $j_{max} = 6$  qui permet d'avoir approximativement un modèle de *broadband* dans chacune des quatre gammes de  $\mu$  de la figure 5.10.

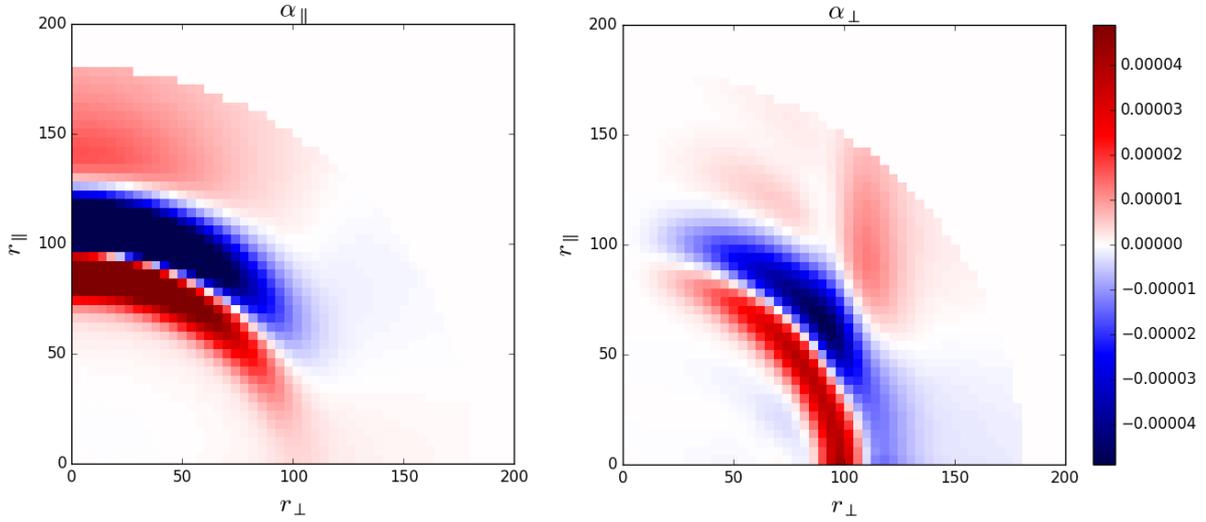


FIGURE 6.13: Contribution des *bins* à l'ajustement des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  à travers les quantités  $\frac{\partial \xi^{mod}}{\partial \alpha_{\parallel}}$  et  $\frac{\partial \xi^{mod}}{\partial \alpha_{\perp}}$  dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ .  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont contraints par les *bins* autour du pic BAO respectivement parallèlement et perpendiculairement à la ligne de visée.

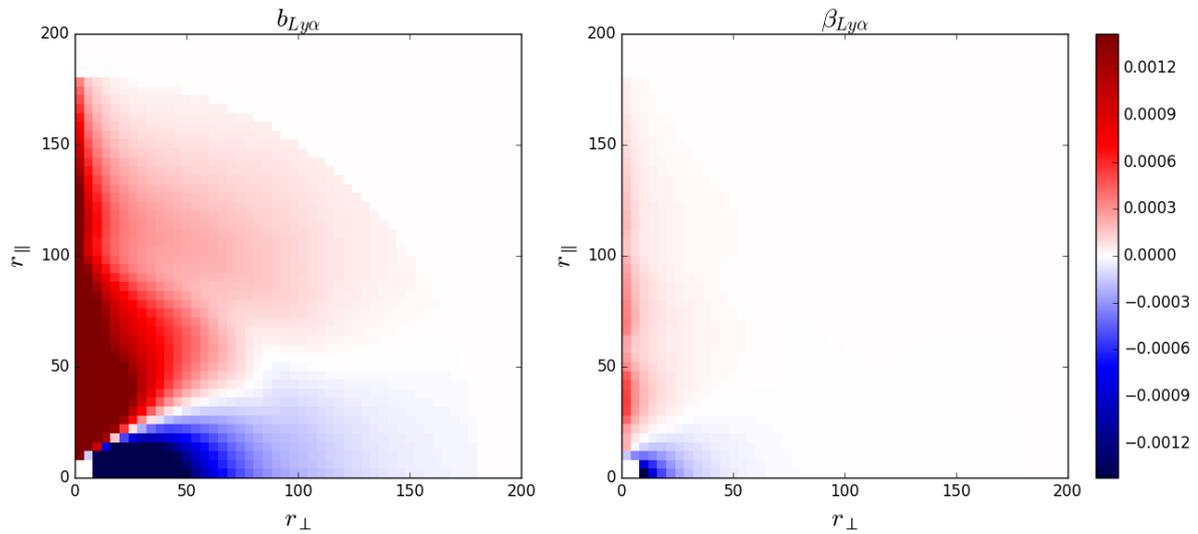


FIGURE 6.14: On représente la contribution des *bins* à l'ajustement des paramètres  $b_{\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$  à travers les quantités  $\frac{\partial \xi^{mod}}{\partial b_{\text{Ly}\alpha}}$  et  $\frac{\partial \xi^{mod}}{\partial \beta_{\text{Ly}\alpha}}$  dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . On voit que l'amplitude de  $\frac{\partial \xi^{mod}}{\partial b_{\text{Ly}\alpha}}$  est supérieure, ce qui signifie que le paramètre  $b_{\text{Ly}\alpha}$  est mieux contraint que  $\beta_{\text{Ly}\alpha}$ .

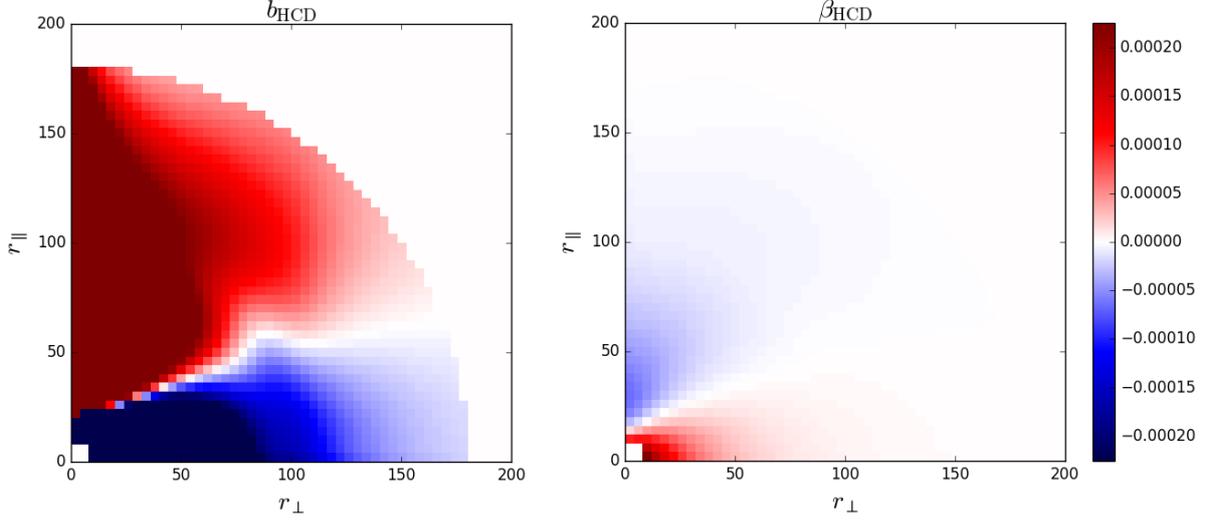


FIGURE 6.15: Contribution des *bins* à l'ajustement des paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  à travers les quantités  $\frac{\partial \xi^{\text{mod}}}{\partial b_{\text{HCD}}}$  et  $\frac{\partial \xi^{\text{mod}}}{\partial \beta_{\text{HCD}}}$  dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . On remarque que l'amplitude des variations dans le cas  $\beta_{\text{HCD}}$  est faible et répartie sur quasiment l'ensemble des *bins*, ce qui montre que ce paramètre est peu contraint.

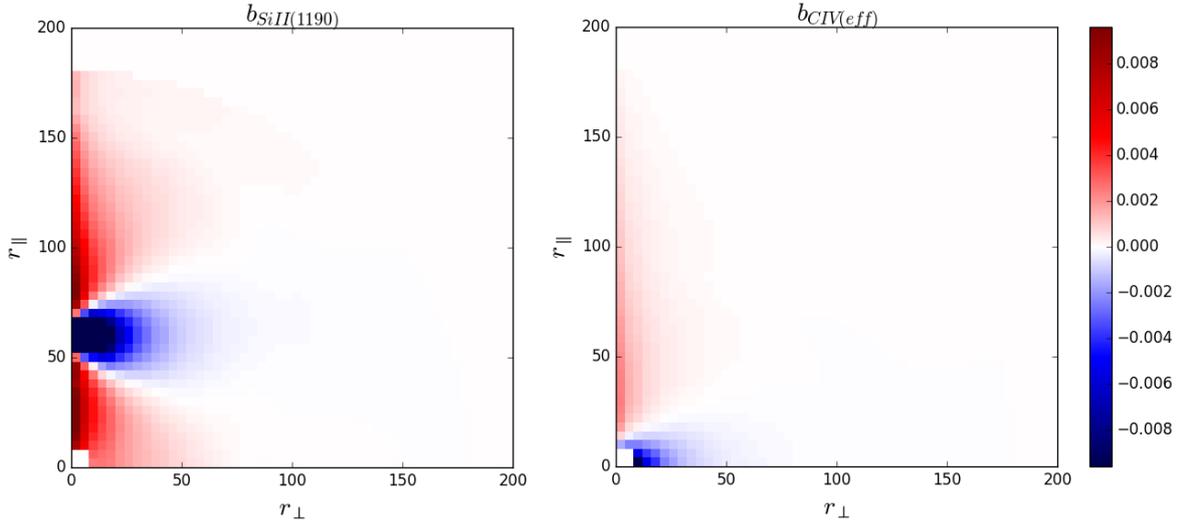


FIGURE 6.16: Contribution des *bins* à l'ajustement des paramètres  $b_{\text{SiII}(1190)}$  et  $b_{\text{CIV}(\text{eff})}$  à travers les quantités  $\frac{\partial \xi^{\text{mod}}}{\partial b_{\text{SiII}(1190)}}$  et  $\frac{\partial \xi^{\text{mod}}}{\partial b_{\text{CIV}(\text{eff})}}$  dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . Dans le cas du biais du Si II(1190), on trouve un motif similaire au biais  $\text{Ly}\alpha$  mais décalé de  $60 h^{-1}\text{Mpc}$  le long de la ligne de visée. Cette distance correspond à celle calculée dans la section 6.2.1

TABLE 6.2: Limites physiques de l’ajustement standard dans le cas modèles avec et sans terme de *broadband*.

Paramètres	Limites
$r_{\parallel}$	$[0,200] h^{-1}\text{Mpc}$
$r_{\perp}$	$[0,200] h^{-1}\text{Mpc}$
$\mu$	$[0,1]$
Modèle physique	
$r$	$[10,180] h^{-1}\text{Mpc}$
Modèle avec <i>broadband</i>	
$r$	$[40,180] h^{-1}\text{Mpc}$

 TABLE 6.3: Liste des paramètres fixés dans le modèle physique et les modèles avec terme de *broadband*.

Paramètres	Définitions	Valeurs	Références
$A_{\text{BAO}}$	Amplitude du pic BAO	1	
$\Sigma_{\perp}$	Élargissement transversal du pic BAO	$3.26 h^{-1}\text{Mpc}$	<a href="#">Kirkby et al. (2013)</a>
$\Sigma_{\parallel}$	Élargissement radial du pic BAO	$6.41 h^{-1}\text{Mpc}$	<a href="#">Kirkby et al. (2013)</a>
$R_{\parallel}, R_{\perp}$	Taille d’un <i>bin</i>	$4 h^{-1}\text{Mpc}$	
$L_{\text{HCD}}$	Taille caractéristique des HCD	$10 h^{-1}\text{Mpc}$	<a href="#">Rogers et al. (2018))</a>
$\beta_m$	Paramètres RSD pour Si II(1190), (1193), (1260) et Si III(1207)	0.5	<a href="#">Font-Ribera &amp; Miralda-Escudé (2012)</a>
$\beta_{\text{CIV}}$	Paramètre RSD pour CIV(eff)	0.27	<a href="#">Blomqvist et al. (2018)</a>
$\gamma_{\text{Ly}\alpha}$	Paramètre d’évolution en <i>redshift</i> de $b_{\text{Ly}\alpha}$	2.9	<a href="#">McDonald (2003)</a>
$\gamma_m$	Paramètre d’évolution en <i>redshift</i> de $b_m$	1	<a href="#">Bautista et al. (2017)</a>
$q_1$	Paramètre correctif de formation des structures	0.856	<a href="#">Arinyo-i-Prats et al. (2015)</a>
$k_v, a_v, b_v$	Paramètre correctif des vitesses particulières	1.11,0.54,1.61	<a href="#">Arinyo-i-Prats et al. (2015)</a>
$k_p$	Paramètre correctif de pression	19.47	<a href="#">Arinyo-i-Prats et al. (2015)</a>

## 6.5 L’ajustement

Je décris dans cette section l’ajustement du modèle défini par l’équation 6.1 sur les fonctions de corrélation mesurées. J’introduirai dans un premier temps l’expression de la fonction de vraisemblance à minimiser pour obtenir le meilleur jeu de paramètres. Puis, en s’appuyant sur cette expression, on verra quels sont les paramètres bien contraints du modèle et quels sont ceux à fixer. Dans la sous-section suivante je décrirai les modèles utilisés et les limites de l’ajustement dans chaque cas.

### 6.5.1 La fonction de vraisemblance

On cherche le jeu de paramètres  $\{p_i\}$  qui minimise l’écart entre les données et le modèle théorique. On définit la fonction de vraisemblance,  $\chi^2$ , dont le minimum est déterminé par les paramètres recherchés :

$$\chi^2(\{p_i\}) = \sum_{A,B} (\hat{\xi}_A - \xi_A^{\text{mod}}(\{p_i\})) C_{AB}^{-1} (\hat{\xi}_B - \xi_B^{\text{mod}}(\{p_i\})), \quad (6.47)$$

où  $C_{AB}$  est la matrice de covariance calculée à la section 5.3.

TABLE 6.4: Liste des paramètres libres dans le modèle physique. Dans le cas du modèle avec terme de *broadband* d'ordre 1, on introduit des priors supplémentaires sur les paramètres  $b_{Ly\alpha}$ ,  $\beta_{Ly\alpha}$  et  $b_{HCD}$ .

Paramètres	Définitions	Priors	Références
$\alpha_{\parallel}, \alpha_{\perp}$	Paramètres de positionnement du pic BAO	non	
$b_{Ly\alpha}$	Biais de $Ly\alpha$	non	
$\beta_{Ly\alpha}$	Paramètre RSD de $Ly\alpha$	non	
$b_{HCD}$	Biais des HCD	non	
$\beta_{HCD}$	Paramètres RSD des HCD	$0.5 \pm 0.2$	Font-Ribera & Miralda-Escudé (2012)
$b_m$	Biais de Si II(1190), (1193), (1260) et Si III(1207)	non	
$b_{CIV(\text{eff})}$	Biais de CIV(eff)	$-0.005 \pm 0.0026$	Blomqvist et al. (2018)
$a_{ij}$	Coefficients de <i>broadband</i>	non	

On verra dans la sous-section suivante que certains paramètres sont peu contraints par les données. On leur impose donc d'être proches d'une valeur théorique déterminée par une mesure externe à cette étude (voir table 6.4). Cette contrainte s'exprime en ajoutant un prior gaussien à la fonction de vraisemblance :

$$\chi^2(\{p_i\}) \longrightarrow \chi^2(\{p_i\}) + \sum_j \frac{(p_j - \mu_{p_j})^2}{\sigma_{p_j}^2}, \quad (6.48)$$

où  $p_j$  désigne les paramètres à contrôler,  $\mu_{p_j}$  la valeur attendue et  $\sigma_{p_j}$  l'erreur de la mesure.

Dans le cas d'un ajustement combiné, c'est-à-dire de plusieurs fonctions de corrélation en même temps, je minimise la somme des  $\chi^2$  des fonctions de corrélations. Les paramètres libres peuvent être communs ou non.

### 6.5.2 Contraintes sur l'ajustement des paramètres

Dans cette sous-section, j'étudie quels sont les *bins* impliqués dans l'ajustement des différents paramètres. Cette étude nous permet de répondre à deux questions :

- Est-ce que le paramètre étudié est bien contraint par le modèle ?
- Quelles sont ses dégénérescences avec les autres paramètres ?

Soit le paramètre libre  $p_i$ . Pour l'ajuster, je dérive la fonction de vraisemblance 6.47 par rapport à  $p_i$  :

$$\frac{\partial \chi^2}{\partial p_i} = \sum_A \frac{\partial \xi_A^{mod}}{\partial p_i} \left\{ \sum_B 2C_{AB}^{-1} (\xi_B^{mod} - \hat{\xi}_B) \right\} \quad (6.49)$$

La contribution du *bin* A à l'ajustement du paramètre  $p_i$  peut être quantifiée par  $\partial \xi_A^{mod} / \partial p_i$ .

Sur les figures 6.13, 6.14, 6.15 et 6.16, je représente les quantités  $\partial \xi^{mod} / \partial p_i$  pour plusieurs paramètres  $p_i$  dans le cas  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha)$ . Le cas  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\beta)$  donne des résultats analogues.

On voit sur la figure 6.13 que les paramètres de position du pic BAO sont, comme attendu, contraints par les *bins* entre 70 et 130  $h^{-1}\text{Mpc}$  et on retrouve l’empreinte du pic BAO vue sur la figure 6.4.  $\alpha_{\parallel}$  est contraint par les *bins* proches de la ligne de visée, c’est-à-dire pour les  $\mu$  grands et  $\alpha_{\perp}$ , à  $\mu$  plus faibles. Les effets des paramètres  $\alpha_{\perp}$  et  $\alpha_{\parallel}$  sont ciblés : c’est-à-dire localisés sur certains *bins*. C’est le signe qu’ils sont bien contraints. On remarque également que  $\alpha_{\parallel}$  est mieux contraint que  $\alpha_{\perp}$  car l’amplitude de variation de  $\partial\xi^{mod}/\partial\alpha_{\parallel}$  est supérieure à celle de  $\partial\xi^{mod}/\partial\alpha_{\perp}$ .

La figure 6.14 montre la contribution des *bins* dans l’ajustement de  $b_{Ly\alpha}$  et  $\beta_{Ly\alpha}$ . Le biais  $Ly\alpha$  est contraint par les *bins* correspondants au premier pic de la fonction de corrélation ( $r < 40h^{-1}\text{Mpc}$ , cf. figure 5.10) ainsi que les *bins* avec  $\mu \sim 1$  à cause de l’amplification de la fonction de corrélation par les RSD.  $\beta_{Ly\alpha}$  est bien moins contraint, car on voit que l’amplitude de variation de  $\partial\xi^{mod}/\partial\beta_{Ly\alpha}$  est plus faible. Les *bins* utilisés lors de l’ajustement sont ceux qui sont les plus impactés par les RSD (voir figure 6.4), c’est-à-dire pour  $\mu \sim 1$  et  $\mu \sim 0$  avec  $r > 20 h^{-1}\text{Mpc}$ .

La figure 6.15 montre que les paramètres de Kaiser des HCD sont ajustés sur les mêmes *bins* que les paramètres  $Ly\alpha$ . En effet, on voit que les motifs de  $\partial\xi^{mod}/\partial p_i$  sont similaires. Ce qui indique que ces paramètres sont dégénérés entre eux :  $b_{Ly\alpha}$  avec  $b_{\text{HCD}}$  et  $\beta_{Ly\alpha}$  avec  $\beta_{\text{HCD}}$ . En revanche, on remarque que l’amplitude des variations est 100 fois plus faible dans le cas des HCD, ce qui veut dire qu’ils sont moins bien contraints. Le paramètre  $\beta_{\text{HCD}}$  est un bon exemple de paramètre aux effets à grande échelle : il est contraint par quasiment l’ensemble des *bins* et l’amplitude des variations de  $\partial\xi^{mod}/\partial\beta_{\text{HCD}}$  est très faible. Dans la section 6.5.3, j’imposerai un prior gaussien sur ce paramètre.

On représente sur la figure 6.16 les contributions à l’ajustement des paramètres  $b_{\text{SiII}(1190)}$  et  $b_{\text{CIV}(\text{eff})}$ . Le motif dessiné par  $\partial\xi^{mod}/\partial b_{\text{SiII}(1190)}$  est similaire au biais  $Ly\alpha$  de la figure 6.14 mais décalé d’environ 60  $h^{-1}\text{Mpc}$  le long de la ligne de visée. Ce décalage correspond à celui calculé dans la section 6.2.1. Le paramètre  $b_{\text{CIV}(\text{eff})}$  est peu contraint, c’est pourquoi je lui imposerai un prior gaussien à la section 6.5.3.

### 6.5.3 Les modèles ajustés

L’objectif de cette étude est la mesure des paramètres de position du pic BAO  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ . Il est essentiel que les valeurs mesurées ne soient pas influencées par le type de modèle et les erreurs systématiques inconnues. C’est pourquoi, je vais ajuster trois modèles différents :

- Le modèle physique : c’est le modèle standard de l’étude à partir duquel je vais déduire les contraintes cosmologiques. Dans ce cas le terme  $\xi^{bb}$  de l’équation 6.1 est nul.
- Le modèle avec terme de *broadband* d’ordre 1 : j’inclus le terme de *broadband* de l’équation 6.1 en imposant des priors gaussiens sur les paramètres  $b_{Ly\alpha}$ ,  $\beta_{Ly\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  correspondants aux valeurs obtenues avec l’ajustement standard. Cela revient à considérer le terme de *broadband* comme une perturbation (et donc un ordre 1) du modèle physique.
- Le modèle avec terme de *broadband* d’ordre 0 : le terme de *broadband* peut dominer, les valeurs des paramètres  $b_{Ly\alpha}$ ,  $\beta_{Ly\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  ne reflètent plus la réalité physique.

À la fin, je vérifie que les trois modèles donnent des valeurs similaires pour  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ .

### 6.5.3.1 Le modèle physique

Puisque le pic BAO est situé à environ  $r \simeq 100 h^{-1}\text{Mpc}$ , je choisis une fenêtre d'étude en  $r$  centrée sur cette valeur. Globalement, je vais ajuster la position du pic sur le domaine  $70 < r < 130 h^{-1}\text{Mpc}$  visible sur la figure 6.13. En revanche, pour l'ajustement du reste de la fonction de corrélation, il faut étendre ce domaine notamment aux  $r$  inférieurs à  $40 h^{-1}\text{Mpc}$  où se trouvent les *bins* qui contraignent le mieux  $b_{Ly\alpha}$  et  $\beta_{Ly\alpha}$ . Les effets non-linéaires se manifestent aux petites échelles, il ne faut pas être trop proche de  $0 h^{-1}\text{Mpc}$ . L'extension au delà de  $130 h^{-1}\text{Mpc}$  permet de contrôler le comportement de la fonction de corrélation aux grandes distances. Pour cette étude, je reprends les limites proposées et testées sur les simulations par [Bautista et al. \(2017\)](#) :  $10 < r < 180 h^{-1}\text{Mpc}$ . Mais j'ai vérifié que  $r_{\min}$  pouvait varier entre 5 et  $40 h^{-1}\text{Mpc}$  et  $r_{\max}$ , entre 160 et  $200 h^{-1}\text{Mpc}$  sans changer significativement les valeurs de  $(\alpha_{\parallel}, \alpha_{\perp})$ .

Les valeurs de  $\mu$  peuvent varier entre 0 et 1 et les distances de séparation  $r_{\perp}$  et  $r_{\parallel}$ , entre 0 et  $200 h^{-1}\text{Mpc}$ . Les bornes de l'ajustement sont répertoriées dans la table 6.2. La limite inférieure en  $r$  supprime 4 *bins*, la limite supérieure, 906. Finalement, les fonctions de corrélation ajustées  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  ont chacune 1590 *bins*.

Certains paramètres du modèle vont être fixés à des valeurs provenant d'études externes. Ils sont répertoriés dans le tableau 6.3 :

- L'amplitude du pic BAO (voir équation 6.23) dépend du paramètre de densité des baryons  $\Omega_b$  ([Sutherland & Mularczyk 2014](#)). Laisser  $A_{\text{BAO}}$  libre revient à utiliser deux valeurs de  $\Omega_b$  différentes : une pour le pic BAO et une pour le reste de la fonction de corrélation<sup>8</sup>. C'est pourquoi  $A_{\text{BAO}}$  sera fixé à 1 dans le modèle standard. Dans des modèles alternatifs (voir chapitre 7), je vérifierai que la position mesurée du pic ne dépend pas de son amplitude.
- Pour simplifier le modèle et diminuer le temps de calcul,  $\Sigma_{\parallel}$  et  $\Sigma_{\perp}$  sont fixés par [Kirkby et al. \(2013\)](#). Mais j'ai vérifié que les laisser libres n'influaient pas significativement sur la position du pic BAO.
- Par définition,  $R_{\parallel}$  et  $R_{\perp}$  sont fixes et égaux à  $4 h^{-1}\text{Mpc}$ .
- Quand il est laissé libre, le paramètre  $L_{\text{HCD}}$  est très dégénéré avec  $\beta_{\text{HCD}}$ ,  $b_{\text{SiII}(1190)}$  et  $b_{\text{SiII}(1193)}$ . C'est pourquoi je le fixe à la valeur de [Rogers et al. \(2018\)](#),  $L_{\text{HCD}} = 10 h^{-1}\text{Mpc}$ .

<sup>8</sup>Puisque le spectre de puissance linéaire est calculé en utilisant les paramètres de la cosmologie fiducielle, cela revient à fixer  $\Omega_b$  dans  $\xi^{\text{lisse}}$ . En revanche, si on fait varier,  $A_{\text{BAO}}$ , alors on modifie artificiellement la partie BAO de la fonction de corrélation (voir équation 6.23) et donc ses paramètres cosmologiques (dont  $\Omega_b$ ). Je rappelle que l'objectif de l'ajustement n'est pas de mesurer directement les paramètres cosmologiques sur la fonction de corrélation mais, supposant une cosmologie fiducielle, de mesurer l'écart entre la réalité et cette cosmologie.

- On a vu dans la section 6.5.2 que les paramètres RSD avaient des effets à grande échelle et étaient ajustés sur l'ensemble de la fonction de corrélation. Or puisque la contamination par les métaux est visible pour  $r_{\perp} \sim 0$ , les paramètres RSD des métaux ne sont pas contraints par la fonction de corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$  mesurée. Dans le cas du CIV, j'utilise  $\beta_{\text{CIV}(\text{eff})} = 0.27$ , valeur obtenue par la corrélation croisée entre les quasars et les absorptions CIV dans les forêts CIV (Blomqvist et al. 2018). Pour les autres métaux, j'utilise  $\beta_m = 0.5$ , valeur déduite de simulations des champs d'absorption par les HCD (Font-Ribera & Miralda-Escudé 2012).
- Le paramètre d'évolution en *redshift*  $\text{Ly}\alpha$  est déterminé par le spectre de puissance à une dimension (McDonald 2003).
- Pour simplifier le modèle, la dépendance en *redshift* des absorptions des métaux est ignorée,  $\gamma_m = 1$  (Bautista et al. 2017).
- Pour simplifier le modèle et éviter les dégénérescences, j'utilise les paramètres correctifs non-linéaires ajustés sur des simulations hydrodynamiques par Arinyo-i-Prats et al. (2015) (Table 7).

Les 11 paramètres libres sont listés dans le tableau 6.4.  $\beta_{\text{HCD}}$  et  $b_{\text{CIV}(\text{eff})}$  sont peu contraints par les données (voir sous-section 6.5.2), je leur ajoute un prior gaussien pour que les valeurs ajustées soient physiques.  $\beta_{\text{HCD}} = 0.5 \pm 0.2$  est donné par Font-Ribera & Miralda-Escudé (2012) grâce à l'utilisation de simulations des champs d'absorption par les HCD et  $b_{\text{CIV}(\text{eff})} = -0.005 \pm 0.0026$  est mesuré par la corrélation croisée entre les quasars et les absorptions CIV dans les forêts CIV par Blomqvist et al. (2018).

### 6.5.3.2 Les modèles avec terme de *broadband*

Le terme de *broadband* est introduit pour décrire l'écart entre la mesure et le modèle et ainsi, vérifier que les paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  ne dépendent pas d'erreurs systématiques inconnues. Pour des raisons expliquées à la section 6.4, on réduit le domaine d'étude à  $r$  dans  $[40, 180] h^{-1} \text{Mpc}$ .

Dans un premier modèle, le modèle d'ordre 1, j'impose un prior gaussien sur les paramètres  $b_{\text{Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$ . Les valeurs du prior sont celles obtenues lors de l'ajustement physique et les écart-types sont égaux à deux fois l'erreur de l'ajustement physique. Le terme de *broadband* peut-être vu comme une perturbation du modèle physique.

Dans un deuxième modèle, le modèle d'ordre 1, je n'impose aucun prior supplémentaire par rapport à l'ajustement standard. Le terme de *broadband* peut devenir dominant et à cause des dégénérescences, les paramètres  $b_{\text{Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  peuvent s'éloigner des valeurs physiques.

Ces deux modèles sont des tests pour vérifier que les paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont compatibles à ceux obtenus avec le modèle physique (voir chapitre 7).

L'ajustement est réalisé avec le code `fitter2` du module `picca` (annexe C).

**Synthèse du chapitre**

Dans ce chapitre, j'ai décrit les trois modèles de la fonction de corrélation  $L_{y\alpha} \times L_{y\alpha}$  qui seront ajustés sur la fonction de corrélation mesurée au chapitre 5.

Dans un premier temps, j'ajusterai le modèle standard (ou modèle physique) qui contient les corrections des effets de RSD, de la contamination par les métaux et les HCD, de l'effet de l'échantillonnage, des effets non-linéaires dus à la formation des structures, et de la projection des  $\delta$ .

Puis j'ajusterai deux modèles avec terme de *broadband* pour vérifier que la position du pic BAO n'est pas influencée par des défauts de modélisation éventuels. Le premier modèle avec terme de *broadband*, dit "d'ordre 1", est ajusté avec des priors gaussiens sur les principaux paramètres. La valeur des priors est définie par les résultats de l'ajustement standard. Le deuxième modèle avec terme de *broadband*, dit "d'ordre 0", est ajusté sans prior supplémentaire. Le terme de *broadband* n'est plus considéré comme une perturbation et peut dominer. Les valeurs des paramètres ne sont plus physiques.



# Chapitre 7

## Résultats

Dans ce chapitre je présente les résultats de l'ajustement des différents modèles sur les données. J'utilise dans un premier temps uniquement la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ , puis j'ajoute  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . Enfin, j'intègre dans l'ajustement la corrélation croisée entre les quasars et les absorptions  $\text{Ly}\alpha$  dans les régions  $\text{Ly}\alpha$  et  $\text{Ly}\beta$ ,  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$ , étudiée par [Blomqvist et al. \(2019\)](#). Dans la deuxième partie de ce chapitre, je m'intéresse aux implications cosmologiques de ces résultats.

### 7.1 Résultats des ajustements

#### 7.1.1 Ajustement simple de $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$

Dans cette sous-section, j'ajuste le modèle de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  sur la fonction de corrélation mesurée en minimisant la fonction de vraisemblance donnée par les équations (6.47) et (6.48).

Je teste divers modèles pour contrôler la stabilité des paramètres de position du pic  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ . J'ajoute petit à petit les ingrédients du modèle décrit dans le chapitre précédent, en commençant par un modèle simple avec uniquement la décomposition du spectre de puissance, le terme de Kaiser et les corrections non-linéaires (appelé "modèle de Kaiser" dans la suite). Ce modèle contient quatre paramètres libres :  $\alpha_{\parallel}, \alpha_{\perp}, b_{\text{Ly}\alpha}, \beta_{\text{Ly}\alpha}$ . Puis j'ajoute les métaux avec cinq paramètres libres supplémentaires (section 6.2), puis la correction des HCD (deux nouveaux paramètres libres). Ce dernier modèle est appelé "modèle physique" ou "standard", duquel je

TABLE 7.1: Les paramètres de position du pic BAO lors de l'ajustement de la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  sont stables quel que soit le modèle utilisé. Les erreurs sur les paramètres sont supposées gaussiennes et correspondent à  $\Delta\chi^2 = 1$ .

Modèles	$\alpha_{\parallel}$	$\alpha_{\perp}$	$\chi^2/\text{DDL}$	probabilité
Kaiser	$1.034 \pm 0.032$	$0.984 \pm 0.041$	1884.27/(1590-4)	0.000
+ Métaux	$1.043 \pm 0.035$	$0.979 \pm 0.042$	1756.90/(1590-9)	0.001
+ HCD (standard)	$1.047 \pm 0.036$	$0.969 \pm 0.042$	1619.77/(1590-11)	0.232
+ <i>broadband</i> ordre 1	$1.055 \pm 0.035$	$0.966 \pm 0.041$	1601.33/(1590-23)	0.267
+ <i>broadband</i> ordre 0	$1.055 \pm 0.035$	$0.967 \pm 0.042$	1597.63/(1590-23)	0.289
+ $A_{\text{BAO}}$ libre	$1.045 \pm 0.026$	$0.985 \pm 0.032$	1593.16/(1590-24)	0.310

TABLE 7.2: Résultats de l'ajustement des paramètres sur les données  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  seules (deuxième colonne) et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  combinées (troisième colonne). Les erreurs sur les paramètres sont supposées gaussiennes et correspondent à  $\Delta\chi^2 = 1$ .

