



**HAL**  
open science

# Détections des oscillations acoustiques de baryons grâce aux forêts Lyman- $\alpha$ des spectres de quasars de l'expérience BOSS

Timothée Delubac

► **To cite this version:**

Timothée Delubac. Détections des oscillations acoustiques de baryons grâce aux forêts Lyman- $\alpha$  des spectres de quasars de l'expérience BOSS. Cosmologie et astrophysique extra-galactique [astro-ph.CO]. Université Paris-Diderot - Paris VII, 2013. Français. NNT : . tel-00936622

**HAL Id: tel-00936622**

**<https://theses.hal.science/tel-00936622v1>**

Submitted on 27 Jan 2014

**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.

# THÈSE

présentée par

**Timothée Delubac**

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS DIDEROT

Spécialité : Champs, Particules, Matières

Ecole doctorale : Particule, Noyaux, Cosmos (ED 517)

---

## **Détections des oscillations acoustiques de baryons grâce aux forêts Lyman- $\alpha$ des spectres de quasars de l'expérience BOSS**

---

Soutenue le 13 Septembre 2013 devant le jury composé de :

J. G.	Bartlett	Président
G.	Chardin	
R.	Croft	Rapporteur
J.-P.	Kneib	
P.	Petitjean	Rapporteur
Ch.	Yèche	Directeur de thèse

Thèse préparée au Service de Physique des Particules de l'IRFU du CEA de Saclay.



# Merci ...

---

... à mon directeur de thèse, Christophe Yèche, pour m'avoir ouvert les portes du Service de Physique des Particules (SPP) ainsi que de la collaboration SDSS-III. Je n'en étais pas pleinement conscient au commencement de ma thèse, je le suis aujourd'hui : évoluer au sein de ces deux organismes fut un privilège. Privilège par la richesse scientifique à laquelle j'ai eu accès et dont, je l'espère, ces travaux de thèse rendent compte. Privilège également par la qualité des personnes qui les composent, à qui certaines des lignes qui suivent sont dédiées. Christophe, je souhaite également te remercier pour le juste équilibre que tu as su trouver dans ton encadrement, entre rigueur et autonomie, qui m'a permis d'évoluer aussi sereinement qu'il m'est possible, au long de ces trois années.

---

... aux chercheurs du groupes BOSS du SPP : Nathalie Palanque-Delabrouille, Jean-Marc LeGoff, James Rich, Eric Armengaud et Etienne Burtin. J'y associe également ceux du groupe de cosmologie, Christophe Magneville, Vanina Ruhlmann-Kleider, Jean-Baptiste Melin et Dominique Yvon. Je n'ai eu l'occasion de travailler qu'auprès d'une partie d'entre vous, mais les nombreux échanges que nous avons eus, scientifiques ou non, ont enrichi cette expérience.

---

... aux membres du jury, Rupert Croft, Patrick Petitjean, James G. Bartlett, Gabriel Chardin et Jean-Paul Kneib pour leur présence lors de ma soutenance ainsi que pour leur précieuse évaluation. Je renouvelle ici mes remerciements à Rupert Croft et Patrick Petitjean pour leurs commentaires sur ce manuscrit. Merci Jean-Paul de me permettre de continuer l'aventure.

---

... à Nicolás Busca et James Rich avec qui j'ai, progressivement, noué une collaboration solide et qui ont laissé une empreinte fondamentale sur ces travaux. Nicolás, Jim, je tiens également à vous remercier pour votre amitié. Jim, sache que j'admire ta capacité d'émerveillement ainsi que ta passion pour la physique, et notamment la cosmologie, restées sans failles au cours de ces trois années. Je te suis reconnaissant quant à ta disponibilité, l'étendue de tes enseignements, ainsi que pour m'avoir mis, et continuer à me mettre en avant.

---

... aux doctorants et "non-permanents" qui ont, par leur amitié, étendu la dimension humaine de cette expérience. Romain, Anais, Céline, Marie-Cécile, Emilie et Homero, vous étiez autant de raisons de venir au laboratoire quand la fatigue se faisait sentir.

---

... à mes proches pour leur soutien durant ces trois années et notamment la dernière qui fut un défi moral autant qu'intellectuel. Elisabeth et Pénélope, merci de m'avoir aidé à préserver, autant que possible, la langue française. Guy, merci d'avoir manifesté ton intérêt pour le contenu de ce manuscrit. Andrée, Yvonne et Marcel, merci pour vos encouragements qui furent une source de motivation. Merci, Olja, toi qui fut la plus exposée à mes agitations, pour ta patience et ta compréhension.

---



# Table des matières

<b>Introduction</b>	<b>9</b>
<b>1 Cosmologie moderne</b>	<b>11</b>
1.1 Du contenant : l'espace . . . . .	11
1.1.1 Coordonnées et métrique . . . . .	11
1.1.2 Principe cosmologique I . . . . .	12
1.1.3 Trois univers possibles . . . . .	12
1.1.4 L'Univers dynamique : la métrique FLRW . . . . .	13
1.1.5 L'Univers en expansion . . . . .	14
1.2 Du contenu : l'énergie . . . . .	14
1.2.1 Conservation de l'énergie . . . . .	14
1.2.2 Matière non-relativiste . . . . .	15
1.2.3 Matière relativiste . . . . .	15
1.3 Du lien entre contenant et contenu . . . . .	16
1.3.1 Equation d'Einstein . . . . .	16
1.3.2 Principe cosmologique II . . . . .	16
1.3.3 Equations de Friedmann . . . . .	17
1.3.4 Energie noire . . . . .	18
1.3.5 Les $\Omega$ s . . . . .	19
1.4 Contraindre le contenant grâce au contenu et <i>vice versa</i> . . . . .	19
1.4.1 Distance comobile . . . . .	20
1.4.2 Distance de luminosité . . . . .	20
1.4.3 Distance angulaire . . . . .	21
1.4.4 Mesure directe de $H(z)$ . . . . .	24
1.4.5 La distance $D_V(z)$ . . . . .	25
<b>2 Oscillations acoustiques de baryons et forêts Lyman-<math>\alpha</math></b>	<b>27</b>
2.1 Musicologie de l'Univers . . . . .	28
2.1.1 Fluctuations primordiales . . . . .	28
2.1.2 Propagation acoustique des fluctuations . . . . .	28
2.1.3 Horizon acoustique . . . . .	29

2.2	Fond diffus cosmologique . . . . .	31
2.2.1	Descriptions des fluctuations . . . . .	31
2.2.2	$\Lambda$ CDM : un modèle à 6 paramètres . . . . .	33
2.2.3	Paramètres dérivés utiles . . . . .	34
2.3	Spectre de puissance de la matière . . . . .	35
2.3.1	Spectre de puissance . . . . .	35
2.3.2	Biais et distortions en <i>redshift</i> . . . . .	36
2.3.3	Spectre de puissance et fonction de corrélation . . . . .	39
2.3.4	Fiabilité de la mesure du pic BAO . . . . .	39
2.3.5	Exemple : le relevé de galaxies CMASS . . . . .	39
2.4	Un nouveau traceur : les forêts Lyman- $\alpha$ . . . . .	41
2.4.1	<i>Quasi-stellar radio sources</i> . . . . .	41
2.4.2	Effet Gunn-Peterson et forêts Lyman- $\alpha$ . . . . .	41
2.4.3	Cas particuliers de spectres : DLA et BAL . . . . .	43
<b>3</b>	<b>The Baryon Oscillation Spectroscopic Survey</b> . . . . .	<b>47</b>
3.1	The Sloan Digital Sky Survey . . . . .	47
3.1.1	SDSS et SDSS-II . . . . .	47
3.1.2	SDSS-III . . . . .	48
3.2	Objectifs de BOSS . . . . .	49
3.3	Stratégie d'observation . . . . .	50
3.3.1	Phase photométrique . . . . .	50
3.3.2	Phase spectroscopique . . . . .	52
3.4	Caractéristiques techniques de BOSS . . . . .	54
3.4.1	Le télescope . . . . .	55
3.4.2	La caméra . . . . .	57
3.4.3	Les spectrographes . . . . .	57
3.5	Réduction des données . . . . .	66
3.5.1	Réduction automatique des données . . . . .	66
3.5.2	Inspection visuelle . . . . .	67
<b>4</b>	<b>Sélection des quasars</b> . . . . .	<b>69</b>
4.1	Sélection par couleurs . . . . .	69
4.2	Sélection par variabilité . . . . .	70
4.2.1	La région du <i>Stripe 82</i> . . . . .	70
4.2.2	Echantillons tests . . . . .	70
4.2.3	Nettoyage des courbes de lumière . . . . .	71
4.2.4	Quantification de la variabilité . . . . .	71
4.2.5	Choix des paramètres . . . . .	75
4.2.6	Sélection des quasars par réseau de neurones . . . . .	76

4.2.7	Performance de la sélection . . . . .	78
4.3	Application à la région du Stripe 82 . . . . .	80
4.3.1	Sélection principale . . . . .	80
4.3.2	Variabilité extrême . . . . .	82
4.3.3	Resultats des sélections . . . . .	83
4.3.4	Comparaison avec une sélection en couleurs . . . . .	84
4.4	Prévisions pour les relevés futurs . . . . .	86
4.4.1	Extrapolation pour le Palomar Transient Factory . . . . .	86
4.4.2	Extrapolation pour Pan-STARRS1 . . . . .	87
4.5	Application à l'ensemble du relevé . . . . .	88
4.5.1	Echantillons . . . . .	88
4.5.2	Tests de variabilité et stratégie d'observation . . . . .	88
4.5.3	Résultats . . . . .	92
<b>5</b>	<b>Fonction de corrélation des forêts Lyman-<math>\alpha</math></b>	<b>95</b>
5.1	Echantillon de données . . . . .	95
5.1.1	Sélection de l'échantillon de quasar . . . . .	96
5.1.2	Définition de la forêt Lyman- $\alpha$ . . . . .	97
5.1.3	Pré-traitement des données . . . . .	98
5.2	Mesure de la fonction de corrélation . . . . .	99
5.2.1	Estimateur . . . . .	100
5.2.2	Estimation du flux moyen attendu . . . . .	101
5.2.3	Champ de $\delta$ . . . . .	104
5.2.4	Détermination des poids . . . . .	104
5.2.5	$\hat{\xi}(r_{\parallel}, r_{\perp})$ . . . . .	107
5.2.6	$\hat{\xi}(r, \mu)$ . . . . .	107
5.3	Etude en multipôles . . . . .	107
5.3.1	Détermination des multipôles . . . . .	107
5.3.2	Estimation de la covariance . . . . .	110
5.4	Comparaison aux simulations . . . . .	114
5.4.1	Production des simulations . . . . .	116
5.4.2	Analyse des simulations . . . . .	117
5.5	Détection et mesure de la position du pic . . . . .	118
5.5.1	Paramétrisation de l'ajustement . . . . .	118
5.5.2	Ajustement des simulations . . . . .	121
5.5.3	Données : ajustement isotrope et niveau de confiance de la détection . . .	121
5.5.4	Données : ajustement en deux dimensions . . . . .	125
5.6	Recherche de systématiques . . . . .	127
5.6.1	Nord - Sud . . . . .	129
5.6.2	Coupure des raies de Balmer . . . . .	129



5.6.3	Plaques . . . . .	129
5.6.4	Conclusion sur la recherche de systématiques . . . . .	131
5.7	Contraintes cosmologiques . . . . .	133
5.7.1	Contraintes sur $D_A/r_s$ et $Hr_s$ . . . . .	133
5.7.2	Contraintes sur les paramètres cosmologiques . . . . .	134
5.7.3	Contraintes combinées avec Planck . . . . .	137
5.8	Perspectives . . . . .	138
<b>Conclusion</b>		<b>141</b>
<b>A Métrique de Friedmann-Lemaître-Robertson-Walker</b>		<b>143</b>
A.1	Les trois espaces possibles . . . . .	143
A.1.1	3-sphère . . . . .	143
A.1.2	3-hyperboloïde . . . . .	144
A.1.3	Espace euclidien . . . . .	144
A.1.4	Généralisation de la métrique . . . . .	145
<b>B Publication DR9</b>		<b>147</b>
<b>Liste des tableaux</b>		<b>167</b>
<b>Table des figures</b>		<b>169</b>
<b>Bibliographie</b>		<b>173</b>

# Introduction

Après une décennie passée dans le monde de l'astronomie amateur, ce travail marque mon entrée dans son pendant professionnel. A proprement parler, cette thèse traite de cosmologie et non d'astronomie. Si l'astronomie est l'étude du mouvement des astres, la distinction peut sembler ténue lorsque la cosmologie s'attache à comprendre le mouvement de récession des galaxies, résultant de la surprenante « énergie noire », ou à expliquer les courbes de rotation de ces mêmes galaxies, mettant en avant la présence de la déroutante « matière noire ».

La cosmologie, telle que nous l'étudions aujourd'hui, trouve ses fondements théoriques au coeur de la relativité générale, introduite par Einstein en 1916. Le chapitre 1 revient sur les concepts généraux de cette théorie bientôt centenaire. Dans un premier temps, il présente le couplage entre géométrie de l'Univers et contenu énergétique, formalisé par la fameuse équation d'Einstein. Il montre ensuite en quoi, comme pour l'astronome, la mesure de distance est une problématique primordiale du cosmologue.

La distinction entre astronomie et cosmologie se trouve certainement dans l'utilisation, par le cosmologue, de nouvelles sondes lui permettant d'appréhender l'Univers. Faisant appel à la physique des particules et à celle des plasmas, ces nouvelles sondes permettent d'étudier l'Univers en des lieux et des époques où il se trouve dépourvu d'astres. La plus connue d'entre elles est sans doute le fond diffus cosmologique, première lumière de l'Univers, émise 300.000 ans après le *Big Bang* et découverte en 1964. L'étude des anisotropies de ce fond diffus, perçues comme les germes des structures (galaxies, amas de galaxies, etc.) que nous observons aujourd'hui, est l'une des branches principales de la cosmologie actuelle. Le chapitre 2 détaille la physique des oscillations acoustiques de baryons, à l'origine de ces anisotropies. Il revient également sur les méthodes et les derniers résultats de l'analyse du fond diffus cosmologique. Il s'intéresse enfin à l'étude de l'hydrogène neutre du milieu intergalactique, grâce à l'effet Gunn-Peterson. Comme nous le verrons, cet effet est à l'origine des forêts Lyman- $\alpha$  des spectres des quasars. La recherche des empreintes des oscillations acoustiques de baryons dans la distribution spatiale d'hydrogène neutre du milieu intergalactique constitue l'objectif principal de cette thèse.

Bien que ce travail de thèse soit, pour moi, un commencement, il est en fait l'un des aboutissements d'une entreprise bien plus grande : le Sloan Digital Sky Survey III. Ce relevé constitue la troisième génération d'un projet d'étude des structures à grandes échelles de l'Univers, pensé au cours des années 80. Il est notamment constitué d'un relevé spectroscopique de galaxies et de quasars, the Baryon Oscillation Spectroscopic Survey (BOSS), dont l'un des objectifs affichés est la détection des oscillations acoustiques de baryons dans la distribution d'hydrogène neutre du milieu intergalactique. Le chapitre 3 présente ce projet, ainsi que les caractéristiques techniques des instruments utilisés pour le mener à bien.

L'une des difficultés d'un tel relevé spectroscopique réside dans la sélection des objets dont on souhaite obtenir les spectres. La méthode courante consiste à sélectionner les objets en fonction de leur magnitude dans différentes bandes photométriques. Le chapitre 4 traite d'une méthode récente de sélection des quasars basée sur la variabilité temporelle de la photométrie de ces objets.

Son application à la sélection des cibles du Baryon Oscillation Spectroscopic Survey représente l'un des travaux de cette thèse.

Le chapitre 5 présente enfin l'analyse de la distribution spatiale d'hydrogène neutre du milieu intergalactique grâce aux forêts Lyman- $\alpha$  des spectres de quasars observés par BOSS. La mise en place de cette analyse constitue l'essentiel du travail effectué au cours de cette thèse. En résulte la première détection des oscillations acoustiques de baryons à grand décalage spectral ( $z \sim 2,38$ ). Ce chapitre s'achève sur les contraintes cosmologiques résultant de cette détection.

# Chapitre 1

## Cosmologie moderne

*« Everything starts somewhere, though many physicists disagree. But people have always been dimly aware of the problem with the start of things. They wonder how the snowplough driver gets to work, or how the makers of dictionaries look up the spelling of words. »*

—T. Pratchett, Hogfather

---

Ce chapitre introduit les concepts généraux nécessaires à la compréhension de l'étude dont cette thèse fait l'objet. Il en est également la motivation. Il est inspiré de trois ouvrages de cosmologie, complémentaires par leur approche du sujet : *Modern Cosmology* de S. Dodelson, *Introduction to the Theory of the Early Univers* de D.S. Gorbunov et V.A. Rubakov, ainsi que *Fundamentals of Cosmology* de J. Rich. Il s'appuie également sur le cours de relativité générale de E.ourgoulhon. Sauf cas particulier, nous travaillerons avec la convention  $c = \hbar = 1$ .

---

### 1.1 Du contenant : l'espace

#### 1.1.1 Coordonnées et métrique

Nous, ainsi que toute la matière qui nous entoure, sommes piégés dans une variété<sup>1</sup> à 4 dimensions, dont trois d'espaces et une de temps. Pour décrire cette variété, il nous faut introduire deux concepts : un système de coordonnées et une métrique. Le système de coordonnées permet d'attribuer à chaque événement un temps  $t_0$  et une position  $x^i = (x^1, x^2, x^3)$ . Il est utile de concaténer les quatre coordonnées dans un seul et même quadri-vecteur  $x^\mu = (x^0, x^1, x^2, x^3)$  où  $x^0$  correspond à la coordonnée temporelle. La métrique permet de calculer le changement du temps

---

1. Une variété est un espace topologique localement euclidien. Une courbe est une variété de dimension 1 tandis qu'une surface est une variété de dimension 2.