Paramètres	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ $+\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$
$\bar{z}$	2.34	2.34
$n_{pairs}$	$5.44 \times 10^{11}$	$6.94 \times 10^{11}$
$\sum w_{pairs}$	$3.56 \times 10^{13}$	$4.20 \times 10^{13}$
$\alpha_{\parallel}$	$1.047 \pm 0.035$	$1.033 \pm 0.031$
$\alpha_{\perp}$	$0.969 \pm 0.041$	$0.953 \pm 0.042$
$\beta_{\text{Ly}\alpha}$	$1.773 \pm 0.066$	$1.933 \pm 0.100$
$b_{\text{Ly}\alpha}(\bar{z})$	$-0.113 \pm 0.036$	$-0.106 \pm 0.033$
$\beta_{\text{HCD}}$	$0.845 \pm 0.157$	$1.031 \pm 0.153$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)}$	$-0.047 \pm 0.003$	$-0.051 \pm 0.004$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)}$	-	$-0.072 \pm 0.005$
$b_{\text{SiII}(1190)}$	$-0.0051 \pm 0.0010$	$-0.0050 \pm 0.0010$
$b_{\text{SiII}(1193)}$	$-0.0046 \pm 0.0010$	$-0.0046 \pm 0.0010$
$b_{\text{SiIII}(1207)}$	$-0.0082 \pm 0.0010$	$-0.0080 \pm 0.0010$
$b_{\text{SiII}(1260)}$	$-0.0025 \pm 0.0013$	$-0.0022 \pm 0.0013$
$b_{\text{CIV}(\text{eff})}$	$-0.0185 \pm 0.0078$	$-0.0163 \pm 0.0089$
$\chi_{min}^2$	1619.77	3258.92
DDL	1590-11	3180-12
Probabilité	0.232	0.127
$\chi^2(\alpha_{\parallel} = \alpha_{\perp} = 1)$	1621.55	3260.54

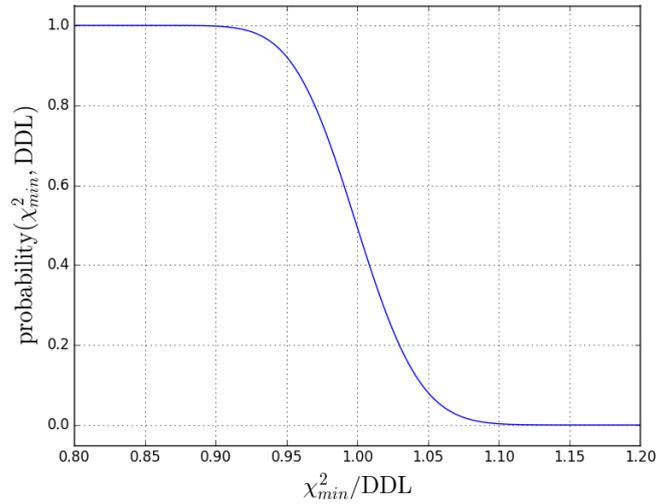


FIGURE 7.1: Probabilité en fonction du  $\chi^2$  réduit. On considère que le modèle n'explique pas les données quand la probabilité est inférieure au pourcent. Ce critère correspond à un  $\chi^2$  réduit supérieur à 1.1.

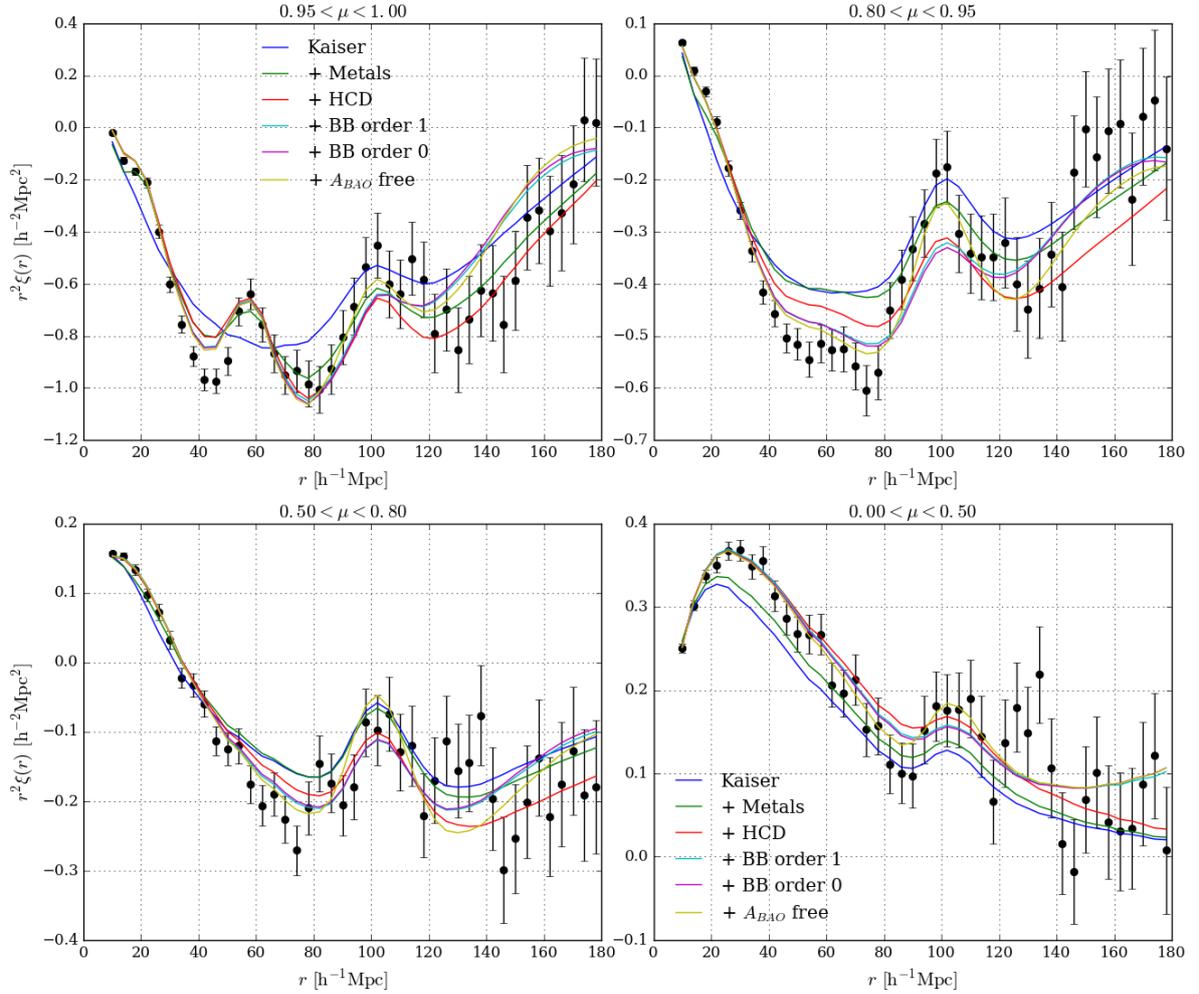


FIGURE 7.2: Fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  mesurée (points noirs) en fonction de la distance de séparation  $r$  entre les paires d'absorptions, pour quatre gammes du paramètre  $\mu = \frac{r_{\parallel}}{r}$ . Je superpose aux données les différents modèles testés. Compte tenu des barres d'erreurs, la position du pic BAO est sensiblement la même.

déduirai à la section 7.2 les contraintes cosmologiques. On étudie également des modèles non physiques avec des termes de *broadband* d'ordre 1 et 0, qui modélisent les erreurs systématiques inconnues. Pour le terme de *broadband* d'ordre 1, je testerai également si l'amplitude du pic BAO influe sur les valeurs de  $(\alpha_{\parallel}, \alpha_{\perp})$ .

Afin de caractériser la qualité de chaque ajustement on introduit la probabilité de  $\chi^2$ . Elle est calculée par l'intégrale de la densité de probabilité entre la valeur ajustée  $\chi_{\min}^2$  et  $\infty$ . La densité de probabilité pour la vraisemblance  $\chi_{\min}^2$  est donnée par :

$$f(\chi_{\min}^2, k) = \frac{1}{2^{k/2}\Gamma(k/2)} (\chi_{\min}^2)^{k/2-1} \exp\left(-\frac{\chi_{\min}^2}{2}\right), \quad (7.1)$$

avec  $k$ , le nombre de degrés de liberté (DDL) égal à la différence entre le nombre de points de mesure (1590 dans le cas présent) et le nombre de paramètres libres du modèle.

Plus le modèle ajusté est proche des données, plus la probabilité s'approche de 1. Au contraire, quand le modèle s'éloigne des points de mesure, la probabilité devient petite. Physiquement, elle désigne la probabilité d'avoir un moins bon ajustement (c'est-à-dire un plus grand  $\chi_{\min}^2$ ) si on répète l'expérience. Une valeur trop faible (typiquement, inférieure au pourcent), nous indique que le modèle ne décrit pas les données. Sur la figure 7.1, je représente la probabilité en fonction du  $\chi^2$  réduit. Le modèle est invalidé quand le  $\chi^2$  réduit est supérieur à 1.1.

Je présente les résultats des différents ajustements sur la figure 7.2 et dans le tableau 7.1. Sur la figure 7.2, on voit que plus le modèle est sophistiqué, mieux il ajuste les points de la mesure (en noir). On remarque néanmoins que les ajustements semblent échouer à expliquer complètement les données quand  $\mu > 0.85$  et  $r$  dans  $[40,60] h^{-1}\text{Mpc}$ . On ne connaît pas l'origine de cet effet, déjà observé par Bautista et al. (2017), mais il pourrait être lié à la contamination par les HCD. Cependant la position du pic BAO ne semble pas être influencée par le modèle, quelque soit la gamme de  $\mu$  considérée. Ces constatations sont confirmées par les valeurs du tableau 7.1. On remarque que la probabilité augmente avec le nombre de paramètres libres. L'ajout du terme de *broadband* montre qu'il reste des erreurs systématiques à comprendre et à intégrer dans le modèle. On note également que les variations des  $(\alpha_{\parallel}, \alpha_{\perp})$  d'un modèle à l'autre sont inférieures aux barres d'erreur. On en déduit que les paramètres de position du pic BAO ne sont pas impactés par notre connaissance incomplète des erreurs systématiques de la mesure.

Je présente dans la deuxième colonne du tableau 7.2 les valeurs ajustées de l'ensemble des paramètres du modèle standard. La première ligne indique le *redshift* effectif de la mesure, qui est défini comme le *redshift* moyen  $\bar{z}$  des paires de Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ).

Dans le modèle, au chapitre 6, j'ai fait la distinction entre les absorptions fortes (les HCD), qui sont caractérisées par les paramètres  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$ , et les autres absorptions, qui sont caractérisées par les paramètres  $b_{\text{Ly}\alpha}$  et  $\beta_{\text{Ly}\alpha}$ . Les valeurs des paramètres de biais et de RSD pour l'ensemble des absorptions Ly $\alpha$  sont données par les paramètres effectifs des équations (6.10) et (6.11), appliquées aux mesures de  $b_{\text{Ly}\alpha}$ ,  $\beta_{\text{Ly}\alpha}$ ,  $b_{\text{HCD}}$  et  $\beta_{\text{HCD}}$  du tableau 7.2. Pris individuellement, chacun de ces quatre paramètres n'a de sens bien défini puisque la frontière entre une absorption classique et un HCD n'est pas parfaitement nette, ce qui induit des dégénérescences entre  $b_{\text{Ly}\alpha}$  et  $b_{\text{HCD}}$ , et entre  $\beta_{\text{Ly}\alpha}$  et  $\beta_{\text{HCD}}$ . C'est pourquoi, il faut utiliser les paramètres effectifs pour décrire la physique du milieu intergalactique. Le long de la ligne de visée,  $F_{\text{HCD}} \simeq 1$  et les paramètres

effectifs de Kaiser (équations 6.10 et 6.11) sont donnés par :

$$\begin{cases} b'_{Ly\alpha} &= -0.161 \pm 0.049 \\ \beta'_{Ly\alpha} &= 1.450 \pm 0.065 \end{cases} \quad (7.2)$$

Ces valeurs sont en accord avec les travaux précédents (Busca et al. 2013, Delubac et al. 2015, Bautista et al. 2017) et les résultats des simulations hydrodynamiques de McDonald (2003).

La figure 5.11 rend compte de l'abondance de certaines transitions atomiques dans les absorbants les plus denses du milieu intergalactique : plus une transition métallique est fréquente, plus le puits correspondant est profond, et donc plus son biais ajusté sur la fonction de corrélation devrait être grand. Ainsi, il est étonnant que le biais de Si II(1260) soit nettement inférieur aux autres transitions du silicium et détecté avec une moins bonne précision. Cet effet est dû à l'éloignement en longueur d'onde de la transitions Si II(1260) par rapport à la transition Ly $\alpha$  : plus la distance de séparation entre les deux transitions est grande, plus le pic métallique apparaît aux grandes séparations, où la précision de la fonction de corrélation est plus faible.

Le CIV(eff) est détecté avec une précision bien moins importante que les autres métaux pour deux raisons :

- On le détecte à travers la corrélation CIV(eff)  $\times$  CIV(eff), qui a une amplitude bien moins grande que les corrélations métal  $\times$  Ly $\alpha$ ,
- Il est détecté sur les mêmes *bins* que  $b_{Ly\alpha}$ , d'où une dégénérescence importante entre les deux paramètres, qui se traduit par l'augmentation de leurs barres d'erreur.

Finalement pour Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ), la probabilité est de 23%, largement au-dessus du critère de 1%, signifiant que le modèle est une bonne reproduction des données. La cosmologie fiducielle ( $\alpha_{\parallel} = \alpha_{\perp} = 1$ ) est proche du meilleur ajustement avec un écart sur le minimum de la fonction de vraisemblance d'environ 0.1 %.

### 7.1.2 Ajustement combiné de Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) et Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\beta$ )

Dans cette sous-section, je présente l'ajustement combiné des fonctions de corrélation mesurées Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ), c'est-à-dire que j'ajuste simultanément les paramètres du modèle comme s'il s'agissait d'un seul jeu de données.

De même que précédemment, je teste plusieurs modèles, du plus simple au plus complet, afin de montrer la stabilité des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ . Je présente les différents ajustements dans le tableau 7.3 et la figure 7.3.

Comme dans le cas de l'ajustement simple de la sous-section précédente, les modèles n'ont pas d'influence (supérieure à la précision de mesure) sur  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  bien que la probabilité augmente avec le nombre de paramètres libres. On remarque que les probabilités des ajustements physiques pour la combinaison Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) et Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) sont plus faibles que pour Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) seul. Cet effet n'est pas encore compris mais il pourrait être dû

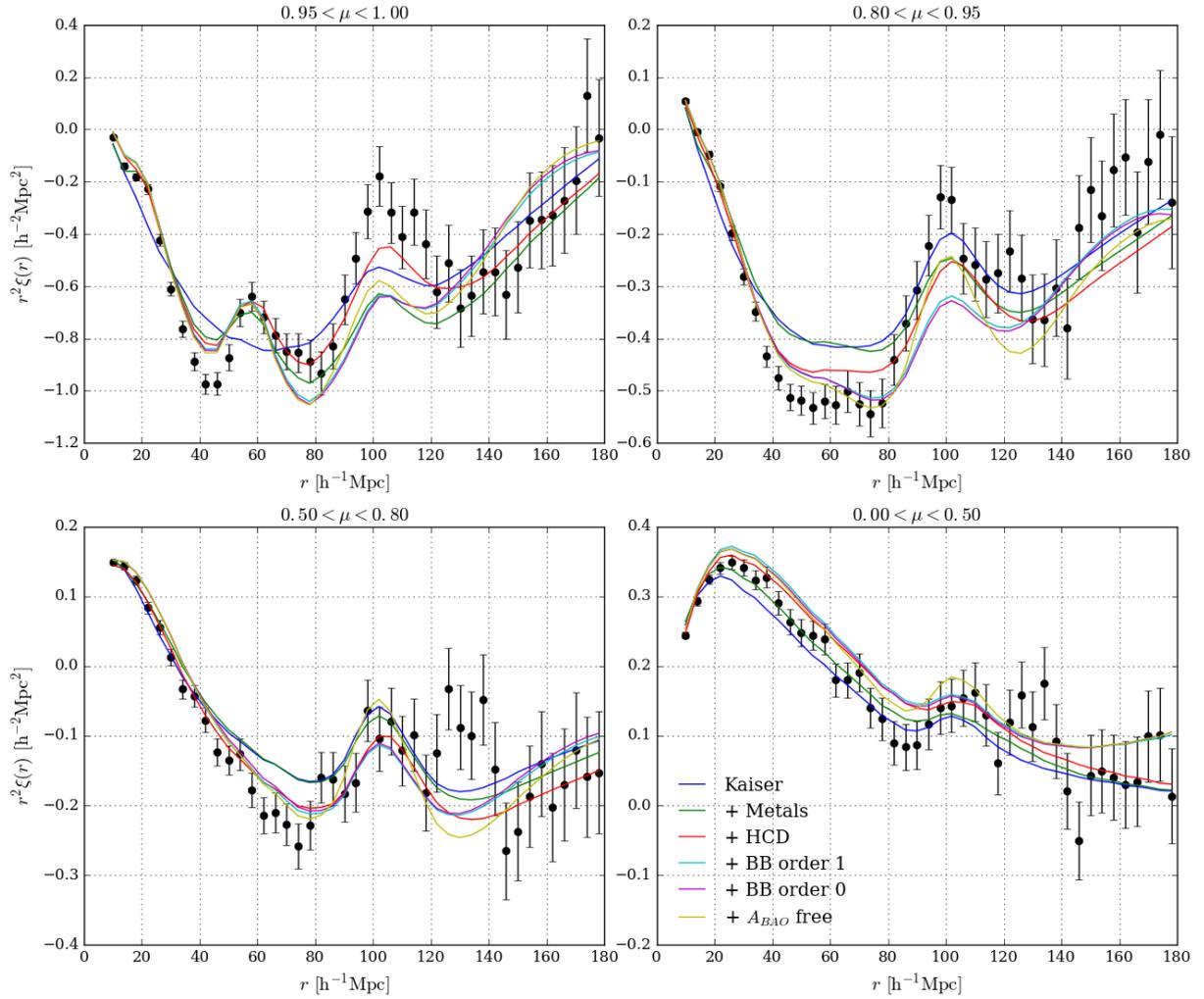


FIGURE 7.3: Moyenne pondérée des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  mesurées (points noirs) en fonction de la distance de séparation  $r$  pour quatre gammes de  $\mu$ . On superpose aux données les différents modèles testés et on remarque que la position du pic BAO est sensiblement la même.

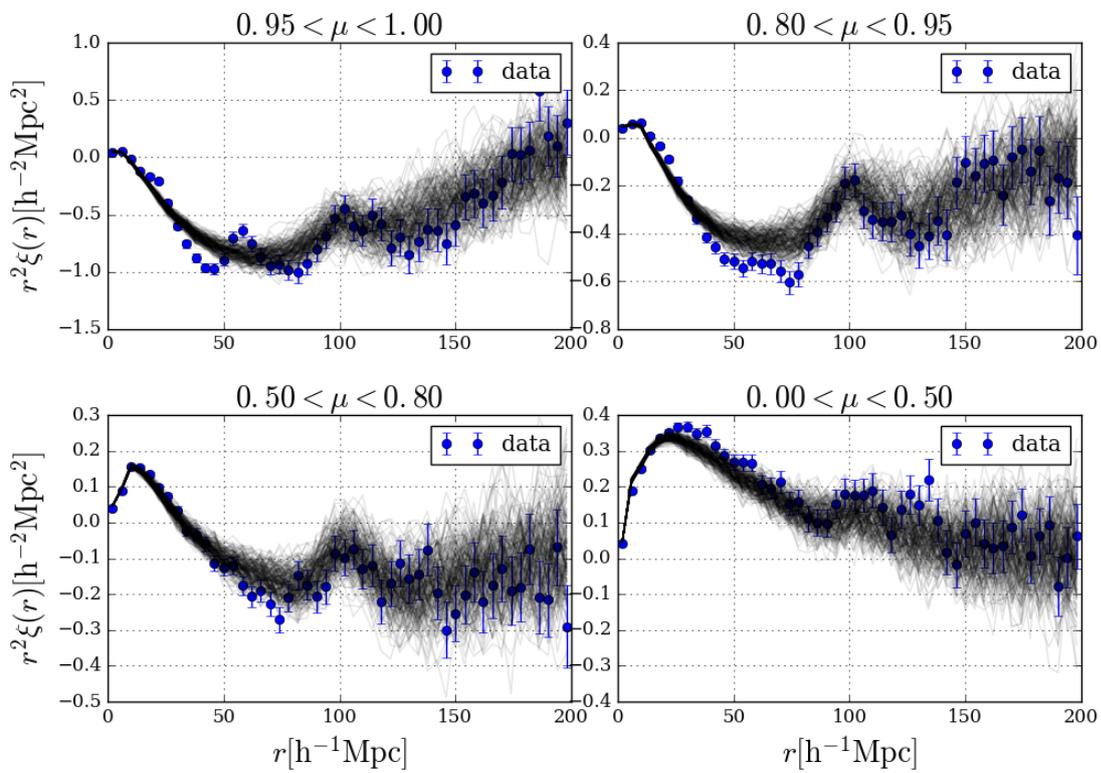


FIGURE 7.4: En bleu, fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  mesurée. En gris, 200 réalisations Monte-Carlo calculées à partir de l'ajustement du modèle de Kaiser. La dispersion des réalisations MC au niveau du pic BAO donne une idée de la précision de mesure de sa position.

TABLE 7.3: Les paramètres de position du pic BAO lors de l’ajustement combiné des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  sont stables quelque soit le modèle utilisé. Les erreurs sur les paramètres sont supposées gaussiennes et correspondent à  $\Delta\chi^2 = 1$ .

Modèles	$\alpha_{\parallel}$	$\alpha_{\perp}$	$\chi^2/\text{DDL}$	probabilité
Kaiser	$1.034 \pm 0.023$	$0.984 \pm 0.029$	3768.53/(3180-4)	$1.07 \times 10^{-12}$
+ Métaux	$1.036 \pm 0.026$	$0.993 \pm 0.034$	3495.05/(3180-9)	$3.96 \times 10^{-5}$
+ HCD (standard)	$1.033 \pm 0.031$	$0.953 \pm 0.042$	3258.92/(3180-12)	0.127
+ <i>broadband</i> ordre 1	$1.055 \pm 0.025$	$0.968 \pm 0.030$	3202.58/(3180-36)	0.229
+ <i>broadband</i> ordre 0	$1.055 \pm 0.025$	$0.967 \pm 0.029$	3200.46/(3180-36)	0.237
+ $A_{BAO}$ libre	$1.045 \pm 0.019$	$0.985 \pm 0.023$	3184.37/(3180-37)	0.299

 TABLE 7.4: Valeurs de  $\Delta\chi^2$  correspondant aux niveaux de confiance (68.27, 95.45%). Ces valeurs sont déduites de l’analyse de 10 000 simulations Monte Carlo des fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ ,  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$ .

Paramètres	$\Delta\chi^2$ (68.27%)	$\Delta\chi^2$ (95.45%)
Auto		
$\alpha_{\parallel}$	$1.19 \pm 0.03$	$4.74 \pm 0.09$
$\alpha_{\perp}$	$1.23 \pm 0.03$	$4.83 \pm 0.08$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.77 \pm 0.04$	$7.33 \pm 0.10$
Cross		
$\alpha_{\parallel}$	$1.15 \pm 0.02$	$4.48 \pm 0.10$
$\alpha_{\perp}$	$1.15 \pm 0.02$	$4.51 \pm 0.07$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.51 \pm 0.03$	$6.67 \pm 0.11$
Auto + Cross		
$\alpha_{\parallel}$	$1.08 \pm 0.02$	$4.29 \pm 0.10$
$\alpha_{\perp}$	$1.08 \pm 0.02$	$4.28 \pm 0.10$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.47 \pm 0.03$	$6.71 \pm 0.13$

à une dépendance des matrices de covariance et/ou de distorsion à la taille des régions étudiées ( $\text{Ly}\alpha$  ou  $\text{Ly}\beta$ ). Celles-ci seraient sous-estimées dans le cas  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , ce qui entraînerait une augmentation du  $\chi_{\min}^2$ . Cette hypothèse est en cours d’investigation dans le cadre de l’analyse finale d’eBOSS.

Les données représentées par les points noirs de la figure 7.3 montrent la moyenne pondérée de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . On leur superpose les différents modèles en fonction de la distance de séparation  $r$  pour différentes gammes de  $\mu = \frac{r_{\parallel}}{r}$ .

La troisième colonne du tableau 7.2 présente les valeurs ajustées des paramètres du modèle standard. Le *redshift* effectif est donné par la moyenne du *redshift* des paires des deux fonctions de corrélation. L’ajout des données de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  augmente de 30% le nombre total de paires mais seulement de 17% leur poids total car la région  $\text{Ly}\beta$  est plus bruitée que

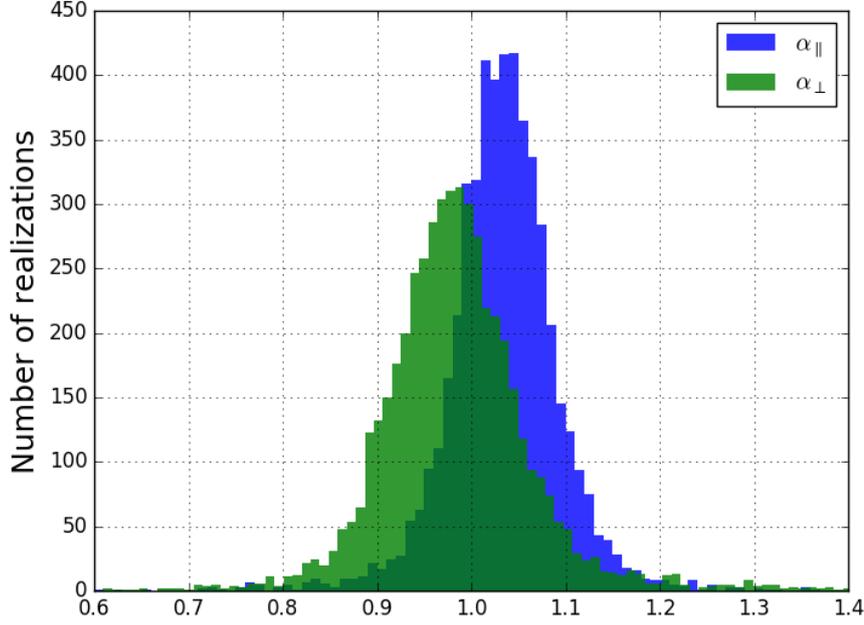


FIGURE 7.5: Histogramme des paramètres  $\alpha_{\parallel}$  (en bleu) et  $\alpha_{\perp}$  (en vert) ajustés sur 10 000 réalisations Monte-Carlo de la fonction de corrélation. Le paramètre  $\alpha_{\parallel}$  est mesuré avec une plus grande précision que  $\alpha_{\perp}$ .

$\text{Ly}\alpha$ <sup>1</sup>. Ainsi, malgré l'augmentation du nombre de données et contrairement aux attentes, les paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont mesurés, dans l'ajustement combiné, avec une précision similaire à l'ajustement simple de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ . Cet effet peut être dû à une fluctuation statistique ou bien à une modélisation imparfaite de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ , qui pourrait être éliminée par l'extension du domaine spectral de la région  $\text{Ly}\beta$ . Cependant, même si l'ajout de la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  n'améliore pas la précision de la mesure, il montre que la position du pic BAO mesurée est indépendante de la région spectrale considérée.

On ajoute le paramètre supplémentaire  $b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)}$  qui décrit le biais des HCD de la fonction de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$ . Il est plus grand que dans le cas de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  indiquant qu'il y a plus de HCD dans la région  $\text{Ly}\beta$  que dans la région  $\text{Ly}\alpha$ . Ce phénomène pourrait être causé par une diminution de la pureté du catalogue de DLA dans la région  $\text{Ly}\beta$ . En effet, puisque celle-ci est plus bruitée, il est plus difficile de détecter les absorptions saturées.

Les biais des métaux ne changent pas avec l'ajustement combiné et sont quasiment les mêmes que dans la deuxième colonne de la table 7.2. Le nombre de degré de liberté a presque doublé avec l'ajout de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et la probabilité est d'environ 13%. L'écart entre la fonction de vraisemblance minimale et la fiducielle est toujours faible et d'environ 0.05%.

<sup>1</sup>Puisque la précision des spectrographes sur le domaine spectral [360, 550]nm diminue avec la longueur d'onde (voir figure 11 de Smees et al. 2013), alors la région  $\text{Ly}\beta$ , située à gauche de la région  $\text{Ly}\alpha$ , est plus bruitée.

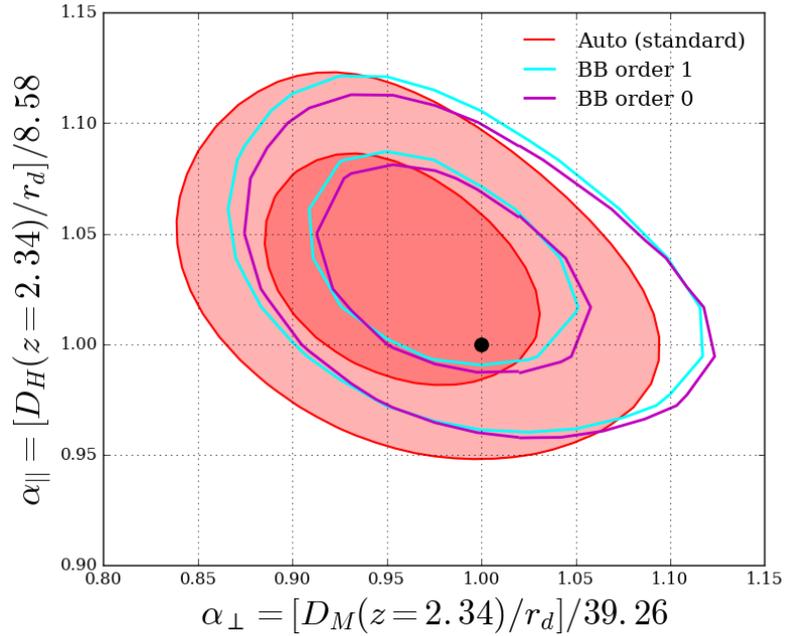


FIGURE 7.6: Niveaux de confiance à 68% et 95% des paramètres  $(\alpha_{\parallel}, \alpha_{\perp})$  dans le cas de l’ajustement combiné de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  en utilisant le modèle standard sans terme de *broadband* (rouge), avec terme de *broadband* d’ordre 1 (cyan) et avec terme de *broadband* d’ordre 0 (magenta). Le point noir représente le modèle fiduciel donné par [Planck Collaboration et al. \(2016\)](#).

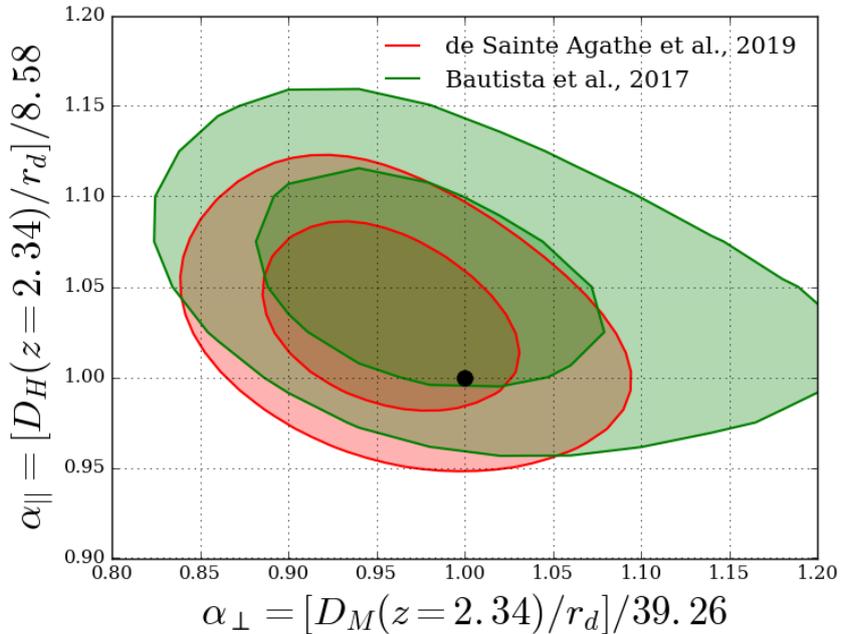


FIGURE 7.7: Niveaux de confiance à 68% et 95% des paramètres  $(\alpha_{\parallel}, \alpha_{\perp})$  obtenus dans cette étude ([de Sainte Agathe et al. 2019](#), en rouge) et dans l’étude précédente ([Bautista et al. 2017](#), en vert). Le point noir représente le modèle fiduciel donné par [Planck Collaboration et al. \(2016\)](#).