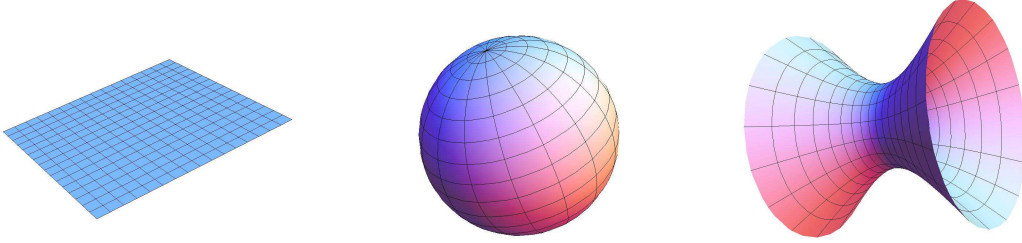


FIGURE 1.1: Représentation graphique des équivalents à deux dimensions des trois espaces vérifiant le principe cosmologique. À gauche le plan, ou espace euclidien, au centre la 2-sphère et à droite le 2-hyperboloïde.

propre d'une horloge idéale, passant de la coordonnée  $x^\mu$  à  $x^\mu + dx^\mu$  :

$$d\tau^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (1.1)$$

où les indices répétés sont sommés. La métrique  $g_{\mu\nu}$  est, par construction, symétrique ( $g_{\mu\nu} = g_{\nu\mu}$ ). Elle dépend à priori de la position  $x^\mu$  à laquelle elle est considérée. Le couple variété-métrique est appelé espace, dans le cas où l'une des dimensions est temporelle, on parle d'espace-temps.

### 1.1.2 Principe cosmologique I

L'idée, philosophique, que l'Univers ne doit pas contenir de position ni de direction privilégiées (c'est-à-dire que chaque observateur doit voir rigoureusement la même chose et ce, indépendamment de la position spatiale  $(x^1, x^2, x^3)$  où il se trouve et de la direction dans laquelle il regarde) nous amène à chercher les espaces<sup>2</sup> homogènes et isotropes. Un espace est dit homogène si et seulement si sa métrique ne change pas en passant d'un point à un autre. Il existe différents espaces homogènes mais tous ne sont pas isotropes. Un exemple d'espace homogène mais non isotrope en dimension deux est le cylindre qui présente une direction privilégiée. L'homogénéité et l'isotropie de l'espace correspondent au principe cosmologique.

### 1.1.3 Trois univers possibles

Il existe trois espaces de dimension 3, vérifiant le principe cosmologique : l'espace euclidien, la sphère de dimension 3 (ou 3-sphère) et l'hyperboloïde de dimension 3 (ou 3-hyperboloïde) (voir figure 1.1). Ces trois espaces possèdent des métriques différentes ; il est toutefois possible (voir annexe A pour la démonstration) d'obtenir une expression généralisée de la métrique, valable pour les trois espaces :

$$d\tau^2 = dt^2 - a^2 (d\chi^2 + f^2(\chi) (d\theta^2 + \sin^2\theta d\phi^2)) \quad (1.2)$$

2. Insistons sur le fait que nous parlons ici d'espace à trois dimensions (les trois dimensions spatiales de l'espace-temps à 4 dimensions).

où

$$f(\chi) = \begin{cases} \sin\chi & \text{3-sphère} \\ \chi & \text{espace plat} \\ \sinh\chi & \text{3-hyperboloïde} \end{cases} \quad (1.3)$$

et  $a$  est le facteur d'échelle.  $a$  représente le rayon de courbure  $R$  pour une 3-sphère ou un 3-hyperboloïde, tandis qu'il est un réel strictement positif dépendant du choix de coordonnées pour un espace plat. Nous avons utilisé les coordonnées sphériques  $(\chi, \theta, \phi)$ , particulièrement adaptées au traitement d'un espace homogène et isotrope. Un univers possédant la géométrie d'une 3-sphère est dit fermé, car est de volume fini, tandis qu'un univers possédant la géométrie d'un 3-hyperboloïde est dit ouvert.

#### 1.1.4 L'Univers dynamique : la métrique FLRW

Le principe cosmologique nous impose l'homogénéité et l'isotropie de l'espace, et non de l'espace-temps. Il n'interdit donc pas que la partie spatiale de la métrique présente une dépendance temporelle à condition que celle-ci préserve l'homogénéité et l'isotropie à chaque instant. Il est ainsi possible de généraliser la métrique 1.2 aux univers dynamiques en permettant au facteur d'échelle  $a$  de varier en fonction du temps. Il s'agit alors de la métrique de Friedmann-Lemaître-Robertson-Walker (FLRW). Pour retrouver l'expression usuelle de cette métrique il suffit d'effectuer le changement de variable suivant :

$$r = \begin{cases} \sin\chi & \text{3-sphère} \\ \chi & \text{espace plat} \\ \sinh\chi & \text{3-hyperboloïde} \end{cases} \quad (1.4)$$

L'équation 1.2 devenant alors :

$$d\tau^2 = dt^2 - a(t)^2 \left( \frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \right) \quad (1.5)$$

où

$$k = \begin{cases} +1 & \text{3-sphère} \\ 0 & \text{espace plat} \\ -1 & \text{3-hyperboloïde} \end{cases} \quad (1.6)$$

Pour un univers courbe, le facteur d'échelle est le rayon de courbure. Un univers courbe dynamique correspond donc à un univers possédant un rayon de courbure variable dans le temps. Un espace plat ne possède pas d'échelle caractéristique, la valeur de  $a$  n'a donc pas de sens physique. La quantité ayant un sens physique pour un tel univers est le paramètre de Hubble qui quantifie la vitesse de variation du facteur d'échelle :

$$H \equiv \frac{\dot{a}}{a} \quad (1.7)$$

De manière générale et dans le but d'alléger les notations, la dépendance temporelle des paramètres n'est pas explicitée, de sorte que  $a(t) \equiv a$ . Il est également habituel de désigner les paramètres évalués au temps présent  $t_0$  par l'indice 0 si bien que  $a(t_0) \equiv a_0$ . Il est toujours possible de changer de système de coordonnées de sorte que  $a_0 = 1$  ; nous travaillerons systématiquement avec cette convention et omettrons généralement le terme  $a_0$  des équations.

### 1.1.5 L'Univers en expansion

Un univers statique correspond au cas particulier pour lequel  $a$  est une constante. Un univers en expansion vérifie  $\dot{a} > 0$  tandis que pour un univers en contraction  $\dot{a} < 0$ . Ces trois types d'univers possèdent des passés ainsi que des futurs entièrement différents, si bien que le signe de  $\dot{a}$  de notre propre Univers fût l'une des questions cruciales de la cosmologie du siècle dernier. Il est aujourd'hui clairement établi que notre Univers est en expansion. Cette conclusion, base de la théorie du *Big Bang*, est supportée par trois « piliers » observationnels :

- La loi de Hubble : Loi résultant de l'observation que les galaxies s'éloignent de la notre avec, en moyenne, une vitesse  $v$  proportionnelle à leur distance physique  $d$ . Le coefficient de proportionnalité  $H_0 = \frac{v}{d}$  est connu sous le nom de constante de Hubble (elle équivaut au paramètre de Hubble pris à l'instant  $t_0$ ).
- Le Fond Diffus Cosmologique : Rayonnement de corps noir, relique d'une transition de phase ayant eu lieu suite au refroidissement de l'Univers, conséquence de l'expansion.
- La nucléosynthèse primordiale : Capacité du modèle d'univers en expansion à expliquer l'abondance observée des éléments légers (hydrogène et hélium).

## 1.2 Du contenu : l'énergie

Jusqu'à présent nous avons considéré les propriétés de l'espace dans lequel nous évoluons. Nous avons notamment décrit les espaces satisfaisant au principe cosmologique et montré que de tels espaces peuvent être dynamiques. Dans cette section nous nous intéressons au comportement de la matière ordinaire se trouvant dans un espace en expansion. Notons que les résultats seraient identiques (au signe de  $\dot{a}$  près) dans un univers en contraction.

### 1.2.1 Conservation de l'énergie

La matière peut être divisée en deux catégories : la matière non-relativiste et la matière relativiste. La matière non-relativiste, parfois simplement dénommée matière, est constituée de particules ayant une vitesse  $v$  négligeable devant la vitesse de la lumière :  $v \ll c$ . La matière relativiste, également appelée rayonnement, possède au contraire une vitesse proche de celle de la lumière :  $v \approx c$ .

Considérons un univers dynamique constitué de matière ordinaire, de densité totale  $\rho = \rho_m + \rho_r$ , où l'indice  $m$  désigne la matière et l'indice  $r$  le rayonnement. D'après la première loi de la thermodynamique, la variation d'énergie d'un système d'énergie interne  $U$  et de pression  $p$  étant soumis à une expansion adiabatique de volume  $dV$  est donnée par :

$$dU = -pdV \quad (1.8)$$

Appliqué à notre univers, cette équation devient :

$$\frac{d(\rho a^3)}{dt} = -p \frac{d(a^3)}{dt} \quad (1.9)$$

ce qui permet d'obtenir l'équation usuelle de conservation locale de l'énergie dans un univers en expansion :

$$\frac{d\rho}{dt} = -3H(\rho + p) \quad (1.10)$$

Cette équation concerne l'évolution de la densité totale d'énergie. Pour connaître l'évolution de la densité d'énergie de chaque composante, il est nécessaire de faire appel à l'équation de Boltzmann qui gouverne les interactions entre les différentes particules et donc l'évolution du nombre

de particules de chaque espèce. Il existe deux cas simples pour lesquels l'équation 1.10 peut être appliquée indépendamment à chaque espèce :

- Cas 1 : Lorsque les temps de réaction des différentes particules sont grands devant le temps caractéristique de l'expansion  $H^{-1}$ . La densité comobile de particules (matière et rayonnement) est alors conservée :

$$\frac{d(n_i a^3)}{dt} = 0 \quad , \quad i = m \text{ ou } r \quad (1.11)$$

où l'indice  $m$  désigne la matière et l'indice  $r$  le rayonnement. Cette équation correspond à l'équation de Boltzmann la plus simple envisageable.

- Cas 2 : Lorsque les différentes particules sont à l'équilibre thermiques. L'expansion adiabatique conserve alors l'entropie impliquant la conservation du nombre de particules.

Dans les autres cas tels qu'au moment de la nucléosynthèse primordiale ( $T \gtrsim 100$  keV) , ou lors de l'annihilation des *Weakly Interactive Massive Particles* (WIMPs), il convient de considérer l'équation de Boltzmann complète. Dans la suite, nous allons étudier l'impact de l'expansion sur la matière et le rayonnement en supposant les cas 1 ou 2 vérifiés.

### 1.2.2 Matière non-relativiste

Pour la matière non-relativiste,  $p_m = 0$ . Nous pouvons réécrire l'équation de conservation de l'énergie :

$$\dot{\rho}_m = -3H\rho_m \quad (1.12)$$

équation qui, une fois intégrée, nous donne l'évolution de la densité de matière avec l'expansion de l'Univers :

$$\rho_m = \rho_{m,0} a^{-3} \quad (1.13)$$

où  $\rho_{m,0}$  désigne la densité d'énergie de la matière aujourd'hui. Cette expression traduit simplement l'idée intuitive que la densité de matière est inversement proportionnelle au volume gagné par l'expansion de l'Univers.

### 1.2.3 Matière relativiste

Pour le rayonnement, l'équation d'état liant pression et densité est  $p_r = \frac{1}{3}\rho_r$ , menant à l'équation de conservation de l'énergie :

$$\dot{\rho}_r = -4H\rho_r \quad (1.14)$$

dont la solution est :

$$\rho_r = \rho_{r,0} a^{-4} \quad (1.15)$$

Comme la matière non-relativiste, les photons subissent une dilution liée à l'expansion de l'Univers, justifiant une diminution de la densité d'énergie en  $a^{-3}$ . Ces mêmes photons subissent toutefois une perte d'énergie supplémentaire, proportionnelle à  $a^{-1}$ . Cette perte d'énergie est différente en nature de la précédente. Il ne s'agit pas de la diminution du nombre de photons dans un volume donné. Il s'agit donc d'une diminution de l'énergie propre à chaque photon. L'énergie d'un photon étant proportionnelle à sa fréquence, une diminution d'énergie correspond à une diminution de fréquence. Cet effet est donc connu sous le nom de décalage spectral ; il correspond à l'effet Doppler. Le décalage spectral d'un photon est quantifié par la variable  $z$  définie par :

$$1 + z = \frac{\nu_r}{\nu_e} = \frac{\lambda_e}{\lambda_r} \quad (1.16)$$



$\nu_e$  (resp.  $\lambda_e$ ) étant la fréquence (resp. la longueur d'onde) du photon au moment de son émission et  $\nu_r$  (resp.  $\lambda_r$ ) est la fréquence (resp. la longueur d'onde) du photon lors de sa réception. Le décalage spectral d'un photon mesuré au moment même de son émission est donc nul. Comme la perte d'énergie est proportionnelle à  $a^{-1}$ , nous obtenons directement la relation :

$$1 + z = \frac{a_r}{a_e} \quad (1.17)$$

$a_e$  étant la valeur du facteur d'échelle à l'instant où le photon a été émis et  $a_r$  sa valeur au moment de la réception du photon. Comme nous ne considérons que des mesures effectuées depuis la Terre (ou son environnement proche) à l'instant  $t_0$ , nous pouvons réécrire l'équation précédente sous sa forme plus conventionnelle :

$$1 + z = \frac{1}{a} \quad (1.18)$$

où nous avons omis l'indice  $e$ , puisqu'il n'y a plus d'ambiguïté. La mesure du décalage spectral d'un photon est donc la mesure directe du facteur d'échelle à l'instant où le photon a été émis.

### 1.3 Du lien entre contenant et contenu

Jusqu'à présent nous avons considéré l'espace dans lequel nous nous trouvons et la matière qui le remplit indépendamment. Dans cette section nous étudions le lien étroit existant entre les deux.

#### 1.3.1 Equation d'Einstein

En 1917, Einstein publie ce qu'il décrit lui même comme « la plus large généralisation concevable de la théorie de la relativité ». Cette nouvelle théorie, aujourd'hui connue sous le nom de « Relativité Générale », décrit la gravitation comme la manifestation de la courbure de l'espace-temps. Elle relie également cette courbure au contenu énergétique de l'espace-temps via l'équation d'Einstein (Einstein, 1917) :

$$G_{\mu\nu} = 8\pi GT_{\mu\nu} - g_{\mu\nu}\Lambda \quad (1.19)$$

où  $G_{\mu\nu}$  correspond au tenseur d'Einstein (terme géométrique, définissant l'espace),  $G$  à la constante de Newton,  $T_{\mu\nu}$  au tenseur énergie-impulsion (terme énergétique),  $g_{\mu\nu}$  à la métrique et où  $\Lambda$  est la constante cosmologique. Cette équation est locale, dans le sens où les expressions des trois tenseurs dépendent a priori de la coordonnée  $x^\mu$  considérée. Elle peut toutefois être généralisée à l'ensemble de l'Univers en invoquant le principe cosmologique. Equation d'Einstein et principe cosmologique permettent ainsi de relier la géométrie de l'Univers dans son ensemble à son contenu énergétique. Comme nous allons le voir, la connaissance de ce contenu énergétique nous permet également de connaître l'évolution de l'Univers. La Relativité Générale, associée au principe cosmologique, constituent le cadre théorique de la cosmologie moderne.

#### 1.3.2 Principe cosmologique II

La densité d'énergie autour de nous étant clairement inhomogène et anisotrope, l'équation d'Einstein nous dit qu'il en est de même pour l'espace qui nous entoure. La formulation du principe cosmologique vue en section 1.1.2 s'avère donc inadaptée à la description de notre propre Univers. Il est toutefois possible d'invoquer une version plus faible du principe cosmologique :

l'Univers est statistiquement homogène et isotrope. Les résultats de la section 1.1 sont alors applicables aux grandes échelles, tandis que les inhomogénéités apparentes aux petites échelles peuvent être traitées par la théorie des perturbations.

Cette version du principe cosmologique est longtemps restée un postulat, jusqu'à ce que l'isotropie puisse-t-elle être testée (au moins depuis notre position dans l'Univers) grâce à l'observation du fond diffus cosmologique (voir section 2.2). Des techniques modernes d'analyse, telles que l'étude du *clustering* des galaxies ou de l'effet Sunyaev-Zeldovich permettent aujourd'hui de tester l'homogénéité et l'isotropie de notre Univers en des coordonnées autres que la notre (voir par exemple Planck Collaboration et al. (2013a); Scrimgeour et al. (2012)).