### 7.1.2.1 Estimation des erreurs

Pour chaque paramètre, l'écart-type associé est relié à la largeur de la fonction de vraisemblance  $\chi^2$  au voisinage du minimum. Soit  $p$  un paramètre libre. On introduit  $\chi_p^2$  qui est la fonction de vraisemblance marginalisée sur  $p$ . Le code d'ajustement utilise le module `iminuit` qui suppose que  $\chi_p^2$  est un polynôme de degré 2 et en déduit l'écart-type de  $p^2$ . Les erreurs des tableaux 7.1, 7.3 et 7.2 sont estimées par cette procédure.

En réalité,  $\chi_p^2$  n'est pas une fonction quadratique des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ , et puisque l'estimation de la précision sur la mesure de  $D_H(\bar{z})/r_d$  et  $D_M(\bar{z})/r_d$  dépend des erreurs de mesure des paramètres de position du pic, il est essentiel de bien les estimer. Pour cela, j'ai généré 10 000 simulations Monte Carlo à partir du modèle des fonctions de corrélation dont j'ai déduit l'erreur d'un paramètre à partir de la dispersion de ce paramètre sur l'ensemble des simulations.

Soient  $\xi^{mod}$  la fonction de corrélation modélisée et  $L$  la matrice telle que la matrice de covariance des données puisse s'écrire :  $C = LL^T$ . Je définis un vecteur  $G$  de la même dimension que la fonction de corrélation dont les composantes sont choisies aléatoirement entre 0 et 1. Ainsi le vecteur  $LG$  contient des écarts aléatoires à la fonction de corrélation inférieurs à  $1\sigma$ . Une réalisation Monte-Carlo de la fonction de corrélation est :

$$\xi^{MC} = \xi^{mod} + LG. \quad (7.3)$$

Pour des raisons de temps de calcul, les réalisations Monte-Carlo sont générées à partir du modèle de Kaiser (qui comprend les corrections non-linéaires et de Kaiser, voir première ligne du tableau 7.3) et non pas du modèle standard. J'ai choisi de négliger les contaminations par les métaux et les HCD car elles n'ont pas d'effet sur la mesure de la position du pic. De plus, [du Mas des Bourboux et al. \(2017\)](#) ont montré que la dispersion des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  était quasiment identique dans les cas du modèle standard et du modèle de Kaiser.

Je présente figure 7.4 les fonctions de corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  mesurées (en bleu) et 200 réalisations Monte-Carlo calculées à partir du modèle de Kaiser (en gris). La dispersion des réalisations MC donne une idée de l'incertitude de l'ajustement. La figure 7.5 montre l'histogramme des paramètres  $\alpha_{\parallel}$  (en bleu) et  $\alpha_{\perp}$  (en vert) pour ces 10 000 réalisations. Plus le pic est fin, plus la mesure est précise. On voit ainsi que le paramètre  $\alpha_{\parallel}$  est mieux mesuré que  $\alpha_{\perp}$ . Ce qui est attendu car les distorsions en *redshift* augmentent l'amplitude du pic BAO le long de la ligne de visée, rendant plus facile sa détection.

On définit le niveau de confiance de  $(\alpha_{\parallel}, \alpha_{\perp})$  à 68% comme la valeur  $\Delta\chi^2$  telle que, pour 68% des réalisations, la quantité  $\chi^2(\alpha_{\parallel} = \alpha_{\perp} = 1) - \chi^2$  est inférieure à  $\Delta\chi^2$ . De façon similaire, on définit les niveaux de confiance d'un paramètre  $p$  en mesurant les écarts entre les minima de la fonction de vraisemblance quand  $p$  est libre et quand il est fixé. Dans le cas d'une distribution gaussienne,  $\Delta\chi^2 = 1$  pour un paramètre et 2.3 pour deux paramètres. Les niveaux de confiance à 68% et 95% (respectivement 1 et 2 écart-type) obtenus lors de l'ajustement combiné de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  sont regroupées dans le tableau 7.4.

<sup>2</sup><http://seal.cern.ch/documents/minuit/mnerror.pdf>

Je corrige, par cette procédure, les erreurs sur la mesure des paramètres BAO données dans le tableau 7.2. Finalement,  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont donnés avec les déviations à *un* et *deux écart-types* :

$$\begin{cases} \alpha_{\parallel} = 1.033^{+0.034}_{-0.034} {}^{+0.071}_{-0.068} \\ \alpha_{\perp} = 0.953^{+0.050}_{-0.045} {}^{+0.108}_{-0.091} \end{cases} \quad (7.4)$$

La mesure de la position du pic perpendiculairement à la ligne de visée est donnée avec une précision de 5.2%. Grâce aux RSD, le signal est amplifié le long de la ligne de visée et la précision de  $\alpha_{\parallel}$  est plus grande, à 3.3%.

En utilisant les équations 6.25 et les valeurs fiducielles données dans l'annexe A, on obtient les rapports  $D_H/r_d$  et  $D_M/r_d$  au *redshift* 2.34 :

$$\begin{cases} D_H(2.34)/r_d = 8.86^{+0.29}_{-0.29} {}^{+0.61}_{-0.58} \\ D_M(2.34)/r_d = 37.41^{+1.96}_{-1.77} {}^{+4.24}_{-3.57} \end{cases} \quad (7.5)$$

La position du pic BAO à la recombinaison,  $r_d$ , est déduite de Planck (Planck Collaboration et al. 2016). La mesure de  $D_H(2.34)$  donne directement la mesure du taux d'expansion au *redshift* 2.34. Soit,

$$H(z = 2.34) = 230 \pm 7 \text{ km/s/Mpc.} \quad (7.6)$$

On peut comparer cette mesure à la cosmologie fiducielle :  $H_{fid}(2.34) = 237 \text{ km/s/Mpc}$ . Les deux valeurs sont compatibles, montrant que l'analyse BAO appliquée aux forêts Ly $\alpha$  est une confirmation à  $z = 2.34$  du modèle  $\Lambda$ CDM plat inspiré par les études de CMB à  $z \simeq 1100$ .

La mesure de  $D_M$  donne une valeur intégrée du taux d'expansion et peut être combinée avec  $D_H$  pour contraindre les paramètres cosmologiques  $\Omega_m$  et  $\Omega_{\Lambda}$  (voir section 7.2).

### 7.1.2.2 Figures de contours

Pour obtenir les contours sur  $(\alpha_{\parallel}, \alpha_{\perp})$ , je fixe le couple  $(\alpha_{\parallel}, \alpha_{\perp})$  pour différentes valeurs d'une grille et j'ajuste tous les autres paramètres du modèle standard. Je représente figure 7.6, les contours de la fonction de vraisemblance aux niveaux de confiance 68% et 95% calculés précédemment, en fonction de  $\alpha_{\parallel}$  et  $\alpha_{\perp}$ . En rouge, je montre le modèle standard que je compare aux modèles avec le terme de *broadband* d'ordre 1 (cyan) et 0 (magenta). On voit que l'effet de l'ajout du terme de *broadband* est plus important dans la direction perpendiculaire à la ligne de visée, ce qui est cohérent avec le fait que le paramètre  $\alpha_{\perp}$  est moins bien mesuré que  $\alpha_{\parallel}$ . Le décalage en  $\alpha_{\perp}$  est inférieur à un quart de l'écart-type, ce qui montre que compte-tenu de la précision statistique, on peut considérer que le modèle d'ajustement utilisé n'influe pas sur la position du pic.

Le point noir représente les résultats de la cosmologie fiducielle donnée par le modèle  $\Lambda$ CDM plat, appliqué aux résultats de la collaboration Planck (Planck Collaboration et al. 2016)<sup>3</sup>. On remarque qu'elle est tout à fait compatible avec la position du pic BAO mesurée.

<sup>3</sup>J'ai choisi de ne pas utiliser les derniers résultats de Planck donnés par la collaboration en 2018 (Planck Collaboration et al. 2018), car ils étaient disponibles seulement en pré-print au moment de la rédaction de cette thèse et pas encore publiés. De plus l'écart très faible (inférieur à 0.1% sur  $r_d$ ,  $\Omega_m$  et  $H_0$ ) entre les nouvelles valeurs et les précédentes aurait eu un impact largement négligeable devant l'incertitude de la mesure de la position du pic BAO.

Sur la figure 7.7, je représente en rouge les mêmes contours  $(\alpha_{\parallel}, \alpha_{\perp})$  que précédemment, que je compare à ceux obtenus dans l'étude antérieure de [Bautista et al. \(2017\)](#), en vert. On remarque que les deux paramètres sont mieux contraints dans la nouvelle analyse car les contours sont plus étroits. Dans le cas  $\alpha_{\parallel}$ , cette amélioration est faible (6.4%). En revanche, elle est très nette pour  $\alpha_{\perp}$  (13%). Ainsi, on voit peu de changement sur la mesure du taux d'expansion à  $z = 2.34$ , mais les contraintes sur  $(\Omega_m, \Omega_{\Lambda})$  sont plus fortes (voir section 7.2). On remarque également que notre résultat se rapproche plus de celui de Planck que le précédent.

Dans la section suivante, je réalise un ajustement combiné des fonctions d'auto-corrélation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  et de la corrélation croisée  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$ , présentée dans [Blomqvist et al. \(2019\)](#). Nous verrons que cet ajustement combiné permet d'augmenter la précision de mesure des paramètres cosmologiques  $\Omega_m$  et  $\Omega_{\Lambda}$ .

### 7.1.3 Ajustement combiné des fonctions de corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ et $\text{Ly}\alpha$ -quasar

#### 7.1.3.1 Modèle de la corrélation croisée

Dans cette sous-section je décris rapidement le modèle de la corrélation croisée  $\text{QSO} \times \text{Ly}\alpha$  détaillé dans [Blomqvist et al. \(2019\)](#).

De façon analogue à la fonction d'auto-corrélation  $\text{Ly}\alpha \times \text{Ly}\alpha$ , le modèle de la corrélation croisée  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  s'écrit comme la somme de plusieurs composantes, sur laquelle on applique une matrice de distorsion  $D$  propre à  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  :

$$\xi_A^{\text{QSO} \times \text{Ly}\alpha, \text{mod}} = \sum_B D_{AB} \left( \xi_B^{\text{QSO} \times \text{Ly}\alpha} + \xi_B^{TP} + \xi_B^{rel} + \xi_B^{\text{QSO} \times \text{metals}} \right). \quad (7.7)$$

$\xi^{\text{QSO} \times \text{Ly}\alpha}$  est la contribution effectivement due aux corrélations  $\text{QSO} \times \text{Ly}\alpha$ . C'est la transformée de Fourier du spectre de puissance quasi-linéaire de la matière (*cf.* section 6.1.1) corrigé par :

- Les facteurs de Kaiser pour les absorptions  $\text{Ly}\alpha$  et les quasars,
- Un terme qui prend en compte les effets non-linéaires,
- La fonction d'échantillonnage,  $G(\vec{k})$ , définie à la sous-section 6.1.4.

Les quasars émettent un jet ionisant qui change la proportion d'hydrogène neutre dans le milieu intergalactique ([Font-Ribera et al. 2013](#)). Cet effet de proximité du quasar, connu sous le nom de *transverse proximity effect* en anglais ([Bajtlik et al. 1988](#)), est modélisé par  $\xi^{TP}$  et dépend de la distance de séparation  $r$  entre le quasar et l'absorbant  $\text{Ly}\alpha$  :

$$\xi^{TP} \propto \xi_0^{TP} \frac{\exp(-r)}{r^2}. \quad (7.8)$$

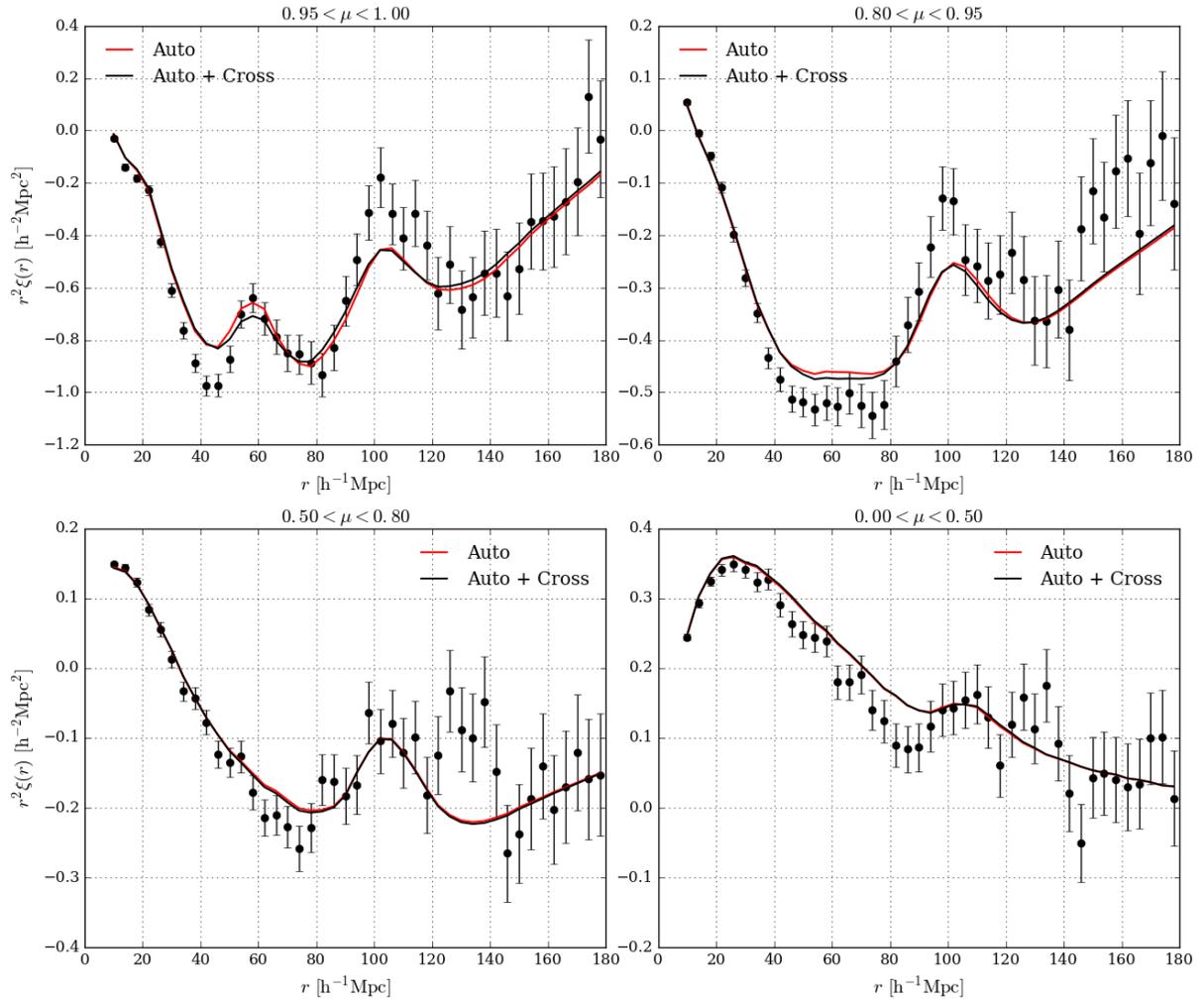


FIGURE 7.8: Moyenne pondérée des fonctions de corrélation mesurées  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  et  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (points noirs). Je superpose aux données le modèle, dans le cas de l'ajustement standard simple (rouge), et combiné avec la corrélation croisée (noir).

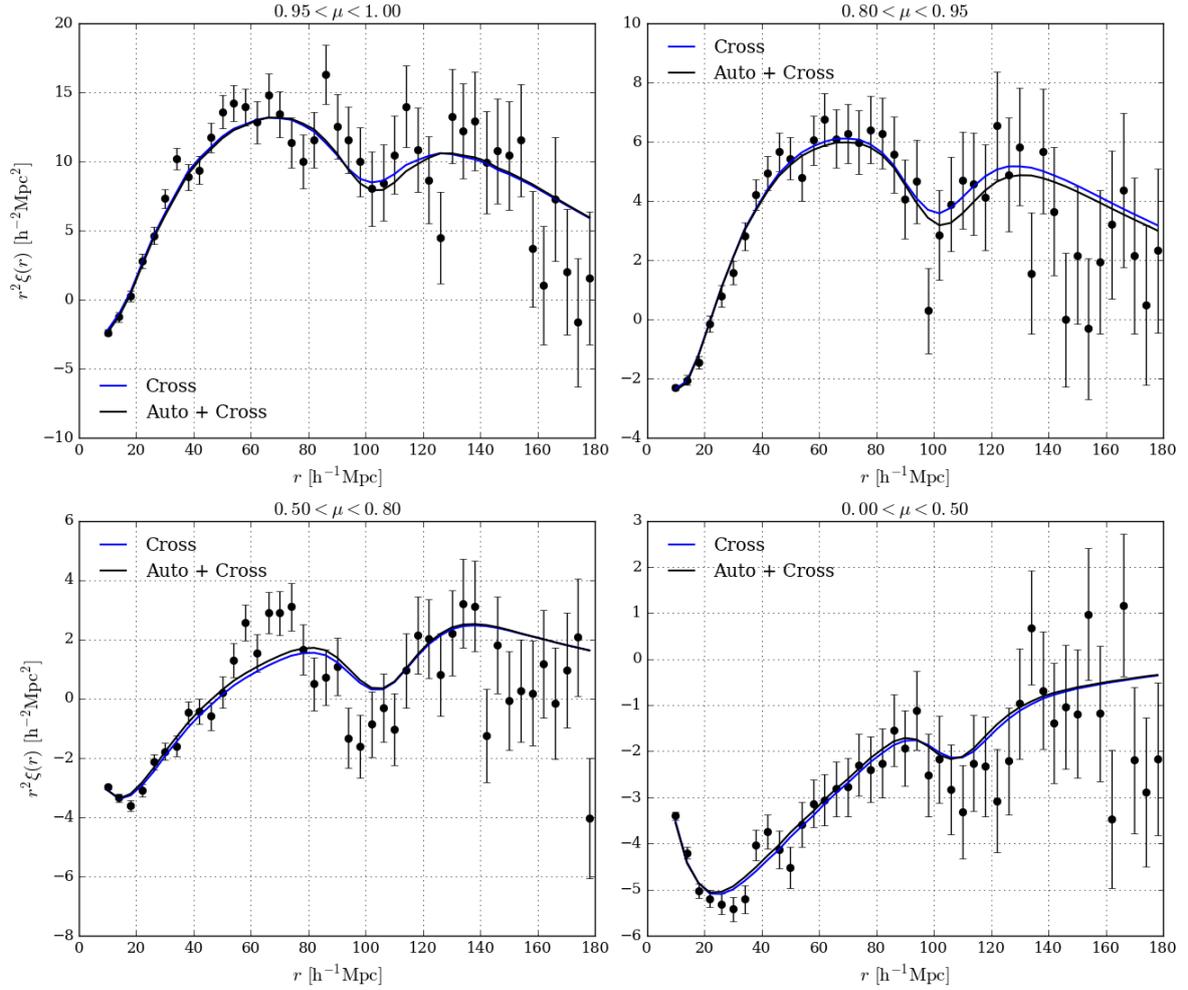


FIGURE 7.9: Fonction de corrélation croisée QSO  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ) mesurée par [Blomqvist et al. \(2019\)](#) (points noirs) en fonction de la distance de séparation  $r$  et de  $\mu = \frac{r_{\parallel}}{r}$ . Je superpose aux données le modèle dans le cas de l'ajustement standard simple (bleu) et combiné avec l'auto-corrélation (noir).

TABLE 7.5: Paramètres ajustés dans le cas de l’auto-corrélation (deuxième colonne), de la corrélation croisée (troisième colonne) et de la combinaison des deux (quatrième colonne). Les erreurs des paramètres  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont ré-ajustées par les simulations Monte-Carlo (voir table 7.4). Pour les autres paramètres, elles correspondent à  $\Delta\chi^2 = 1$ .

Paramètres	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$	QSO $\times$ $\text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$ QSO $\times$ $\text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$
$\alpha_{\parallel}$	$1.033 \pm 0.034$	$1.076 \pm 0.042$	$1.049 \pm 0.026$
$\alpha_{\perp}$	$0.953 \pm 0.048$	$0.923 \pm 0.046$	$0.942 \pm 0.031$
$\beta_{\text{Ly}\alpha}$	$1.933 \pm 0.101$	$2.28 \pm 0.31$	$1.994 \pm 0.099$
$b_{\text{Ly}\alpha}(\bar{z})$	$-0.106 \pm 0.033$	$-0.113 \pm 0.062$	$-0.104 \pm 0.032$
$\beta_{\text{QSO}}$	-	0.257	$0.209 \pm 0.006$
$\beta_{\text{HCD}}$	$1.031 \pm 0.153$	$0.500 \pm 0.200$	$0.972 \pm 0.150$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)}$	$-0.051 \pm 0.004$	-	$-0.052 \pm 0.004$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)}$	$-0.072 \pm 0.005$	-	$-0.073 \pm 0.005$
$b_{\text{HCD}}^{\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)}$	-	$-0.000 \pm 0.004$	$-0.000 \pm 0.004$
$b_{\text{SiII}(1190)}$	$-0.0050 \pm 0.0010$	$-0.0057 \pm 0.0024$	$-0.0043 \pm 0.0009$
$b_{\text{SiII}(1193)}$	$-0.0046 \pm 0.0010$	$-0.0015 \pm 0.0024$	$-0.0034 \pm 0.0009$
$b_{\text{SiIII}(1207)}$	$-0.0080 \pm 0.0010$	$-0.0117 \pm 0.0024$	$-0.0083 \pm 0.0009$
$b_{\text{SiIII}(1260)}$	$-0.0022 \pm 0.0013$	$-0.0022 \pm 0.0017$	$-0.0019 \pm 0.0009$
$b_{\text{CIV}(\text{eff})}$	$-0.0163 \pm 0.0089$	-	$-0.0167 \pm 0.0090$
$\sigma_{\nu}[h^{-1}\text{Mpc}]$	-	$7.60 \pm 0.61$	$7.053 \pm 0.357$
$\Delta r_{\parallel}[h^{-1}\text{Mpc}]$	-	$-0.22 \pm 0.32$	$-0.169 \pm 0.284$
$\xi_0^{TP}$	-	$0.276 \pm 0.158$	$0.478 \pm 0.112$
$A_{\text{rel1}}$	-	$-13.5 \pm 5.8$	$-13.573 \pm 4.721$
$\chi_{\text{min}}^2$	3258.91	3231.61	6499.31
DDL	3180-12	3180-14	6360-18
Probabilité	0.13	0.20	0.08
$\chi^2(\alpha_{\parallel} = \alpha_{\perp} = 1)$	3260.54	3235.79	6504.30

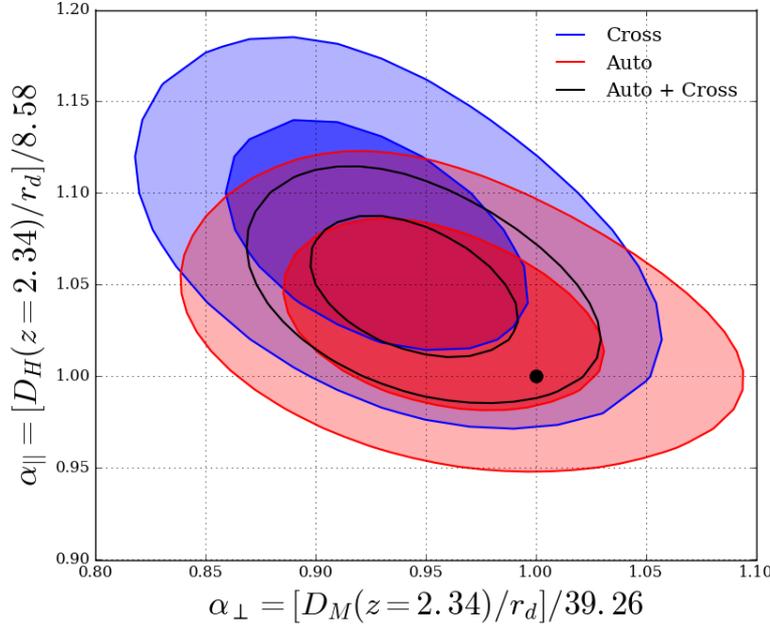


FIGURE 7.10: Niveaux de confiance à 68% et 95% des paramètres  $(\alpha_{\parallel}, \alpha_{\perp})$  dans le cas de l’ajustement de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha) + \text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (rouge),  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  (bleu) et de la combinaison des deux (noir). Le point noir représente le modèle fiduciel donné par [Planck Collaboration et al. \(2016\)](#).

Le paramètre libre  $\xi_0^{TP}$  permet d’ajuster l’amplitude de  $\xi^{TP}$ .

Dans le cas d’un traceur de matière discret comme les quasars ou les galaxies, des effets relativistes peuvent perturber la fonction de corrélation à deux points ([Bonvin 2014](#)). Ils sont modélisés par la composante  $\xi^{rel}$  qui fait intervenir un paramètre libre.

$\xi^{\text{QSO} \times \text{metals}}$  modélise la contamination due aux métaux du milieu intergalactique.

Les erreurs systématiques de l’estimation du *redshift* des quasars créent un décalage de la fonction de corrélation le long de la ligne de visée. On introduit le paramètre libre  $\Delta r_{\parallel}$  qui permet de corriger cet effet via la transformation :

$$\xi^{\text{QSO} \times \text{Ly}\alpha, \text{mod}}(r_{\parallel}^A, r_{\perp}^A) \longrightarrow \xi^{\text{QSO} \times \text{Ly}\alpha, \text{mod}}(r_{\parallel}^A + \Delta r_{\parallel}, r_{\perp}^A) \quad (7.9)$$

Finalement, le modèle standard de la fonction de corrélation croisée  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  possède six paramètres libres supplémentaires.

### 7.1.3.2 Résultats

Dans cette sous section, je combine  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ ,  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  et  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  pour améliorer la précision de la mesure de la position du pic BAO. Le *red-*

*shift* effectif de la mesure est le *redshift* moyen  $\bar{z}$  des paires des trois fonctions de corrélation. Il est toujours égal à 2.34.

Comme précédemment, je réalise 10 000 simulations de la fonction de corrélation croisée en utilisant  $\xi^{mod} \simeq \sum_B D_{AB} \xi_B^{QSO \times Ly\alpha}$ . Pour chacune d'elles, je fais l'ajustement simple de  $QSO \times Ly\alpha(Ly\alpha + Ly\beta)$  et l'ajustement combiné avec  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha)$  et  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\beta)$ . Dans la suite du manuscrit, je noterai  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$ , l'ajustement combiné de  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha)$  et  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\beta)$ . Les niveaux de confiance obtenus sont répertoriés dans le tableau 7.4.

Je présente dans le tableau 7.5 les résultats de l'ajustement de l'auto-corrélation  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$  (deuxième colonne), de la corrélation croisée (troisième colonne) et de la combinaison des deux (quatrième colonne). Les barres d'erreur sur la mesure de  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  sont corrigées grâce au tableau 7.4. On remarque que les paramètres de position du pic ajustés indépendamment sur  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$  et sur  $QSO \times Ly\alpha(Ly\alpha + Ly\beta)$  sont décalés, mais qu'ils sont compatibles avec l'ajustement combiné. Cette petite tension sur la mesure de  $\alpha_{\parallel}$  et  $\alpha_{\perp}$  s'illustre par une probabilité pour l'ajustement combiné de 8%, plus faible que pour les ajustements simples. Les paramètres propres à la corrélation croisée sont peu influencés par l'ajustement combiné, et les autres paramètres (à l'exception des biais des métaux) sont gouvernés par la mesure de la fonction d'auto-corrélation.

La figure 7.8 présente la fonction d'auto-corrélation pondérée  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$ <sup>4</sup> ainsi que les modèles pour l'ajustement simple (en rouge) et combiné (en noir). On voit que les courbes sont très proches. La plus grande différence concerne le pic Si II(1193) autour de 58  $h^{-1}$ Mpc. La transition Si II(1193) est moins bien détectée dans la corrélation croisée, ce qui a pour effet de diminuer la valeur de son biais dans l'ajustement combiné.

Je représente figure 7.9, la fonction de corrélation croisée  $QSO \times Ly\alpha(Ly\alpha + Ly\beta)$  mesurée ainsi que les modèles pour l'ajustement simple (en bleu) et combiné (en noir). Les ajustements sont similaires mais on observe que la position du pic BAO est légèrement décalée d'un ajustement à l'autre près de la ligne de visée ( $\mu > 0.8$ ). Cela correspond à un décalage du paramètre  $\alpha_{\parallel}$ , lié aux RSD. En effet, on remarque dans la table 7.5 que le paramètre de RSD est plus grand pour les absorptions  $Ly\alpha$  que pour les quasars, indiquant que la fonction de corrélation  $Ly\alpha$  est dominée par les effets des vitesses particulières. L'amplification par les RSD du signal le long de la ligne de visée est donc supérieure pour  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$  que pour  $QSO \times Ly\alpha(Ly\alpha + Ly\beta)$ , d'où une meilleure détection du paramètre  $\alpha_{\parallel}$ . On remarque également que l'amplitude du pic BAO semble avoir varié le long de la ligne de visée. Cette évolution est due à la diminution du paramètre  $\beta_{Ly\alpha}$  dans l'ajustement combiné par rapport à l'ajustement simple de  $QSO \times Ly\alpha(Ly\alpha + Ly\beta)$ .

Finalement, les paramètres de position du pic lors de l'ajustement combiné sont donnés, avec une précision à un et deux écart-types, par :

$$\begin{cases} \alpha_{\parallel} = 1.049 \begin{matrix} +0.026 & +0.052 \\ -0.025 & -0.051 \end{matrix} \\ \alpha_{\perp} = 0.942 \begin{matrix} +0.032 & +0.067 \\ -0.030 & -0.059 \end{matrix} \end{cases} \quad (7.10)$$

<sup>4</sup>La fonction de corrélation pondérée  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha + Ly\beta)$  est définie par la somme des fonctions de corrélations  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha)$  et  $Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\beta)$ , pondérée par leurs poids respectifs pour chaque *bin*.

La précision sur les mesures des paramètres de position du pic BAO est améliorée par l'ajout de la corrélation croisée :  $\alpha_{\parallel}$  est alors mesuré à 2.5% et  $\alpha_{\perp}$  à 3.4%. On remarque que ces paramètres se sont éloignés de la cosmologie fiducielle par rapport à l'ajustement de l'équation (7.4). Néanmoins, cette évolution n'est pas significative puisqu'elle reste inférieure à la précision de la mesure.

La figure 7.10 représente les contours à 1 et 2  $\sigma$  des paramètres mesurés ( $\alpha_{\parallel}, \alpha_{\perp}$ ) dans le cas de l'ajustement de l'auto-corrélation seule (en rouge), de la corrélation croisée QSO  $\times$  Ly $\alpha$  seule (en bleu) et de la combinaison des deux (en noir). On voit que la corrélation croisée est décalée par rapport à l'auto-corrélation, cet effet avait déjà été observé dans l'analyse précédente de [du Mas des Bourboux et al. \(2017\)](#). Néanmoins, compte-tenu des précisions des mesures, les deux résultats sont largement compatibles. L'ajustement combiné réduit la largeur des contours et est situé à 1.7 $\sigma$  des résultats de Panck ([Planck Collaboration et al. 2016](#)).

En utilisant les valeurs fiducielles de l'annexe A, les paramètres de position du pic BAO peuvent s'écrire :

$$\begin{cases} D_H(2.34)/r_d = 9.00^{+0.22}_{-0.22} {}^{+0.45}_{-0.43} \\ D_M(2.34)/r_d = 36.98^{+1.26}_{-1.18} {}^{+2.63}_{-2.32} \end{cases} \quad (7.11)$$

J'utilise ces résultats dans la section suivante pour contraindre les paramètres cosmologiques  $\Omega_m$  et  $\Omega_{\Lambda}$ .

## 7.2 Contraintes cosmologiques

Dans cette section, j'étudie les rapports  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  mesurés par la position du pic BAO à différents *redshifts*. Puis j'utilise l'ensemble de ces résultats pour contraindre les densité de matière et d'énergie noire ( $\Omega_m, \Omega_{\Lambda}$ ).

### 7.2.1 Contraintes sur les paramètres ( $D_H(z)/r_d, D_M(z)/r_d, D_V(z)/r_d$ )

Les rapports  $D_H/r_d$  et  $D_M/r_d$  sont donnés par la mesure de la position du pic BAO respectivement parallèlement et perpendiculairement à la ligne et de visée (équation 6.25). Quand le volume de données est insuffisant pour mesurer avec précision  $D_H/r_d$  et  $D_M/r_d$ , on utilise  $D_V$ , qui est une combinaison de  $D_M$  et  $D_H$ , optimisée pour minimiser la variance. Elle est définie par [Beutler et al. \(2011\)](#) :

$$D_V(z) = (zD_H(z))^{1/3} D_M(z)^{2/3}. \quad (7.12)$$

Dans la table 7.6, je répertorie des mesures de  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  réalisées grâce à la mesure de la position du pic par les expériences 6dFGS ([Beutler et al. 2011](#)), BOSS ([Ross et al. 2015](#), [Alam et al. 2017](#)) et eBOSS ([Bautista et al. 2018](#), [Ata et al. 2018](#), [de Sainte Agathe et al. 2019](#), [Blomqvist et al. 2019](#)).

Pour  $z \leq 0.72$ , les galaxies sont utilisées comme traceurs de masse. Elles ont pour avantage d'avoir une densité importante et permettent donc d'accéder à un volume de données très large. Cependant, elles sont de moins en moins visibles à mesure que le *redshift* augmente. Dans le cas du relevé eBOSS, la densité de galaxies observées pour les *redshifts* supérieurs à 1 est quasi nulle ([Bautista et al. 2018](#)).

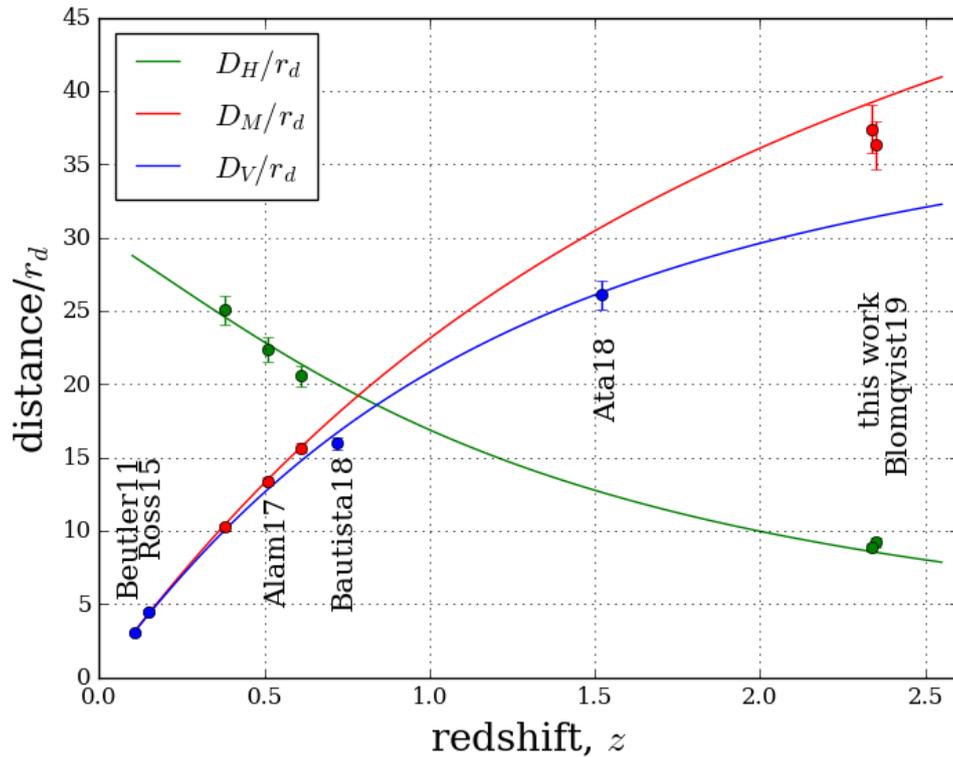


FIGURE 7.11: Mesure de  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  en utilisant la mesure de la position du pic BAO dans les fonctions de corrélation de plusieurs traceurs de la matière. Les mesures aux *redshifts* inférieurs à 1 sont obtenues grâce aux galaxies (Beutler et al. 2011, Ross et al. 2015, Alam et al. 2017, Bautista et al. 2018). À plus haut *redshift*, Ata et al. (2018) utilisent les quasars. L'utilisation des absorptions Ly $\alpha$  permet d'accéder à des *redshifts* supérieurs à 2 (de Sainte Agathe et al. 2019, Blomqvist et al. 2019). Quand le volume de données n'est pas suffisant pour mesurer simultanément  $D_H/r_d$  et  $D_M/r_d$ , on mesure  $D_V = (zD_H)^{1/3}D_M^{2/3}$ . Les courbes désignent le modèle  $\Lambda$ CDM plat avec les paramètres de la collaboration Planck (Planck Collaboration et al. 2016). Source : de Sainte Agathe et al. (2019)

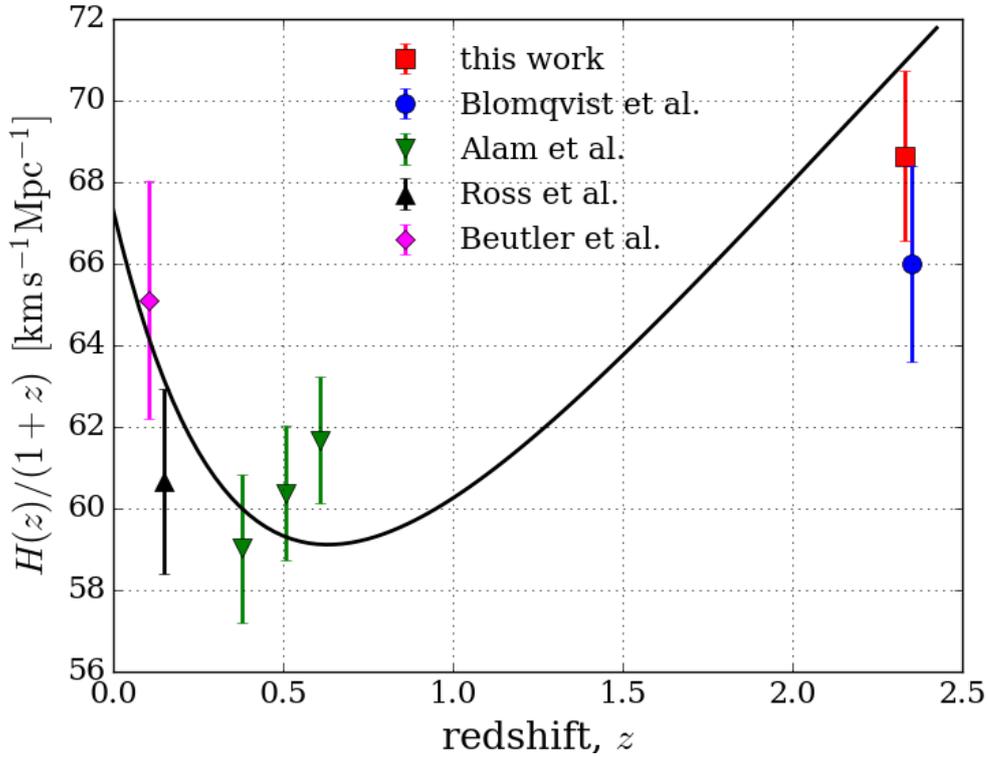


FIGURE 7.12: Mesure du taux d'expansion,  $H(z)/(1+z)$ , à partir des mesures BAO. On prend  $r_d = 147$  Mpc (Planck Collaboration et al. 2016). Le carré rouge montre le résultat de l'ajustement simple de la fonction d'auto-corrélation présenté dans cette thèse et publié dans de Sainte Agathe et al. (2019). Le rond bleu concerne la corrélation croisée QSO  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ) (Blomqvist et al. 2019) et les autres points sont mesurés grâce aux corrélations de galaxies (Beutler et al. 2011, Ross et al. 2015, Alam et al. 2017). Les points aux *redshifts* 0.106 (Beutler et al. 2011) et 0.15 (Ross et al. 2015) sont convertis en utilisant les mesures de  $q_0$  de Betoule et al. (2014). Le modèle fiduciel (annexe A) est représenté par la ligne noire.

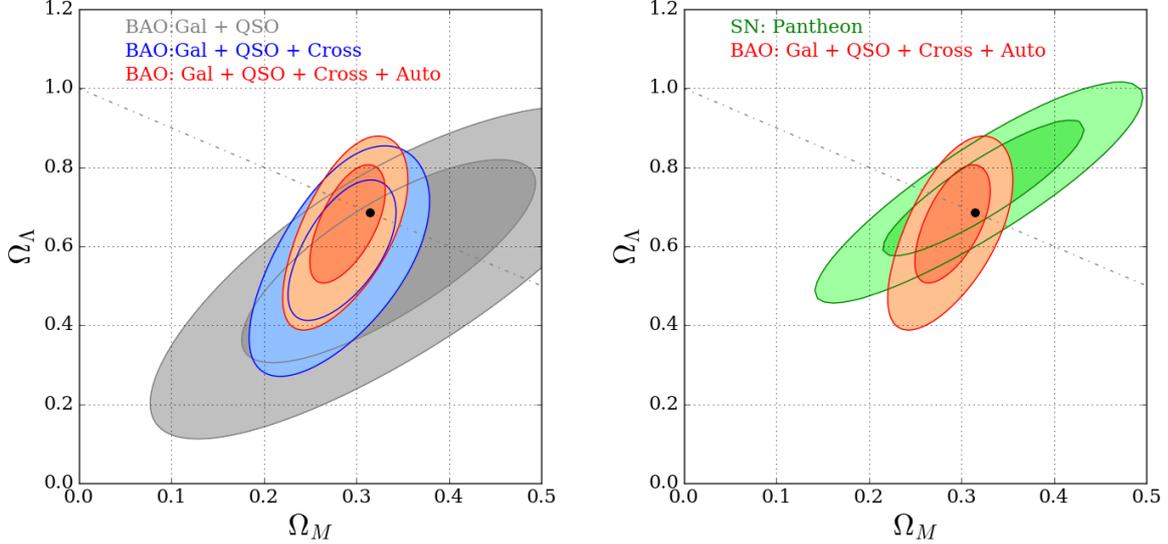


FIGURE 7.13: Contours en  $(\Omega_m, \Omega_\Lambda)$  pour les niveaux de confiance à 68% ( $\Delta\chi^2 = \chi^2(\Omega_m, \Omega_\Lambda) - \chi_{\min}^2 = 2.3$ ) et 95% ( $\Delta\chi^2 = 6.0$ ). Sur la figure de gauche, je montre l'amélioration des contraintes lorsqu'on utilise les mesures BAO, réalisées au cours de ma thèse, avec les absorptions Ly $\alpha$ . En gris, on utilise seulement les résultats BAO issus des galaxies et des quasars. En bleu, on ajoute la corrélation QSO  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ), puis Ly $\alpha$  (Ly $\alpha$ )  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ), en rouge. Sur la figure de droite, je compare les contraintes obtenues avec la sonde BAO (en rouge) ou les supernovae de type Ia mesurées par Scolnic et al. (2018) (en vert).

TABLE 7.6: Rapports  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  obtenus en mesurant la position du pic BAO dans les fonctions de corrélation de différents traceurs de la matière. La première série de mesures est réalisée avec des traceurs de masse discrets (galaxies ou quasars). La deuxième série concerne les absorptions Ly $\alpha$  discutées dans cette thèse. L'auto-corrélation Ly $\alpha$  (Ly $\alpha$ )  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ) est présentée dans de Sainte Agathe et al. (2019), la corrélation croisée QSO  $\times$  Ly $\alpha$  (Ly $\alpha$  + Ly $\beta$ ) dans Blomqvist et al. (2019) et les résultats de l'ajustement combiné des deux est renseigné dans la dernière ligne.

	<i>redshift</i>	Traceurs	$D_H/r_d$ [Mpc]	$D_M/r_d$ [Mpc]	$D_V/r_d$ [Mpc]
Beutler et al. (2011)	0.106	Galaxies	-	-	$3.047 \pm 0.137$
Ross et al. (2015)	0.15	Galaxies	-	-	$4.466 \pm 0.168$
Alam et al. (2017)	0.38	Galaxies	$24.89 \pm 0.58$	$10.27 \pm 0.15$	-
	0.51	Galaxies	$22.43 \pm 0.48$	$13.38 \pm 0.18$	-
	0.61	Galaxies	$20.86 \pm 0.45$	$15.45 \pm 0.22$	-
Bautista et al. (2018)	0.72	Galaxies	-	-	$23.73 \pm 0.62$
Ata et al. (2018)	1.52	Quasars	-	-	$26.08 \pm 1.00$
de Sainte Agathe et al. (2019)	2.34	Absorptions Ly $\alpha$	$8.86 \pm 0.29$	$37.41 \pm 1.86$	-
Blomqvist et al. (2019)	2.35	Absorptions Ly $\alpha$ $\times$ Quasars	$9.20 \pm 0.36$	$36.3 \pm 1.8$	-
Combiné	2.34	Absorptions Ly $\alpha$ + Quasars	$9.00 \pm 0.22$	$36.98 \pm 1.22$	-

Au delà, on peut utiliser des quasars comme traceurs car ceux-ci sont extrêmement lumineux. Grâce à la mesure de leur spectre de puissance, [Ata et al. \(2018\)](#) détecte le pic BAO au *redshift* intermédiaire  $z = 1.52$ . Néanmoins, le nombre de quasars observés n'est pas suffisant pour mesurer  $D_H/r_d$  et  $D_M/r_d$  indépendamment.

Plus le *redshift* augmente, plus la densité de quasars observés diminue ([Pâris et al. 2018](#)). Mais en utilisant les quelques centaines d'absorptions Ly $\alpha$  visibles sur les spectres des quasars, on peut détecter le pic BAO à haut *redshift* malgré la densité plus faible. Comme on l'a vu dans la section précédente, on peut mesurer  $D_H/r_d$  et  $D_M/r_d$  grâce aux corrélations Ly $\alpha \times$  Ly $\alpha$  et QSO  $\times$  Ly $\alpha$  ([Blomqvist et al. 2019](#)).

L'ensemble de ces mesures est visible sur la figure 7.11 où les courbes représentent le modèle  $\Lambda$ CDM plat appliqué aux mesures de CMB ([Planck Collaboration et al. 2016](#)).

Puisque  $D_H(z) = c/H(z)$ , les mesures d'[Alam et al. \(2017\)](#), de [Sainte Agathe et al. \(2019\)](#) et [Blomqvist et al. \(2019\)](#) nous permettent d'accéder directement au taux d'expansion aux *redshifts* correspondants. Pour les *redshifts* faibles ([Beutler et al. 2011](#), [Ross et al. 2015](#)), on peut faire un développement limité et déduire  $H(z)$  de  $D_V(z)/r_d$  et du paramètre de décélération  $q_0$  ([Weinberg 1972](#)) mesuré grâce aux supernovae de type Ia ([Betoule et al. 2014](#)). On utilise  $q_0 = \frac{1}{2}\Omega_m - \Omega_\Lambda = 0.542$ . Je représente les valeurs des taux d'expansion obtenus par des mesures de BAO, sur la figure 7.12, à comparer avec le modèle fiduciel de l'annexe A, donné par la ligne noire.

Depuis la première détection du pic BAO avec les absorptions Ly $\alpha$  ([Busca et al. 2013](#)), la mesure du taux d'expansion est passée de  $H(z = 2.3) = 219 \pm 24$  km/s/Mpc à  $H(z = 2.34) = 230 \pm 7$  km/s/Mpc (équation 7.6) et la précision de la mesure a été considérablement améliorée. La figure 7.12 montre que les absorptions Ly $\alpha$  sont les seuls traceurs capables de mesurer la position du pic BAO juste avant l'accélération de l'expansion. En augmentant le nombre de données, les relevés à venir (*e.g.*, DESI) devraient permettre de réaliser des mesures tomographiques du taux d'expansion pour  $z > 2.3$  et ainsi, d'avoir davantage de contraintes lors de la phase de décélération.

Dans la sous-section suivante, on utilise l'ensemble des mesures de  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  pour contraindre les paramètres cosmologiques  $(\Omega_m, \Omega_\Lambda)$ .

## 7.2.2 Contraintes sur les paramètres $(\Omega_m, \Omega_\Lambda)$

D'après la section 1.3, pour calculer les valeurs fiducielles de  $D_H(\bar{z})$ ,  $D_M(\bar{z})$  et  $D_V(\bar{z})$  on n'a besoin que du taux d'expansion  $H(z)$  pour  $z$  variant entre 0 et  $\bar{z}$ . Or, dans les gammes de *redshift* de la table 7.6,  $z$  est suffisamment petit pour négliger le terme de rayonnement (*cf.* figure 1.5). Ainsi, d'après la première équation de Friedmann-Lemaître,  $H(z)$  s'écrit :

$$H(z) \simeq H_0 \sqrt{\Omega_m(1+z)^3 + (1 - \Omega_m - \Omega_\Lambda)(1+z)^2 + \Omega_\Lambda}. \quad (7.13)$$

Par conséquent, les modèles de  $D_H/r_d$ ,  $D_M/r_d$  et  $D_V/r_d$  ne dépendent que des paramètres  $(\Omega_m, \Omega_\Lambda, H_0 r_d)$ .

Pour déterminer le couple  $(\Omega_m, \Omega_\Lambda)$  qui explique au mieux les données, je vais minimiser, en marginalisant analytiquement sur  $H_0 r_d$ , la fonction de vraisemblance suivante :

$$\chi^2(\Omega_m, \Omega_\Lambda, H_0 r_d) = \left( D - M(\Omega_m, \Omega_\Lambda, H_0 r_d) \right)^T C^{-1} \left( D - M(\Omega_m, \Omega_\Lambda, H_0 r_d) \right) \quad (7.14)$$

avec :

- $D$  le vecteur qui contient les données  $\frac{D_M}{r_d}$ ,  $\frac{D_H}{r_d}$  et  $\frac{D_V}{r_d}$  du tableau 7.6,
- $M(\Omega_m, \Omega_\Lambda, H_0 r_d)$  le vecteur dont les éléments sont les distances cosmologiques  $\left( \frac{D_M}{r_d} \right)_{fid}$ ,  $\left( \frac{D_H}{r_d} \right)_{fid}$  et  $\left( \frac{D_V}{r_d} \right)_{fid}$  calculées aux *redshifts* du tableau 7.6,
- $C^{-1}$  est une matrice diagonale par bloc composée des matrices de covariance inverses des mesures.

On peut facilement montrer que  $M(\Omega_m, \Omega_\Lambda, H_0 r_d)$  est inversement proportionnel à  $H_0 r_d$  et peut s'écrire sous la forme :

$$M(\Omega_m, \Omega_\Lambda, H_0 r_d) = \frac{m(\Omega_m, \Omega_\Lambda)}{H_0 r_d}. \quad (7.15)$$

Finalement, à  $(\Omega_m, \Omega_\Lambda)$  fixés, le minimum (*best fit* en anglais) sur  $H_0 r_d$  de la fonction de vraisemblance, noté  $\left( H_0 r_d \right)^{BF}$ , est obtenu pour :

$$\left( H_0 r_d \right)^{BF} = \frac{m(\Omega_m, \Omega_\Lambda)^T C^{-1} m(\Omega_m, \Omega_\Lambda)}{m(\Omega_m, \Omega_\Lambda)^T C^{-1} D}. \quad (7.16)$$

Je représente sur la figure 7.13 les contours à un et deux écart-types de  $\chi^2(\Omega_m, \Omega_\Lambda, \left( H_0 r_d \right)^{BF})$  en fonction de  $\Omega_m$  et  $\Omega_\Lambda$ . En gris, j'utilise la première série de mesures du tableau 7.6 et la fonction d'auto-corrélation, en rouge, la première série de mesures et l'ajustement combiné de  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  et  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$ . Je compare ces contours à ceux obtenus par Scolnic et al. (2018) grâce aux mesures de supernovae de type Ia (en vert). Les deux sondes donnent des résultats compatibles. Le modèle fiduciel de l'annexe A est désigné par le point noir et la ligne en pointillés représente le cas d'un univers plat, c'est-à-dire quand  $\Omega_\kappa = 0$ .

On déduit des contours rouges le couple  $(\Omega_m, \Omega_\Lambda)$  qui minimise  $\chi^2$  :

$$\Omega_m = 0.293 \pm 0.027 \quad \Omega_\Lambda = 0.675 \pm 0.099 \quad (7.17)$$

qui correspond à  $\Omega_\kappa = 0.032 \pm 0.117$ , compatible avec un univers plat. Dans ce cas, le minimum sur  $H_0 r_d$  de la fonction de vraisemblance est donné par :

$$\left( H_0 r_d \right)^{BF} = (68.3 \pm 1.3) \times 147.33 \text{ Mpc}. \quad (7.18)$$

Dans le cas où  $\Omega_\kappa \neq 0$ , on compare les paramètres  $\Omega_m$  et  $\Omega_\Lambda$  mesuré de l'équation 7.17 avec ceux obtenus par Scolnic et al. (2018) grâce à l'utilisation des SNIa, qui donnent une moins bonne précision :

$$\Omega_m = 0.319 \pm 0.070 \qquad \Omega_\Lambda = 0.733 \pm 0.113 \qquad (7.19)$$

Le modèle favorisé par les mesures de CMB (Planck Collaboration et al. 2016) a  $\chi^2 = 13.76$  pour 12 degrés de liberté et diffère de l'ajustement équation 7.17 de moins d'un écart-type, comme indiqué sur la figure 7.13. On en déduit que les mesures BAO sont une confirmation, pour les *redshift* entre 0.1 et 2.4, du modèle  $\Lambda$ CDM favorisé par les mesures du CMB.

La mesure de l'empreinte des BAO dans la distribution de matière est une méthode récente, dont la première détection a été réalisée par Eisenstein et al. (2005), en mesurant la fonction de corrélation des galaxies de type LRG. Depuis, l'utilisation de nouveaux traceurs (galaxies de type ELG, quasars, absorptions Ly $\alpha$ ) a permis d'obtenir des mesures des rapports  $D_M(z)/r_d$  et  $D_H(z)/r_d$  pour  $z \in [0.11, 2.25]$  et donc de contraindre les paramètres  $\Omega_m$  et  $\Omega_\Lambda$  indépendamment des SNIa ou du CMB. Actuellement, les mesure de position du pic BAO apportent les contraintes les plus fortes, à bas *redshift*, sur ces paramètres.



# Conclusion

L'utilisation des absorptions Ly $\alpha$  pour la mesure de l'échelle acoustique est une technique récente mais robuste. Elle permet d'accéder à des *redshifts* beaucoup plus grands que pour les galaxies ou les quasars, et ainsi d'obtenir des informations sur la période qui précède l'accélération de l'expansion de l'univers.

Dans cette thèse, j'ai présenté l'analyse des fonctions de corrélation à deux points des absorptions Ly $\alpha$ , mesurées en utilisant le relevé DR14 de SDSS-IV. En combinant ces résultats avec la corrélation croisée entre les absorptions Ly $\alpha$  et les quasars, j'ai obtenu les précisions 2.5% et 3.4% sur la mesure des rapports  $D_H(z)/r_d$  et  $D_M(z)/r_d$  au *redshift*  $z = 2.34$ . La principale amélioration de mon analyse a été la prise en compte de la région Ly $\beta$ , entre les pics d'émission Ly $\gamma$  et Ly $\beta$ . L'ajout de cette nouvelle région a confirmé la robustesse de la méthode en montrant que la mesure de la position du pic BAO n'était pas influencée par la région spectrale. Ce qui ouvre la voie à une utilisation systématique de la région Ly $\beta$  dans les analyses à venir, notamment pour les résultats finaux d'eBOSS (du Mas des Bourboux, in prep) et pour DESI (DESI Collaboration et al. 2016). En effet, l'extension du domaine spectral étudié permet d'exploiter des données déjà enregistrées par l'instrument et donc d'augmenter la précision statistique sans augmenter le temps d'observation. De plus, les absorptions Ly $\alpha$  de la région Ly $\beta$  sont créées par des absorbants plus lointains que ceux de la région Ly $\alpha$ , permettant ainsi d'accroître les données à plus haut *redshift*. Dans cette analyse, le nombre des spectres où la forêt Ly $\beta$  était visible n'était pas assez grand pour pouvoir détecter le pic BAO dans la fonction de corrélation Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\beta$ (Ly $\beta$ ). Mais le formalisme et les codes du module `picca` développés au cours de cette thèse permettront, une fois que le nombre de spectres sera suffisant, de mesurer la position du pic BAO dans la fonction de corrélation Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\beta$ (Ly $\beta$ ) et de réaliser des ajustements combinés avec d'autres fonctions de corrélation.

La mesure de l'échelle acoustique dans la distribution de la matière permet, en utilisant des traceurs variés, d'obtenir les rapports  $D_H(z)/r_d$  et  $D_M(z)/r_d$ , pour  $z$  compris entre 0.11 et 2.34 environ, d'où on peut déduire des contraintes sur les paramètres cosmologiques  $\Omega_m$  et  $\Omega_\Lambda$ . En combinant les résultats de cette thèse avec ceux donnés par les autres traceurs de la matière, j'ai obtenu les contraintes actuelles les plus fortes à haut *redshift* sur ces paramètres. Ces résultats devraient encore s'améliorer avec la publication finale d'eBOSS et l'expérience DESI à venir. Comme indiqué par les projections de la figure 7.14, les 30 millions de spectres de galaxies et de quasars qui seront enregistrés par l'instrument DESI au cours de la décennie à venir (DESI Collaboration et al. 2016), permettront de réaliser une mesure tomographique du taux d'expansion de l'univers entre les *redshifts* 0 et 3.5 environ, ce qui devrait permettre d'améliorer considérablement les contraintes sur le modèle d'énergie noire.

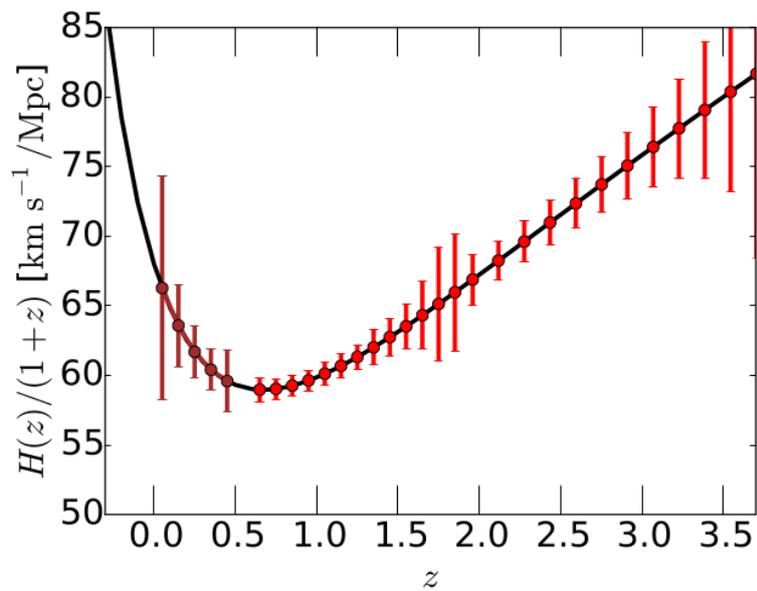


FIGURE 7.14: Projections des mesures par DESI du taux d’expansion de l’univers en fonction du *redshift*. L’utilisation des forêts Ly $\alpha$  permettra d’obtenir des mesures tomographiques de  $H(z)$  pour  $z \in [2.0, 3.5]$ . Source : [DESI Collaboration et al. \(2016\)](#)

## Annexe A

# Cosmologie fiducielle

TABLE A.1: Liste des paramètres du modèle  $\Lambda$ CDM ajustés sur les mesures du CMB réalisées par la collaboration Planck (Planck Collaboration et al. 2016, PlanckTT+lowP, table 3). Ces six paramètres sont décrits à la sous-section 1.1.3.

Paramètres	Valeurs	Description
$\Omega_b h^2$	0.02222	Mesure la densité d'énergie actuelle des baryons
$\Omega_c h^2$	0.1197	Mesure la densité d'énergie actuelle de matière noire
$100\theta_s$	1.04085	Mesure l'angle apparent du premier pic acoustique
$n_s$	0.9655	Indice spectral des perturbations primordiales scalaires
$\ln(10^{10} A_s)$	3.089	Mesure l'amplitude des perturbations primordiales scalaires
$\tau$	0.078	Épaisseur optique de réionisation

TABLE A.2: Liste des paramètres dérivés utiles.

Paramètres	Valeurs
$\Omega_m$	0.3147
$\Omega_\Lambda$	0.685
$h$	0.6731
$r_d$	147.33 Mpc (99.17 $h^{-1}$ Mpc)
$D_H(z = 2.34)/r_d$	8.581
$D_M(z = 2.34)/r_d$	39.26

La cosmologie fiducielle est donnée par le modèle  $\Lambda$ CDM plat. Les six paramètres du modèle (section 1.1) sont ajustés sur les mesures de CMB réalisées par la collaboration Planck (Planck Collaboration et al. 2016) et donnés par la table A.1. L'ensemble des autres paramètres cosmologiques peuvent être déduits de ces six paramètres. Je repertorie dans la table A.2, les paramètres qui seront utiles pour cette étude.



## Annexe B

# Transitions atomiques utiles

Dans la table B.1 je dresse la liste des transitions atomiques utilisées dans ce manuscrit. Les valeurs des transitions sont données par NIST<sup>1</sup>.

TABLE B.1: Liste des transitions atomiques utiles mesurée dans le vide.

Transitions	Notations	$\lambda$ [Å]
Lyman- $\gamma$	Ly $\gamma$	972.52
Lyman- $\beta$	Ly $\beta$	1025.73
Silicium II	Si II(1190)	1190.42
Silicium II	Si II(1193)	1193.28
Silicium III	Si III(1207)	1206.51
Lyman- $\alpha$	Ly $\alpha$	1215.67
Silicium II	Si II(1260)	1260.42
Carbone IV	C IV(1548)	1548.20
Carbone IV (effectif)	C IV(eff)	1549.06
Carbone IV	C IV(1551)	1550.77

<sup>1</sup>[https://physics.nist.gov/PhysRefData/ASD/lines\\_form.html](https://physics.nist.gov/PhysRefData/ASD/lines_form.html)



## Annexe C

# Module `picca`

Le module python "`picca`" (`Package for Igm Cosmological-Correlations Analyses`) a été développé par le groupe Ly $\alpha$  de l'expérience eBOSS pour permettre l'analyse des fonctions de corrélation entre les absorptions (identiques ou différentes) des forêts (identiques ou différentes) et de corrélation croisée entre des absorptions et des quasars. Tous les résultats de cette thèse sont reproductibles en utilisant la version "`v3.DR14`"<sup>1</sup> de ces codes appliqués aux données DR14 d'eBOSS. Les analyses présentées dans [Blomqvist et al. \(2018\)](#), [du Mas des Bourboux et al. \(2019\)](#) et [Blomqvist et al. \(2019\)](#) utilisent également ce logiciel.

Je présente dans cette annexe, la procédure à suivre pour reproduire l'analyse présentée dans cette thèse :

- calcul de  $(\delta_q(\lambda), w_q(\lambda))$  à partir des données de sortie de la chaîne d'extraction SDSS,  $(f_q(\lambda), \sigma_{f,q}(\lambda))$ , et du catalogue de quasars de [Pâris et al. \(2018\)](#) (*cf.* chapitres 4 et 3),
- calcul de la fonction de corrélation à partir de  $(\delta_q(\lambda), w_q(\lambda))$  (*cf.* chapitre 5),
- calcul de la matrice de distorsion à partir de  $(\delta_q(\lambda), w_q(\lambda))$  (*cf.* chapitre 6),
- calcul de la matrice de covariance à partir de  $(\delta_q(\lambda), w_q(\lambda))$  (*cf.* chapitre 5),
- calcul de la matrice des métaux à partir de  $(\delta_q(\lambda), w_q(\lambda))$  (*cf.* chapitre 6) et de la liste des métaux contaminants (*cf.* chapitre 5),
- ajustement du modèle sur la fonction de corrélation mesurée (*cf.* chapitre 6).

### C.1 Calcul du champ de $\delta$

Le passage du flux traité par la chaîne d'extraction SDSS aux champ de  $\delta$  est détaillé dans le chapitre 4. Concrètement, pour calculer les champs de  $\delta$  de la région Ly $\alpha$ , j'ai utilisé le code

---

<sup>1</sup><https://github.com/igmhub/picca/tree/v3.DR14>

"`do_deltas.py`" en lançant la commande :

```
do_deltas.py --drq DR14Q.fits --in-dir v5-DR14/ --mode spplate --zqso-max 3.5 -  
-lambda-rest-min 1040. --lambda-rest-max 1200. --lambda-max 5400 --mask-file  
list-veto-line.txt --dla-vac DLA-DR14.dat --bi-max 0. --out-dir deltas-lya
```

Avec :

- `DR14Q.fits` le catalogue de quasars de [Pâris et al. \(2018\)](#),
- `flux-dr14/` le répertoire qui contient les flux et les variances en sortie de la chaîne d'extraction SDSS. L'option `--mode` désigne le format de ces données,
- `--zqso-max`, `--zqso-min`, `--lambda-rest-min`, `--lambda-rest-max`, `--lambda-min`, `--lambda-max` les options qui définissent les domaines de la table [4.2](#),
- `list-veto-line.txt` la liste (optionnelle) des raies du ciel à masquer,
- `DLA-DR14.dat` le fichier (optionnel) de DLA fourni par Pasquier Noterdaeme,
- `--bi-max` l'option qui donne la valeur (optionnelle) de la coupure sur les indices de Balmer. Les spectres dont l'indice de Balmer est supérieur à cette valeur, sont exclus de l'analyse,
- `deltas-lya/` le répertoire de sortie qui contient un fichier `fits`<sup>2</sup> par plaque. Dans un fichier, une ligne regroupe (entre autres) les informations  $(\lambda, z_q, \delta_q(\lambda), w_q(\lambda))$  relatives à un spectre.