### 1.3.3 Equations de Friedmann

#### Première équation de Friedmann

L'équation d'Einstein fait apparaître des tenseurs de dimension  $4 \times 4$  tous symétriques, elle contient donc 10 équations. Dans le cas d'un univers isotrope, il est possible de trouver un repère dans lequel ces tenseurs sont diagonaux et les trois composantes spatiales de chaque tenseur sont identiques. Notre système de 10 équations se réduit donc à un système de deux équations. La première est la composante temporelle ( $\mu = \nu = 0$ ) de l'équation 1.19 et peut s'écrire en fonction du facteur d'échelle comme :

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \quad (1.20)$$

où  $\rho = \rho_m + \rho_r$  est la somme des densités de matière non-relativiste et relativiste et où  $k = -1, 0$  ou  $+1$  en fonction que l'Univers soit respectivement ouvert, plat ou fermé. Cette équation est connue sous le nom de première équation de Friedmann. Elle relie l'évolution du facteur d'échelle au contenu énergétique de l'Univers. Pour un univers plat, la première équation de Friedmann peut alors se réécrire :

$$H^2 = \frac{8\pi G}{3} \left( \rho + \frac{\Lambda}{8\pi G} \right) \quad (1.21)$$

Il est possible d'interpréter la constante cosmologique comme une densité d'énergie  $\rho_\Lambda = \frac{\Lambda}{8\pi G}$ . Comme nous le verrons en section 1.3.4, une telle densité d'énergie possède des propriétés surprenantes. Du fait de son caractère particulier et de sa nature inconnue, cette énergie est communément appelée énergie noire. Nous obtenons alors une expression pour la densité d'énergie d'un univers plat :

$$\rho_{\text{plat}} = \frac{3H^2}{8\pi G} \quad (1.22)$$

qui est donc simplement proportionnelle à  $H^2$ . Il est courant de définir la densité critique comme la densité d'énergie d'un Univers plat au temps  $t_0$  :

$$\rho_c \equiv \rho_{\text{plat}}(t_0) = \frac{3H_0^2}{8\pi G} \quad (1.23)$$

ce qui permet de réécrire l'équation 1.20 en introduisant la densité critique  $\rho_c$  :

$$\left(\frac{H}{H_0}\right)^2 = \frac{\rho_{\text{tot}}}{\rho_c} - \frac{k}{H_0^2 a^2} \quad (1.24)$$

où  $\rho_{\text{tot}} = \rho_m + \rho_r + \rho_\Lambda$ . Il apparaît ainsi qu'un univers ouvert ( $k = -1$ ) possède au temps  $t_0$  une densité  $\rho_{\text{tot}}(t_0) < \rho_c$  tandis qu'un univers fermé ( $k = +1$ ) possède  $\rho_{\text{tot}}(t_0) > \rho_c$ .

### Seconde équation de Friedmann

La deuxième équation de Friedmann correspond à la trace de l'équation d'Einstein. Elle peut également s'écrire en fonction du facteur d'échelle :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \quad (1.25)$$

où, de nouveau,  $\rho = \rho_m + \rho_r$  et  $p = p_m + p_r$ . En interprétant la constante cosmologique comme résultant de l'action de l'énergie noire, nous pouvons réarranger cette équation sous la forme :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho_{tot} - 3\rho_\Lambda + 3p) \quad (1.26)$$

Pour une description complète de la constante cosmologique en terme d'énergie noire, nous sommes donc amenés à introduire un terme de pression :

$$p_\Lambda = -\rho_\Lambda \quad (1.27)$$

L'équation précédente s'écrit alors :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho_{tot} + 3p_{tot}) \quad (1.28)$$

Notons que la seconde équation de Friedmann peut être combinée avec la première pour retrouver l'équation de la conservation de l'énergie 1.10.

#### 1.3.4 Energie noire

Alors que la première équation de Friedmann (équation 1.24) contraint la dérivée première du facteur d'échelle, la seconde équation (1.28) gère l'évolution de la dérivée seconde de  $a$ . Il est intéressant de noter que  $\ddot{a} \propto -(\rho_{tot} + 3p_{tot})$ . La matière ordinaire, relativiste ou non, ayant à la fois  $\rho > 0$  et  $p > 0$ , elle ne peut entraîner que la décélération de l'expansion de l'Univers. Puisque nous pouvons affirmer, avec certitude, qu'il y a de la matière ordinaire autour de nous, notre théorie semble favoriser un univers soit en expansion décélérée, soit en contraction.

Il est connu qu'Einstein n'aimait pas l'idée d'un univers dynamique et qu'il introduisit la constante cosmologique afin d'obtenir un univers statique. En traitant cette constante cosmologique comme une forme d'énergie, nous avons vu que cette énergie, dite noire, possède une pression de signe opposé à sa densité. Si  $\Lambda > 0$ ,  $\rho_\Lambda$  est alors une densité d'énergie positive possédant une pression  $p_\Lambda$  négative ! En introduisant cette composante dans l'équation 1.28, nous voyons qu'il est maintenant possible d'obtenir un  $\ddot{a}$  positif ou nul.

La constante cosmologique fut abandonnée suite à la découverte de l'expansion de l'Univers ; le paradigme étant alors un univers en expansion décélérée. Elle a été réintroduite récemment, lors de la découverte de l'accélération de l'expansion de l'Univers par l'observation des supernovae de type Ia (Riess et al., 1998; Perlmutter et al., 1999). Que cette constante résulte effectivement de l'action d'une énergie noire aux propriétés particulières est une des interprétations possibles des équations et n'a rien d'évident.

Il est possible d'élargir le modèle d'énergie noire en autorisant son paramètre d'équation d'état  $w$  (rapport de la densité d'énergie sur la pression) à varier au cours du temps. Une paramétrisation fréquemment utilisée pour tenir compte d'une éventuelle dépendance temporelle de  $w$  est donnée par (Chevallier and Polarski, 2001; Linder, 2003) :

$$w(z) = w_0 + \frac{z}{1+z} w_a \quad (1.29)$$

qui permet d'obtenir une forte dépendance temporelle à bas décalage spectral. D'autres modèles, faisant appel à des concepts de physique des particules, ont également été proposés pour expliquer l'accélération de l'expansion de l'Univers. Le modèle de quintessence, par exemple, y voit l'effet d'un champ scalaire. Par extension, on appelle modèle d'énergie noire l'intégralité de ces modèles. Ils sont à opposer aux modèles de gravité modifiée qui tentent d'expliquer l'accélération de l'expansion par la modification de la relativité générale aux grandes échelles. La mission première de l'expérience BOSS et de contraindre ces différents modèles.

### 1.3.5 Les $\Omega$ s

Au vu des équations précédentes nous pouvons réécrire la première équation de Friedmann sous la forme :

$$\left(\frac{H}{H_0}\right)^2 = \frac{1}{\rho_c} (\rho_{m,0}a^{-3} + \rho_{r,0}a^{-4} + \rho_\Lambda) - \frac{k}{H_0^2 a^2} \quad (1.30)$$

expression qui est simplifiée grâce à la définition des paramètres « oméga » :

$$\Omega_i = \frac{\rho_{i,0}}{\rho_c}, \quad i = m, r \text{ ou } \Lambda \quad (1.31)$$

Nous définissons également un terme « oméga » pour la courbure :

$$\Omega_k = -\frac{k}{H_0^2} \quad (1.32)$$

et obtenons ainsi l'expression usuelle de la première équation de Friedmann :

$$\left(\frac{H}{H_0}\right)^2 = \Omega_k a^{-2} + \Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_\Lambda \quad (1.33)$$

Cette équation possède cinq paramètres :  $H_0$  et les quatre  $\Omega$ . A l'instant  $t_0$  cette équation se simplifie :

$$\Omega_m + \Omega_r + \Omega_\Lambda = 1 - \Omega_k \quad (1.34)$$

faisant apparaître que sur les quatre  $\Omega$ , seuls 3 sont indépendants. Nous retrouvons le fait que pour un univers plat ( $\Omega_k = 0$ ) la densité totale de l'univers équivaut à la densité critique :  $\Omega_m + \Omega_r + \Omega_\Lambda = 1$  ou de manière équivalente  $\rho_{m,0} + \rho_{r,0} + \rho_\Lambda = \rho_c$ . Il est courant de définir  $\Omega_{tot} = \Omega_m + \Omega_r + \Omega_\Lambda$  de sorte que :  $\Omega_k = 1 - \Omega_{tot}$ .

## 1.4 Contraindre le contenant grâce au contenu et *vice versa*

Les sections précédentes présentent l'espace, son contenu, ainsi que le lien fort qui les relie l'un à l'autre. La cosmologie moderne consiste à exploiter au maximum ce lien étroit pour contraindre tantôt le contenant, tantôt le contenu. En tant que cosmologues expérimentateurs, notre seul outil de travail est le télescope. Cet outil peut se décliner sous des aspects très différents en fonction de la gamme du spectre lumineux à laquelle il est dédié. Dans tous les cas, son seul but est de mesurer un flux de photons dans une direction donnée. Pour contraindre notre modèle, nous disposons donc de deux types de mesures : des flux de photons (éventuellement en fonction de la longueur d'onde ou du temps), et des angles mesurés sur le ciel. De ces deux observations, comment peut-on contraindre notre modèle d'Univers ?

### 1.4.1 Distance comobile

L'observation d'un flux de photons nous permet de contraindre le contenu de l'Univers. Par exemple, la mesure du spectre d'une étoile nous permet de déterminer sa composition et sa température. De même, l'observation de la lumière provenant d'une galaxie nous renseigne sur cette galaxie. Considérons un photon émis par cette galaxie en  $(t_e, \chi_e, \theta_e, \phi_e)$  que nous observons en  $(t_0, \chi = 0)$ . Les photons se propagent à  $(\theta, \phi)$  constant, et vérifient  $d\tau^2 = 0$ , si bien que d'après l'équation 1.2 :

$$\frac{dt}{a} = d\chi \quad (1.35)$$

La coordonnée  $\chi_e$  vérifie donc l'équation :

$$\chi_e = \int_{t_e}^{t_0} \frac{dt}{a} \quad (1.36)$$

Bien que  $\chi_e$  soit une coordonnée, elle est très fréquemment appelée « distance comobile » de la galaxie. A proprement parler, la distance comobile est en fait  $a_0\chi$ , qui est bien homogène à une distance. En effectuant le changement de variable  $a \rightarrow z$  nous obtenons l'expression suivante :

$$\chi(z_e) = \int_0^{z_e} \frac{dz}{H(z)} \quad (1.37)$$

En connaissant à la fois le décalage spectral d'une galaxie et sa distance comobile, nous pouvons mettre des contraintes sur l'intégrale du paramètre de Hubble.  $H$  dépend lui même de  $H_0$  ainsi que de la valeurs des  $\Omega$ , c'est-à-dire du contenu énergétique de l'univers. Obtenir une mesure de  $\chi$  à un  $z$  donné, c'est donc obtenir une contrainte sur les paramètres cosmologiques. Mesurer  $\chi$  à différents décalages spectraux permet ensuite d'obtenir différentes contraintes et donc de lever les dégénérescences entre les différents paramètres.

Bien que la mesure du décalage spectral soit aisée<sup>3</sup>, la détermination de la distance comobile est bien plus délicate, l'information n'étant pas contenue dans le seul flux de photons que nous recevons.

### 1.4.2 Distance de luminosité

Considérons une source lumineuse située à la distance comobile  $\chi$  de nous. Supposons qu'à l'instant  $t_1$  et pendant un temps infinitésimal  $dt$ , cette source émette  $N$  photons d'énergie moyenne  $\nu_1$ , dans toutes les directions. La luminosité de cette source est alors :

$$L_1 = \frac{N\nu_1}{dt} \quad (1.38)$$

Supposons également qu'une partie de cette lumière nous parvienne à l'instant  $t_0$ . La surface de la sphère centrée sur la source à  $t_0$  est alors :

$$S(\chi) = 4\pi f^2(\chi) \quad (1.39)$$

où  $f$  est définie en 1.3. Le flux  $F_s$ , provenant de cette source, que nous recevons dans un univers statique, vaut donc :

$$F_s = \frac{L_1}{4\pi f^2(\chi)} \quad (1.40)$$

---

3. Pour une galaxie, par exemple, il suffit de chercher les raies d'émission dont la longueur d'onde dans le référentiel du laboratoire est connue.

Ce n'est pas le cas dans un univers en expansion, pour lequel la luminosité de la source dépend des coordonnées de l'observateur. Cette dépendance a deux origines. Premièrement l'énergie des photons est proportionnelle au facteur d'échelle :  $\nu \propto a$ . Deuxièmement, l'expansion de l'univers entraîne l'étalement de la durée d'arrivée des photons : des photons émis pendant l'intervalle de temps  $dt$  à un décalage spectral  $z$  sont reçus à  $t_0$  pendant un intervalle de temps  $dt' = (1+z)dt = \frac{dt}{a}$ . La luminosité de la source après propagation des photons est donc :

$$L_0 = \frac{N\nu_1 a}{dt'} = L_1 a^2 \quad (1.41)$$

Le flux nous parvenant dans notre Univers en expansion est donc :

$$F_e = \frac{L_0}{4\pi f^2(\chi)} = F_s a^2 \quad (1.42)$$

Il est courant de définir la distance de luminosité  $D_L$  :

$$D_L = (1+z)f(\chi) \quad (1.43)$$

de sorte que l'équation 1.42 puisse-t-êre réécrite sous la forme :

$$F_e = \frac{L_1}{4\pi D_L^2} \quad (1.44)$$

Par comparaison avec l'équation 1.40, nous constatons que la distance de luminosité correspond à la distance à laquelle il faudrait placer la source dans un univers statique et plat, pour que son flux soit celui que nous observons dans notre Univers dynamique et courbe.

L'équation 1.44 est particulièrement intéressante puisqu'elle montre que mesurer le flux lumineux  $F_e$  d'une source dont la luminosité intrinsèque  $L_1$  est connue, revient à mesurer sa distance de luminosité  $D_L$ . Si le décalage spectral de la source est en plus connu, il est alors possible de contraindre à la fois la courbure de l'Univers et l'intégrale du paramètre de Hubble.

Un objet astrophysique dont nous sommes capables d'estimer la luminosité intrinsèque est appelé « chandelle standard ». Sans entrer dans les détails, il existe principalement deux types de chandelles standards : les céphéides (étoiles variables dont la période est corrélée à la luminosité) et les supernovae de type Ia (SNIa). Les SNIa sont la fin de vie cataclysmique d'un couple d'étoiles dont l'une au moins est une naine blanche. La naine blanche accrète la matière de son compagnon, jusqu'à atteindre la masse de Chandrasekhar, limite au-delà de laquelle elle explose. L'explosion de la naine blanche systématiquement aux environs de la masse de Chandrasekhar, justifie que les SNIa possèdent des luminosités intrinsèques proches. Par rapport aux céphéides, les SNIa présentent l'avantage d'être extrêmement lumineuses et donc visibles jusqu'à des décalages spectraux élevés.

La constitution de « diagrammes de Hubble » (voir figure 1.2), c'est-à-dire de graphiques représentant l'évolution de  $D_L$  en fonction de  $z$ , a permis d'exclure les modèles ne contenant que de la matière ordinaire, prouvant l'existence de l'énergie noire.

### 1.4.3 Distance angulaire

Considérons un objet de taille comobile transverse à la ligne de visée  $l$ , situé à la distance comobile  $\chi_1$  de nous. Supposons que chaque extrémité de l'objet émette un photon dans notre direction, comme visible sur la figure 1.3. Dans un univers plat, en supposant  $\chi_1 \gg l$ , les photons nous arriveraient sous un angle :

$$\theta_p = \frac{l}{\chi_1} \quad (1.45)$$

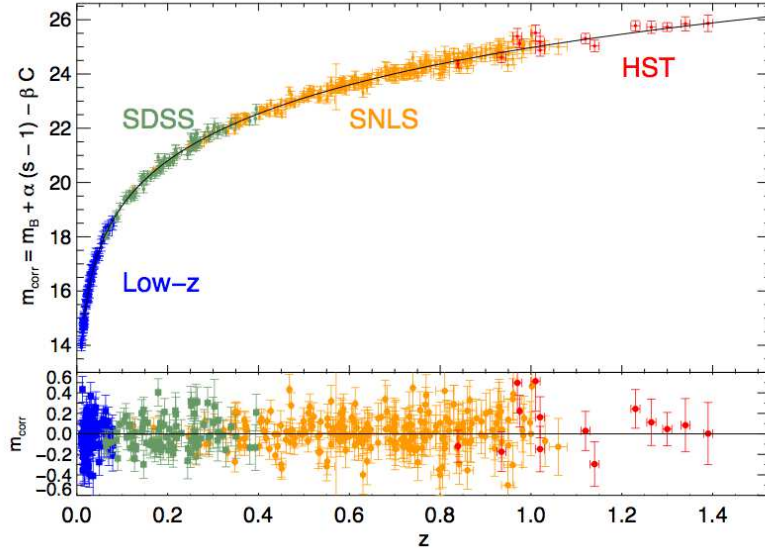


FIGURE 1.2: Diagramme de Hubble représentant la magnitude des SNIa  $m_{corr} \propto \log(D_L)$  en fonction du décalage spectral (Conley et al., 2011). La ligne noire correspond au meilleur ajustement pour un modèle cosmologique avec énergie noire pour lequel  $w$  est laissé libre. L'ajustement donne  $w = -0,90^{+0,17}_{-0,24}$ , compatible avec une constante cosmologique.

Dans un univers ouvert (ou fermé), les photons voyagent en ligne droite, c'est-à-dire suivant des courbes de coordonnées angulaires  $(\theta, \phi)$  constantes. A la réception des photons, nous observons donc l'objet sous l'angle :

$$\theta = \frac{l}{f(\chi_1)} \quad (1.46)$$

Il est courant de définir la distance angulaire  $D_A$  :

$$D_A = \frac{f(\chi_1)}{1 + z_1} \quad (1.47)$$

de sorte que l'équation 1.46 puisse-t-êtré réécrite sous la forme :

$$\theta = (1 + z_1) \frac{l}{D_A} \quad (1.48)$$

Il apparaît ainsi que  $D_A$  correspond au rapport entre la taille physique de l'objet  $\frac{l}{a}$  et son diamètre apparent. Elle permet de transformer les angles observés dans un télescope en distance propre à la source.

La figure 1.4 compare l'évolution des distances comobile, de luminosité et angulaire avec le décalage spectral  $z$  pour les valeurs de  $\Omega_m$  et  $\Omega_\Lambda$  mesurée par Planck. Notons que pour  $z \gtrsim 1.6$ , la distance angulaire est une fonction décroissante du décalage spectral, ce qui signifie que la taille apparente d'un objet de taille comobile donnée augmente avec le décalage spectral de l'objet.