## C.2 Calcul de la fonction de corrélation

Le calcul de la fonction de corrélation Ly $\alpha$  et de sa matrice de covariance est décrit en détail dans le chapitre [5](#). Pour ce calcul, j'utilise les fichiers de sortie de l'étape précédente, auquel j'applique la commande suivante :

```
do_cf.py --out cf.fits --in-dir deltas-1/ --in-dir2 deltas-2/ --lambda-abs LYA -  
-lambda-abs2 LYA --z-evol 2.9 --z-evol2 --np 50 --nt 40 --rp-min 0
```

Avec :

- `cf.fits` le fichier de sortie en format `fits`. Dans ce fichier, chaque ligne regroupe, pour un HEALPix  $h$ , la fonction de corrélation  $\hat{\xi}_A^h$  et son poids  $W_A^h$  pour l'ensemble des *bins*  $A$ ,
- `--in-dir` et `--in-dir-2` les répertoires contenant les champs de  $\delta$  de l'étape précédente. Ils peuvent être identiques ou différents.

---

<sup>2</sup><https://heasarc.gsfc.nasa.gov/docs/heasarc/fits.html>

- `--lambda-abs` et `--lambda-abs2` les types d'absorptions à considérer. Par défaut,  $Ly\alpha$ .
- `--z-evol1` et `--z-evol2` les valeurs de  $\gamma$  utilisées pour ré-étalonner les poids (*cf.* section 4.5.2).
- `--np` et `--nt` les options désignant le nombre de *bins* respectivement parallèles et perpendiculaires à la ligne de visée,
- `--rp-min`, `--rp-max`, `--rt-min` et `--rt-max` les options contrôlant les bornes de la fonction de corrélation.

### C.3 Calcul de la matrice de distorsion

Pour calculer la matrice de distorsion (*cf.* section 6.3), j'utilise la commande :

```
do_dmat.py --out dmat.fits --in-dir deltas-1/ --in-dir2 ../deltas-2/ --lambda-abs
LYA --lambda-abs2 LYA --z-evol 2.9 --z-evol2 --np 50 --nt 40 --rp-min 0 --rej 0.95
```

Les options sont les mêmes que pour le calcul de la fonction de corrélation, à l'exception de `--rej` qui désigne la proportion des paires à rejeter. Ici, je ne considère que 5% des paires pour le calcul. En sortie, j'obtiens un fichier `dmat.fits` qui contient les  $2500 \times 2500$  entrées de la matrice de distorsion.

### C.4 Calcul de la matrice de covariance

Dans cette étape, je moyenne la fonction de corrélation sur l'ensemble des `HEALPixs` et je calcule sa fonction de covariance lissée (*cf.* section 5.3) en utilisant la commande :

```
export.py --cf cf.fits --dmat dmat.fits --out exported_cf.fits
```

Avec :

- `cf.fits` le fichier de sortie de l'étape C.2,
- `dmat.fits` le fichier de sortie de l'étape C.3,
- `exported_cf.fits` le fichier produit, qui sera utilisé par le code d'ajustement.

### C.5 Calcul de la matrice des métaux

Pour calculer la matrice des métaux (*cf.* section 6.2.2), j'utilise la commande :

```
do_metal_dmat.py --out metal-dmat.fits --in-dir deltas-1/ --in-dir2 ../deltas-2/
--lambda-abs LYA --lambda-abs2 LYA --z-evol 2.9 --z-evol2 --np 50 --nt 40 --rp-min
```

```
0 --rej 0.999 --abs-igm "SiII(1190)" "SiII(1193)" "SiIII(1207)" "SiII(1260)" "CIV(eff)"
```

Les options sont les mêmes que pour le calcul de la matrice de distorsion, à l'exception de `--abs-igm` qui donne la liste des métaux à considérer. On peut également ajouter l'option `--abs-igm2` si les contaminants sont différents dans les deux régions considérées (*e.g.*, pour  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$ ).

## C.6 L'ajustement

L'ajustement est réalisé avec le code `fitter2` du module `picca`. Le fichier d'entrée se compose :

- des fonctions de corrélations à ajuster et de leurs matrices de covariance, de distorsion, et des métaux, calculées lors des étapes précédentes,
- de la liste des paramètres fixes,
- de la liste des paramètres libres avec leurs éventuelles limites,
- de la liste des paramètres avec priors gaussiens avec leurs valeurs, écarts-types, et limites éventuelles.

Un exemple de fichier d'entrée est donné à l'adresse [https://github.com/igmhub/picca/blob/master/py/picca/test/data/config\\_cf.ini](https://github.com/igmhub/picca/blob/master/py/picca/test/data/config_cf.ini).

En sortie du code d'ajustement, j'obtiens un fichier avec toutes les valeurs ajustées et leurs erreurs gaussiennes ainsi que le modèle et la valeur du minimum de vraisemblance pour chaque *bin*.

## Annexe D

# Publication

Je reproduis ici la publication [de Sainte Agathe et al. \(2019\)](#), acceptée pour publication dans la revue *Astronomy and Astrophysics*.

## Baryon acoustic oscillations at $z = 2.34$ from the correlations of Ly $\alpha$ absorption in eBOSS DR14

Victoria de Sainte Agathe<sup>1\*</sup>, Christophe Balland<sup>1</sup>, Hélión du Mas des Bourboux<sup>2</sup>, Nicolás G. Busca<sup>1</sup>, Michael Blomqvist<sup>3</sup>, Julien Guy<sup>4</sup>, James Rich<sup>5</sup>, Andreu Font-Ribera<sup>6</sup>, Matthew M. Pieri<sup>3</sup>, Julian E. Bautista<sup>7</sup>, Kyle Dawson<sup>2</sup>, Jean-Marc Le Goff<sup>5</sup>, Axel de la Macorra<sup>8</sup>, Nathalie Palanque-Delabrouille<sup>5</sup>, Will J. Percival<sup>9</sup>, Ignasi Pérez-Ràfols<sup>3</sup>, Donald P. Schneider<sup>10,11</sup>, Anže Slosar<sup>12</sup>, and Christophe Yèche<sup>5</sup>

<sup>1</sup> Sorbonne Université, Université Paris Diderot, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies, LPNHE, 4 Place Jussieu, F-75252 Paris, France

<sup>2</sup> Department of Physics and Astronomy, University of Utah, 115 S. 1400 E., Salt Lake City, UT 84112, U.S.A.

<sup>3</sup> Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France

<sup>4</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.

<sup>5</sup> IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

<sup>6</sup> University College London, Gower St, Kings Cross, London WC1E6BT

<sup>7</sup> Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK

<sup>8</sup> Instituto de Astronomía, Universidad Nacional Autónoma de México, A.P. 70-264, 04510, México, D.F., México

<sup>9</sup> Department of Physics and Astronomy, University of Waterloo, 200 University Ave. W., Waterloo ON N2L 3G1, Canada

<sup>10</sup> Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802

<sup>11</sup> Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802

<sup>12</sup> Brookhaven National Laboratory, 2 Center Road, Upton, NY 11973, USA

Received / Accepted

### ABSTRACT

We measure the imprint of primordial baryon acoustic oscillations (BAO) in the correlation function of Ly $\alpha$  absorption in quasar spectra from the Baryon Oscillation Spectroscopic Survey (BOSS) and the extended BOSS (eBOSS) in Data Release 14 (DR14) of the Sloan Digital Sky Survey (SDSS)-IV. In addition to 179,965 spectra with absorption in the Lyman- $\alpha$  (Ly $\alpha$ ) region, we use, for the first time, Ly $\alpha$  absorption in the Lyman- $\beta$  region of 56,154 spectra. We measure the Hubble distance,  $D_H$ , and the comoving angular diameter distance,  $D_M$ , relative to the sound horizon at the drag epoch  $r_d$  at an effective redshift  $z = 2.34$ . Using a physical model of the correlation function outside the BAO peak, we find  $D_H(2.34)/r_d = 8.86 \pm 0.29$  and  $D_M(2.34)/r_d = 37.41 \pm 1.86$ , within  $1\sigma$  from the flat- $\Lambda$ CDM model consistent with CMB anisotropy measurements. With the addition of polynomial “broadband” terms, the results remain within one standard deviation of the CMB-inspired model. Combined with the quasar-Ly $\alpha$  cross-correlation measurement presented in a companion paper (Blomqvist et al. 2019), the BAO measurements at  $z = 2.35$  are within  $1.7\sigma$  of the predictions of this model.

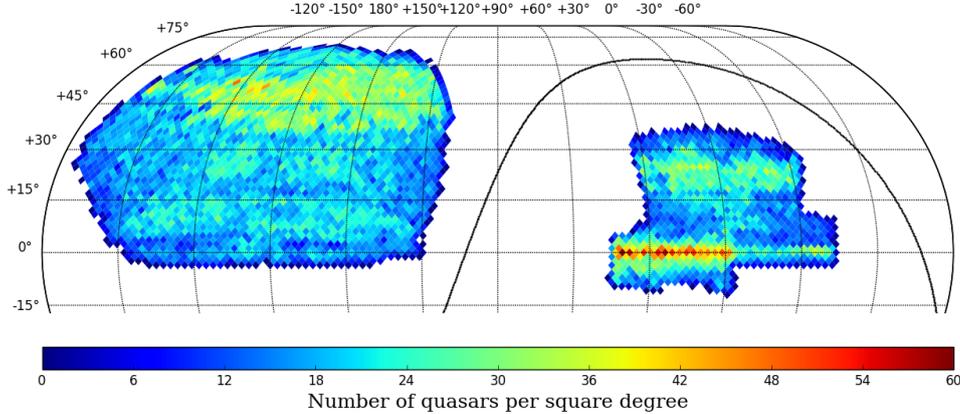
**Key words.** cosmological parameters - dark energy - cosmology : observations

### 1. Introduction

Since the first observations of the imprint of primordial baryonic acoustic oscillations (BAO) as a peak in the galaxy correlation function (Eisenstein et al. 2005) or as a periodic modulation in the corresponding power spectrum (Cole et al. 2005), the BAO signal has led to significant constraints on cosmological parameters. The BAO peak in the radial direction at a redshift  $z$  yields  $D_H(z)/r_d = c/(r_d H(z))$ , where  $H(z)$  is the Hubble parameter and  $r_d$  is the sound horizon at the drag epoch (Eisenstein & Hu 1998). The transverse measurement constrains the quantity  $D_M(z)/r_d = (1+z)D_A(z)/r_d$ , where  $D_A(z)$  is the angular diameter distance. Because of its sensitivity to both the distance and the expansion rate, the ensemble of BAO measurements yield tight constraints on  $\Lambda$ CDM parameters (Aubourg et al. 2015) even without the use of Cosmic Microwave Background (CMB) data.

\* Contact e-mail: victoria.de.sainte.agathe@lpnhe.in2p3.fr

Most BAO measurements have employed discrete objects like galaxies (Percival et al. 2010; Reid et al. 2010; Beutler et al. 2011; Blake et al. 2011; Anderson et al. 2012, 2014a,b; Ross et al. 2015; Alam et al. 2017; Bautista et al. 2018) or quasars (Ata et al. 2018; Gil-Marín et al. 2018; Hou et al. 2018; Zarrouk et al. 2018). An alternative tracer of the density is the intergalactic medium (IGM), itself traced by Ly $\alpha$  absorption in quasar spectra. Such measurements at  $z \approx 2.4$  were suggested by McDonald (2003) and McDonald & Eisenstein (2007). The first detections of a BAO peak in the Ly $\alpha$  auto-correlation function (Busca et al. 2013; Slosar et al. 2013) used data from the Baryon Oscillation Spectroscopic Survey (BOSS) in the Sloan Digital Sky Survey (SDSS) data-release 9 (DR9), at an effective redshift of  $z = 2.3$ . Delubac et al. (2015), using BOSS in SDSS-DR11, confirmed the detection of a BAO acoustic peak in the Ly $\alpha$  auto-correlation function at the  $5\sigma$  level. Most recently, Bautista et al. (2017) (B17 hereafter) used Ly $\alpha$  forests from BOSS DR12 data and provided a measurement of  $D_H/r_d$  at 3.4% precision level (or of the optimal combination  $D_H^{0.7} D_M^{0.3}/r_d$  at the 2.5% level).



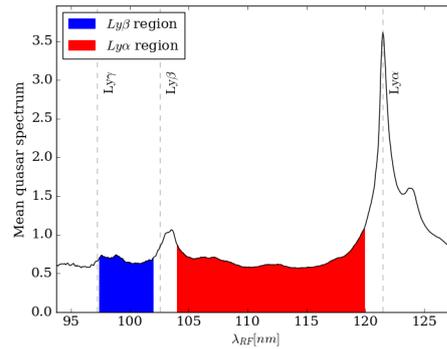
**Fig. 1.** Sky distribution of the 216,163 quasars with redshift in the  $[2.0, 3.5]$  range in the DR14 footprint of the BOSS and eBOSS surveys. The high-density regions are the eBOSS and SEQUELS observations (for the highest declinations in the two Galactic caps) and SDSS-stripe 82 (on the celestial equator in the south galactic cap).

The results were within  $1\sigma$  of the prediction of the flat  $\Lambda$ CDM model favored by CMB anisotropies (Planck Collaboration et al. 2016). However, when combined with the BAO imprint on the cross-correlation of the  $\text{Ly}\alpha$  forest with BOSS DR12 quasars (du Mas des Bourboux et al. 2017), the values of  $D_H/r_d$  and  $D_M/r_d$  at  $z \sim 2.3$  differ by  $2.3\sigma$  from this model. This mild tension was already present in the combined constraints of the cross-correlation measurement of Font-Ribera et al. (2014) and the auto-correlation of Delubac et al. (2015).

In the present paper, we use quasar spectra from the BOSS survey and from its extended version eBOSS in the SDSS DR14 to study BAO in the  $\text{Ly}\alpha$  auto-correlation function. The quasar- $\text{Ly}\alpha$  cross-correlation is studied in a companion paper (Blomqvist et al. 2019). As in previous measurements, we use  $\text{Ly}\alpha$  absorption in the “ $\text{Ly}\alpha$  region” of quasar spectra, i.e., quasar rest-frame wavelengths in the range  $104 < \lambda_{\text{RF}} < 120$  nm. We call the auto-correlation function using only this region the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  correlation<sup>1</sup>. To increase the statistical power we also include  $\text{Ly}\alpha$  absorption in the  $\text{Ly}\beta$  regions of quasars,  $97.4 < \lambda_{\text{RF}} < 102$  nm, correlated with the  $\text{Ly}\alpha$  absorption in  $\text{Ly}\alpha$  regions, i.e., the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlation function. The  $\text{Ly}\beta$  region was previously used by Iršič et al. (2013) to investigate the flux transmission power spectrum within individual spectra.

Besides the use of the  $\text{Ly}\beta$  region, the analysis presented here differs in several ways from that of Bautista et al. (2017) based on DR12 data. For each quasar, we now use all observations instead of just the best one. We analyze  $\sim 15\%$  more  $\text{Ly}\alpha$  regions. We have refined the modelling of the weights (§3.2), the way we take into account the effect of unmasked High Column Density (HCD) systems, and the modeling of nonlinearities in the power spectrum (§4.1). We have not developed new mock spectra beyond those used in the DR12 analysis though this is being done for the final eBOSS analysis.

<sup>1</sup> We use the notation “absorption(spectral region)” to distinguish the nature of absorption from the wavelength interval where the absorption is observed. Hence, the notation  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  denotes  $\text{Ly}\alpha$  absorption in  $\text{Ly}\alpha$  regions correlated with  $\text{Ly}\alpha$  absorption in  $\text{Ly}\alpha$  regions.



**Fig. 2.** The  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  spectral regions defined in Table 1.

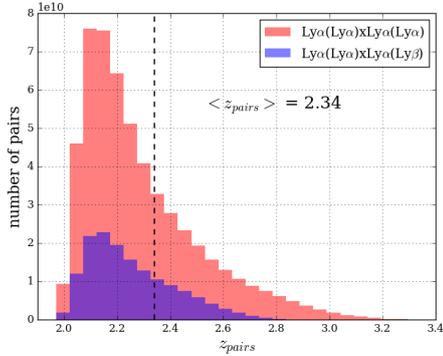
The layout of the paper is the following. In §2, we present the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  spectral region samples used in the present study. We compute correlation function of  $\text{Ly}\alpha$  absorption for the DR14 data in §3 and present our physical model for this function in §4. The results of fitting the data are presented and discussed in §5. We draw cosmological conclusions in §6 and summarize our results in §7.

The computations of the correlation functions presented in this paper have been performed using a dedicated software package, Package for Igm Cosmological-Correlations Analyses (picca), developed by our team<sup>2</sup>.

## 2. Data sample and reduction

The extended Baryon Oscillation Spectroscopic Survey (eBOSS; Dawson et al. 2016) is the extension of the BOSS experiment (Dawson et al. 2013) which aims to measure cosmology with

<sup>2</sup> Available at <https://github.com/igmhub/picca>.



**Fig. 3.** Weighted distribution of the redshift of pairs used to measure the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlation functions. The mean redshift of the combined sample is  $\langle z_{pairs} \rangle = 2.34$ .

BAO using optical spectra from quasars, emission line galaxies and luminous red galaxies. It is one of the four projects of the fourth stage of the Sloan Digital Sky Survey (SDSS-IV; Blanton et al. 2017).

The optical spectra are collected from 1000 fibers, attached to the focal plane of a 2.5 m telescope in Apache Point Observatory (Gunn et al. 2006), by two spectrographs in the wavelength range [360, 1000] nm (Smee et al. 2013). The spectral resolution of the spectrographs is  $\approx 2000$ .

In this paper, we use the forests of the high-redshift quasar sample from the SDSS Data Release 14 (DR14; Abolfathi et al. 2017). This sample contains the first two years of eBOSS data and the five years of BOSS observations reprocessed using the eBOSS pipeline. It also includes data from the ancillary programs Time-Domain Spectroscopic Survey (TDSS) and SPectroscopic IDentification of ERosita Sources (SPIDERS). The quasar target selection is presented in Myers et al. (2015). Note that eBOSS also targets quasars at low redshifts (where the Ly $\alpha$  region is not observable) to be used in other programs (Ata et al. 2018; Wang et al. 2018; Blomqvist et al. 2018; Zhao et al. 2019; du Mas des Bourboux et al. 2019).

The automated data reduction is organized in two steps (Dawson et al. 2016). The pipeline initially extracts the two-dimensional raw data into a one-dimensional flux-calibrated spectrum. During this procedure, the spectra are wavelength and flux calibrated and the individual exposures of one object are coadded into a rebinned spectrum with  $\Delta \log(\lambda) = 10^{-4}$ . The spectra are then classified as STAR or GALAXY or QSO, and their redshift is estimated. Objects that cannot be automatically classified are visually inspected (P aris et al. 2017) and a quasar catalog is produced, which contains 526,356 quasar spectra with redshift  $0 < z < 7$ . Among these objects, 144,046 were not in DR12. The coverage footprint of DR14 quasars is presented in Fig. 1

In this study, we examine both Ly $\alpha$  and Ly $\beta$  regions (see Fig. 2). The Ly $\alpha$  region in the quasar spectrum lies between the Ly $\alpha$  and the Ly $\beta$  emission peaks. We limit its coverage to the rest-frame wavelength range [104, 120] nm in order not to

**Table 1.** Definitions of the Ly $\alpha$  and Ly $\beta$  regions in terms of restframe wavelength range. Also listed are the allowed observer frame wavelength ranges, the corresponding quasar redshift ranges and the number of forests available in our sample.

Regions	$\lambda_{\text{RF}}$ [nm]	$\lambda_{\text{obs}}$ [nm]	$z_q$	# forests
Ly $\beta$	[97.4, 102]	[360, 459]	[2.53, 3.5]	56,154
Ly $\alpha$	[104, 120]	[360, 540]	[2.0, 3.5]	179,965

include the emission peaks, whose shape depends on the environment of the quasar. This approach minimizes the variance of the flux-transmission field defined in section 3.1. Similarly, we define the Ly $\beta$  region as the rest-frame wavelength range [97.4, 102] nm (Table 1, Fig. 2).

In the DR14 quasar catalog, selecting quasar redshifts in the range [2.0, 3.5] yields to 216,162 spectra containing, at least partially, the Ly $\alpha$  region, and selecting quasar redshifts in the range  $z_q \in [2.53, 3.5]$  yields 86,245 spectra containing the Ly $\beta$  region. We choose  $z_q = 3.5$  as an upper limit, as beyond this redshift the quasar density is insufficient to measure correlations and the rate of redshift misidentification is large (Busca & Balland 2018). The requirement that the observed wavelength must be greater than 360 nm is due to the low CCD response and atmospheric transmission in the UV region.

In order to mask damped Ly $\alpha$  systems (DLA), we use the updated DR14 DLA catalog of Noterdaeme et al. (2009, 2012), which contains 34,541 DLA in 27,212 forests. The absorption of the identified DLAs are modeled with a Voigt profile and the regions with more than 20% of absorbed flux are masked. For the Ly $\beta$  regions, we apply this procedure both for Ly $\alpha$  and Ly $\beta$  strong absorbers. We also mask the sky emission and absorption lines listed on the SDSS website<sup>3</sup>. The Broad Absorption Line (BAL) quasars are automatically identified (P aris et al. 2017) and excluded from the data, leaving a sample of 201,286 objects for the Ly $\alpha$  regions and 80,443 for the Ly $\beta$  regions.

For the determination of the correlation function, we divide spectra into “analysis pixels” that are the inverse-variance-weighted flux average over three adjacent pipeline pixels. Throughout the rest of this paper, “pixel” refers to analysis pixels unless otherwise stated. Spectral regions with less than 50 such pixels in regions or which have failed the continuum-fitting procedure (Sec. 3.1) are discarded. These selection criteria produce 179,965 Ly $\alpha$  regions (compared to 157,783 in B17) and 56,154 Ly $\beta$  regions (see Table 1).

The analysis procedure described in the next section assigns redshifts to the observed pixel wavelengths by assuming that flux decrements, in both the Ly $\alpha$  and Ly $\beta$  regions, are due to Ly $\alpha$  absorption. The effect of non-Ly $\alpha$  absorption is taken into account in the correlation-function model presented in §4. The weighted distribution of the redshifts of pairs of pixels used to measure  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlations are presented in Fig. 3. The mean redshift of the combined set of pixel pairs is 2.34.

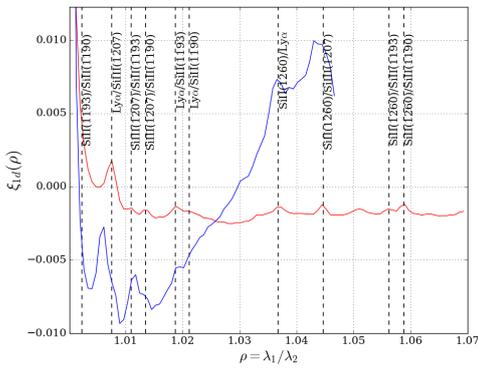
### 3. Computing the Ly $\alpha$ correlation function from the data

This section describes, first, the measurement of the flux-transmission field and then its correlation function and associated covariance matrix.

<sup>3</sup> <http://classic.sdss.org/dr6/algorithms/linestable.html>

**Table 2.** Parameters of the “Pl2015 model”, i.e. the flat  $\Lambda$ CDM model of Planck Collaboration et al. (2016) that we use here to transform redshifts and angular separations into radial and transverse separations.

Parameters	Values
$\Omega_M h^2$	0.1426
$= \Omega_C h^2 + \Omega_B h^2 + \Omega_V h^2$	0.1197+ 0.02222+ 0.0006
$h$	0.6731
$N_s, \sigma_8, n_s$	3, 0.8298, 0.9655
$\Omega_m$	0.3147
$r_d$	147.33 Mpc (99.17 $h^{-1}$ Mpc)
$D_H(2.34)/r_d, D_M(2.34)/r_d$	8.581, 39.26



**Fig. 4.** The one-dimensional correlation functions,  $\xi_{1d}$ , in the Ly $\alpha$  (red curve) and Ly $\beta$  (blue curve) regions as a function of the ratio of transition wavelengths. Peaks are due to absorption by the two labeled elements at zero physical separation (Table 3).

**Table 3.** The Ly $\alpha$ /metal and metal/metal pairs contributing to the flux correlation function. The table shows the ratio of transition wavelengths and the corresponding apparent separation,  $r_{||}^{\text{app}}$ , for pairs at vanishing physical separation, computed at an average redshift of 2.34 using eqn. (23)

Transitions	$\lambda_1/\lambda_2$	$r_{  }^{\text{app}} [h^{-1} \text{Mpc}]$
Si II(1193)/Si II(1190)	1.002	7
Ly $\alpha$ (1216)/Si II(1207)	1.008	21
Si II(1207)/Si II(1193)	1.011	31
Si II(1207)/Si II(1190)	1.014	38
Ly $\alpha$ (1216)/Si II(1193)	1.019	52
Ly $\alpha$ (1216)/Si II(1190)	1.021	59
Si II(1260)/Ly $\alpha$ (1216)	1.037	105
Si II(1260)/Si II(1207)	1.045	126
Si II(1260)/Si II(1193)	1.056	157
Si II(1260)/Si II(1190)	1.059	164

### 3.1. The flux-transmission field $\delta_q(\lambda)$

The computation of the correlation function requires an estimation of the transmission field along the line-of-sight (LOS) towards surveyed quasars. This field arises due to the presence along the LOS of intergalactic gas. More precisely, for the correlation calculation, we only need to know the flux fluctuations around the average transmitted flux spectrum in the forests of

quasar  $q$  at wavelength  $\lambda$ . We thus define the field  $\delta_q(\lambda)$ , for each quasar  $q$  under investigation, as:

$$\delta_q(\lambda) \equiv \frac{f_q(\lambda)}{C_q(\lambda)\bar{F}(z)} - 1, \quad (1)$$

where  $f_q(\lambda)$  denotes the observed flux of quasar  $q$  at observed wavelength  $\lambda$ ,  $C_q(\lambda)$  is the continuum flux and  $\bar{F}(z)$  the mean transmission at the absorber redshift  $z$ .

We estimate the quantity  $C_q\bar{F}(z)$  from the average of the transmitted flux of all forest spectra in the sample:

$$\bar{f}(\lambda_{\text{RF}}) = \frac{\sum_q w_q(\lambda_{\text{RF}})f_q(\lambda_{\text{RF}})}{\sum_q w_q(\lambda_{\text{RF}})}. \quad (2)$$

where  $\lambda_{\text{RF}}$  is the rest-frame wavelength and  $w_q$  a weight (see § 3.2). For each quasar,  $\bar{f}(\lambda_{\text{RF}})$  is then multiplied by a linear polynomial function of  $\Lambda \equiv \log(\lambda)$  to account for the diversity of quasar luminosity and spectral shape:

$$C_q(\lambda)\bar{F}(z) = \bar{f}(\lambda_{\text{RF}})(a_q + b_q\Lambda). \quad (3)$$

There are thus two adjustable parameters per quasar,  $a_q$  and  $b_q$ .

Those forests with identified DLAs are given a special treatment. All pixels where the absorption due to the DLA is higher than 20% are not used. The absorption in the wings is corrected using a Voigt profile following the procedure of Noterdaeme et al. (2012).

The fitting procedure to determine  $(a_q, b_q)$  forces to zero the mean and spectral slope of  $\delta_q(\lambda)$  for each quasar, thus introducing spurious correlations in the measured field. To make it easier to deal with this distortion in the analysis, we follow B17 by transforming the measured  $\delta_q(\lambda)$  to  $\hat{\delta}_q$ :

$$\hat{\delta}_q(\lambda_i) = \sum_j \eta_{ij} \delta_q(\lambda_j), \quad (4)$$

where

$$\eta_{ij} = \delta_{ij}^K - \frac{w_q(\lambda_j)}{\sum_k w_q(\lambda_k)} - (\Lambda_i - \bar{\Lambda}_q) \frac{w_q(\lambda_j)(\Lambda_j - \bar{\Lambda}_q)}{\sum_k w_q(\lambda_k)(\Lambda_k - \bar{\Lambda}_q)^2}, \quad (5)$$

where  $\delta_{ij}^K$  denotes the Kronecker symbol. The advantage of this transformation is that it makes the distortion of the true field introduced by the continuum fit procedure explicit, and, as a consequence, simplifies the link between the true correlation function and the measured, distorted one (see § 4.3).

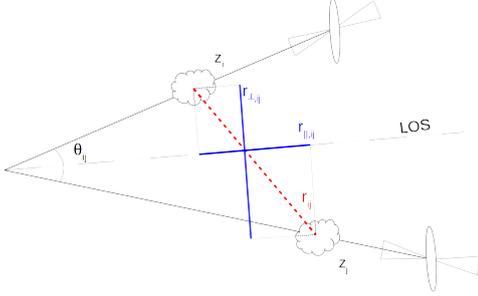
The statistics of  $\hat{\delta}_q(\lambda)$  within individual forests are described (in part) by the so-called one-dimensional correlation function,  $\xi_{1d}(\lambda_1/\lambda_2) = \langle \hat{\delta}(\lambda_1)\hat{\delta}(\lambda_2) \rangle$ . Fig. 4 presents this function for the Ly $\alpha$  and Ly $\beta$  forests. The peaks are due to absorption by different transitions at the same physical position. Table 3 lists the important observed transition pairs. (See also Pieri et al. (2014)).

### 3.2. Pixel weights

The pixel weights are proportional to the inverse of the variance of  $\delta_q(\lambda)$ . Following Blomqvist et al. (2018), the variance is modeled as the sum of three terms:

$$\sigma_q^2(\lambda) = \eta(\lambda)\sigma_{\text{noise}}^2 + \sigma_{\text{LSS}}^2(\lambda) + \epsilon(\lambda)/\sigma_{\text{noise}}^2. \quad (6)$$

The noise pixel variance is  $\sigma_{\text{noise}}^2 = \sigma_{\text{pip}}^2/(C_q\bar{F})^2$  where  $\sigma_{\text{pip}}^2$  is the pipeline estimate of the pixel variance. The intrinsic, redshift dependent, contribution of the density fluctuations underlying Ly $\alpha$  regions is  $\sigma_{\text{LSS}}^2$ . The third term,  $\epsilon(\lambda)/\sigma_{\text{noise}}^2$ , takes into



**Fig. 5.** Definition of the coordinates of pixels used in the computation of the correlation function. Absorbers  $i$  and  $j$  have angular separation  $\theta_{ij}$  and distance separation  $r_{ij}$ . The radial separation  $r_{\parallel,ij}$  is the projection of  $r_{ij}$  on the median LOS and the transverse separation  $r_{\perp,ij}$  is the LOS perpendicular component of  $r_{ij}$ , assuming the flat  $\Lambda$ CDM model (Table 2).

account differences between the fitted quasar spectrum and the individual spectrum of quasar  $q$  (these differences appear at high signal-to-noise). The functions  $\eta(\lambda)$  and  $\epsilon(\lambda)$  correct for imperfections of the pipeline estimates and differences between the average and individual spectra, respectively.

Following Busca et al. (2013), the weights are corrected to take into account the expected redshift dependence of the correlation function amplitude:

$$w_q(\lambda) = \frac{(\lambda/\lambda_\alpha)^{\gamma_\alpha - 1}}{\sigma_q^2(\lambda)} \quad (7)$$

where the Ly $\alpha$  bias redshift-evolution parameter,  $\gamma_\alpha = 2.9$  (McDonald et al. 2006) and  $\lambda_\alpha$  is the Ly $\alpha$  restframe wavelength.

In practice, one starts with an initial estimate of the weights, allowing a first estimate of the mean spectrum  $f(\lambda_{\text{RF}})$  (eqn. 2) and the quasar parameters  $a_q$  and  $b_q$  (eqn. 3). The functions  $\eta(\lambda)$ ,  $\epsilon(\lambda)$  and  $\sigma_{\text{LSS}}(\lambda)$  are then fit and the mean spectrum is then recalculated with the new weights. This process is repeated until stable values are obtained after about five iterations.

### 3.3. The correlation function

To compute the correlation function, we correlate absorption at an observed wavelength  $\lambda_i$  in the LOS of a given quasar  $q$ , with absorption at an observed wavelength  $\lambda_j$  in the LOS of another quasar  $q'$ . Assuming the absorption is due to the Ly $\alpha$  transition, one can compute, from the values of  $\lambda_i$  and  $\lambda_j$ , the redshifts  $z_i$  and  $z_j$  of the matter absorbing these lines. Each pair of absorbers ( $z, q$ ) entering the computation defines a "pixel" in real space and we call  $r_{ij}$  the physical separation between two such pixels  $i$  and  $j$  (see Fig. 5). This distance is calculated assuming the  $\Lambda$ CDM cosmology (Table 2). The distance  $r_{ij}$  can be projected on the radial and the transverse directions, leading to two components  $r_{\parallel,ij}$  and  $r_{\perp,ij}$ . These components can be expressed in terms of the comoving distances  $D(z_i)$  and  $D(z_j)$  from us to absorbers  $i$  and  $j$  and the subtended angle between the two LOS,  $\theta_{ij}$ , as:

$$\begin{cases} r_{\parallel,ij} = (D(z_j) - D(z_i)) \cos\left(\frac{\theta_{ij}}{2}\right) \\ r_{\perp,ij} = (D(z_i) + D(z_j)) \sin\left(\frac{\theta_{ij}}{2}\right) \end{cases} \quad (8)$$

We then define bins of  $(r_{\parallel,ij}, r_{\perp,ij})$  on a 2D grid. In practice, the grid uses 2500 bins of dimensions  $4h^{-1}\text{Mpc} \times 4h^{-1}\text{Mpc}$  over  $0 < r_{\perp} < 200h^{-1}\text{Mpc}$  and  $0 < r_{\parallel} < 200h^{-1}\text{Mpc}$ . For a given bin in this grid,  $A$ , we consider each pair of pixels  $(i, j)$  whose  $r_{\parallel}$  and  $r_{\perp}$  coordinates fall on this bin. The measured correlation function in bin  $A$  reads:

$$\hat{\xi}(A) = \frac{\sum_{(i,j) \in A} w_i w_j \hat{\delta}_i \hat{\delta}_j}{\sum_{(i,j) \in A} w_i w_j}, \quad (9)$$

with  $w_k \equiv w_{q_k}(\lambda_k)$  and  $\hat{\delta}_k \equiv \hat{\delta}_{q_k}(\lambda_k)$ .

We discard from the computation all pixel pairs belonging to the same LOS, since two pixels belonging to the same quasar spectrum are affected in a correlated way by the fitting procedure described in §3.1. Likewise, pixels belonging to the same half plate at the same wavelength are excluded, to avoid unphysical correlations induced by the extraction pipeline.

### 3.4. The covariance matrix

The covariance between two bins  $A$  and  $B$  is defined as:

$$C_{AB} = \langle \hat{\xi}_A \hat{\xi}_B \rangle - \langle \hat{\xi}_A \rangle \langle \hat{\xi}_B \rangle, \quad (10)$$

where  $\langle \dots \rangle$  denotes an ensemble average. Following Delubac et al. (2015) and B17, we estimate equation (10) by dividing the eBOSS footprint in  $N_h = 876$  sky pixels, using the HEALPix tessellation scheme (see Górski et al. 2005), and by equating the ensemble averages of equation (10) with the weighted mean over these sky pixels:

$$\langle \hat{\xi}_A \rangle \approx \frac{\sum_h W_A^h \hat{\xi}_A^h}{\sum_h W_A^h}, \quad (11)$$

and

$$\langle \hat{\xi}_A \hat{\xi}_B \rangle \approx \frac{\sum_h W_A^h W_B^h \hat{\xi}_A^h \hat{\xi}_B^h}{\left(\sum_h W_A^h\right) \left(\sum_h W_B^h\right)} = W_A^{-1} W_B^{-1} \sum_h W_A^h W_B^h \hat{\xi}_A^h \hat{\xi}_B^h, \quad (12)$$

with  $W_A^h$  the sum of the weights of pairs in sky pixels  $h$  contributing to bin  $A$ . Similarly,  $\hat{\xi}_A^h$  is the correlation function of pairs in sky pixels  $h$  that contribute to bin  $A$ .

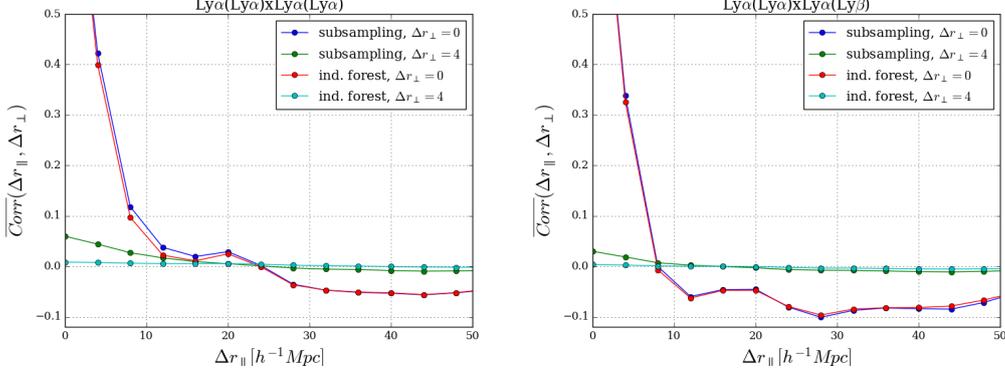
In practice, for the computation of the correlation function, a pair  $(i, j)$  is attributed to the sky pixel of the first quasar of the pair, and the pair  $(j, i)$  is never considered, insuring that a pair is not counted twice in the calculation.

In this approximation, we assume that each sky pixel provides an independent realization of the  $\delta$  field. This statement is not exactly true as correlations do exist between pairs in different sky pixels, but these correlations are small (e.g., Delubac et al. 2015).

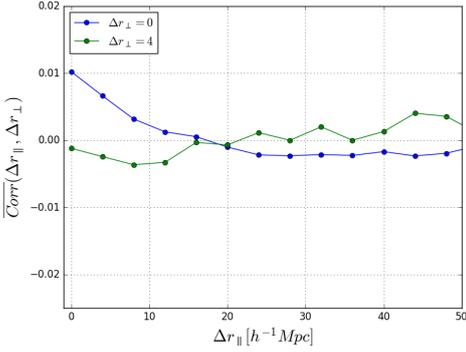
We thus compute the covariance matrix defined in equation (10) using the following expression:

$$C_{AB} = \frac{\sum_h W_A^h W_B^h (\hat{\xi}_A^h \hat{\xi}_B^h - \hat{\xi}_A \hat{\xi}_B)}{\left(\sum_h W_A^h\right) \left(\sum_h W_B^h\right)}, \quad (13)$$

where  $\hat{\xi}_A$  is given by (9). Due to the finite number of sky pixels, the estimate (13) is noisy and must be smoothed before it can be used in fits. We perform the smoothing by approximating the correlation,  $\text{Corr}_{AB} = C_{AB} / \sqrt{C_{AA} C_{BB}}$ , as a function of  $\Delta r_{\parallel} = |r_{\parallel}^A - r_{\parallel}^B|$  and  $\Delta r_{\perp} = |r_{\perp}^A - r_{\perp}^B|$  only, ignoring the small dependence on  $r_{\parallel}$  and  $r_{\perp}$ .



**Fig. 6.** Averaged correlations,  $Corr_{AB} = C_{AB} / \sqrt{C_{AA}C_{BB}}$ , vs.  $\Delta r_{\parallel}$  for the two lowest intervals of  $\Delta r_{\perp}$ , for the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (left) and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  (right) correlation functions. The subsampling covariances are calculated using (13) and the independent-forest estimates by (14).



**Fig. 7.** Averaged correlations,  $Corr_{AB} = C_{AB} / \sqrt{C_{AA}C_{BB}}$ , vs.  $\Delta r_{\parallel}$  for the two lowest intervals of  $\Delta r_{\perp}$ , for the cross-covariance matrix between  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlation functions.

As a check of the subsampling method, the covariance can also be estimated by neglecting inter-forest correlations, in which case the four-point function vanishes unless the four pixels are drawn from just two spectra:

$$C_{AB} = \frac{1}{W_A W_B} \sum_{i \in A} \sum_{k \in B} w_i w_j w_k w_l \xi_{1d}(\lambda_i / \lambda_k) \xi_{1d}(\lambda_j / \lambda_l) \quad (14)$$

where  $\xi_{1d}$  is the intra-forest correlation function shown in Fig. 4. The sum can then be estimated from a random sample of forest pairs. Because neighboring forests are nearly parallel, the sum necessarily produces  $C_{AB} = 0$  unless  $r_{\perp}^A \sim r_{\perp}^B$ .

Because the  $\text{Ly}\alpha$  and  $\text{Ly}\beta$  forests have different  $\xi_{1d}$ , we expect differences between the covariances for  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlations. These differences are illustrated in Fig. 6 showing, for the two lowest values of  $\Delta r_{\perp}$ , the correlation,  $Corr_{AB}$ . For  $\Delta r_{\perp} = 0$ , there is good agreement between the subsampling (13) and independent-forest (14) calculations.

**Table 4.** Parameters of the model of the correlation function. The standard-fit parameters are given in the first section of the table. The second section lists parameters that are fixed in the standard fit, together with their values.

Parameter	Description
$\alpha_{\parallel}, \alpha_{\perp}$	BAO peak-position parameters
$b_{\text{Ly}\alpha}, \beta_{\text{Ly}\alpha}$	Bias parameters for $\text{Ly}\alpha$ absorption
$b_{\text{HCD}}, \beta_{\text{HCD}}$	Bias parameters of HCD systems
$b_m$	Bias of metal species
$L_{\text{HCD}}$ $= 10 h^{-1} \text{Mpc}$	Smoothing scale of HCD systems
$\Sigma_{\perp} = 3.26 h^{-1} \text{Mpc}$	Transverse broadening of BAO peak
$\Sigma_{\parallel} = 6.41 h^{-1} \text{Mpc}$	Radial broadening of BAO peak
$\beta_m = 0.5$	Redshift-space distortion for $\text{Si III}(1190), (1193), (1260), \text{Si III}(1207)$
$\beta_{\text{CIV(eff)}} = 0.27$	C IV(eff) redshift-space distortion
$R_{\parallel}, R_{\perp}$ $= 4 h^{-1} \text{Mpc}$	Binning smoothing parameter
$A_{\text{peak}} = 1$	BAO peak amplitude
$\gamma_{\alpha} = 2.9$	$\text{Ly}\alpha$ bias evolution exponent
$\gamma_m = 1$	Metal bias evolution exponent

Figure 7 displays the  $Corr_{AB}$  between the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlation functions. The  $Corr_{AB}$  are less than a percent, and will be ignored in the fits of the correlation functions.

#### 4. Modeling the correlation function

This section details the model of the  $\text{Ly}\alpha$  auto-correlation function that will be fitted against the estimator  $\hat{\xi}_A$  of equation (9). Table 4 lists the different parameters of the model. The model includes two components: one is the  $\text{Ly}\alpha$ -only correlation function computed from  $\text{Ly}\alpha$  absorption only; the other component incorporates the contribution to the correlation function of absorption by metals, for which the nominal separation for the pixel is not the true separation (see §4.2). We thus write:

$$\xi_{\text{mod}} = \xi_{\text{mod}}^{\text{Ly}\alpha\text{-Ly}\alpha} + \xi_{\text{mod}}^{\text{metals}}. \quad (15)$$

The following subsections describe these two components. Section 4.3 explains how the model is “distorted” to fit the data.

#### 4.1. The baseline model for $\xi_{\text{mod}}^{\text{Ly}\alpha\text{-Ly}\alpha}$

We start from the CAMB linear power spectrum (Lewis et al. 2000) which is decomposed into a smooth and a peak components, following the side-band technique described by Kirkby et al. (2013). This allows one to constrain the position of the BAO peak independently of the correlation function at scales much smaller or much larger than the BAO distance scale. We thus model the matter power spectrum as the sum of two terms corresponding to the smooth and the peak terms in the correlation function. Moreover, in order to incorporate the effects of the non-linear growth of matter that lead to broadening of the BAO peak, the peak term is corrected by a Gaussian factor (Eisenstein et al. 2007). The “quasi-linear” power spectrum hence reads:

$$P_{\text{QL}}(\mathbf{k}, z) = P_{\text{smooth}}(\mathbf{k}, z) + \exp\left(-\frac{k_{\parallel}^2 \Sigma_{\parallel}^2}{2} - \frac{k_{\perp}^2 \Sigma_{\perp}^2}{2}\right) P_{\text{peak}}(\mathbf{k}, z), \quad (16)$$

where  $P_{\text{smooth}}(\mathbf{k}, z)$  and  $P_{\text{peak}}(\mathbf{k}, z)$  are the power spectra of the smooth and peak components, and  $\Sigma_{\parallel}$  and  $\Sigma_{\perp}$  represent the RMS displacements in the parallel and transverse directions, respectively. We adopt the values of Kirkby et al. (2013) for these parameters:  $\Sigma_{\parallel} = 6.41 h^{-1} \text{Mpc}$  and  $\Sigma_{\perp} = 3.26 h^{-1} \text{Mpc}$ .

The power spectrum is obtained from  $P_{\text{QL}}$  as

$$P_{\text{Ly}\alpha\text{-Ly}\alpha}(\mathbf{k}, z) = P_{\text{QL}}(\mathbf{k}, z) d_{\text{Ly}\alpha}^2(\mathbf{k}, z) D_{\text{NL}}(\mathbf{k}) G(\mathbf{k}), \quad (17)$$

where  $d_{\text{Ly}\alpha}$  is the Kaiser factor (Kaiser 1987) for the Ly $\alpha$  absorption and  $D_{\text{NL}}(\mathbf{k})$  takes into account non-linear effects. The function  $G(\mathbf{k})$  models the effect of binning of the correlation function on the separation grid.

The Kaiser factor can be written as:

$$d_{\text{Ly}\alpha} = b'_{\text{Ly}\alpha}(z)(1 + \beta'_{\text{Ly}\alpha} \mu_k^2), \quad (18)$$

where  $b'_{\text{Ly}\alpha}$  is the effective bias of Ly $\alpha$  absorbers with respect to the underlying matter density field,  $\beta'_{\text{Ly}\alpha}$  is the effective redshift space distortion (RSD) parameter, and  $\mu_k = k_{\parallel}/k$ . The two effective parameters ( $b'_{\text{Ly}\alpha}$  and  $\beta'_{\text{Ly}\alpha}$ ) combine Ly $\alpha$  absorption in the IGM and in unmasked high-column density (HCD) systems, i.e., HI absorbers with column densities  $N_{\text{HI}} > 10^{17.2} \text{cm}^{-2}$ :

$$\begin{cases} b'_{\text{Ly}\alpha} = b_{\text{Ly}\alpha} + b_{\text{HCD}} F_{\text{HCD}}(k_{\parallel}) \\ b'_{\text{Ly}\alpha} \beta'_{\text{Ly}\alpha} = b_{\text{Ly}\alpha} \beta_{\text{Ly}\alpha} + b_{\text{HCD}} \beta_{\text{HCD}} F_{\text{HCD}}(k_{\parallel}) \end{cases}, \quad (19)$$

where  $(b_{\text{Ly}\alpha}, \beta_{\text{Ly}\alpha})$  and  $(b_{\text{HCD}}, \beta_{\text{HCD}})$  are the bias parameters associated with the IGM and HCD systems and  $F_{\text{HCD}}$  is a function defined below.

Following McDonald et al. (2006), we assume that the product of  $b_{\text{Ly}\alpha}$  and the growth factor of structures varies with redshift as  $(1+z)^{\gamma_{\alpha}-1}$ , with  $\gamma_{\alpha} = 2.9$ , while we make use of the approximation that  $\beta_{\text{Ly}\alpha}$  does not depend on redshift.

HCD absorbers are expected to trace the underlying density field and their effect on the flux-transmission field depends on whether they are identified and given the special treatment described in Sect. 2. If they are correctly identified with the total absorption region masked and the wings correctly modeled,

they can be expected to have no significant effect on the field. Conversely, if they are not identified, the measured correlation function will be modified because their absorption is spread along the radial direction. This broadening effect introduces a  $k_{\parallel}$  dependence of the effective bias (Font-Ribera & Miralda-Escudé 2012). Following the study of Rogers et al. (2018), we adopt a simple exponential form,  $F_{\text{HCD}} = \exp(-L_{\text{HCD}} k_{\parallel})$ , where  $L_{\text{HCD}}$  is a typical length scale for these systems. DLA identification is possible if their width (wavelength interval for absorption greater than 20%) is above  $\sim 2.0 \text{ nm}$ , corresponding to  $\sim 14 h^{-1} \text{Mpc}$  in our sample. Based on results from Rogers et al. (2018) results, we impose  $L_{\text{HCD}} \sim 10 h^{-1} \text{Mpc}$  while fitting for the bias parameters  $b_{\text{HCD}}$  and  $\beta_{\text{HCD}}$ . Fixing  $L_{\text{HCD}}$  is necessary because otherwise the model becomes too unconstrained. We have verified that setting  $L_{\text{HCD}}$  in the range  $7 < L_{\text{HCD}} < 13 h^{-1} \text{Mpc}$  does not significantly change the inferred BAO peak position.

We focus on the minimal model able to reproduce the data, designated as “baseline model”. This baseline model does not include the correction of the UV background fluctuations (Pontzen et al. 2014; Gontcho A Gontcho et al. 2014) used in B17. We discuss the improvement of the fit when this UV correction is added, in section 5.

The function  $D_{\text{NL}}(\mathbf{k})$  accounts for non-linear effects such as thermal broadening, peculiar velocities and non-linear structure growth. A fitting formula for  $D_{\text{NL}}$  is given by equation (21) of McDonald (2003) and has been extensively used in previous studies. More recently, Arinyo-i-Prats et al. (2015) proposed a new fitting formula involving 6 free parameters given by their equation (3.6). Besides reducing number of free parameters with respect to McDonald (2003), it has the correct behavior at small wavenumber  $k$  and an explicit dependence on  $P_{\text{QL}}(k)$ , whereas this dependence is only implicit in the McDonald (2003) formula. In practice, the two approaches yield similar results but for the above reasons, we adopt the formula of Arinyo-i-Prats et al. (2015) in the present work and linearly interpolate the parameter values from their Table 7 at the effective redshift  $z = 2.34$ .

To account for the effect of the binning of the correlation function on the separation grid, we assume the distribution to be homogeneous on each bin<sup>4</sup> and compute the function  $G(\mathbf{k})$  as the product of the Fourier transforms of the rectangle functions that model a uniform square bin:

$$G(\mathbf{k}) = \text{sinc}\left(\frac{k_{\parallel} R_{\parallel}}{2}\right) \text{sinc}\left(\frac{k_{\perp} R_{\perp}}{2}\right), \quad (20)$$

where  $R_{\parallel}$  and  $R_{\perp}$  are the radial and transverse widths of the bins, respectively.

The two terms in  $P_{\text{QL}}(\mathbf{k}, z)$  (eqn. 16) are Fourier transformed to the smooth and peak components of the correlation function:

$$\xi_{\text{mod}}^{\text{Ly}\alpha\text{-Ly}\alpha}(r_{\parallel}, r_{\perp}, \alpha_{\parallel}, \alpha_{\perp}) = \xi_{\text{smooth}}(r_{\parallel}, r_{\perp}) + A_{\text{peak}} \xi_{\text{peak}}(\alpha_{\parallel} r_{\parallel}, \alpha_{\perp} r_{\perp}). \quad (21)$$

The amplitude of the peak,  $A_{\text{peak}}$ , is fixed to unity in the standard fit. In the peak component we have introduced the parameters  $(\alpha_{\parallel}, \alpha_{\perp})$  which allow us to fit for the peak position independently of the smooth component:

$$\alpha_{\parallel} = \frac{D_{\text{H}}(\bar{z})/r_d}{[D_{\text{H}}(\bar{z})/r_d]_{\text{fid}}}, \quad \alpha_{\perp} = \frac{D_{\text{M}}(\bar{z})/r_d}{[D_{\text{M}}(\bar{z})/r_d]_{\text{fid}}}, \quad (22)$$

<sup>4</sup> In fact, in the perpendicular direction the distribution is approximately proportional to  $r_{\perp}$ ; however, assuming homogeneity produces a sufficiently accurate correlation function (B17).

where  $\bar{z}$  is the effective redshift of the measurement and the suffix 'fid' denotes the PI2015 cosmology from Table 2.

#### 4.2. The contamination by metals

The second term in the model correlation function (eq. 15) accounts for absorption by metals along the quasar LOS. Such absorption is correlated with Ly $\alpha$  absorption (Pieri et al. 2014) and can be used as a tracer of the density field (Blomqvist et al. 2018; du Mas des Bourboux et al. 2019). Here, it is a complicating factor in the analysis because the redshifts of pixels are calculated assuming Ly $\alpha$  absorption.

The important metals can be seen in the 1D correlation function,  $\xi_{1d}(\lambda_1/\lambda_2)$ , shown in Fig. 4. Column 2 of Table 3 lists the wavelength ratios for the main metal/metal and metal/Ly $\alpha$  absorption correlations, relevant for the Ly $\alpha$  auto-correlation function computation. The corresponding apparent radial separation at vanishing physical separation is

$$r_{\parallel}^{\text{app}} \approx (1 + \bar{z}) D_H(\bar{z}) \frac{\lambda_1 - \lambda_2}{\lambda_\alpha} \quad (23)$$

where  $\bar{z}$  is the mean redshift of the pair. Values are given in Table 3 for  $\bar{z} = 2.34$ .

We model the power spectrum of each pair of absorbers,  $(m, n)$ , with the same form as that for Ly $\alpha$ -Ly $\alpha$  absorption (17) except that HCD effects are neglected:

$$P_{mn}(\mathbf{k}, z) = b_m b_n (1 + \beta_m \mu_k^2)(1 + \beta_n \mu_k^2) G(\mathbf{k}) P_L(\mathbf{k}, z). \quad (24)$$

Since the  $b_m$  and  $\beta_m$  are mostly determined near  $(r_{\perp}, r_{\parallel}) \sim (0, r_{\parallel}^{\text{app}})$ , they cannot be determined separately. We therefore fix  $\beta_{\text{CIV(eff)}} = 0.27$  (Blomqvist et al. 2018). For the other metal species we keep the value  $\beta = 0.50$  used in Bautista et al. (2017) which comes from DLA measurements (Font-Ribera & Miralda-Escudé 2012).

The Fourier transform of  $P_{mn}(\mathbf{k}, z)$  is then the model correlation function of the pair  $(m, n)$ :  $\xi_{\text{mod}}^{m-n}(\bar{r}_{\parallel}, \bar{r}_{\perp})$ , where  $(\bar{r}_{\parallel}, \bar{r}_{\perp})$  are the separations calculated using the correct restframe wavelengths,  $(\lambda_m, \lambda_n)$ .

Since we assign a redshift,  $z_\alpha$ , assuming Ly $\alpha$  absorption, the rest-frame wavelength we ascribe to a metal transition  $m$  observed at wavelength  $\lambda_i$  is not equal to the true rest-frame wavelength  $\lambda_i/(1 + z_m)$ , where  $z_m$  is the true redshift of the metal absorber. This misidentification results in a shift of the model contaminant correlation function. For each pair  $(m, n)$  of contaminants, we compute the shifted model correlation function with respect to the unshifted model metal correlation function  $\xi_{\text{mod}}^{m-n}$  by introducing a metal matrix  $M_{AB}$  (Blomqvist et al. 2018), such that:

$$\xi_{\text{mod}}^{m-n}(A) = \sum_B M_{AB} \xi_{\text{mod}}^{m-n}(\bar{r}_{\parallel}(B), \bar{r}_{\perp}(B)), \quad (25)$$

where:

$$M_{AB} = \frac{1}{W_A} \sum_{(m,n) \in A, (m,n) \in B} w_m w_n, \quad (26)$$

and  $(m, n) \in A$  refers to pixel separation computed assuming  $z_\alpha$ , and  $(m, n) \in B$  to pixel separation computed using the redshifts of the  $m$  and  $n$  absorbers,  $z_m$  and  $z_n$ . We take into account the redshift dependence of the weights, equation (7), in the computation of  $w_m$  and  $w_n$ .

The total metal contaminant correlation function,  $\xi_{\text{mod}}^{\text{metals}}$ , is the sum of all the  $\xi_{\text{mod}}^{m-n}$  contributions, where  $(m, n)$  runs over all

the involved transition pairs for the Ly $\alpha$  auto-correlation function, see Table 3:

$$\xi_{\text{mod}}^{\text{metals}}(A) = \sum_{m,n} \xi_{\text{mod}}^{m-n}(A). \quad (27)$$

#### 4.3. The distorted model

The model correlation function,  $\xi_{\text{mod}}$  of eqn. (15), cannot be fit directly to the estimated correlation function (9) because the measured  $\hat{\delta}(\lambda)$  are only related to the true  $\delta(\lambda)$  through the transformation (4). Following B17, we can account for this effect in the fit by using a distorted model:

$$\hat{\xi}_{\text{mod}}(A) = \sum_B D_{AB} \xi_{\text{mod}}(B), \quad (28)$$

where  $D_{AB}$  is the distortion matrix which, following equation (5), is given by

$$D_{AB} = W_A^{-1} \sum_{ij \in A} w_i w_j \left( \sum_{i'j' \in B} \eta_{i'i} \eta_{j'j} \right). \quad (29)$$

The accuracy of this method of accounting for the distortion of the correlation function was tested with mock data sets by Bautista et al. (2017).

In practice, to avoid prohibitive computational time, the distortion matrix is computed using only a random 5% portion of the total number of pairs.

### 5. Fitting the BAO peak position

Table 5 presents the best-fit parameters for the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) correlation function alone and those including the Ly $\beta$  region, i.e. the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ ) correlation function. Figure 8 displays data for the latter in four ranges of  $\mu$  along with the best fits. The BAO peak is apparent for  $\mu > 0.8$  and is suggested for  $0.5 < \mu < 0.8$ .

The BAO parameters for the fit using both Ly $\alpha$  and Ly $\beta$  regions are

$$\begin{cases} \alpha_{\parallel} = 1.033 \begin{smallmatrix} +0.034 & +0.071 \\ -0.034 & -0.068 \end{smallmatrix} \\ \alpha_{\perp} = 0.953 \begin{smallmatrix} +0.050 & +0.108 \\ -0.045 & -0.091 \end{smallmatrix} \end{cases}, \quad (30)$$

Using the  $D/r_d$  values for the PI2015 cosmology in Table 2, these values yield

$$\begin{cases} D_H(2.34)/r_d = 8.86 \begin{smallmatrix} +0.29 & +0.61 \\ -0.29 & -0.58 \end{smallmatrix} \\ D_M(2.34)/r_d = 37.41 \begin{smallmatrix} +1.96 & +4.24 \\ -1.77 & -3.57 \end{smallmatrix} \end{cases}. \quad (31)$$

These results can be compared with those of B17, who found  $\alpha_{\parallel} = 1.053 \pm 0.036$  and  $\alpha_{\perp} = 0.965 \pm 0.055$  at  $z = 2.33$  using only the Ly $\alpha$  region. These values are very near the present results using only the Ly $\alpha$  region:  $\alpha_{\parallel} = 1.047 \pm 0.035$  and  $\alpha_{\perp} = 0.960 \pm 0.041$ . Our use of the Ly $\beta$  region produces consistent results, given the increase in the data set; the main improvement is that on the precision on  $D_M/r_d$  by  $\sim 25\%$ .

Constraints on the BAO parameters  $(\alpha_{\parallel}, \alpha_{\perp})$  are presented in Fig. 9. Following the method introduced and described in detail in du Mas des Bourboux et al. (2017), we estimate the relation between  $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}^2$  and confidence levels (CLs) for the BAO parameters using a large number of simulated correlation functions generated from the best-fit model and the covariance

**Table 5.** The parameters of the model of the correlation function and the best fit values of the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) data (third column) and to the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) and Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) data (fourth column). Errors on parameters correspond to  $\Delta\chi^2 = 1$ .

Parameter description		Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ )	Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) + Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\beta$ )
	$n_{pairs}$	$5.44 \times 10^{11}$	$6.94 \times 10^{11}$
	$\sum w_{pairs}$	$3.56 \times 10^{13}$	$4.20 \times 10^{13}$
Radial BAO peak-position	$\alpha_{  }$	$1.047 \pm 0.035$	$1.033 \pm 0.031$
Transverse BAO peak-position	$\alpha_{\perp}$	$0.969 \pm 0.041$	$0.953 \pm 0.042$
Ly $\alpha$ redshift-space distortion	$\beta_{Ly\alpha}$	$1.773 \pm 0.066$	$1.933 \pm 0.100$
Ly $\alpha$ velocity bias $b_{\eta Ly\alpha} = b_{Ly\alpha} f / \beta_{Ly\alpha}$	$b_{\eta Ly\alpha}$	$-0.208 \pm 0.004$	$-0.211 \pm 0.004$
HCD redshift-space distortion	$\beta_{HCD}$	$0.845 \pm 0.157$	$1.031 \pm 0.153$
HCD bias Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ )	$b_{HCD}^{Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\alpha)}$	$-0.047 \pm 0.003$	$-0.051 \pm 0.004$
HCD bias Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\beta$ )	$b_{HCD}^{Ly\alpha(Ly\alpha) \times Ly\alpha(Ly\beta)}$	-	$-0.072 \pm 0.005$
Metal absorption bias	$b_{SII(1190)}$	$-0.0051 \pm 0.0010$	$-0.0050 \pm 0.0010$
	$b_{SII(1193)}$	$-0.0046 \pm 0.0010$	$-0.0046 \pm 0.0010$
	$b_{SII(1207)}$	$-0.0082 \pm 0.0010$	$-0.0080 \pm 0.0010$
	$b_{SII(1260)}$	$-0.0025 \pm 0.0013$	$-0.0022 \pm 0.0013$
	$b_{CIV(ef)}$	$-0.0185 \pm 0.0078$	$-0.0163 \pm 0.0089$
	$\chi^2_{min}$	1619.77	3258.92
	DOF	1590-11	3180-12
	Probability	0.232	0.127
	$\chi^2(\alpha_{  } = \alpha_{\perp} = 1)$	1621.55	3260.54

**Table 6.** Best fit values of  $(\alpha_{||}, \alpha_{\perp})$  for the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ ) correlation function fit with various models. The first group includes physical models starting with the basic Kaiser redshift-space model and then including, progressively, metals, HCD, and UV corrections. Fits in the second group include polynomial broadband terms, as described in the text.

Models	$\alpha_{  }$	$\alpha_{\perp}$	$\chi^2/DOF$	Probability
Kaiser	$1.021 \pm 0.028$	$0.977 \pm 0.040$	3624.74/(3180-4)	$3.46 \times 10^{-8}$
+Metals	$1.025 \pm 0.032$	$0.979 \pm 0.044$	3607.96/(3180-9)	$7.14 \times 10^{-8}$
+HCD (baseline)	$1.033 \pm 0.031$	$0.953 \pm 0.042$	3258.92/(3180-12)	0.127
+UV	$1.033 \pm 0.031$	$0.953 \pm 0.042$	3258.84/(3180-13)	0.125
BB				
Physical priors on $(b_{Ly\alpha}, \beta_{Ly\alpha}, b_{HCD})$	$1.037 \pm 0.028$	$0.972 \pm 0.040$	3006.25/(3030-36)	0.434
No additional priors	$1.032 \pm 0.027$	$0.980 \pm 0.039$	3001.00/(3030-36)	0.460

matrix measured with the data. The results of the study, summarized in Table D.1, indicate that the (68,27,95.45%) confidence levels for  $(\alpha_{||}, \alpha_{\perp})$  correspond to  $\Delta\chi^2 = (2.74, 7.41)$  (instead of the nominal values  $\Delta\chi^2 = (2.3, 6.18)$ ). These levels are shown as the red contours in Fig. 9 for the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ ) fit. The best fit is within one standard deviation of the PI2015 model.

In addition to the baseline fits, we performed a variety of non-standard fits to verify that our BAO results are robust and independent of the model. The results of this exercise are given in Table 6. The first group of fits starts with the simple Kaiser redshift-space model and then includes progressively metals, HCD absorption, and UV background fluctuations. Including HCD absorption is necessary to obtain a good  $\chi^2$  but adding UV fluctuations such as those characterized in Gontcho A Gontcho et al. (2014) does not improve the fit, justifying our choice of ignoring the UV issue in the baseline fit. In the standard fit, the physical parameters  $(L_{HCD}, \Sigma_{\perp}, \Sigma_{||}, \beta_m)$  are fixed (see table 4) in order to avoid degeneracies with other parameters and non-physical values. We have verified that letting them free has no impact on the  $\alpha_{||}$  and  $\alpha_{\perp}$  parameters.

An important test of systematic effects in the position of the BAO peak is performed by adding polynomial ‘‘broadband’’ terms to the correlation function (before distortion). We follow the procedure and choice of broadband forms used by B17 and

adopt the form

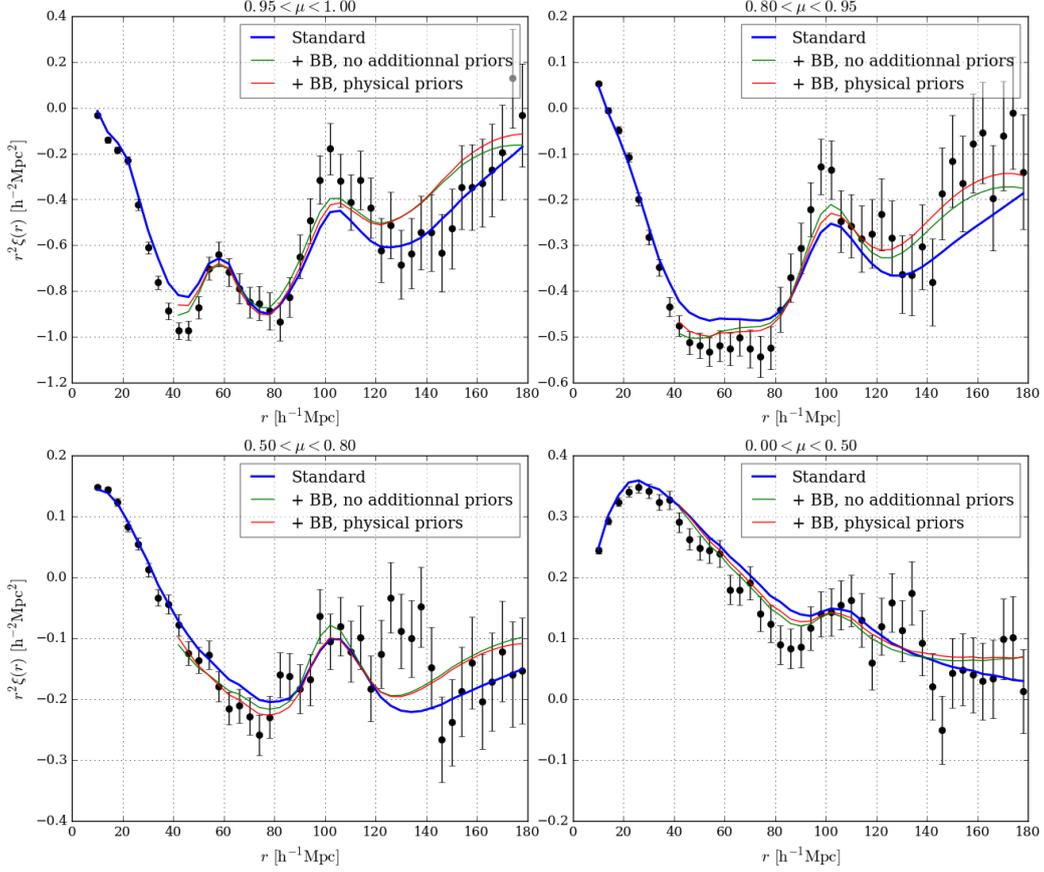
$$B(r, \mu) = \sum_{j=0}^{j_{max}} \sum_{i=i_{min}}^{i_{max}} a_{ij} \frac{L_j(\mu)}{r^i} \quad (j \text{ even}), \quad (32)$$

where the  $L_j$  are Legendre polynomials.