L'équation 1.48 nous montre qu'en mesurant le diamètre apparent d'un objet dont nous connaissons la taille comobile et le décalage spectral, nous mesurons sa distance angulaire. Comme pour les chandelles standard, il est possible de contraindre  $f(\chi)$  et donc la courbure et les paramètres

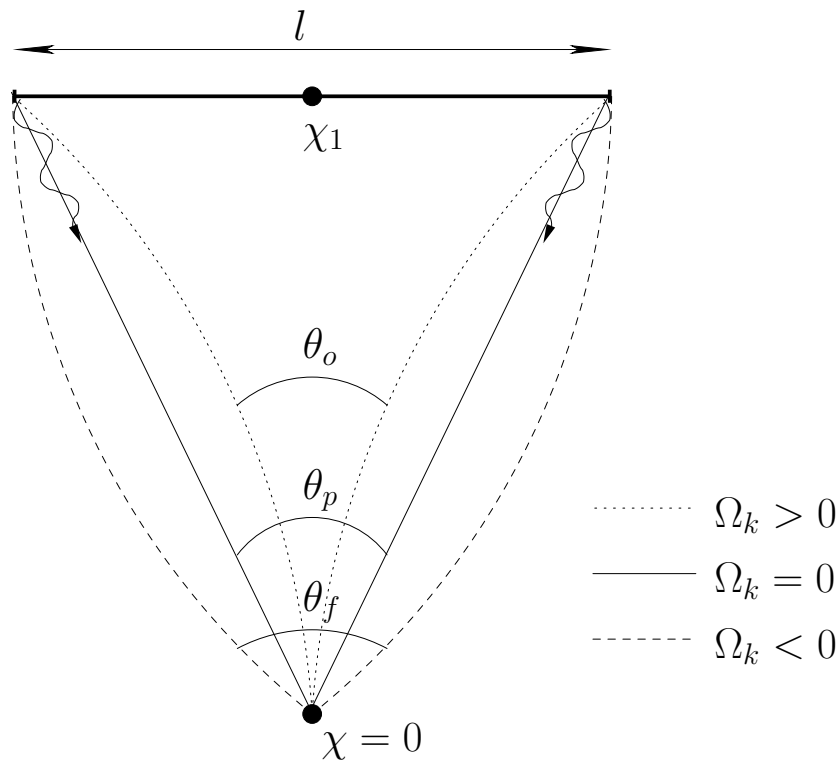


FIGURE 1.3: Représentation d'une règle standard située en  $\chi_1$ , transverse à la ligne de visée. La règle est de taille comobile  $l$ . Le parcours des photons avant de nous parvenir en  $\chi = 0$  dépend du type d'univers, si bien que l'angle apparent  $\theta_o$  de la règle dans un univers ouvert ( $\Omega_k > 0$ ) est inférieur à l'angle apparent  $\theta_p$  pour un univers plat ( $\Omega_k = 0$ ), lui-même inférieur à l'angle apparent pour un univers fermé ( $\Omega_k < 0$ ).



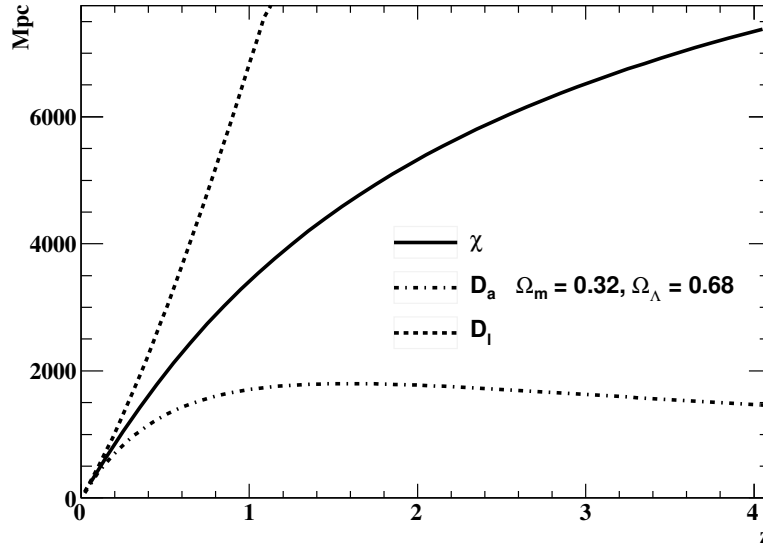


FIGURE 1.4: Comparaison des distances comobile, de luminosité et angulaire pour le modèle  $\Lambda$ CDM avec les valeurs des paramètres cosmologiques mesurés par le satellite Planck ( $\Omega_\Lambda = 0.68$  et  $\Omega_m = 0.32$ ). Notons que pour  $z \gtrsim 1.6$ , la distance angulaire est une fonction décroissante du décalage spectral.

cosmologiques. Un objet dont la taille comobile est connue est appelé règle standard. Nous disposons d'une règle standard : les oscillations acoustiques de baryons (voir chapitre 2). Le travail réalisé pour cette thèse consiste essentiellement en la mesure de cette règle standard. Dans la prochaine section nous montrons comment l'utilisation d'une règle standard permet de mesurer directement le paramètre de Hubble  $H(z)$ .

#### 1.4.4 Mesure directe de $H(z)$

Considérons toujours notre règle standard de taille comobile  $l$  mais cette fois disposée parallèlement à la ligne de visée à une distance comobile  $\chi \gg l$  (voir figure 1.5). Supposons qu'à l'instant  $t_e$  chaque extrémité de la règle standard émette un photon dans notre direction. Soit  $\chi_1$  la distance comobile de l'extrémité de la règle située la plus proche de nous et  $\chi_2$  celle de la seconde extrémité. Nous recevrons alors les deux photons séparés par l'intervalle de temps :

$$\Delta t = \frac{1}{1 + \bar{z}} (\chi_2 - \chi_1) = \frac{1}{1 + \bar{z}} \Delta \chi \quad (1.49)$$

où  $\bar{z} = \frac{z_2 + z_1}{2}$  est la moyenne des décalages spectraux de chaque extrémité de la règle standard. D'après l'expression 1.37 de la distance comobile, nous pouvons écrire :

$$\Delta \chi = l = \frac{\Delta z}{H(\bar{z})} \quad (1.50)$$

Cette équation nous montre que la mesure de la différence du décalage spectral entre les deux extrémités d'une règle standard disposée le long de la ligne de visée à un  $z$  connu est une mesure directe du paramètre de Hubble  $H(z)$ .

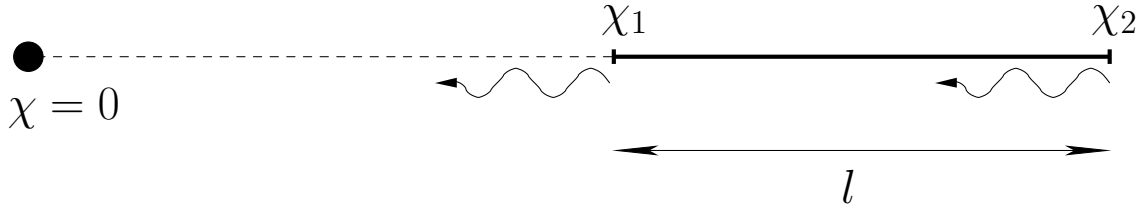


FIGURE 1.5: Représentation d'une règle standard disposée le long de la ligne de visée. Les photons sont émis par chaque extrémité de la règle, dans notre direction, à l'instant  $t_e$ . La règle est de taille comobile  $l$ . Comme détaillé dans le texte, cette configuration permet d'évaluer le paramètre de Hubble au décalage spectral de la règle.

#### 1.4.5 La distance $D_V(z)$

Comme nous le verrons au chapitre 5, la mesure d'une règle standard peut se faire en deux dimensions (et non seulement suivant la direction transverse ou le long de la ligne de visée). La mesure de la règle obtenue en moyennant sur les angles permet alors de contraindre la distance « volumique » définie par :

$$D_V(z) \equiv [cz(1+z)^2 D_A^2(z) H^{-1}(z)]^{1/3} \quad (1.51)$$

Le chapitre 2 étudie d'un point de vue théorique la règle standard que représente les oscillations acoustiques de baryons. Le chapitre 5 présente quant à lui les mesures expérimentales de cette règle standard, qui constitue le travail principale de cette thèse ; il présente également les contraintes cosmologiques correspondantes.



## Chapitre 2

# Oscillations acoustiques de baryons et forêts Lyman- $\alpha$

*« I see skies of blue,  
And clouds of white.  
The bright blessed day,  
The dark sacred night.  
And I think to myself,  
What a wonderful world. »*

—B. Thiele & G. D. Weiss

---

Après avoir introduit les bases de la cosmologie moderne, nous nous intéressons maintenant aux oscillations acoustiques de baryons résultant de la propagation de perturbations du champ de densité de matière peu après le *Big Bang*. Ces oscillations ont laissé une échelle caractéristique décelable dans les anisotropies du fond diffus cosmologique ainsi que dans la distribution de matière. La mesure de cette échelle caractéristique est l'un des outils les plus puissants dont nous disposons actuellement pour étudier le contenu de notre Univers.

---

Les résultats du chapitre précédent ont été obtenus dans le cadre de la relativité générale, en invoquant le principe cosmologique. Nous avons donc considéré un tenseur énergie-impulsion  $T_{\mu\nu}$  ainsi qu'une métrique  $g_{\mu\nu}$  homogènes et isotropes. Il est clair que l'Univers n'est pas parfaitement homogène est isotrope ; La suite logique de l'étude est donc de perturber ces tenseurs et d'étudier l'évolution des perturbations. Lorsque l'Univers est fortement dominé par la matière, l'évolution des sur-densités peut être traitée, de manière simplifiée, dans un cadre galiléen. Au contraire, lorsque il est dominé par la radiation et au début de l'ère de domination de la matière, les perturbations se déplacent à des vitesses relativistes ; Il est alors nécessaire d'avoir recours aux équations de la relativité générale. Les calculs qui en découlent sont particulièrement longs et n'ont que peu d'intérêt pour la compréhension des travaux réalisés au cours de cette thèse ; Ils ne sont donc pas présentés ici. Ils sont traités en détails dans différents manuels tel que *Modern Cosmology* de S. Dodelson. Pour une approche plus intuitive, voir *Fundamentals of Cosmology* de J. Rich.

## 2.1 Musicologie de l'Univers

Cette section décrit qualitativement l'évolution des perturbations, à l'origine des oscillations acoustiques de baryons, dans le plasma primordial emplissant l'Univers après le *Big Bang*.

### 2.1.1 Fluctuations primordiales

Avant de décrire l'évolution des perturbations, il est nécessaire de préciser de quelles perturbations nous parlons. Les fluctuations « primordiales » sont des perturbations (des champs de densité, vitesse, etc.) qui sont à l'origine de la formation des structures (galaxies, amas de galaxies, etc.) que nous observons aujourd'hui. Les modèles d'inflations décrivent ces fluctuations comme provenant de fluctuations quantiques dans le plasma emplissant l'Univers à son début. Une phase d'expansion particulièrement intense, l'inflation, aurait alors projeté ces fluctuations quantiques à des échelles cosmologiques où elles ont ensuite évolué pendant l'ère de domination de la radiation puis celle de la matière.

Dans le cadre de la relativité générale, il y a trois types de perturbations possibles : les perturbations scalaires, vectorielles et tensorielles. Dans la suite de ce chapitre nous ne nous intéressons qu'aux perturbations scalaires, qui sont à l'origine des BAO. Ces perturbations scalaires peuvent être « adiabatiques », correspondant à des fluctuations de la densité totale d'énergie, ou « d'isocourbure », c'est-à-dire des fluctuations de composition chimique telles que de la fraction de baryon sur photon  $\eta$  :

$$\begin{aligned} \text{fluctuation adiabatique} & : \Delta\rho/\rho \neq 0 \quad \Delta\eta/\eta = 0 \\ \text{fluctuation d'isocourbure} & : \Delta\rho/\rho = 0 \quad \Delta\eta/\eta \neq 0 \end{aligned}$$

Les fluctuations observées dans le fond diffus cosmologique sont en accord avec des fluctuations primordiales adiabatiques. Ces fluctuations adiabatiques sont également celles naturellement générées par les modèles d'inflation. Nous considérons donc des fluctuations de densité d'énergie scalaires adiabatiques :

$$\delta(t, \mathbf{r}) = \frac{\rho(t, \mathbf{r})}{\bar{\rho}(t)} - 1 \quad (2.1)$$

$\bar{\rho}(t)$  étant la densité moyenne de l'Univers à l'instant  $t$ .

Dans cette première section nous décrivons les mécanismes à l'origine des oscillations acoustiques de baryons. Nous commençons par décrire qualitativement l'évolution des fluctuations dans l'espace des configurations. Nous calculons ensuite la taille de l'horizon acoustique, dont l'emprunte dans la distribution de matière nous servira de règle standard au chapitre 5.

### 2.1.2 Propagation acoustique des fluctuations

La recombinaison définit le moment où l'Univers a suffisamment refroidi pour que les électrons et les protons forment des atomes d'hydrogène. Il n'y a plus suffisamment d'électrons libres dans l'Univers sur lesquels pourraient diffuser les photons ; L'Univers est transparent. Avant la recombinaison, les interactions Thomson entre photons et électrons libres sont par contre très fréquentes. Le libre parcours moyen des photons est alors bien inférieur à la taille de l'horizon  $H^{-1}(z)$ . La fréquence des interactions photons-électrons crée un fort couplage entre ces deux particules. Les interactions coulombiennes entre les électrons libres et les protons entraînent également le couplage des protons aux électrons et donc aux photons. Ces trois particules se comportent ainsi comme un seul et même fluide, appelé plasma baryons-photons.

La figure 2.1 montre l'évolution temporelle d'une sur-densité « adiabatique » dans l'espace des configurations. Cette perturbation est constituée initialement de matière noire froide, de neutrinos,

ainsi que de plasma baryons-photons. La gravitation tend à faire croître la sur-densité de plasma, mais la forte pression de radiation provoquée par les photons s'y oppose et amorce la propagation d'une onde acoustique. Cette onde se propage dans le plasma de baryons et de photons à la vitesse relativiste  $c_s(z) \sim c/\sqrt{3}$  que nous calculerons précisément en section 2.1.3. La sur-densité de matière noire ne subit, quant à elle, que la gravitation, et croît donc à sa position originelle. Les neutrinos présents dans la perturbation initiale ne sont pas couplés aux autres espèces et se dispersent à la vitesse de la lumière. Ils ne sont pas représentés sur la figure.

A la recombinaison ( $z \sim 1080$ ), les électrons et protons forment des atomes d'hydrogènes, libérant ainsi les photons, qui se dispersent alors à la vitesse de la lumière. Dès lors, il n'y a plus de pression de radiation entre les baryons. La sur-densité de baryons se trouve figée, formant une coquille de 150 Mpc comobiles de rayon autour de la sur-densité originelle. Le fort puits de potentiel gravitationnel créé par la matière noire rappelle la plus grande partie des baryons au centre de cette coquille, tandis qu'une petite partie de la matière noire tombe dans le puits de potentiel créé par les baryons à 150 Mpc. Vers un décalage spectral de 10, les distributions de matière noire et de baryons se sont uniformisées. La distribution obtenue sert de condition initiale à la formation des structures.

Il est intéressant de noter que la position de la sur-densité de baryons continue légèrement de se déplacer, après la libération des photons. Ce phénomène est lié à l'asymétrie des nombres de baryons et de photons. Le rapport du nombre de baryons sur le nombre de photons est  $\eta \sim 10^{-9}$ . Cette large asymétrie fait que les photons se découplent des baryons avant que les baryons soient découplés des photons. Le décalage spectral marquant le découplage des photons, souvent noté  $z_*$  est défini comme le moment où les photons ne subiront en moyenne qu'une seule interaction jusqu'à notre époque<sup>1</sup> et vaut  $z_* = 1.090,43$  (Planck Collaboration et al., 2013b). Le découplage des baryons est défini comme le moment où les baryons ne subissent en moyenne qu'une seule interaction Compton avec les photons jusqu'à notre époque<sup>2</sup> et vaut  $z_{drag} = 1.059,32$  (Planck Collaboration et al., 2013b).

### 2.1.3 Horizon acoustique

La vitesse du son dans le plasma baryons-photons, considéré comme fluide parfait, est donnée par :

$$c_s^2 = \frac{dp_{b\gamma}}{d\rho_{b\gamma}} \quad (2.2)$$

où l'indice  $b\gamma$  désigne le plasma baryons-photons. La densité  $\rho_{b\gamma}$  et la pression  $p_{b\gamma}$  vérifient :

$$\rho_{b\gamma} = \rho_\gamma + \rho_b \quad \text{et} \quad p_{b\gamma} = p_\gamma + p_b \quad (2.3)$$

En se rappelant que les équations d'état pour la radiation (donc pour les photons) et pour la matière (donc pour les baryons) sont respectivement  $\rho_\gamma = 3p_\gamma$  et  $p_b = 0$ , nous obtenons :

$$c_s^2 = \frac{1}{3 \left( 1 + \frac{d\rho_b}{d\rho_\gamma} \right)} \quad (2.4)$$

L'équation de conservation de l'énergie 1.10 permet alors de réécrire cette équation sous la forme :

$$c_s^2 = \frac{1}{3 \left( 1 + \frac{3\rho_b}{4\rho_\gamma} \right)} \quad (2.5)$$

1. Et cela sans tenir compte de la ré-ionisation. Ce nombre moyen d'interactions est quantifié par la profondeur optique  $\tau(z)$ . Nous reviendrons plus en détails sur ce sujet lorsque nous parlerons du fond diffus cosmologique en section 2.2.

2. Une fois encore, sans tenir compte de la ré-ionisation.

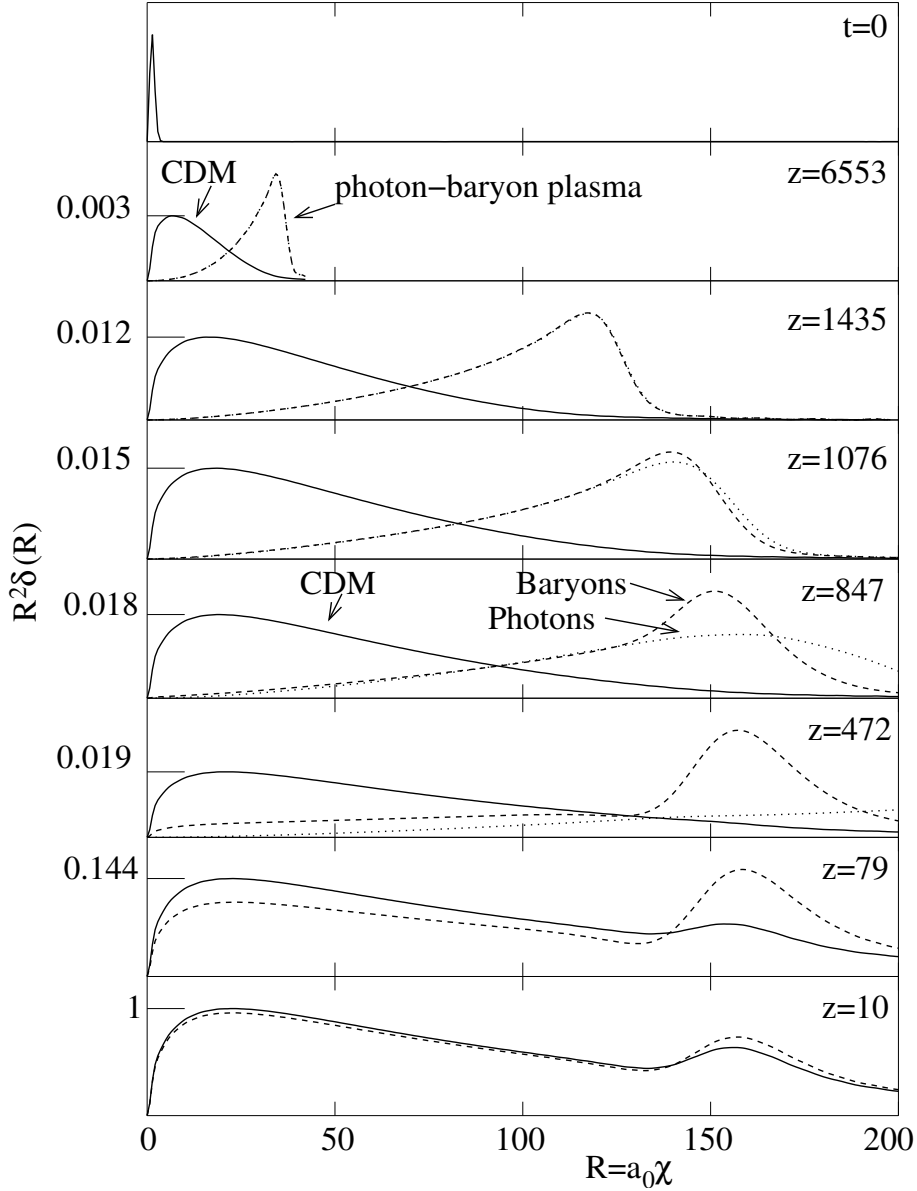


FIGURE 2.1: Propagation d'une sur-densité adiabatique dans un univers composé de matière noire, de photons, de baryons et de neutrinos (non représentés). A  $t = 0$ , toutes les espèces sont présentes dans une sur-densité située à l'origine du graphique. L'évolution temporelle se déroule ensuite comme décrit dans le texte. *Source : Fundamentals of Cosmology, 2<sup>nd</sup> edition, J. Rich.*

A partir de l'expression de  $c_s$  nous pouvons calculer la taille de l'horizon acoustique en fonction du décalage spectral :

$$r_s(z) = \int_z^{+\infty} \frac{c_s(z')}{H(z')} dz' \quad (2.6)$$

Ce calcul n'est correct qu'avant le découplage des baryons, c'est-à-dire pour  $z > z_{drag}$ , puisque par la suite il n'y a plus de propagation acoustique des sur-densités. Deux valeurs de l'horizon acoustique  $r_s$  sont particulièrement intéressantes :  $r_s(z_*)$  et  $r_s(z_{drag})$ . La première,  $r_s(z_*)$ , représente la taille de l'horizon acoustique au moment du découplage des photons. Cette échelle est, aujourd'hui encore, visible dans le fond diffus cosmologique. La seconde,  $r_s(z_{drag})$ , est la taille de l'horizon acoustique au moment du découplage des baryons. C'est cette échelle que nous observons par l'étude des forêts Lyman-alpha au chapitre 5. Les valeurs les plus précises de  $r_s(z_*)$  et  $r_s(z_{drag})$  disponibles aujourd'hui ont été obtenues grâce aux données du satellite Planck (Planck Collaboration et al., 2013b) et sont :

$$\begin{aligned} r_s^{Planck}(z_*) &= 144,71 \pm 0,61 \text{ Mpc} \\ r_s^{Planck}(z_{drag}) &= 147,49 \pm 0,59 \text{ Mpc} \end{aligned} \quad (2.7)$$

les erreurs étant données à un écart standard. La différence entre les deux horizons est donc de 2%.

Eisenstein and Hu (1998) donnent une formule simple permettant d'estimer à quelques pourcents près l'horizon acoustique lors du découplage des baryons :

$$r_s^{EH}(z_{drag}) = \frac{44,5 \ln(9,83/\Omega_m h^2)}{\sqrt{1 + 10(\Omega_b h^2)^{3/4}}} \text{ Mpc} \quad (2.8)$$

où nous voyons clairement apparaître l'influence des densités de matière et de baryons. Pour les valeurs des paramètres données par la collaboration Planck ( $\Omega_m, \Omega_b h^2, h$ ) = (0,32 ; 0,022 ; 0,67), l'approximation donne  $r_s^{EH}(z_{drag}) = 150,0$  Mpc, soit un désaccord de 1,7% avec  $r_s^{Planck}(z_{drag})$ .

## 2.2 Fond diffus cosmologique

Le fond diffus cosmologique est l'une des sources les plus riches en information pour la contrainte des modèles cosmologiques. Cette section en présente brièvement les principales caractéristiques en se focalisant sur celles d'intérêt pour l'étude des BAO.

### 2.2.1 Descriptions des fluctuations

Le taux de diffusion Thomson  $\dot{\tau}$  pour un photon sur les électrons du plasma est donné par :

$$\dot{\tau}(t) = -n_e \sigma_T \quad (2.9)$$

où  $n_e$  représente la densité d'électrons libres et  $\sigma_T$  la section efficace de la diffusion Thomson. La profondeur optique de l'Univers est alors définie comme le nombre moyen de collisions subi par un photon entre les temps  $t$  et  $t_0$  :

$$\tau(t) = \int_{t_0}^t \dot{\tau} dt' \quad (2.10)$$

Avant la recombinaison, la densité d'électrons libres est importante ; Le taux de diffusion Thomson des photons est donc élevé et l'Univers est opaque à la lumière. Après la recombinaison, il n'y a presque plus d'électrons libres dans l'Univers ( $n_e \approx 0$ ) ; Le taux de diffusion Thomson est donc



nul et l'Univers transparent. Les photons initialement piégés dans le plasma sont alors libérés et se propagent en ligne droite jusqu'à nous parvenir aujourd'hui<sup>3</sup>. Ce rayonnement électromagnétique est appelé fond diffus cosmologique (FDC dans la suite); Il est encore observable aujourd'hui.

Le FDC est un rayonnement de corps noir presque parfait. L'intensité du rayonnement reçu, à une fréquence  $\nu$  donnée suit donc la loi de Planck :

$$I(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \quad (2.11)$$

où  $T$  est la température du corps noir, également appelée température du FDC. Cette température présente de légères fluctuations en fonction de la position sur le ciel ( $\frac{\Delta T}{T} \sim 10^{-5}$ ). Pour étudier ces fluctuations il est possible de décomposer cette température en harmoniques sphériques :

$$T(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \phi) \quad (2.12)$$

où  $(\theta, \phi)$  représentent les coordonnées équatoriales. L'amplitude  $a_{lm}$ , associée à l'harmonique  $Y_{lm}$ , quantifie les variations de températures aux échelles angulaires  $\Delta\theta$  vérifiant :

$$\Delta\theta \sim \frac{\pi}{l} \quad (2.13)$$

L'harmonique  $Y_{00}$  correspond à la moyenne sur le ciel. L'amplitude  $a_{00}$  correspond donc à la température moyenne du FDC mesurée par Fixsen (2009) à :

$$a_{00} = 2,7255 \pm 0,0006 \text{ K} \quad (2.14)$$

L'information contenue par le dipôle ( $l = 1$ ) est largement dominée par la vitesse de déplacement du système solaire par rapport au FDC. Lineweaver et al. (1996) ont mesuré :

$$a_{10} = 3,358 \pm 0,023 \text{ mK} \quad (2.15)$$

où  $Y_{10}$  est choisi dans la direction du mouvement. Cette température correspond à une vitesse de  $369,0 \pm 2,5 \text{ km.s}^{-1}$  par rapport au référentiel du FDC.

La variance des fluctuations de température  $\langle \Delta T^2 \rangle$  permet de quantifier l'importance des différentes harmoniques. Les harmoniques sphériques étant orthogonales, nous obtenons :

$$\langle \Delta T^2 \rangle = \frac{1}{4\pi} \sum_{l=2}^{\infty} \sum_{m=-l}^l |a_{lm}|^2 \quad (2.16)$$

Il est alors courant d'introduire les coefficients  $C_l$  définis par :

$$C_l = \langle |a_{lm}|^2 \rangle_m \quad (2.17)$$

la moyenne étant calculée sur les  $(2l + 1)$  valeurs de  $a_{lm}$ . La variance de la température peut ainsi être réécrite sous la forme :

$$\langle \Delta T^2 \rangle = \frac{1}{4\pi} \sum_{l=2}^{\infty} (2l + 1) C_l \quad (2.18)$$

3. Nous négligeons pour l'instant la ré-ionisation. Nous reviendrons sur ce point en section 2.2.2

Nous définissons enfin le spectre des fluctuations de température par :

$$\Delta_T^2 \equiv \frac{l(l+1)C_l}{2\pi} \quad (2.19)$$

Le spectre obtenu par la collaboration Planck est reproduit en figure 2.2 (Planck Collaboration et al., 2013b). Les anisotropies sont le reflet des conditions au moment du découplage des photons. Trois types de fluctuations du plasma baryons-photons sont à prendre en compte pour décrire l'intégralité des anisotropies :

- Les fluctuations de densité
- Les fluctuations du potentiel gravitationnel
- Les fluctuations du champs de vitesse

L'impact des fluctuations de densité sur les anisotropies du FDC est peut-être le plus intuitif. Avant le découplage des photons, la température du plasma baryons-photons est directement liée à sa densité, les zones plus denses étant plus chaudes et les zones moins denses plus froides. Au moment du découplage, le rayonnement de corps noir émis par les zones de sur-densité à donc une température légèrement supérieure à la moyenne, tandis que les zones de sous-densité ont une température légèrement inférieure.

A cette effet s'ajoute l'impact des fluctuations du potentiel gravitationnel. Un photon présent dans une sur-densité de plasma au moment du découplage, subit un décalage spectral vers le rouge pour s'extraire du potentiel gravitationnel de la sur-densité. A l'inverse, un photon présent dans une sous-densité subit un décalage spectral vers le bleu. Cet effet est dénommé effet Sachs-Wolfe d'après les noms des deux premiers cosmologues à l'avoir calculé (Sachs and Wolfe, 1967). Il brise la relation directe entre la densité du plasma au moment du découplage et la température observée.

Le dernière effet à avoir un impact sur les anisotropies du FDC est un effet Doppler. La propagation acoustique des sur-densités produit un champs de vitesse non-uniforme. Au moment du découplage, la vitesse du plasma le long de la ligne de visée augmente ou diminue la température apparente du FDC, suivant qu'elle soit positive (la sur-densité se dirige vers nous) ou négative (la sur-densité s'éloigne de nous).

Différents logiciels libres tels que CAMB (Lewis et al., 2000) ou CMBFAST (Seljak and Zaldarriaga, 1996) permettent de calculer la contribution de ces différents effets en fonction de l'échelle angulaire et des paramètres cosmologiques.

## 2.2.2 $\Lambda$ CDM : un modèle à 6 paramètres

La courbe en rouge de la figure 2.2 correspond au meilleur ajustement des  $C_l$  obtenu pour le modèle  $\Lambda$ CDM plat. Ce modèle ne comporte que six paramètres, à savoir :

- L'angle apparent de la position du premier pic acoustique  $\theta_s$
- La densité de baryons au temps présent  $\Omega_b h^2$
- La densité de matière noire froide (non baryonique) au temps présent  $\Omega_c h^2$
- La profondeur optique liée à la ré-ionisation  $\tau$
- La normalisation du spectre de puissance de courbure  $A_s$
- L'indice spectral du spectre de puissance de courbure  $n_s$

La table 2.1 présente les mesures des six paramètres obtenues pas Planck.  $A_s$  et  $n_s$  définissent le spectre de puissance de courbure qui définit lui-même le spectre des fluctuations primordiales à la sortie de l'inflation, fluctuations qui donneront naissance aux oscillations acoustiques de baryons.  $\theta_s$  dépend de la taille de l'horizon acoustique lors du découplage des photons  $r_s(z_*)$  et de la distance angulaire à la surface de dernière diffusion  $D_A(z_*)$ . D'après la formule 1.48 nous pouvons

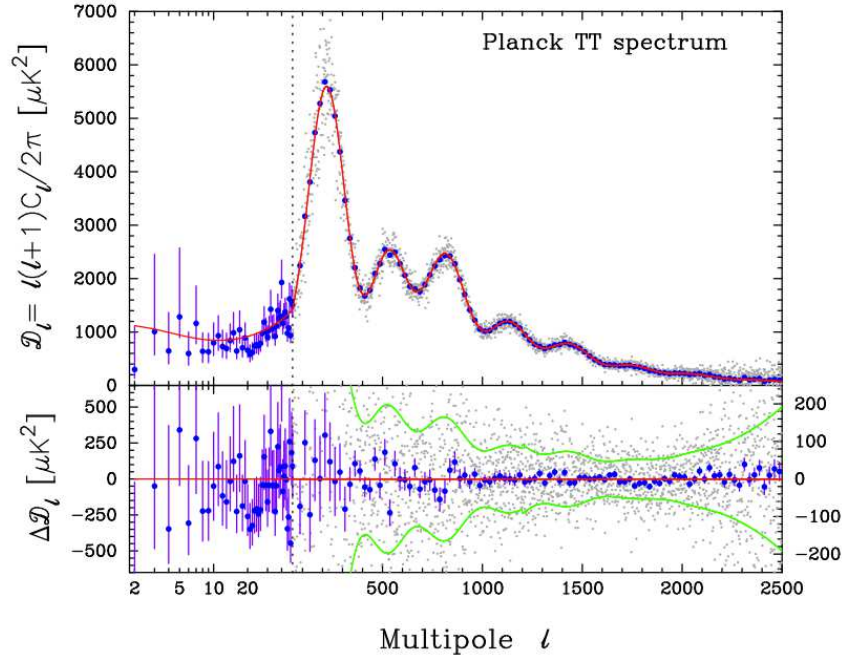


FIGURE 2.2:  $C_l$  du FDC obtenus par le satellite Planck (Planck Collaboration et al., 2013b).

écrire :

$$\theta_s = \frac{D_A(r_*)}{(1+z_*)r_s(z_*)} \quad (2.20)$$

$\theta_s$  détermine également l'écartement des pics des  $C_l$ . Les paramètres  $\Omega_b h^2$  et  $\Omega_c h^2$  définissent les densités moyennes de baryons et de matière noire, influant notamment sur l'amplitude des pics.

Le paramètre restant,  $\tau$ , quantifie la profondeur optique jusqu'à la surface de dernière diffusion (c'est-à-dire jusqu'à la surface d'où ont été émis les photons du FDC nous parvenant aujourd'hui). Il est donné par l'équation 2.10 appliquée au temps  $t_* = t(z_*)$  :

$$\tau = \int_{t_0}^{t_*} \dot{\tau} dt' \quad (2.21)$$

Ce paramètre est nécessaire car contrairement à ce que nous avons supposé précédemment, l'Univers n'est pas parfaitement transparent jusqu'à nos jours. Aux alentours de  $z = 11$ , la formation des premières étoiles et des premiers AGN (voir section 2.4.1) ré-ionise une partie de l'hydrogène du milieu intergalactique (MIG), rendant l'Univers légèrement opaque. Planck mesure  $\tau \sim 0,1$  signifiant qu'environ un photon du FDC sur dix subit une diffusion Thomson avant de nous parvenir.

### 2.2.3 Paramètres dérivés utiles

Il est possible de dériver différents paramètres cosmologiques des six principaux présentés précédemment. L'angle  $\theta_s$  est une fonction de  $D_A(z_*)$  et  $r_s(z_*)$  qui dépendent elles mêmes de  $\Omega_m$ ,

	Planck	Planck + WMAP pol.
$\Omega_b h^2$	$0,02207 \pm 0,00033$	$0,02205 \pm 0,00028$
$\Omega_c h^2$	$0,1196 \pm 0,0031$	$0,1199 \pm 0,0027$
$100\theta_s$	$1,04132 \pm 0,00068$	$1,04131 \pm 0,00063$
$\tau$	$0,097 \pm 0,038$	$0,089^{+0,012}_{-0,014}$
$n_s$	$0,9616 \pm 0,0094$	$0,9603 \pm 0,0073$
$\ln(10^{10} A_s)$	$3,103 \pm 0,072$	$3,089^{+0,024}_{-0,027}$

TABLE 2.1: Valeurs des 6 paramètres cosmologiques du modèle  $\Lambda$ CDM estimés à partir des  $C_l$  de Planck seuls (deuxième colonne) et des  $C_l$  de Planck combinés avec la polarisation de WMAP (troisième colonne). Tableau tiré de Planck Collaboration et al. (2013b).