We want to ensure that the power-law terms model variations of the slowly-varying part of the correlation function under the BAO peak. We therefore perform these fits only over the restricted range  $40 < r < 180 h^{-1} \text{Mpc}$ , avoiding introducing undue influence of the  $10 < r < 40 h^{-1} \text{Mpc}$  range on the amplitudes of the power laws. Following B17 we fit with  $(i_{min}, i_{max}) = (0, 2)$  corresponding to a parabola in  $r^2 \xi_{smooth}$  underneath the BAO peak. We set  $j_{max} = 6$ , giving four values of  $j$  corresponding to approximately independent broadbands in each of the four angular ranges in Fig. 8.

We performed the broadband fits in two ways. The first placed ‘‘physical priors’’ on  $(b_{Ly\alpha}, \beta_{Ly\alpha}, b_{HCD})$  in the form of a Gaussian of mean and width of the fit without broadband terms. Such priors ensured that the broadband terms were relatively small perturbations to the physical model. The second type of fit placed no priors on  $(b_{Ly\alpha}, \beta_{Ly\alpha}, b_{HCD})$ .

The results of these fits are given in Table 6 and Fig. 9. We see that the addition of such terms does not change significantly  $\alpha_{||}$  but does shift  $\alpha_{\perp}$  by  $0.5\sigma$  or  $0.7\sigma$  for fits with and without



**Fig. 8.** Weighted combination between measured  $\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\alpha(\text{Ly}\beta)$  correlation functions along with the model best fits in four ranges of  $\mu = r_{\parallel}/r$ . The curves show the standard fit and the two fits with broadband terms defined by eqn. 32 with  $(i_{\min}, i_{\max}, j_{\max}) = (0, 2, 6)$  with and without additional priors, as described in the text.

physical priors. This effect was already seen in B17 but, with less significance. Figure 9 shows that in all cases the BAO peak position is within one standard deviation of the prediction of the Planck 2015 model.

The fits described above of the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  and  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$  correlation functions are the primary results of this paper. We also performed fits with two redshift bins, as described in Appendix B. Each of the two redshifts yielded values of  $(\alpha_{\parallel}, \alpha_{\perp})$  that are within  $1.2\sigma$  of the Planck 2015 model (Fig. B.2). We also fit the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$  correlation as described in Appendix C. Adding the  $\text{Ly}\beta$  absorption data does not add a significant signal to the BAO peak, but it does allow us to measure the  $\text{Ly}\beta$  bias parameters.

Finally, we combine the measurement of  $\text{Ly}\alpha$  auto-correlation function of the present analysis with the  $\text{Ly}\alpha$  - quasar cross-correlation measurement of Blomqvist et al. (2019) by performing a joint fit of the two correlation functions. We use the baseline models of the two analyses and consider the errors to

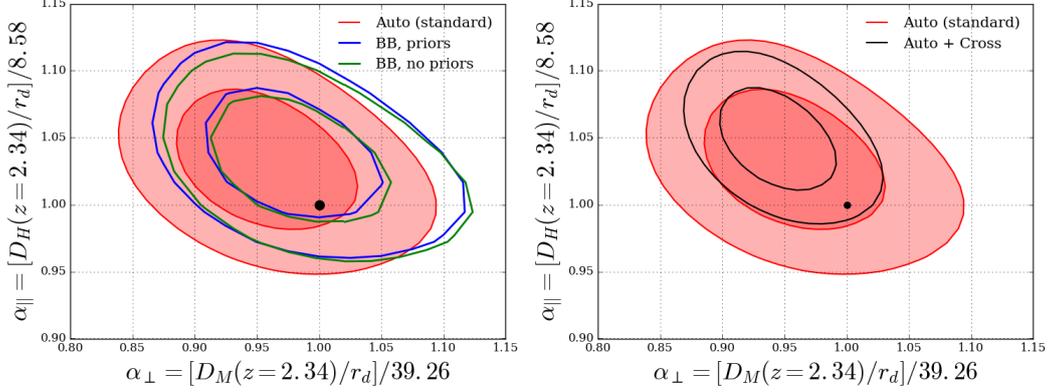
be independent. The joint fit has 18 free parameters and the effective redshift is  $z = 2.34$ . The results are given in the column four of Table 7 and the constraints on  $(\alpha_{\perp}, \alpha_{\parallel})$  in the right panel of Fig. 9. From this combined fit, we obtain:

$$\begin{cases} \alpha_{\parallel} = 1.049^{+0.026}_{-0.025} {}^{+0.052}_{-0.051} \\ \alpha_{\perp} = 0.942^{+0.032}_{-0.030} {}^{+0.067}_{-0.059} \end{cases}, \quad (33)$$

corresponding to:

$$\begin{cases} D_H(2.34)/r_d = 9.00^{+0.22}_{-0.22} {}^{+0.45}_{-0.43} \\ D_M(2.34)/r_d = 36.98^{+1.26}_{-1.18} {}^{+2.63}_{-2.32} \end{cases}. \quad (34)$$

The value of  $\chi^2$  for  $(\alpha_{\parallel} = 1, \alpha_{\perp} = 1)$  is 4.99 greater than the best fit. Using the confidence levels of Table D.1, we conclude that the results of the combined fit are  $1.7\sigma$  from the predictions of the Planck 2015 model (Planck Collaboration et al. 2016).



**Fig. 9.** The left panel shows the 68% and 95% confidence level contours in the  $(\alpha_{\parallel}, \alpha_{\perp})$  plane from the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  auto-correlation function for the standard fit and for fits with polynomial broadband (BB) terms with and without additional priors, as described in the text. The right panel shows the contours for  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  auto-correlation standard fit and those from the combined fit of the auto-correlation and the quasar-Ly $\alpha$  cross-correlation of Blomqvist et al. (2019). In both panels the value for the Planck 2015 model (Planck Collaboration et al. 2016) is shown as a black point.

**Table 7.** Best fit results of the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  correlation function (second column), of the  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  correlation function given by Blomqvist et al. (2019) (third column) and of the two correlation functions (fourth column). Errors on BAO parameters correspond to CL = 68.27%, while the other parameters have errors corresponding to  $\Delta\chi^2 = 1$ . The  $\sigma_v$ ,  $\Delta r_{\parallel}$ ,  $\xi_0^{TP}$  and  $A_{rel1}$  parameters are fit on the  $\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$  correlation function and fully described in Blomqvist et al. (2019)

Parameters	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$	$\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$	$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta) + \text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$
$\alpha_{\parallel}$	$1.033 \pm 0.034$	$1.076 \pm 0.042$	$1.049 \pm 0.026$
$\alpha_{\perp}$	$0.953 \pm 0.048$	$0.923 \pm 0.046$	$0.942 \pm 0.031$
$\beta_{\text{Ly}\alpha}$	$1.933 \pm 0.101$	$2.28 \pm 0.31$	$1.994 \pm 0.099$
$b_{\eta\text{Ly}\alpha}$	$-0.211 \pm 0.004$	$-0.267 \pm 0.014$	$-0.214 \pm 0.004$
$\beta_{\text{QSO}}$	-	0.257	$0.209 \pm 0.006$
$\beta_{\text{HCD}}$	$1.031 \pm 0.153$	$0.500 \pm 0.200$	$0.972 \pm 0.150$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)}$	$-0.051 \pm 0.004$	-	$-0.052 \pm 0.004$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)}$	$-0.072 \pm 0.005$	-	$-0.073 \pm 0.005$
$b_{\text{HCD}}^{\text{QSO} \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)}$	-	$-0.000 \pm 0.004$	$-0.000 \pm 0.004$
$b_{\text{SIII}(1190)}$	$-0.0050 \pm 0.0010$	$-0.0057 \pm 0.0024$	$-0.0043 \pm 0.0009$
$b_{\text{SIII}(1193)}$	$-0.0046 \pm 0.0010$	$-0.0015 \pm 0.0024$	$-0.0034 \pm 0.0009$
$b_{\text{SIII}(1207)}$	$-0.0080 \pm 0.0010$	$-0.0117 \pm 0.0024$	$-0.0083 \pm 0.0009$
$b_{\text{SIII}(1260)}$	$-0.0022 \pm 0.0013$	$-0.0022 \pm 0.0017$	$-0.0019 \pm 0.0009$
$b_{\text{CIV}(e\text{ff})}$	$-0.0163 \pm 0.0089$	-	$-0.0167 \pm 0.0090$
$\sigma_v [h^{-1} \text{Mpc}]$	-	$7.60 \pm 0.61$	$7.053 \pm 0.357$
$\Delta r_{\parallel} [h^{-1} \text{Mpc}]$	-	$-0.22 \pm 0.32$	$-0.169 \pm 0.284$
$\xi_0^{TP}$	-	$0.276 \pm 0.158$	$0.478 \pm 0.112$
$A_{rel1}$	-	$-13.5 \pm 5.8$	$-13.573 \pm 4.721$
$\chi_{min}^2$	3258.91	3231.61	6499.31
$DOF$	3180-12	3180-14	6360-18
Probability	0.13	0.20	0.08
$\chi^2(\alpha_{\parallel} = \alpha_{\perp} = 1)$	3260.54	3235.79	6504.30

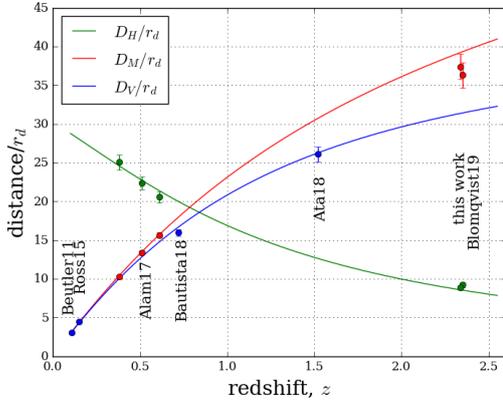
## 6. Cosmological constraints

BAO data over the redshift range  $0.1 < z < 2.4$  is in overall good agreement with the predictions of the flat  $\Lambda$ CDM models consistent with CMB anisotropies, as illustrated in Figure 10. A striking illustration of the expansion history can be made by transforming  $D_H(z)/r_d$  to  $H(z)r_d$ . The measurement presented

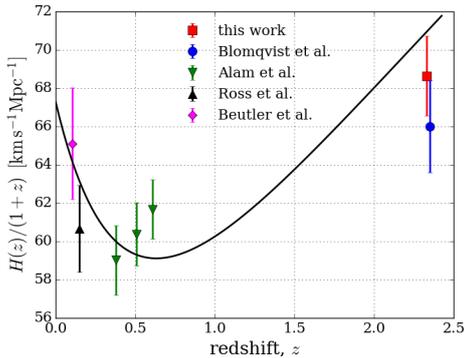
here gives

$$H(2.34) \frac{r_d}{r_d(\text{Planck 2015})} = (227 \pm 8) \text{ km s}^{-1} \text{ Mpc}^{-1}. \quad (35)$$

Fig. 11 plots this value along with other measurements. The data are consistent with the expected behavior of deceleration at high redshift followed by acceleration at low redshift.



**Fig. 10.** BAO measurement of  $D_H/r_d$  and  $D_M/r_d$  using BOSS galaxies (Alam et al. 2017), Ly $\alpha$  absorption in BOSS-eBOSS quasars (this work) and correlation between BOSS-eBOSS quasars and Ly $\alpha$  absorption (Blomqvist et al. 2019). Other measurements give  $D_V/r_d$ , with  $D_V = D_M^{2/3}(zD_H)^{1/3}$ , using galaxies (Beutler et al. (2011), Ross et al. (2015), Bautista et al. (2018)) and BOSS-eBOSS quasars (Ata et al. (2018)). Solid lines show the Planck 2015 values (Planck Collaboration et al. 2016).

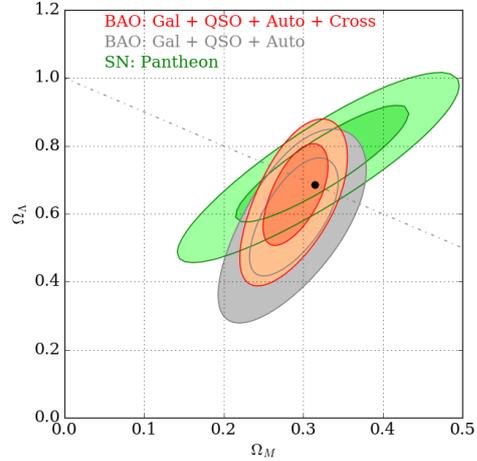


**Fig. 11.** BAO measurement of the comoving expansion rate,  $H(z)/(1+z)$ , measured with BAO with  $r_d = 147.3$  Mpc. The red square is the present measurement at  $z = 2.34$ . The measurement by Blomqvist et al. (2019) is the blue dot. The other points are computed using galaxy measurements (Beutler et al. (2011), Ross et al. (2015), Alam et al. (2017)). The points at  $z = 0.106$  (Beutler et al. 2011) and  $z = 0.15$  (Ross et al. 2015) are converted from  $D_V$  to  $H(z)$  using the SNIa measurement of  $q_0$  given by Betoule et al. (2014). Solid black line shows the Planck 2015 values (Planck Collaboration et al. 2016).

Independent of CMB data and without assuming flatness, the BAO data by themselves constrain the parameters ( $\Omega_m, \Omega_\Lambda, H_0/r_d$ ) of the ( $\Lambda$ )CDM model. Using the combined fit (eqn. 34), the galaxy data of Beutler et al. (2011), Ross et al. (2015), Alam et al. (2017) and Bautista et al. (2018) and the quasar data of Ata et al. (2018) yields

$$\Omega_M = 0.293 \pm 0.027 \quad \Omega_\Lambda = 0.675 \pm 0.099 \quad (36)$$

Article number, page 12 of 19



**Fig. 12.** One and two standard deviation constraints on  $(\Omega_m, \Omega_\Lambda)$ . The red contours use BAO measurements of  $D_M/r_d$  and  $D_H/r_d$  of this work, of Blomqvist et al. (2019) and Alam et al. (2017), and the measurements of  $D_V/r_d$  of Beutler et al. (2011), Ross et al. (2015), Ata et al. (2018) and Bautista et al. (2018). The gray contours do not use the Ly $\alpha$ -quasar cross-correlation measurement of Blomqvist et al. (2019). The green contours show the constraints from SN Ia Pantheon sample (Scolnic et al. 2018). The black point indicates the values for the Planck (2016) best-fit flat  $\Lambda$ CDM cosmology.

corresponding to  $\Omega_k = 0.032 \pm 0.117$ . The best fit gives  $(c/H_0)/r_d = 29.78 \pm 0.55$  corresponding to  $hr_d = (0.683 \pm 0.013) \times 147.33$  Mpc. The Planck 2015 model has  $\chi^2 = 13.76$  for 12 degrees of freedom and is within one standard deviation of the best fit, as illustrated in Figure 12.

## 7. Conclusions

We have used Ly $\alpha$  and Ly $\beta$  spectral regions from the BOSS and eBOSS DR14 data sample to study BAO. Following B17, we have built a model for the Ly $\alpha$  auto-correlation function that we have then fit to the data. Our model incorporates the effects of redshift space distortions, the non-linear growth of matter, the contamination by metals and the modeling of high column density systems along lines-of-sight to quasars. Including UV fluctuations has only a minor impact on the fit results. We measure the ratios  $D_H/r_d$  and  $D_M/r_d$  at the average redshift of pixel pairs,  $z = 2.34$ . We have also performed a measurement of these ratios from the Ly $\alpha$  auto-correlation function in two redshift bins, at  $z = 2.19$  and  $z = 2.49$ .

The  $D_H/r_d$  ratio is measured with a precision of  $\sim 3.3\%$ , a slight improvement over the precision obtained by B17 for this ratio. The  $D_M/r_d$  ratio is measured with a precision of  $\sim 4.4\%$ , which represents an improvement of about 25% with respect to B17. The cosmological measurements obtained in this analysis are in agreement with the predictions of the flat  $\Lambda$ CDM model (Planck 2015) favored by the measurement of CMB anisotropies by Planck.

We have also combined the measurements of the present analysis with the ones obtained from the cross-correlation of Ly $\alpha$

absorption and quasars by Blomqvist et al. (2019). The latter alone favors a value of the  $D_H/r_d$  ratio  $\sim 3\%$  higher than the one favored by the Ly $\alpha$  auto-correlation. As a result, the best-fit value of  $D_H/r_d$  for the combined fit is shifted towards a higher value than the best-fit from the Ly $\alpha$  auto-correlation alone. Combining the measurement of Ly $\alpha$  auto-correlation (this paper) with the quasar - Ly $\alpha$  cross-correlation of Blomqvist et al. (2019), the BAO measurements at  $z = 2.34$  are within  $1.7\sigma$  of the predictions of the  $\Lambda$ CDM model.

The ensemble of BAO measurements is in good agreement with the  $\Lambda$ CDM model (Planck Collaboration et al. 2016). They provide an independent way of determining cosmological parameters that is based only on low-redshift measurements. As illustrated in Fig. 12, the BAO results are also consistent with the recent Pantheon SNIa results (Scolnic et al. 2018).

The present measurements will be much improved by the greater statistical power of the upcoming DESI (DESI Collaboration et al. 2016) and WEAVE-QSO (Pieri et al. 2016) projects. The challenge will be to improve the physical modeling of the correlation function in order to fully profit from the improved data.

*Acknowledgements.* We thank Pasquier Noterdaeme for providing the DLA catalog for eBOSS DR14 quasar. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is <http://www.sdss.org/>. MB, MP and IPR were supported by the A\*MIDEX project (ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French Government program, managed by the French National Research Agency (ANR), and by ANR under contract ANR-14-ACHN-0021 SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

## References

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2017, ArXiv e-prints [arXiv:1707.09322]
- Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
- Anderson, L., Aubourg, E., Bailey, S., et al. 2014a, MNRAS, 441, 24
- Anderson, L., Aubourg, E., Bailey, S., et al. 2014b, MNRAS, 439, 83
- Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435
- Arinyo-i-Prats, A., Miralda-Escudé, J., Viel, M., & Cen, R. 2015, J. Cosmology Astropart. Phys., 12, 017
- Ata, M., Baumgarten, F., Bautista, J., et al. 2018, MNRAS, 473, 4773
- Aubourg, É., Bailey, S., Bautista, J. E., et al. 2015, Phys. Rev. D, 92, 123516
- Bautista, J. E., Busca, N. G., Guy, J., et al. 2017, A&A, 603, A12 (B17)
- Bautista, J. E., Vargas-Magaña, M., Dawson, K. S., et al. 2018, ApJ, 863, 110
- Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
- Beutler, F., Blake, C., Colless, M., et al. 2011, MNRAS, 416, 3017
- Blake, C., Kazin, E. A., Beutler, F., et al. 2011, MNRAS, 418, 1707
- Blanton, M. R., Bershad, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28
- Blomqvist, M., Pieri, M. M., du Mas des Bourboux, H., et al. 2018, J. Cosmology Astropart. Phys., 5, 029
- Blomqvist et al., M. 2019, ArXiv e-prints [arXiv:1901.01950]
- Busca, N. & Balland, C. 2018, ArXiv e-prints [arXiv:1808.09955]
- Busca, N. G., Delubac, T., Rich, J., et al. 2013, A&A, 552, A96
- Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10
- Delubac, T., Bautista, J. E., Busca, N. G., et al. 2015, A&A, 574, A59
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints [arXiv:1611.00036]
- Dijkstra, M., Lidz, A., & Hui, L. 2004, ApJ, 605, 7
- du Mas des Bourboux, H., Dawson, K. S., Busca, N. G., et al. 2019, arXiv e-prints [arXiv:1901.01950]
- du Mas des Bourboux, H., Le Goff, J.-M., Blomqvist, M., et al. 2017, A&A, 608, A130
- Eisenstein, D. J. & Hu, W. 1998, ApJ, 496, 605
- Eisenstein, D. J., Seo, H.-J., & White, M. 2007, ApJ, 664, 660
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560
- Font-Ribera, A., Kirkby, D., Busca, N., et al. 2014, J. Cosmology Astropart. Phys., 5, 027
- Font-Ribera, A. & Miralda-Escudé, J. 2012, J. Cosmology Astropart. Phys., 7, 028
- Gil-Marín, H., Guy, J., Zarrouk, P., et al. 2018, MNRAS, 477, 1604
- Gontcho A Gontcho, S., Miralda-Escudé, J., & Busca, N. G. 2014, MNRAS, 442, 187
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, AJ, 131, 2332
- Hou, J., Sánchez, A. G., Scocimarro, R., et al. 2018, MNRAS, 480, 2521
- Iršič, V., Slosar, A., Bailey, S., et al. 2013, J. Cosmology Astropart. Phys., 9, 016
- Iršič, V. & Viel, M. 2014, J. Cosmology Astropart. Phys., 12, 024
- Kaiser, N. 1987, MNRAS, 227, 1
- Kirkby, D., Margala, D., Slosar, A., et al. 2013, J. Cosmology Astropart. Phys., 3, 024
- Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 743
- McDonald, P. 2003, ApJ, 585, 34
- McDonald, P. & Eisenstein, D. J. 2007, Phys. Rev. D, 76, 063009
- McDonald, P., Seljak, U., Burles, S., et al. 2006, ApJS, 163, 80
- Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, ApJS, 221, 27
- Noterdaeme, P., Petitjean, P., Carithers, W. C., et al. 2012, A&A, 547, L1
- Noterdaeme, P., Petitjean, P., Ledoux, C., & Srianand, R. 2009, A&A, 505, 1087
- Pâris, I., Petitjean, P., Aubourg, E., et al. 2017, ArXiv e-prints [arXiv:1712.05029]
- Percival, W. J., Reid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148
- Pieri, M. M., Bonoli, S., Chaves-Montero, J., et al. 2016, in SF2A-2016: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé, J. Richard, L. Cambresy, M. Deleuil, E. Pécontal, L. Tresse, & I. Vauglin, 259–266
- Pieri, M. M., Mortonson, M. J., Frank, S., et al. 2014, MNRAS, 441, 1718
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
- Pontzen, A., Bird, S., Peiris, H., & Verde, L. 2014, ApJ, 792, L34
- Reid, B. A., Percival, W. J., Eisenstein, D. J., et al. 2010, MNRAS, 404, 60
- Rogers, K. K., Bird, S., Peiris, H. V., et al. 2018, MNRAS, 476, 3716
- Ross, A. J., Samushia, L., Howlett, C., et al. 2015, MNRAS, 449, 835
- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, J. Cosmology Astropart. Phys., 4, 26
- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
- Wang, D., Zhao, G.-B., Wang, Y., et al. 2018, MNRAS, 477, 1528
- Zarrouk, P., Burtin, E., Gil-Marín, H., et al. 2018, MNRAS, 477, 1639
- Zhao, G.-B., Wang, Y., Saito, S., et al. 2019, MNRAS, 482, 3497

## Appendix A: Effective redshift of the fitted parameters

In this section we present a method to determine the region in  $(r_\perp, r_\parallel, z)$  space that is most constraining for the various parameters in the fits of the correlation function. We can expect that the parameters  $(\alpha_\perp, \alpha_\parallel)$  are mostly determined by  $(r_\perp, r_\parallel)$  bins near the BAO peak and at a redshift near the mean redshift of the pixel pairs used in the BAO region. In fact, previous studies (B17, Busca et al. (2013); Delubac et al. (2015)) defined the effective redshift of the BAO measurement in this way. Here, we make this intuitive conclusion more precise by using a Fisher matrix analysis.

We use the Fisher matrix formalism as follows: given a parameter  $p$  varying linearly with redshift, we define the effective redshift  $z_0$  at which it is measured by:

$$p(z) = p_0 + p_1(z - z_0), \quad (\text{A.1})$$

where  $p_0$  is the value given by the fit at  $z = z_0$ . The covariance matrix  $C_p$  between two parameters  $p_0$  and  $p_1$  is given by:

$$C_p \equiv \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix}, \quad (\text{A.2})$$

where  $\sigma_i^2$  is the variance of the parameter  $p_i$  and  $\rho$  is the correlation coefficient between  $p_0$  and  $p_1$ . By definition,  $C_p$  is the inverse of the Fisher matrix  $F_p$ :

$$C_p^{-1} \equiv F_p \equiv \begin{pmatrix} \sum_{ij} \frac{\partial m_i}{\partial p_0} C_{ij}^{-1} \frac{\partial m_j}{\partial p_0} & \sum_{ij} \frac{\partial m_i}{\partial p_0} C_{ij}^{-1} \frac{\partial m_j}{\partial p_1} \\ \sum_{ij} \frac{\partial m_i}{\partial p_1} C_{ij}^{-1} \frac{\partial m_j}{\partial p_0} & \sum_{ij} \frac{\partial m_i}{\partial p_1} C_{ij}^{-1} \frac{\partial m_j}{\partial p_1} \end{pmatrix}, \quad (\text{A.3})$$

with  $m_i$  the model at bin  $i$ . In the case of the linear redshift dependency (A.1), the Fisher matrix  $F_p$  at redshift  $z$ , computed using the set of all the fitted parameter values,  $\{\lambda_0\}$ , is given by:

$$F_p(z) = \sum_{ij} \frac{\partial m_i}{\partial p} \Big|_{\{\lambda_0\}} C_{ij}^{-1} \frac{\partial m_j}{\partial p} \Big|_{\{\lambda_0\}} \begin{pmatrix} 1 & (z_j - z) \\ (z_i - z) & (z_i - z)(z_j - z) \end{pmatrix}, \quad (\text{A.4})$$

with  $z_i$  the mean redshift of the pairs in bin  $i$ .

We represent the quantities  $\frac{\partial m_i}{\partial p}$  for 4 of the 12 fitted parameters of the  $\text{Ly}\alpha$  auto-correlation function in Fig..1. The covariance matrix  $C_p(z)$  then reads:

$$C_p(z) = \frac{1}{|F_p|} \sum_{ij} M_{ij} \begin{pmatrix} (z_i - z)(z_j - z) & -(z_i - z) \\ -(z_j - z) & 1 \end{pmatrix}, \quad (\text{A.5})$$

with

$$M_{ij} \equiv \frac{\partial m_i}{\partial p} \Big|_{\{\lambda_0\}} C_{ij}^{-1} \frac{\partial m_j}{\partial p} \Big|_{\{\lambda_0\}}. \quad (\text{A.6})$$

Since  $M_{ij}$  is symmetric, the determinant of the Fisher matrix,  $|F_p|$ , does not depend on redshift and is given by:

$$|F_p| = \sum_{i,j,k,l} M_{ij} M_{kl} z_i \times (z_j - z_k). \quad (\text{A.7})$$

The variance of  $p_0$  at redshift  $z$  becomes:

$$\sigma_0^2(z) = \frac{1}{|F_p|} \sum_{ij} M_{ij} (z_i - z)(z_j - z). \quad (\text{A.8})$$

Article number, page 14 of 19

**Table A.1.** Effective redshifts at which the  $\alpha_\parallel$  and  $\alpha_\perp$  parameters are measured. The average redshift of pairs is also given.  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  (low  $z$ ) and (high  $z$ ) are introduced in Appendix B.

Correlation functions	$\bar{z}$	$z_{\alpha_\parallel}$	$z_{\alpha_\perp}$
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$	2.35	2.34	2.34
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\beta)$	2.29	2.29	2.28
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha + \text{Ly}\beta)$	2.34	2.33	2.33
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ (low $z$ )	2.19	2.19	2.18
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$ (high $z$ )	2.49	2.49	2.49
$\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\beta(\text{Ly}\beta)$	2.76	2.77	2.78

The effective redshift  $z_0$  is the value which minimizes the error on  $p_0$ :

$$\frac{d\sigma_0^2}{dz} \Big|_{z_0} = -\frac{2}{|F_p|} \sum_{ij} M_{ij} (z_i - z_0) = 0, \quad (\text{A.9})$$

i.e.,

$$z_0 = \frac{\sum_{ij} M_{ij} z_i}{\sum_{ij} M_{ij}}. \quad (\text{A.10})$$

In the case of a combined fit, we compute one matrix  $M^d$  for each correlation function entering the fit:

$$M_{ij}^d \equiv \frac{\partial m_i^d}{\partial p} \Big|_{\{\lambda_0\}} (C_{ij}^d)^{-1} \frac{\partial m_j^d}{\partial p} \Big|_{\{\lambda_0\}}, \quad (\text{A.11})$$

where  $m_i^d$  is the model for the correlation function  $d$  at bin  $i$ . In this case,  $z_0$  reads:

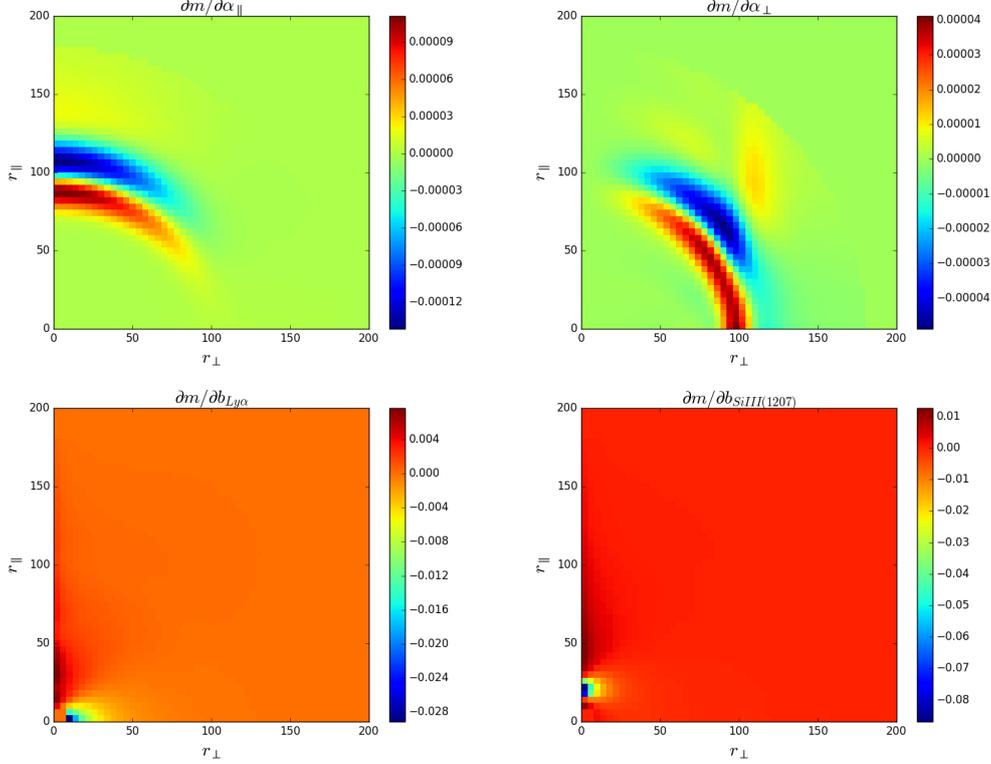
$$z_0 = \frac{\sum_d \sum_{ij} M_{ij}^d z_i}{\sum_d \sum_{ij} M_{ij}^d}. \quad (\text{A.12})$$

Table A.1 presents the effective redshifts at which the  $\alpha_\parallel$  and  $\alpha_\perp$  parameters are measured for the different correlation functions computed in this paper. The effective redshift values differ by less than 0.5% for  $\alpha_\parallel$  and  $\alpha_\perp$ . Figure .1 shows the quantities  $\partial m/\partial p$  in the  $(r_\perp, r_\parallel)$  plane for the fitted parameters  $p \in [\alpha_\parallel, \alpha_\perp, b_{\text{Ly}\alpha}, b_{\text{SIII}(1207)}]$ .  $m$  is the baseline model for the  $\text{Ly}\alpha$  auto-correlation.