$h$  et  $\Omega_b h^2$ . La mesure de cet angle permet d'obtenir une contrainte sur  $\Omega_m h^3$  (Planck Collaboration et al., 2013b) :

$$\Omega_m h^3 = 0,0959 \pm 0,0006 \quad (2.22)$$

Combinée avec la mesure de  $\Omega_m h^2 = (\Omega_b + \Omega_c) h^2$ , nous obtenons alors une mesure de  $H_0 = 67,4 \pm 1,4$ , mesure qui est améliorée en ajoutant les contraintes venant de la polarisation du FDC mesurée par WMAP.

$$H_0 = 67,3 \pm 1,2 \quad (2.23)$$

La combinaison de  $H_0$  et de  $(\Omega_b + \Omega_c) h^2 = (1 - \Omega_\Lambda) h^2$ , permet alors de mesurer  $\Omega_\Lambda = 0,686 \pm 0,020$ , mesure qui peut également être sensiblement améliorée en utilisant la polarisation mesurée par WMAP :

$$\Omega_\Lambda = 0,685^{+0,016}_{-0,018} \quad (2.24)$$

## 2.3 Spectre de puissance de la matière

Nous nous intéressons maintenant à la distribution de la matière dans l'Univers.

### 2.3.1 Spectre de puissance

Tant que les fluctuations définies en section 2.1.1 sont faibles ( $\delta(t, \mathbf{r}) \ll 1$ ), il est plus aisé de traiter l'évolution des perturbations dans l'espace de Fourier, puisque les différents modes évoluent indépendamment. Dans la suite, pour simplifier les calculs, nous nous restreignons au cas d'un Univers plat. Le système de coordonnées spatiales  $\mathbf{r}$  est alors le système de coordonnées comobiles cartésiennes.

Pour passer dans l'espace de Fourier, nous effectuons la transformation :

$$\delta(t, \mathbf{r}) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \delta_{\mathbf{k}}(t) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (2.25)$$

Puisque nous utilisons des coordonnées  $r$  comobiles, le vecteur  $k$  est également comobile, dans le sens où sa longueur d'onde  $\lambda_k$  associée grandit avec l'expansion :

$$\lambda_k(t) = \frac{2\pi}{ak} \quad (2.26)$$

Le spectre de puissance des fluctuations est défini comme :

$$P(k) \equiv \langle |\delta_{\mathbf{k}}(t)|^2 \rangle \quad (2.27)$$

où la moyenne est effectuée sur tous les modes  $\mathbf{k}$  tels que  $|\mathbf{k}| = k$ . Il caractérise donc la variance des  $\delta_{\mathbf{k}}$ . La variance des  $\delta(t, \mathbf{r})$  est reliée au spectre de puissance par :

$$\langle \delta^2(t, \mathbf{r}) \rangle = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} |\delta_{\mathbf{k}}(t)|^2 = \frac{1}{2\pi^2} \int_0^\infty k^2 dk P(k) \quad (2.28)$$

Le spectre de puissance  $P(k)$  a la dimension d'un volume. Il est courant de définir le spectre des fluctuations relatives de densité  $\Delta(k)$  :

$$\Delta^2(k) \equiv \frac{k^3 P(k)}{2\pi^2} \quad (2.29)$$

qui sont elles sans dimension. La variance des  $\delta(t, \mathbf{r})$  peut alors être exprimée sous la forme :

$$\langle \delta^2(t, \mathbf{r}) \rangle = \int_0^\infty \frac{dk}{k} \Delta^2(k) \quad (2.30)$$

La figure 2.3 montre le spectre de puissance ainsi que le spectre des fluctuations relatives de densité au temps présent, pour  $(\Omega_\Lambda; \Omega_m) = (0,73; 0,27)$  (traits pleins) et pour  $(\Omega_\Lambda; \Omega_m) = (0; 1)$  (traits pointillés). Cette figure montre également les fluctuations de potentiel  $\Delta_\phi(k)$  obtenues en appliquant l'équation de Poisson  $\nabla^2\phi = 4\pi G\rho$  à l'équation 2.28.

Le spectre de puissance  $P(k)$  croit proportionnellement à  $k$  jusqu'à atteindre un maximum en  $k = k_{eq}$ . Il décroît alors proportionnellement à  $k^{-3}$ . Pour  $(\Omega_\Lambda; \Omega_m) = (0,73; 0,27)$ , l'échelle caractéristique  $k_{eq}$  vaut environ  $0,2 h^{-1} \text{Mpc}$ . Le fait que les modes soient supprimés à grand  $k$  (petites échelles) est lié au fait que ces modes sont entrés dans l'horizon pendant l'ère de domination de la radiation. A cette époque la gravitation étaient dominée par les photons, dont les oscillations ont empêché les modes de croître. Un mode est donc d'autant plus supprimé qu'il a passé de temps dans l'horizon pendant l'ère de domination de la radiation. Les modes en dehors de l'horizon ont quant à eux pu croître normalement.

Lorsque l'Univers est entré dans l'ère de domination de la matière, tous les modes du spectre de puissance de la matière noire se sont mis à croître de manière identique. L'échelle caractéristique  $k_{eq}$  correspond donc à la taille de l'horizon au moment de l'équivalence matière-radiation. Les modes des baryons situés dans l'horizon ont quant à eux oscillé jusqu'à la recombinaison, moment à partir duquel ils ont rapidement crû pour rattraper les modes de matière noire.

Notons que le spectre des fluctuations de potentiel  $\Delta_\phi(k)$  est plat pour  $k < k_{eq}$  ce qui justifie l'appellation de spectre « invariant d'échelle » souvent attribué aux spectres des fluctuations primordiales.

### 2.3.2 Biais et distortions en *redshift*

La mesure du spectre de puissance (ou de la fonction de corrélation) de la matière représente une alternative à l'observation du FDC pour obtenir des contraintes sur les paramètres cosmologiques. Toute la difficulté résulte dans la mesure de la densité de matière, puisque l'essentielle de cette densité se trouve sous forme de matière noire que nous ne pouvons observer directement. La méthode actuelle consiste à sonder la distribution de densité des baryons, supposée suivre la distribution de la matière. Le traceur le plus utilisé pour sonder cette densité de baryon est les

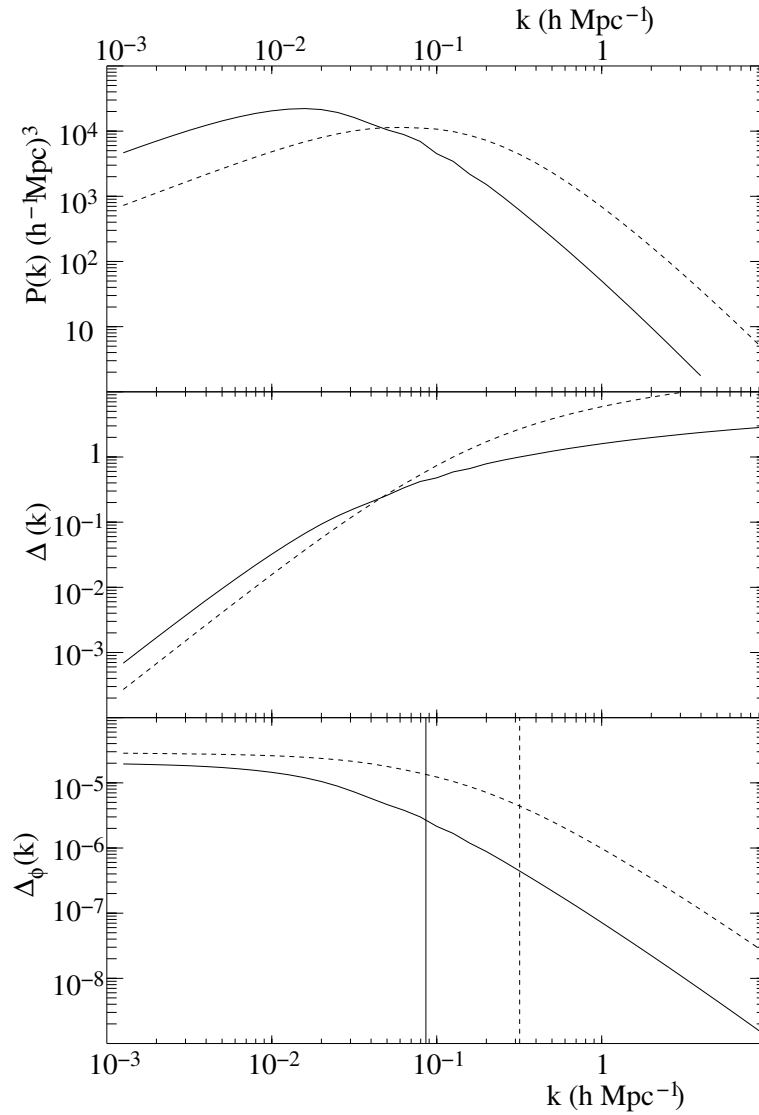


FIGURE 2.3: Spectre de puissance  $P(k)$ , fluctuations de densité  $\Delta(k)$  et fluctuations de potentiel gravitationnel, au temps présent  $t_0$ , pour  $(\Omega_\Lambda; \Omega_m) = (0,73; 0,27)$  (traits pleins) et pour  $(\Omega_\Lambda; \Omega_m) = (0; 1)$  (traits pointillés). Ces courbes supposent que le régime linéaire est valable jusqu'à aujourd'hui.

Source : *Fundamentals of Cosmology, 2<sup>nd</sup> edition, J. Rich.*

traceur	$z$	$b$	$\beta$	$b(1 + \beta)$
CMASS	0,57	$1,75^{+0,047}_{-0,046}$	$0,338^{+0,028}_{-0,027}$	—
Quasar	2,39	$3,8 \pm 0,3$	—	—
Forêts Lyman- $\alpha$	2,25	0,13	1,47	$0,336 \pm 0,012$

TABLE 2.2: Mesure du biais  $b$  et de  $\beta$  pour différents traceurs à différents décalages spectraux  $z$ . Les mesures pour les galaxies massives (échantillon CMASS du Sloan Digital Sky Survey) sont tirées des contraintes de  $b\sigma_8$  et  $b\beta\sigma_8$  de Reid et al. (2012) combinée avec  $\sigma_8 = 0,7$ . Le biais des quasars est tiré de White et al. (2012). Les valeurs de  $b$  et  $\beta$  pour les forêts Lyman- $\alpha$  sont tirées des simulations de McDonald (2003) et n'ont pas d'erreurs. La valeur bien contrainte de  $b(1 + \beta)$  obtenue pour les forêts Lyman- $\alpha$  par Slosar et al. (2011) est également renseignée. Les erreurs correspondent à une déviation standard.

galaxies. Il est toutefois peu probable que les fluctuations de densité des galaxies soient rigoureusement égales aux fluctuations de densité de matière. Ce désaccord entre les deux distributions est généralement modélisé par un paramètre  $b$  appelé biais défini par :

$$\frac{\Delta n_{gal}}{n_{gal}} = b \frac{\Delta \rho}{\rho} \quad (2.31)$$

où  $n_{gal}$  est la densité de galaxies et  $\rho$  la densité de matière. Ce biais peut, a priori, être une fonction compliquée de l'échelle, du type de galaxie et de leur environnement, etc. A cause de la complexité de son estimation, le biais  $b$  est souvent supposé être une simple constante. Les erreurs systématiques sur l'estimation du spectre de puissance de la matière sont donc dominées par l'estimation du biais  $b$ .

Un autre effet est à prendre en compte lors de la mesure du spectre de puissance : les distortions en *redshift*. Ces distortions sont un effet observationnel résultant de l'utilisation du décalage spectral (*redshift* en anglais) pour déterminer la position radiale des objets. Le décalage spectral d'une galaxie résulte de la composition d'un décalage spectral cosmologique, résultant de l'expansion, et de l'effet Doppler, lié à la vitesse propre de la galaxie le long de la ligne de visée. Cet effet rend anisotrope le spectre de puissance mesuré. Cette anisotropie est quantifiée par le paramètre  $\beta$  introduit par Kaiser (1987) :

$$P(\mathbf{k}_\perp, k_\parallel) = (1 + \beta \mu_{\mathbf{k}}^2)^2 P(k) \quad (2.32)$$

où  $\mu_{\mathbf{k}}$  est l'angle entre la ligne de visée et le vecteur d'onde  $\mathbf{k}$ ,  $\mathbf{k}_\perp$  sa composante transverse à la ligne de visée et  $k_\parallel$  sa composante le long de la ligne de visée.

Les valeurs des biais et de  $\beta$  pour différents traceurs sont présentées en table 2.2. Il est intéressant de noter que le biais des forêts Lyman- $\alpha$  (que nous détaillerons en section 2.4) est très faible, tandis que celui des quasars est très élevé. Cette différence s'explique par le fait que les quasars se situent dans des régions où la densité de matière est plus élevée que la moyenne, tandis que les forêts Lyman- $\alpha$  sondent des régions où la densité de matière est particulièrement faible. Comme nous le verrons en section 2.4, à cet effet s'ajoute le fait que les forêts Lyman- $\alpha$  sondent l'hydrogène neutre qui est lui-même une faible partie de l'hydrogène du milieu intergalactique qui est principalement ionisé. Enfin, l'observable des forêts Lyman- $\alpha$  n'est pas directement la densité d'hydrogène neutre, mais la fraction de flux transmise (définie en équation 2.39) qui est une fonction non-linéaire de la densité d'hydrogène, qui tend également à diminuer le biais.

### 2.3.3 Spectre de puissance et fonction de corrélation

Considérons deux points de l'espace  $\mathbf{r}_1$  et  $\mathbf{r}_2$  tels que  $\mathbf{r}_2 - \mathbf{r}_1 = \mathbf{r}$ . La fonction de corrélation à deux points est définie par :

$$\xi(r) = \langle \delta(\mathbf{r}_1)\delta(\mathbf{r}_2) \rangle_r \quad (2.33)$$

où  $r = |\mathbf{r}|$  et la moyenne est effectuée sur tous les points  $\mathbf{r}_1$  et  $\mathbf{r}_2$  vérifiant  $|\mathbf{r}_2 - \mathbf{r}_1| = r$ . En reprenant l'expression 2.25 des  $\delta(t, \mathbf{r})$ , nous pouvons réécrire la fonction de corrélation sous la forme :

$$\xi(r) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} |\delta_{\mathbf{k}}|^2 e^{i\mathbf{k}\cdot\mathbf{r}} \quad (2.34)$$

ou encore, en faisant apparaître le spectre de puissance :

$$\xi(r) = \frac{1}{2\pi^2} \int_0^\infty k^2 dk P(k) \frac{\sin(kr)}{kr} \quad (2.35)$$

Le spectre de puissance est donc directement relié à la fonction de corrélation par transformée de Fourier. De même nous pouvons écrire :

$$P(k) = \frac{1}{2\pi^2} \int_0^\infty r^2 dr \xi(r) \frac{\sin(kr)}{kr} \quad (2.36)$$

Le pic résultant des oscillations acoustiques de baryons dans la fonction de corrélation (visible dans les simulations de la figure 2.1) se manifeste sous forme d'oscillations dans le spectre de puissance.

### 2.3.4 Fiabilité de la mesure du pic BAO

Nous souhaitons insister sur le fait que le biais  $b$  et  $\beta$  sont de simples facteurs multiplicatifs du spectre de puissance de la distribution de matière (nous verrons en section 5.5.1 qu'il en est de même pour la fonction de corrélation). S'ils ont un impact important sur la normalisation ainsi que l'anisotropie des spectres de puissance ou des fonctions de corrélations mesurées expérimentalement, ils n'ont pas d'influence sur la position du pic BAO dans la fonction de corrélation ou des maxima des oscillations dans le spectre de puissance.

### 2.3.5 Exemple : le relevé de galaxies CMASS

Nous prenons comme exemple les résultats obtenus par Anderson et al. (2012) sur le catalogue CMASS du Sloan Digital Sky Survey (décrit en chapitre 3), contenant 264.283 galaxies massives à un décalage spectral moyen  $z = 0,57$ . Les figures 2.4 et 2.5 représentent respectivement la fonction de corrélation et le spectre de puissance calculés sur l'échantillon CMASS. Comme attendu, la fonction de corrélation présente un pic tandis que le spectre de puissance présente des oscillations, permettant la détection des BAO avec un niveau de confiance supérieur à  $5\sigma$ . Ces résultats ont permis une mesure de  $D_V(0,57)$  à 1,7% près :

$$D_V(0,57) = 2094 \pm 34 \text{ Mpc} \quad (2.37)$$

ce qui correspond actuellement à la mesure de distance la plus précise basée sur un relevé de galaxies.



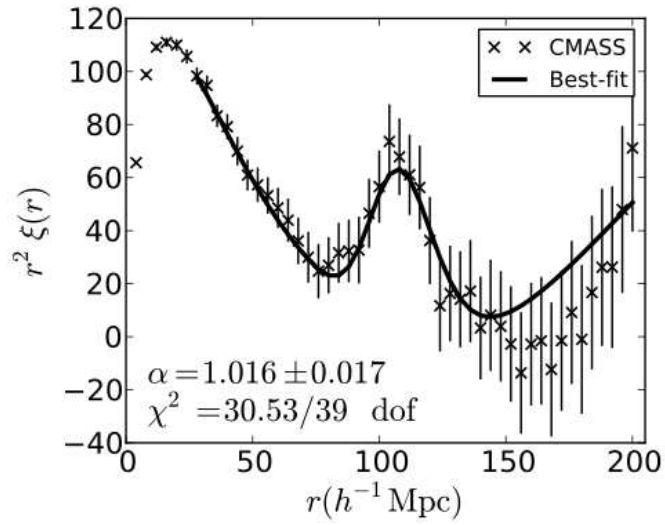


FIGURE 2.4: Fonction de corrélation des galaxies de l'échantillon CMASS (Anderson et al., 2012)

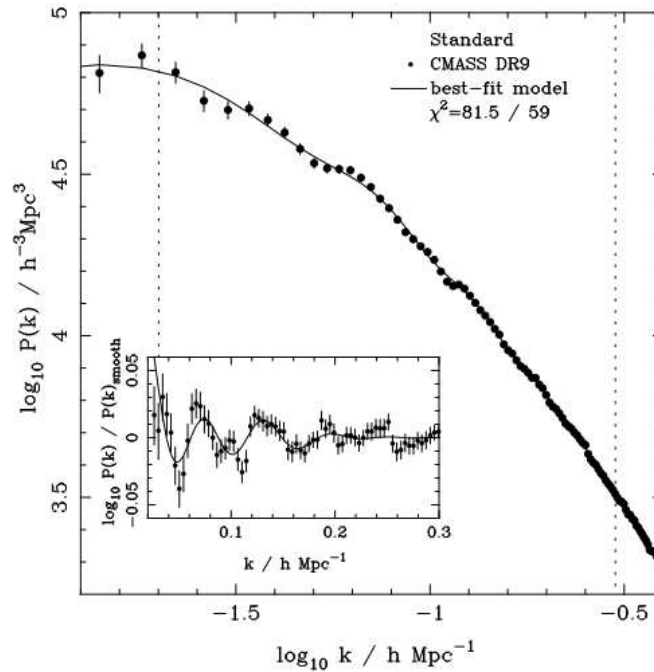


FIGURE 2.5: Spectre de puissance des galaxies de l'échantillon CMASS (Anderson et al., 2012)

## 2.4 Un nouveau traceur : les forêts Lyman- $\alpha$

Cette section présente la relation entre les « forêts Lyman- $\alpha$  » des spectres de quasars et la distribution de l'hydrogène neutre du milieu intergalactique. Nous utiliserons cette relation au chapitre 5 pour estimer la fonction de corrélation de la distribution de matière. La compréhension des quasars, en tant qu'objet astrophysique, a peu d'intérêt pour cette étude. Nous nous limitons donc à présenter leurs principales caractéristiques. Nous étudions ensuite l'effet Gunn-Peterson, à l'origine des forêts Lyman- $\alpha$ , puis terminons par présenter des cas particuliers de spectres de quasars.

### 2.4.1 *Quasi-stellar radio sources*

Le premier quasar a été observé par Schmidt (1963). Il s'agissait d'un objet ressemblant à une étoile, car ponctuel dans un télescope, mais, contrairement aux étoiles, présentant également une forte émission radio. Plusieurs objets du même type ont rapidement été observés. En l'absence d'explication quant à la nature astrophysique de ces objets, ils ont été surnommés *quasi-stellar radio sources* (signifiant « sources radio presque stellaire » en anglais). Cette dénomination a par la suite été abrégée en « quasars ».

Les quasars sont aujourd'hui compris comme étant une sous-catégorie de la population de noyaux actifs de galaxie (AGN pour l'anglais *active galactic nuclei*) (Antonucci, 1993; Urry and Padovani, 1995). La figure 2.6 montre une représentation schématique d'un AGN. Ces objets sont constitués en leur centre d'un trou noir super-massif. Directement autour du trou noir se trouve un disque d'accrétion constitué par la matière environnante tombant dans le puits gravitationnel du trou noir. En tombant, la matière libère son énergie potentielle par friction et au travers de phénomènes turbulents générant des particules de très haute énergie, canalisée par les lignes de champ magnétique de l'AGN. Ces particules sont ensuite libérées sous forme de deux jets collimatés aux niveaux des pôles magnétiques. Autour du disque d'accrétion se trouve un tore de poussière. La géométrie particulière des AGN fait qu'ils nous apparaissent de manières différentes selon la direction des pôles magnétiques par rapport à la ligne de visée. Un AGN dont le jet de particules pointe dans notre direction est appelé « blazar ». Si le jet pointe dans une autre direction et que la galaxie hôte du trou noir est visible, l'AGN est une galaxie Seyfert. Dans les cas contraires, l'AGN est un quasar.

### 2.4.2 Effet Gunn-Peterson et forêts Lyman- $\alpha$

Un spectre typique de quasar est reproduit en figure 2.7. Ce spectre présente une très forte raie d'émission à la longueur d'onde du Lyman- $\alpha$  ( $\lambda_{Ly-\alpha} = 1.215 \text{ \AA}$ ), caractéristique des quasars. Les raies d'émission du Lyman- $\beta$  ( $\lambda_{Ly-\beta} = 1.025 \text{ \AA}$ ), du carbone 4 (CIV) et du silicium 4 (SiIV) sont également visibles.

La partie du spectre située à plus faibles longueurs d'onde du pic Lyman- $\alpha$  est très différente de la partie située à plus grandes longueurs d'onde. Alors que la partie droite du spectre ( $\lambda > \lambda_{Ly-\alpha}$ ) présente peu de raies d'absorption, la partie de gauche ( $\lambda < \lambda_{Ly-\alpha}$ ) en compte un grand nombre. Ces raies d'absorption sont produites par l'hydrogène neutre situé le long de la ligne de visée du quasar. Cet effet est connu sous le nom d'effet Gunn-Peterson, d'après les noms des deux chercheurs ayant les premiers étudié ces raies d'absorption pour en tirer des informations sur le milieu intergalactique Gunn and Peterson (1965).

Les raies d'absorption comprises entre le pic Lyman- $\alpha$  et le pic Lyman- $\beta$  ont principalement été produites par l'hydrogène neutre du MIG à la longueur d'onde du Lyman- $\alpha$ . Viennent également s'ajouter des raies d'absorptions causées par les métaux présents dans le MIG, mais en

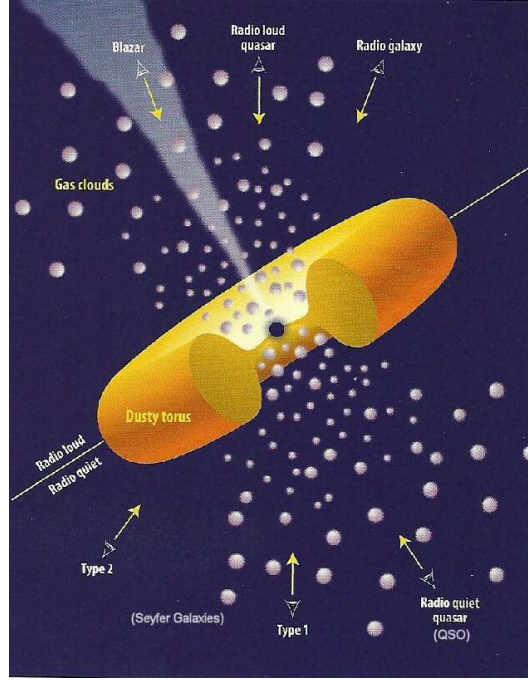


FIGURE 2.6: Schéma unifiant différents objets astrophysiques (blazars, quasars, galaxies seyfert) comme résultant de l'observation d'AGN sous différents angles. *Source : (Urry and Padovani, 1995).*

bien plus faible quantité que l'hydrogène. Le fait que les raies d'absorption se retrouvent décalées dans le spectre du quasar est lié au décalage vers le rouge de la lumière tout au long de sa propagation entre le quasar et nous. L'ensemble de ces raies d'absorption forment la forêt Lyman- $\alpha$ . Prenons l'exemple d'un quasar situé à un décalage spectral  $z_q = 3$ . Supposons que la lumière de ce quasar traverse un nuage d'hydrogène situé à  $z_{HI} = 2,8$ , avant de nous parvenir. Le nuage produit une raie d'absorption dans le spectre du quasar à la longueur d'onde du Lyman- $\alpha$ . Le nuage étant à  $z_{HI} = 2,8$ , cette raie est visible dans le spectre du quasar à la longueur d'onde  $\lambda_{HI} = \lambda_{Ly-\alpha}(z_{HI} + 1) = 4.617 \text{ \AA}$  tandis que le pic Lyman- $\alpha$  du spectre du quasar se situe à la longueur d'onde  $\lambda_p = \lambda_{Ly-\alpha}(z_q + 1) = 4.860 \text{ \AA}$ . La raie d'absorption se situe bien à plus faible longueur d'onde que le pic d'émission Lyman- $\alpha$ .

La partie située entre le pic Lyman- $\beta$  et le pic Lyman- $\gamma$  est appelée forêt Lyman- $\beta$ . Elle est constituée de la superposition de raies d'absorption causées par des nuages d'hydrogène proches du quasar à la longueur d'onde du Lyman- $\beta$  et de raies d'absorption dues à des nuages d'hydrogène à plus bas décalages spectraux à longueur d'onde du Lyman- $\alpha$ . Il est également possible de définir une forêt Lyman- $\gamma$  et ainsi de suite jusqu'au continu de Lyman à  $911 \text{ \AA}$ .

Le flux  $f_{obs}$  que nous observons dans la forêt Lyman- $\alpha$  correspond à la fraction du flux émis par le quasar  $f_q$  non absorbée par l'hydrogène neutre du MIG. Nous pouvons définir la fraction de flux transmis  $F$  comme suit :

$$F = \frac{f_{obs}}{f_q} \quad (2.38)$$

Par définition, la fraction de flux transmise est également donnée par :

$$F = e^{-\tau} \quad (2.39)$$

L'étude de simulations hydrodynamiques (Cen et al., 1994; Zhang et al., 1995; Hernquist et al., 1996; Croft et al., 1997) et le développement de modèle analytique (Bi, 1993; Miralda-Escude and Rees, 1993; Hui and Gnedin, 1997) a permis de montrer l'existence d'une relation de proportionnalité entre la profondeur optique  $\tau$  et la densité d'hydrogène neutre ( $\tau \propto n_{HI}$ ) qui trace elle-même la densité de baryon d'après la relation  $n_{HI} \propto \rho_b^\beta$ , avec  $\beta \sim 1,5 - 2,0$ . Nous pouvons donc réécrire l'équation 2.39 comme :

$$F = e^{-k \cdot \rho_b^\beta} \quad (2.40)$$

où  $k$  est la constante de proportionnalité entre la profondeur optique  $\tau$  et  $\rho_b^\beta$ . La fraction de flux transmise  $F$  est donc reliée de manière monotone à la densité de baryon qui à son tour trace la densité de matière.

En chapitre 5, nous exploiterons la relation entre la densité de matière le long de la ligne de visée des quasars et les fluctuations de leur forêt Lyman- $\alpha$ , pour calculer la fonction de corrélation de la distribution de matière. Pour cela, nous estimerons la fraction de flux transmise  $F$ , pour chaque point de la ligne de visée de chaque quasar. Il ne nous sera pas possible de retrouver directement la densité de baryon  $\rho_b$  qui est une fonction non-linéaire de cette fraction de flux transmise et qui dépend du coefficient  $k$  mal connu. Du fait de ce couplage non-linéaire, il est notamment particulièrement difficile d'estimer  $\rho_b$  lorsque la fraction de flux transmise est proche de 0 ( $F \sim 0$ ).

Les sections efficaces des transition de Lyman sont très élevées. Le fait que le flux des forêts Lyman- $\alpha$  des quasars ne soit pas complètement absorbé vient de la forte ionisation du MIG (Gunn and Peterson, 1965). Cette ionisation est causée par le flux de photons ultra-violet émit par les galaxies formant des étoiles et les AGN et dépend donc du décalage spectral. Plus l'Univers forme d'étoiles et d'AGN, plus le MIG est ionisé. Par ailleurs, l'expansion de l'Univers entraîne une diminution de la densité moyenne d'hydrogène avec le temps, proportionnelle à  $(1+z)^3$ . Ces dépendances font que le biais des forêts Lyman- $\alpha$  varie avec le décalage spectral. Nous quantifierons cette dépendance en équation 5.26.

### 2.4.3 Cas particuliers de spectres : DLA et BAL

On distingue différents types d'absorbeurs Lyman- $\alpha$  en fonction de leur colonne de densité (Weymann et al., 1981) :

- En-dessous de  $10^{17} \text{ cm}^{-2}$  les absorbeurs constituent la forêt Lyman- $\alpha$ .
- Entre  $10^{17}$  et  $2 \times 10^{20} \text{ cm}^{-2}$  on parle de systèmes à la limite de Lyman (*Lyman limit system* ou LLS en anglais). Ces systèmes sont suffisamment denses pour que tous les photons capables d'ioniser l'hydrogène soit absorbés. Le spectre d'un quasar dont la lumière a traversé un LLS a donc un flux nul pour une longueur d'onde inférieure à la limite de Lyman, dans le référentiel du LLS, c'est-à-dire pour une longueur d'onde observée  $\lambda_{obs} < 911,3 \times (1 + z_{LLS}) \text{ \AA}$ .
- Au-delà de  $2 \times 10^{20} \text{ cm}^{-2}$  les systèmes sont trop denses pour être ionisés par les photons ultra-violet provenant des galaxies à formation d'étoiles et des AGN. L'essentiel de l'hydrogène y est donc neutre. Ces systèmes absorbent l'intégralité du flux du quasar aux longueurs d'ondes de la série de Lyman (dans leur propre référentiel) et au-delà de la limite de Lyman. Ils présentent également des absorptions causées par des métaux. Ils sont appelés *damped Lyman- $\alpha$  systems* (DLA) en anglais. Les DLA sont perçus comme étant associés à des galaxies ou proto-galaxies. Un spectre de quasar présentant un DLA est visible en figure 2.8. L'absorption causée par le DLA dans la forêt Lyman- $\alpha$  est clairement visible, ainsi que la limite de Lyman associée.

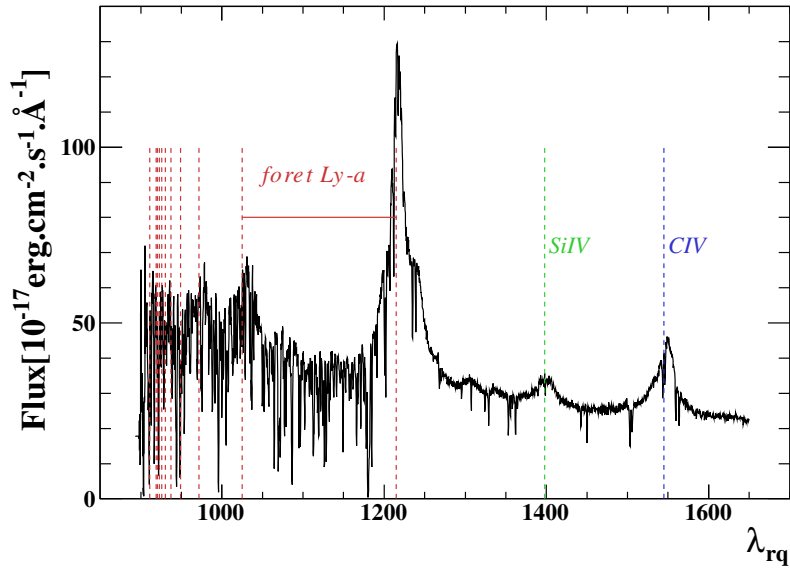


FIGURE 2.7: Spectre d'un quasar de BOSS à un décalage spectral de 3,29. Le spectre est tracé dans la longueur d'onde du référentiel du quasar  $\lambda_{rq}$ . Le signal sur bruit dans la forêt est de 42,5 ce qui est particulièrement élevé pour un spectre de BOSS (voir figure 5.2). Les pointillés verticaux rouges représentent la série Lyman- $\alpha$ , avec de la droite vers la gauche, la raie Lyman- $\alpha$ , puis la raie Lyman- $\beta$ , puis la raie Lyman- $\gamma$  et ainsi de suite jusqu'au continu de Lyman à 911,3 Å.

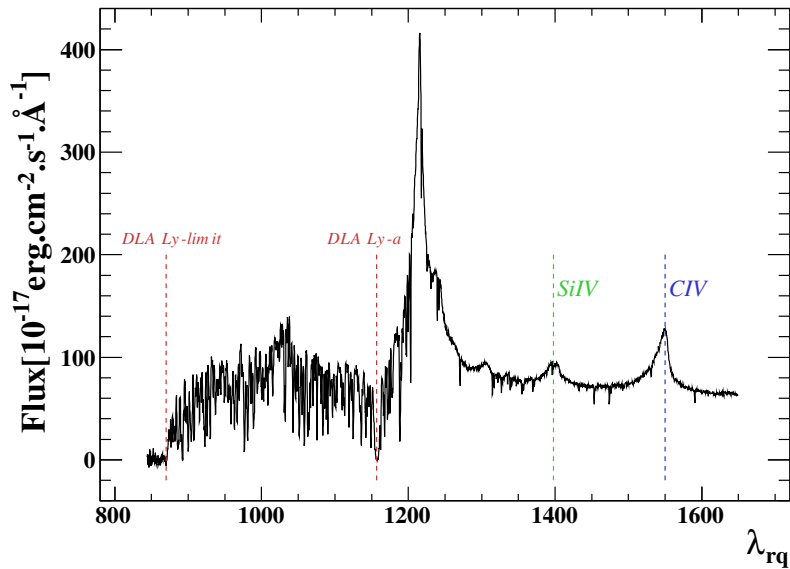


FIGURE 2.8: Spectre d'un quasar de BOSS présentant un système DLA. Le signal sur bruit moyen du spectre dans la forêt est particulièrement élevé pour BOSS (SNR  $\sim 53$ ).