## Appendix B: Fits in two redshift bins

The present data set is large enough to constrain the BAO parameters in two independent redshift bins, in a way similar to what will be done in forthcoming cosmological surveys, such as the DESI project (DESI Collaboration et al. 2016). To simplify the analysis, we consider only the  $\text{Ly}\alpha(\text{Ly}\alpha) \times \text{Ly}\alpha(\text{Ly}\alpha)$  correlation function.

A straightforward way of defining a high- and low-redshift sample of pixel pairs would be to simply use pixel pairs of mean redshift less than or greater than an appropriately chosen value,  $z_{\text{cut}}$ . The drawback of such an approach is that a given pair of forests could belong to both bins, as some pixels in a given forest would be associated with some pixels in the other forest, either in pixel pairs with mean redshift less than  $z_{\text{cut}}$ , or in pixel pairs with mean redshift greater than  $z_{\text{cut}}$ . The fact that some pairs of forests belong to both redshift bins introduces unwanted correlations when correcting for the distortions introduced by our continuum fitting procedure. To circumvent this problem, we choose



**Fig. 1.** The quantity  $\partial m / \partial p$  in the  $(r_{\perp}, r_{\parallel})$  plane for the fitted parameters  $p \in [\alpha_{\parallel}, \alpha_{\perp}, b_{\text{Ly}\alpha}, b_{\text{SiIII}(1207)}]$ .  $m$  is the baseline model of the Ly $\alpha$  auto-correlation function. The graphs show which pixels contribute the most to the constraints on the considered parameter. BAO parameters  $\alpha_{\parallel}$  and  $\alpha_{\perp}$  are constrained by the bins around the location of the BAO peak, while the Ly $\alpha$  bias is mostly constrained by the bins at  $\sim$  zero separation. We also note that the Si III(1207) bias is mostly constrained by the bins  $r_{\perp} \sim 0$ ,  $r_{\parallel} \sim 21 h^{-1} \text{Mpc}$ , in agreement with the  $r_{\parallel}^{app}$  apparent separation given in Table 3.

to assign forest pairs to the high or low-redshift sample by cutting on the mean of the maximum  $z$  of the two forests.

We thus evaluate, for all pairs of forests  $(i, j)$ , the following quantity:

$$z_{ij} = \frac{z_{\text{max,abs}}^i + z_{\text{max,abs}}^j}{2}, \quad (\text{B.1})$$

where  $1 + z_{\text{max,abs}}^k = \max(\lambda_{\text{obs}}^k) / \lambda_{\text{abs}}$ , with  $\max \lambda_{\text{obs}}^k$  the last pixel of forest  $k$ , and  $\lambda_{\text{abs}}$  the rest-frame wavelength of the considered transition. The condition  $z_{ij} < z_{\text{cut}}$  defines the low redshift bin, while the opposite condition defines the high redshift one.

The value of  $z_{\text{cut}}$  is tuned so that the sum of the weights of all absorber pairs in the high redshift correlation function equals the one of all pairs in the low redshift correlation function. This ensures the two bins have a comparable statistical power. This process leads us to select  $z_{\text{cut}} = 2.5$ . The redshift distribution of absorber pairs obtained in this way are shown on Figure B.1. The average pair redshift is  $\bar{z} = 2.19$  and  $\bar{z} = 2.49$  for the low and high redshift bin, respectively.

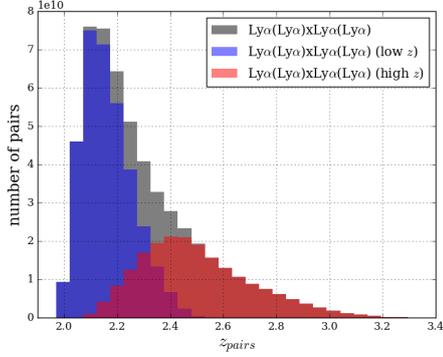
Figure B.3 presents the result of fitting the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) correlation baseline model to the data in the low (blue

points) and high (red points) redshift bins in the usual four  $\mu$  wedges.

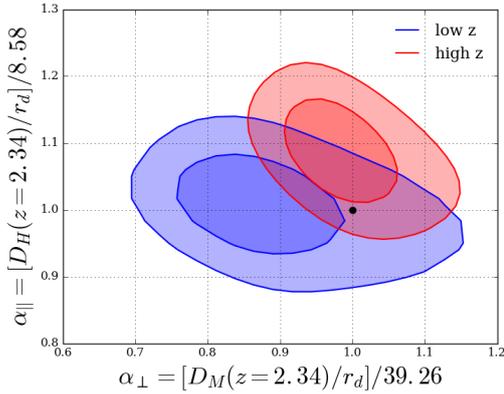
Table B.1 shows the associated best-fit parameters. From the table, we note that  $\beta_{\text{Ly}\alpha}$  is notably different at low and high redshift. When fitting the full sample, we assumed, following Kirkby et al. (2013),  $\beta_{\text{Ly}\alpha}$  to be constant. A redshift dependent  $\beta_{\text{Ly}\alpha}$  could thus be an improvement in future analyses.

From the lower right panel of Figure B.3, we see that the amplitude of the high-redshift correlation function is higher than the amplitude of the low-redshift one. This is expected, as  $b_{\text{Ly}\alpha}$  increases with redshift (Kirkby et al. 2013).

Figure B.2 presents the constraints obtained, in the  $(\alpha_{\perp}, \alpha_{\parallel})$  parameter space, from fitting the Ly $\alpha$  auto-correlation Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) in the low (blue contours) and high (red contours) redshift bins. The values of  $\Delta\chi^2$  corresponding to a given confidence level were taken from Table D.1. Both the high and low redshift measurements are within  $1.2\sigma$  of the P11015 model.



**Fig. B.1.** Pixel pair redshift distribution of the subsamples used in the present analysis: full Ly $\alpha$  auto-correlation function (gray), low-redshift Ly $\alpha$  auto-correlation function (blue), high redshift Ly $\alpha$  auto-correlation function (red). The two latter subsamples are used to produce a measurement of  $H(z)$  at  $z = 2.19$  and  $z = 2.49$ .



**Fig. B.2.** The 68% and 95% confidence level contours in the  $(\alpha_{\parallel}, \alpha_{\perp})$  plane from the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) computed with the low and high redshift bins. The  $\Delta\chi^2$  values corresponding to confidence levels are taken from Table D.1. The black dot corresponds to the PI2015 model.

### Appendix C: The Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\beta$ (Ly $\beta$ ) cross correlation

As an extension of our main analysis, we compute the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ) correlation function, following the procedure previously described. We computed the 1D correlation function in the Ly $\beta$  region (Fig. C.1) to identify the contaminating metals (see Table C.1).

The results of this analysis are presented in Table C.2. The Ly $\beta$  absorption signal is clearly detected as  $b_{\text{Ly}\beta}$  is non zero at at the  $2.9\sigma$  level. We also see the signal due to O vi(1032) and O vi(1036). Note that the correlation between Ly $\beta$  and Si ii(1260) occurs near the BAO peak (last line of Table C.1). Due to the small Ly $\beta$  absorption cross-section and to the small wavelength extent of the Ly $\beta$  region, it is harder to detect the BAO peak than for the Ly $\alpha$  auto-correlation function. Moreover, the Ly $\beta$  -Si ii

**Table B.1.** Results in two redshift bins from fitting the baseline model to the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) correlation function.

Parameters	low z	high z
$\alpha_{\parallel}$	$1.008 \pm 0.043$	$1.088 \pm 0.046$
$\alpha_{\perp}$	$0.861 \pm 0.062$	$0.977 \pm 0.044$
$\beta_{\text{Ly}\alpha}$	$2.083 \pm 0.160$	$1.585 \pm 0.113$
$b_{\eta\text{Ly}\alpha}$	$-0.218 \pm 0.006$	$-0.201 \pm 0.005$
$\beta_{\text{HCD}}$	$0.745 \pm 0.174$	$0.678 \pm 0.179$
$b_{\text{HCD}}$	$-0.058 \pm 0.007$	$-0.040 \pm 0.006$
$b_{\text{SiII}(1190)}$	$-0.0049 \pm 0.0015$	$-0.0045 \pm 0.0016$
$b_{\text{SiII}(1193)}$	$-0.0067 \pm 0.0015$	$-0.0021 \pm 0.0015$
$b_{\text{SiII}(1207)}$	$-0.0105 \pm 0.0016$	$-0.0055 \pm 0.0017$
$b_{\text{SiII}(1260)}$	$-0.0039 \pm 0.0018$	$-0.0021 \pm 0.0017$
$b_{\text{CIV}(\text{eff})}$	$-0.0178 \pm 0.0095$	$-0.0186 \pm 0.0095$
$\chi^2_{\text{min}}$	1580.95	1737.15
$\text{DOF}$	1590 - 11	1590 - 11
Probability	0.481	0.003
$\chi^2(\alpha_{\parallel} = \alpha_{\perp} = 1)$	1584.30	1740.89

**Table C.1.** Same as Table 3 for the main metal/metal, metal/Ly $\beta$  and Ly $\alpha$ /metal correlations relevant to the computation of the Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ) correlation function. The apparent separation  $r_{\parallel}^{\text{app}} = (1 + \Xi)H(z)\left(\frac{\lambda_{\alpha}}{\lambda_{\beta}} - \frac{\lambda_{\alpha}}{\lambda_{\beta}}\right)$  is computed at the average redshift of 2.76.

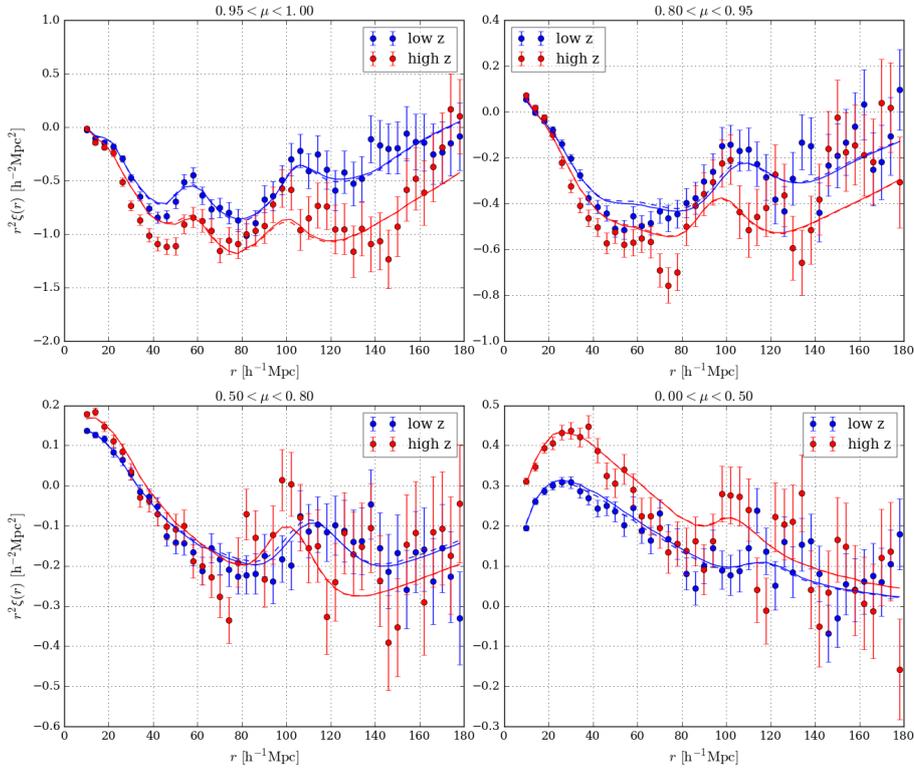
Transitions	$\lambda_1/\lambda_2$	$r_{\parallel}^{\text{app}}[h^{-1}\text{Mpc}]$
Si ii(1190)/O vi(1038)	1.147	-88
Si ii(1193)/O vi(1038)	1.150	-81
Si ii(1190)/O vi(1032)	1.154	-73
Si ii(1193)/O vi(1032)	1.156	-66
Si ii(1190)/Ly $\beta$ (1026)	1.161	-56
Si iii(1207)/O vi(1038)	1.163	-52
Si ii(1193)/Ly $\beta$ (1026)	1.163	-50
Si iii(1207)/O vi(1032)	1.169	-37
Ly $\alpha$ (1216)/O vi(1038)	1.172	-31
Si iii(1207)/Ly $\beta$ (1026)	1.176	-20
Ly $\alpha$ (1216)/O vi(1032)	1.178	-16
Si ii(1260)/O vi(1038)	1.215	68
Si ii(1260)/O vi(1032)	1.221	83
Si ii(1260)/Ly $\beta$ (1026)	1.229	100

correlation further overlaps the BAO signal at small  $r_{\perp}$ , hampering its detection in our data.

Figure C.2 presents for the first time the 2D Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ) correlation function. It is shown in the usual four wedges of  $\mu$  values, as a function of  $r = \sqrt{r_{\parallel}^2 + r_{\perp}^2}$  multiplied by the sign of  $r_{\parallel}$ . Note that the model is not symmetric around zero separation.

The oscillator strength of Ly $\beta$  absorption is a fifth of that of Ly $\alpha$ , and consequently there are far fewer Ly $\beta$  HCD systems than Ly $\alpha$  HCD systems. On the other hand Ly $\beta$  absorption in our analysis occurs at a systematically higher redshift. Overall we find that  $b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\beta(\text{Ly}\beta)}$  is consistent with zero.

In summary, there are not enough data at present to constrain the BAO peak position with Ly $\beta$  absorption only. However, Ly $\beta$  absorption could be used to access to physical IGM parameters at redshifts for which the Ly $\alpha$  absorption is saturated (Dijkstra et al. (2004), Iršič & Viel (2014)).



**Fig. B.3.** Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) function in four ranges of  $\mu = r_{\parallel}/r$  values, computed in a low redshift (blue dots) and in a high redshift (red dots) bin obtained by splitting our sample so that the two bins have equivalent statistical power. The dashed lines correspond to the simple fits to the data of the Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) correlation function, the solid line, to the combining fits to the data of the Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) and Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\beta$ ) correlation functions.

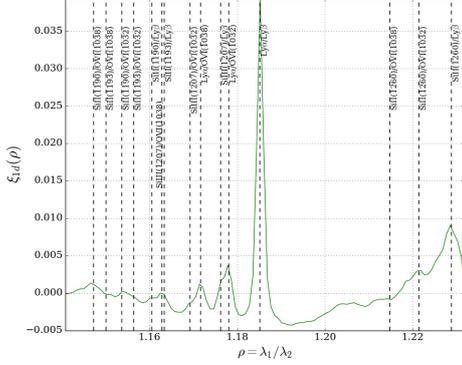
#### Appendix D: Confidence levels

To make a precise estimate of the relation between  $\Delta\chi^2$  and confidence level, we followed closely the procedure of du Mas des Bourboux et al. (2017). We generated a large number of simulated correlation functions using the fiducial cosmological model and the best-fit values of non-BAO parameters, randomized using the covariance matrix measured with the data. Each simulated correlation function was then fit for the model parameters and the  $\chi^2$  for the best-fit parameters compared with the best  $\chi^2$  with one or more parameters set to the known input values. Confidence levels are the fractions of the generated data sets that have best fits below the  $\Delta\chi^2$  limit. The uncertainties are estimated using a bootstrap technique.

The analysis of du Mas des Bourboux et al. (2017) followed this procedure using models that incorporated only Ly $\alpha$  absorption and models that incorporated also HCDs and metals. Since no significant differences were seen in the two

methods, we use here only Ly $\alpha$  absorption, considerable simplifying the analysis.

The results are summarized in Table D.1 for various correlation functions. In all cases the  $\Delta\chi^2$  values corresponding to a given confidence level are increased above the standard values. For example, for the Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ ) correlation, the one- and two-standard deviation contours for  $(\alpha_{\parallel}, \alpha_{\perp})$  correspond to  $\Delta\chi^2 = 2.77$  and  $\Delta\chi^2 = 7.33$  to be compared with the standard values of 2.29 and 6.18.



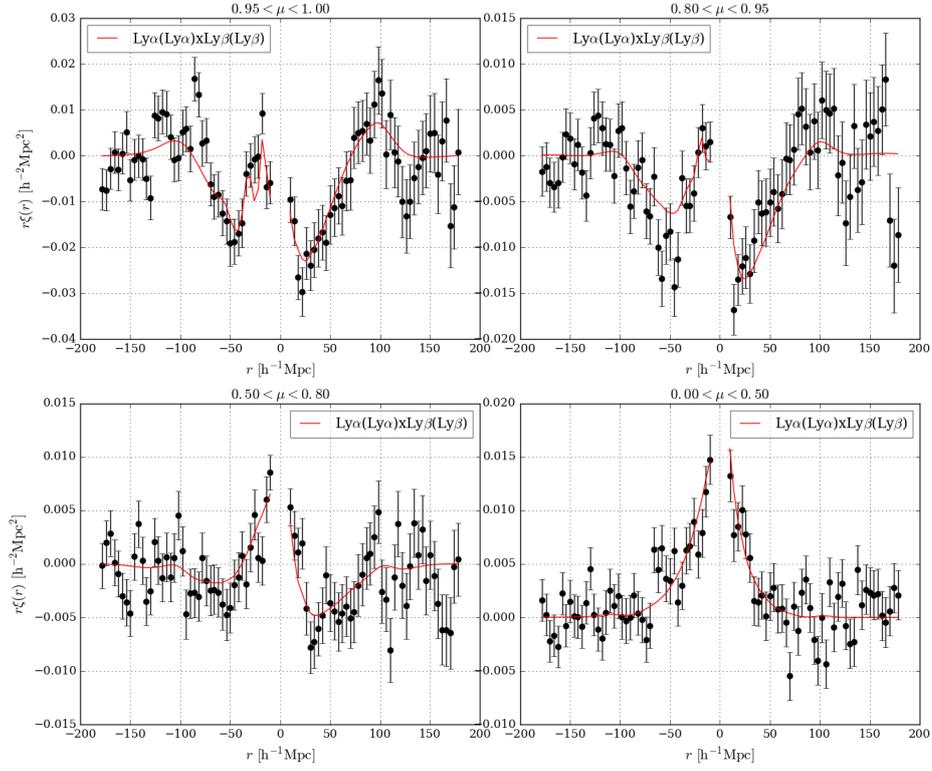
**Fig. C.1.** Same as Fig. 4 for cross-correlation function of Ly $\alpha$  with Ly $\beta$  regions, as a function of the ratio of transition wavelengths.

**Table C.2.** Results of the combining fit on Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ) , Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) and Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ) correlation functions with the BAO parameters ( $\alpha_{\parallel}$ ,  $\alpha_{\perp}$ ) fixed to 1.

Parameters	Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ ) + Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\beta$ ) + Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\beta$ (Ly $\beta$ )
$\beta_{\text{Ly}\alpha}$	$1.840 \pm 0.084$
$b_{\eta\text{Ly}\alpha}$	$-0.212 \pm 0.004$
$\beta_{\text{Ly}\beta}$	$1.123 \pm 0.384$
$b_{\eta\text{Ly}\beta}$	$-0.098 \pm 0.018$
$\beta_{\text{HCD}}$	$1.116 \pm 0.152$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\alpha(\text{Ly}\alpha)}$	$-0.050 \pm 0.004$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\alpha(\text{Ly}\beta)}$	$-0.073 \pm 0.005$
$b_{\text{HCD}}^{\text{Ly}\alpha(\text{Ly}\alpha)\times\text{Ly}\beta(\text{Ly}\beta)}$	$-0.002 \pm 0.032$
$b_{\text{OVI}(1032)}$	$-0.0081 \pm 0.0015$
$b_{\text{OVI}(1038)}$	$-0.0055 \pm 0.0014$
$b_{\text{SiIII}(1190)}$	$-0.0025 \pm 0.0005$
$b_{\text{SiIII}(1193)}$	$-0.0022 \pm 0.0005$
$b_{\text{SiIII}(1207)}$	$-0.0035 \pm 0.0005$
$b_{\text{SiIII}(1260)}$	$-0.0011 \pm 0.0006$
$b_{\text{CIV}(eff)}$	$-0.0047 \pm 0.0025$
$\chi^2_{\text{min}}/DOF$	$6469.77/(6360-15)$
Probability	0.134

**Table D.1.** Values of  $\Delta\chi^2$  corresponding to confidence levels (CLs) (68.27, 95.45%). Values are derived from 10,000 Monte Carlo simulations of the correlation function that are fit using the model containing only Ly $\alpha$  absorption.

Parameter	$\Delta\chi^2$ (68.27%)	$\Delta\chi^2$ (95.45%)
Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ )		
$\alpha_{\parallel}$	$1.19 \pm 0.03$	$4.74 \pm 0.09$
$\alpha_{\perp}$	$1.23 \pm 0.03$	$4.83 \pm 0.08$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.77 \pm 0.04$	$7.33 \pm 0.10$
Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ ) + QSO $\times$ Ly $\alpha$ (Ly $\alpha$ +Ly $\beta$ )		
$\alpha_{\parallel}$	$1.08 \pm 0.02$	$4.29 \pm 0.10$
$\alpha_{\perp}$	$1.08 \pm 0.02$	$4.28 \pm 0.10$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.47 \pm 0.03$	$6.71 \pm 0.13$
Ly $\alpha$ (Ly $\alpha$ ) $\times$ Ly $\alpha$ (Ly $\alpha$ )		
$\alpha_{\parallel}$	$1.19 \pm 0.02$	$4.65 \pm 0.09$
$\alpha_{\perp}$	$1.17 \pm 0.02$	$4.32 \pm 0.07$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.65 \pm 0.04$	$6.99 \pm 0.10$
low z		
$\alpha_{\parallel}$	$1.28 \pm 0.02$	$5.09 \pm 0.08$
$\alpha_{\perp}$	$1.35 \pm 0.02$	$5.09 \pm 0.10$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.89 \pm 0.04$	$7.55 \pm 0.12$
high z		
$\alpha_{\parallel}$	$1.28 \pm 0.03$	$4.92 \pm 0.09$
$\alpha_{\perp}$	$1.23 \pm 0.03$	$4.74 \pm 0.08$
$(\alpha_{\parallel}, \alpha_{\perp})$	$2.83 \pm 0.04$	$7.44 \pm 0.11$



**Fig. C.2.** Measured Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ) correlation function in four ranges of  $\mu$ . In order to see the asymmetry of this correlation function,  $r = (r_{\parallel}^2 + r_{\perp}^2)^{1/2}$  is multiplied by the sign of  $r_{\parallel}$  which is positive if the Ly $\alpha$  absorber is farther than the Ly $\beta$  absorber, and negative in the opposite configuration. The model for this correlation function (red solid line) is fitted on Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\alpha$ ), Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\alpha$ (Ly $\beta$ ) and Ly $\alpha$ (Ly $\alpha$ )  $\times$  Ly $\beta$ (Ly $\beta$ ).

---

M



# Bibliographie

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42
- Aihara, H., Allende Prieto, C., An, D., et al. 2011, *ApJS*, 193, 29
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, 219, 12
- Alam, S., Ata, M., Bailey, S., et al. 2017, *MNRAS*, 470, 2617
- Arinyo-i-Prats, A., Miralda-Escudé, J., Viel, M., & Cen, R. 2015, *J. Cosmology Astropart. Phys.*, 2015, 017
- Astier, P. 2012, arXiv e-prints, arXiv :1211.2590
- Astier, P., Guy, J., Regnault, N., et al. 2006, *A&A*, 447, 31
- Ata, M., Baumgarten, F., Bautista, J., et al. 2018, *MNRAS*, 473, 4773
- Ata, M., Kitaura, F.-S., Chuang, C.-H., et al. 2017, *MNRAS*, 467, 3993
- Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, *ApJ*, 327, 570
- Bautista, J. E., Bailey, S., Font-Ribera, A., et al. 2015, *J. Cosmology Astropart. Phys.*, 2015, 060
- Bautista, J. E., Busca, N. G., Guy, J., et al. 2017, *A&A*, 603, A12 (B17)
- Bautista, J. E., Vargas-Magaña, M., Dawson, K. S., et al. 2018, *ApJ*, 863, 110
- Betoule, M., Kessler, R., Guy, J., et al. 2014, *A&A*, 568, A22
- Beutler, F., Blake, C., Colless, M., et al. 2011, *MNRAS*, 416, 3017
- Blanton, M. R., Dalcanton, J., Eisenstein, D., et al. 2001, *AJ*, 121, 2358
- Blomqvist, M., du Mas des Bourboux, H., Busca, N. G., et al. 2019, arXiv e-prints, arXiv :1904.03430
- Blomqvist, M., Pieri, M. M., du Mas des Bourboux, H., et al. 2018, *J. Cosmology Astropart. Phys.*, 2018, 029
- Bolton, A. S., Schlegel, D. J., Aubourg, É., et al. 2012, *AJ*, 144, 144
- Bonvin, C. 2014, *Classical and Quantum Gravity*, 31, 234002
- Busca, N. & Balland, C. 2018, arXiv e-prints, arXiv :1808.09955

- Busca, N. G., Delubac, T., Rich, J., et al. 2013, *A&A*, 552, A96
- Chabanier, S., Palanque-Delabrouille, N., Yèche, C., et al. 2018, arXiv e-prints, arXiv :1812.03554
- Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, *MNRAS*, 362, 505
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, 151, 44
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
- de Sainte Agathe, V., Balland, C., du Mas des Bourboux, H., et al. 2019, arXiv e-prints, arXiv :1904.03400
- Delubac, T., Bautista, J. E., Busca, N. G., et al. 2015, *A&A*, 574, A59
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv :1611.00036
- Dodelson, S. 2003, *Modern cosmology* (San Diego, CA : Academic Press)
- du Mas des Bourboux, H., Dawson, K. S., Busca, N. G., et al. 2019, *ApJ*, 878, 47
- du Mas des Bourboux, H., Le Goff, J.-M., Blomqvist, M., et al. 2017, *A&A*, 608, A130
- Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, *ApJ*, 664, 675
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
- Font-Ribera, A., Arnau, E., Miralda-Escudé, J., et al. 2013, *J. Cosmology Astropart. Phys.*, 2013, 018
- Font-Ribera, A. & Miralda-Escudé, J. 2012, *J. Cosmology Astropart. Phys.*, 2012, 028
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, 622, 759
- Gunn, J. E., Carr, M., Rockosi, C., et al. 1998, *AJ*, 116, 3040
- Gunn, J. E. & Peterson, B. A. 1965, *ApJ*, 142, 1633
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al. 2006, *AJ*, 131, 2332
- Hamilton, A. J. S. 1998, in *Astrophysics and Space Science Library*, Vol. 231, *The Evolving Universe*, ed. D. Hamilton, 185
- Huterer, D. & Shafer, D. L. 2018, *Reports on Progress in Physics*, 81, 016901
- Kaiser, N. 1984, *ApJ*, 284, L9
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kirkby, D., Margala, D., Slosar, A., et al. 2013, *J. Cosmology Astropart. Phys.*, 2013, 024
- Kolb, E. W. & Turner, M. S. 1990, *The early universe*, Vol. 69
- Lahav, O., Lilje, P. B., Primack, J. R., & Rees, M. J. 1991, *MNRAS*, 251, 128
- Lee, K.-G., Bailey, S., Bartsch, L. E., et al. 2013, *AJ*, 145, 69

- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
- Madrid, J. P. & Macchetto, D. 2009, in *BAAS*, Vol. 41, 913–914
- McDonald, P. 2003, *ApJ*, 585, 34
- McDonald, P. & Eisenstein, D. J. 2007, *Phys. Rev. D*, 76, 063009
- McDonald, P., Seljak, U., Burles, S., et al. 2006, *ApJS*, 163, 80
- McQuinn, M. & White, M. 2011, *MNRAS*, 415, 2257
- Murdin, P. 2001, *Encyclopedia of astronomy and astrophysics*
- Noterdaeme, P., Petitjean, P., Carithers, W. C., et al. 2012, *A&A*, 547, L1
- O’Raifeartaigh, C. & Mitton, S. 2018, *Physics in Perspective*, 20, 318
- Palanque-Delabrouille, N., Yèche, C., Borde, A., et al. 2013, *A&A*, 559, A85
- Pâris, I., Petitjean, P., Aubourg, É., et al. 2012, *A&A*, 548, A66
- Pâris, I., Petitjean, P., Aubourg, É., et al. 2018, *A&A*, 613, A51
- Peebles, P. J. E. 1980, in *Ninth Texas Symposium on Relativistic Astrophysics*, Vol. 336, 161–171
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565
- Pieri, M. M., Frank, S., Weinberg, D. H., Mathur, S., & York, D. G. 2010, *ApJ*, 724, L69
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, *A&A*, 594, A13
- Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv :1807.06209
- Prakash, A. & SDSS-IV/eBOSS. 2016, in *American Astronomical Society Meeting Abstracts*, Vol. 227, *American Astronomical Society Meeting Abstracts #227*, 439.01
- Raichoor, A., Comparat, J., Delubac, T., et al. 2017, *MNRAS*, 471, 3955
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, 116, 1009
- Robertson, H. P. 1936, *ApJ*, 83, 257
- Rogers, K. K., Bird, S., Peiris, H. V., et al. 2018, *MNRAS*, 474, 3032
- Ross, A. J., Samushia, L., Howlett, C., et al. 2015, *MNRAS*, 449, 835
- Schmidt, M. 1963, *Nature*, 197, 1040
- Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, *ApJ*, 859, 101
- Seljak, U. 2012, *Journal of Cosmology and Astro-Particle Physics*, 2012, 004
- Shields, G. A. 1999, *PASP*, 111, 661
- Slosar, A., Font-Ribera, A., Pieri, M. M., et al. 2011, *J. Cosmology Astropart. Phys.*, 9, 001
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, *J. Cosmology Astropart. Phys.*, 4, 26

## Bibliographie

---

- Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, AJ, 146, 32
- Sullivan, M., Guy, J., Conley, A., et al. 2011, ApJ, 737, 102
- Sutherland, W. & Mularczyk, L. 2014, MNRAS, 438, 3128
- Walker, A. G. 1937, Proceedings of the London Mathematical Society, 42, 90
- Weinberg, S. 1972, Gravitation and Cosmology : Principles and Applications of the General Theory of Relativity
- Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, ApJ, 373, 23
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579
- Zwicky, F. 1933, Helvetica Physica Acta, 6, 110
- Zwicky, F. 1937, ApJ, 86, 217



---

**Résumé :** La propagation des oscillations acoustiques baryoniques a laissé une empreinte dans la répartition de la matière de l'univers, visible sous la forme d'un excès de probabilité pour deux objets d'être séparés par une distance égale à l'échelle acoustique. La mesure de l'échelle acoustique dans la fonction de corrélation à deux points de la matière au *redshift*  $z$ , parallèlement et perpendiculairement à la ligne de visée donne accès, respectivement, aux rapports  $D_H(z)/r_d$  et  $D_M(z)/r_d$  où  $D_H$  est la distance de Hubble,  $D_M$  la distance angulaire comobile et  $r_d$  l'horizon acoustique. On peut suivre la répartition de la matière en utilisant les absorptions Lyman- $\alpha$  visibles, sous la forme de forêts, dans les spectres des quasars à haut *redshift*. Puisque chaque spectre donne accès une centaine d'absorptions, cela permet de mesurer l'échelle acoustique même quand la densité de quasars observés est faible. Dans cette thèse, je décris le processus d'analyse qui, en utilisant environ 200 000 spectres de quasars du relevé eBOSS-SDSS IV, aboutit aux mesures  $D_H(2.34)/r_d = 8.86 \pm 0.29$  et  $D_M(2.34)/r_d = 37.41 \pm 1.86$ . En combinant ces résultats avec les mesures de l'échelle acoustique à d'autres *redshifts*, j'obtiens la plus forte contrainte actuelle, à bas *redshift*, sur les paramètres  $\Omega_m$  et  $\Omega_\Lambda$  dans le cadre du modèle  $\Lambda$ CDM.

---

**Mots-clés :** cosmologie, expansion de l'univers, énergie noire, oscillations acoustiques baryoniques, forêts Lyman- $\alpha$

---

**Abstract:** The propagation of the baryonic acoustic oscillations has been unprinted in the matter distribution in the Universe as a probability excess for two objects to be separated by the acoustic scale. Measuring the acoustic scale in the matter 2 point correlation function at *redshift*  $z$ , along and transversally to the line-of-sight, gives access to the  $D_H(z)/r_d$  et  $D_M(z)/r_d$  ratios, with  $D_H$  the Hubble distance,  $D_M$  the comoving angular distance and  $r_d$  the acoustic horizon. We are able to trace the matter in the Universe by using the Lyman- $\alpha$  absorptions which shape the spectra of the high *redshift* quasars. Since each spectrum contain hundreds of absorption, this allow us to measure the acoustic scale even if the observed quasar density is low. In this thesis, I describe the analysis of about 200,000 spectra from the eBOSS-SDSS IV survey which conducts to the measurements  $D_H(2.34)/r_d = 8.86 \pm 0.29$  et  $D_M(2.34)/r_d = 37.41 \pm 1.86$ . By combining these results with measurements of the acoustic scale at other *redshifts*, I obtain the strongest current constraints at low *redshift* on the  $\Omega_m$  and  $\Omega_\Lambda$   $\Lambda$ CDM parameters.

---

**Key words:** cosmology, expansion of the universe, dark energy, baryonic acoustic oscilations, Lyman- $\alpha$  forests

---