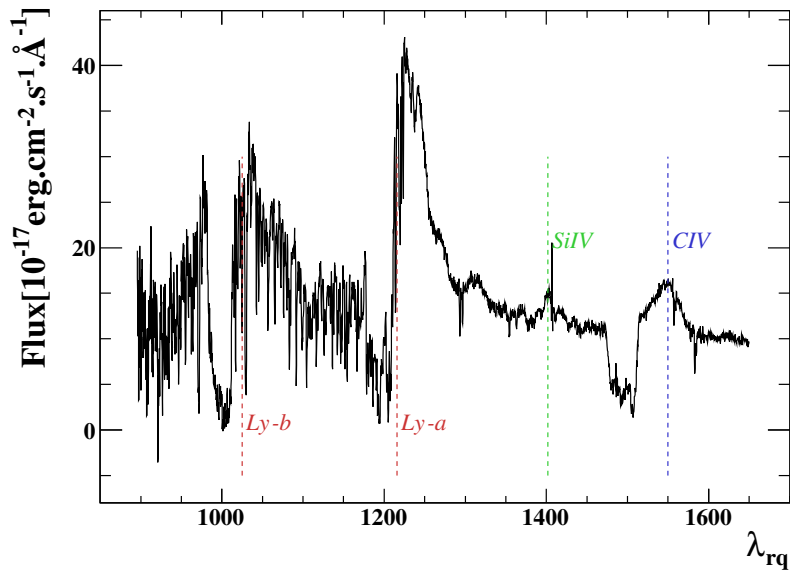


FIGURE 2.9: Spectre d'un quasar de BOSS présentant des BAL. Les BAL sont clairement identifiables à gauche des raies d'émission CIV, Lyman- $\alpha$  et Lyman- $\beta$ .

La figure 2.9 présente un cas particulier de spectre de quasar présentant des raies d'absorption particulièrement larges (BAL pour l'anglais *Broad Absorption Lines*) qui ne sont pas associées qu'à la série de Lyman. Sur l'exemple donné, des BAL se situent à gauche des raies d'émission Lyman- $\alpha$  et Lyman- $\beta$  du quasar, mais également des raies CIV (particulièrement visible) et SiIV (plus faible). Les BAL sont comprises comme résultant de l'absorption du flux du quasar par la matière éjectée par le quasar lui-même, notamment composée de métaux (d'où les absorptions liées aux SiIV et CIV). La vitesse de la matière éjectée par rapport au quasar fait que les BAL se situent dans la majorité des cas à plus faibles longueurs d'onde que les raies d'émissions associées. Dans de rares cas, elles peuvent toutefois se situer à plus grandes longueurs d'onde.



## Chapitre 3

# The Baryon Oscillation Spectroscopic Survey

*« I like nonsense, it wakes up the brain cells. Fantasy is a necessary ingredient in living, it's a way of looking at life through the wrong end of a telescope. »*

—Dr. Seuss

---

Le relevé spectrographique BOSS est dédié à la mesure des oscillations acoustiques de baryons dans la distribution spatiale des galaxies et grâce aux forêts Lyman- $\alpha$  des quasars. Ce chapitre détaille les principales caractéristiques du relevé : le télescope utilisé, le fonctionnement de la caméra et des spectrographes ainsi que la réduction des données.

---

### 3.1 The Sloan Digital Sky Survey

#### 3.1.1 SDSS et SDSS-II

Le Sloan Digital Sky Survey a été pensé durant les années 80, son but premier étant la compréhension des structures à grandes échelles de l'Univers. Le développement rapide des capteurs CCDs joint à l'explosion des capacités en traitement de données, ont rendu possible l'idée d'un relevé du ciel, à la fois profond et étendu. Le projet a abouti à la construction d'un télescope de 2,5 m de diamètre de type Ritchey-Chrétien basé à l'Apache Point Observatory au Nouveau Mexique (Etats-Unis). Situé à 2.800 mètres d'altitude, ce télescope a débuté les observations en 2000. Il est étudié en détails en section 3.4.1. De 2000 à 2005 puis de 2005 à 2008, les deux premières générations du SDSS ont mené un relevé photométrique multi-bandes couvrant environ  $11.600 \text{ deg}^2$  du ciel de l'hémisphère Nord, associé à un relevé spectroscopique. La combinaison des données astrométriques du relevé photométrique et des décalages spectraux des objets, déterminés grâce aux données spectroscopiques, permet la création d'une carte en trois dimensions d'une partie de notre Univers. A la fin du SDSS-II, cette carte contenait plus de 930.000 galaxies, 120.000 quasars



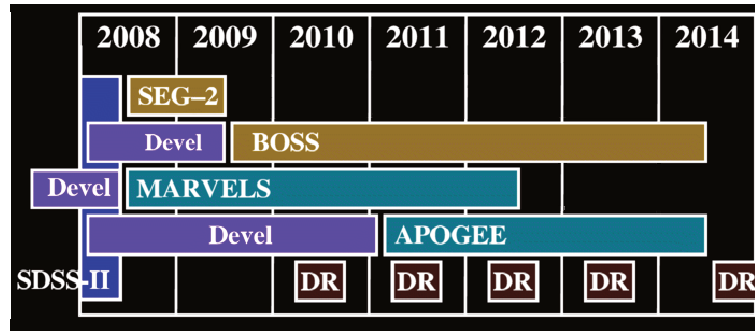


FIGURE 3.1: Déroulement prévu du SDSS-III. Les mentions « DR » (*Data Release*) indiquent les périodes de prise de données rendues publiques un an après leur fin. La mention « Devel » correspond aux phases de développement. *Source* : <http://www.sdss3.org>

et 460.000 étoiles. Cette carte, ainsi que les spectres des objets ont été rendus publiques lors de la septième publication de données du SDSS (*Data Release 7* ou DR7) et sont accessibles via le site web du SDSS<sup>1</sup>. L'étude de cette carte et notamment de la distribution des objets qu'elle contient, a eu un impact considérable sur nos connaissances en astrophysique et en cosmologie. Au moment d'écrire ces lignes, le nombre de publications faisant référence au SDSS dans leur titre vient de dépasser le nombre de 5.000 (incluant les publications concernant la troisième génération du SDSS) et cumulent plus de 200.000 citations. Parmi les résultats scientifiques majeurs du SDSS se trouve la première détection à  $3,6\sigma$  des oscillations acoustiques de baryons par Eisenstein et al. (2005). Effectuée à l'aide d'un sous échantillon de plus de 46.000 galaxies lumineuses rouges, cette détection a permis la mesure de la distance absolue à un décalage spectral  $z = 0,35$  avec une incertitude de 5%.

### 3.1.2 SDSS-III

La troisième génération du SDSS a commencé à l'automne 2008, à la suite du SDSS-II. Basée sur le même télescope que les générations précédentes, elle est constituée de quatre relevés différents : The Apache Point Observatory Galactic Evolution Experiment (APOGEE), the Multi-object APO Radial Velocity Exoplanet Large-area Survey (MARVELS), the Sloan Extension for Galactic Understanding and Exploration 2 (SEGUE-2) et the Baryon Oscillation Spectroscopique Survey (BOSS). Le déroulement prévu du SDSS-III est illustré en figure 3.1. Dans cette section, les trois premiers relevés cités sont brièvement présentés. Une description plus complète du relevé BOSS, sur lequel a été effectué cette thèse, est donnée dans la suite de ce chapitre.

#### SEGUE-2

The Sloan Extension for Galactic Understanding and Exploration a produit les spectres d'environ 118.000 étoiles appartenant au halo stellaire de notre galaxie. Possédant des magnitudes apparentes inférieures à 19, ces étoiles appartiennent à différentes populations mais sont toutes situées entre 10 kpc et 60 kpc du centre galactique. Combinées avec les 230.000 spectres d'étoiles obtenus par SEGUE-1, les données de SEGUE-2 révèlent la complexité des sous-structures cinématiques et chimiques du halo stellaire de la Voie Lactée, nous renseignant sur la façon dont

1. <http://www.sdss.org/dr7/>

notre galaxie s'est assemblée et enrichie en métaux. SEGUE-2 a enregistré ses données pendant les périodes sombres, entre l'automne 2008 et le printemps 2009, s'achevant ainsi avant le démarrage de BOSS à l'automne 2009. Il a utilisé un spectrographe couvrant une gamme en longueur d'onde allant de 385 nm à 920 nm. La résolution moyenne des spectres obtenus est de 2.000 pour un signal sur bruit typique de 25.

### APOGEE

The APO Galactic Evolution Experiment est un relevé visant à observer environ 100.000 étoiles géantes rouges de magnitudes apparentes inférieures à 12,5 et situées en différents endroits de notre galaxie (bulbe, disque, bar et halo). En mesurant précisément les vitesses particulières ainsi que la composition chimique des étoiles, APOGEE étudie la dynamique ainsi que l'histoire chimique de notre galaxie. La prise de données a commencé au printemps 2011 et devrait durer jusqu'au printemps 2014. Tandis que BOSS observe durant les périodes « sombres » (quand la partie éclairée du disque lunaire est inférieure à 60% et la Lune se situe sous l'horizon), APOGEE observe durant les périodes « claires ». Le signal sur bruit moyen des spectres d'APOGEE (S/B  $\sim$  100) ainsi que la résolution type de 20.000 sont bien supérieurs à ceux de SEGUE-2, ce qui est évidemment lié à la plus faible magnitude des étoiles observées.

### MARVELS

The Multi-object APO Radial Velocity Exoplanet Large-area Survey était un relevé spectroscopique visant à observer 11.000 étoiles brillantes de notre galaxie. MARVELS devait observer chaque étoile entre 25 et 35 fois sur une période de 18 mois, étudiant leur vitesse radiale à la recherche d'exoplanètes gazeuses géantes. Démarré à l'automne 2008, ce relevé visait à contraindre les modèles théoriques de formation et d'évolution des systèmes à planètes géantes. Il était mené en parallèle de BOSS, pendant les périodes claires d'observation, et devait être achevé au printemps 2014. La résolution des spectres requise pour détecter les exoplanètes géantes n'ayant pas été atteinte, ce projet a été arrêté en 2012.

## 3.2 Objectifs de BOSS

Comme son nom l'indique, BOSS est un relevé destiné à l'observation des BAO. Le but du relevé est d'atteindre une sensibilité suffisante pour mesurer de manière indépendante le paramètre de Hubble  $H(z)$ , grâce à une observation du pic acoustique le long de la ligne de visée, et la distance angulaire  $D_A(z)$ , grâce à une observation du même pic mais transversalement à la ligne de visée. Ces mesures peuvent alors être transformées en contraintes sur les paramètres cosmologiques tels que  $\Omega_M$ ,  $\Omega_\Lambda$ ,  $w_0$ , etc. Pour obtenir de telles contraintes, BOSS mène en parallèle deux relevés spectroscopiques couvrant plus de 10.000 deg<sup>2</sup> de ciel.

Le premier est dédié à l'observation de 1,5 millions de galaxies ayant des décalages spectraux compris dans l'intervalle  $0,15 < z < 0,7$  avec une densité moyenne de 150 deg<sup>-2</sup>. Ce relevé peut lui-même être subdivisé en deux échantillons. Un premier échantillon à bas décalage spectral ( $0,15 < z < 0,43$  avec un décalage moyen  $\bar{z} = 0,3$ ) vise à compléter celui commencé par SDSS puis repris par SDSS-II. L'objectif est d'augmenter la densité de LRG passant ainsi de  $\sim 10$  deg<sup>-2</sup> pour SDSS-II à 30 deg<sup>-2</sup>. Un second échantillon à plus grand décalage spectral ( $0,43 < z < 0,7$  pour un décalage moyen  $\bar{z} = 0,57$ ), appelé CMASS (pour *constant mass*), élargit la sélection des galaxies, atteignant une densité de surface de 120 deg<sup>-2</sup>. La figure 3.2 présente la précision attendue sur la distance volumique  $D_V$  (définie en section 1.4.5) par l'échantillon CMASS à la fin

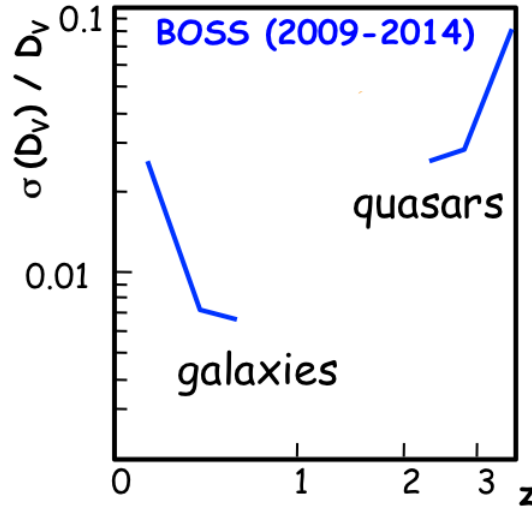


FIGURE 3.2: Prévisions de sensibilité sur la distance volumique  $D_V$  à la fin de BOSS pour le relevé de galaxies et celui de quasars.

du relevé. Elle est l'ordre de 7 pour mille. Pour la première fois, cette précision ne sera plus limité par le bruit de grenaille (en anglais *shot noise*) mais par la variance cosmique.

Le deuxième relevé a pour but d'observer un minimum de 150.000 quasars à grands décalages spectraux ( $2,15 < z < 3,5$ ) avec une densité de surface de  $17 \text{ deg}^{-2}$ . L'objectif est de sonder le milieu intergalactique le long de la ligne de visée de chaque quasar, via l'absorption en Lyman- $\alpha$  visible dans leur spectre. Il est alors possible de calculer la fonction de corrélation de l'hydrogène neutre du milieu interstellaire à un décalage spectral moyen  $\bar{z} \approx 2,3$ . La précision attendue sur la distance  $D_V(2,3)$  à la fin du relevé, visible en figure 3.2, est d'environ 3%.

En parallèle de ces deux relevés principaux, environ 5% des cibles observées par les spectrographes sont décernées à des programmes scientifiques de plus faibles envergures appelés *ancillary programs*. Ces programmes sont proposés par les membres du SDSS-III et traitent d'une grande variété de sujets tels que l'étude des blazars de hautes énergies, l'étude des étoiles de très faibles masses ou de la variabilité des BAL (voir section 2.4.3). La sélection des quasars par variabilité extrême décrite en section 4.3.2, faisant partie du travail réalisé pour cette thèse, est un programme *ancillary*.

### 3.3 Stratégie d'observation

La phase de spectroscopie, actuellement en cours, a été précédée d'une phase de photométrie multi-bandes, aujourd'hui achevée, venant compléter les données photométriques de SDSS et SDSS-II. La carte du ciel en deux dimensions ainsi obtenue sert de base à la sélection des objets à observer durant la phase spectroscopique. Les principales caractéristiques des deux phases du relevé sont ici décrites.

#### 3.3.1 Phase photométrique

Entre 2000 et 2005, SDSS a produit la photométrie de  $11.600 \text{ deg}^2$  du ciel de l'hémisphère Nord terrestre, dont  $7.600 \text{ deg}^2$  formant une région continue de l'hémisphère Nord galactique

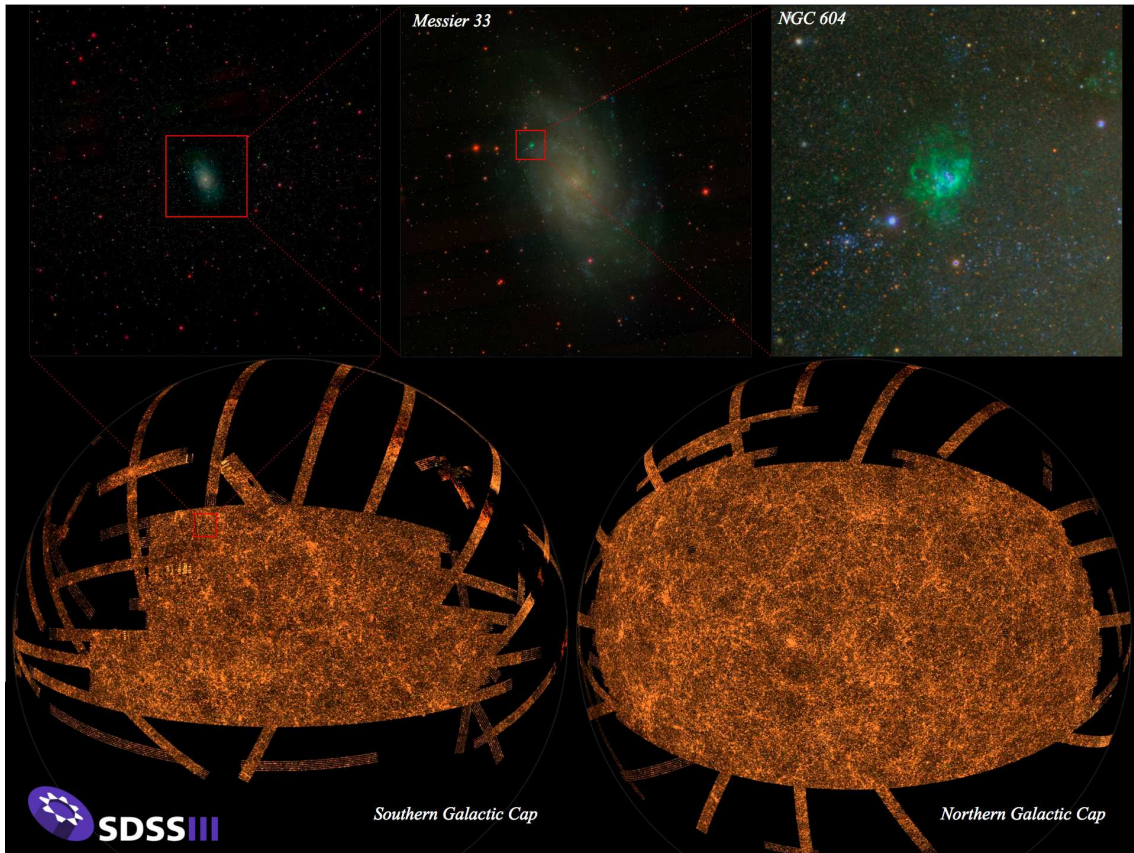


FIGURE 3.3: Image des  $14.555 \text{ deg}^2$  de ciel observés en photométrie par SDSS. L'agrandissement sur la galaxie Messier 33 donne un aperçu de la qualité du relevé. Les deux régions du pôle Sud galactique et pôle Nord galactique sont représentées. Les bandes, résultant de la stratégie d'observation en TDI *drift scanning* sont en partie visibles.

(NGC pour l'anglais North Galactic Cap). La photométrie a été obtenue dans cinq bandes différentes ( $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$ ) à l'aide d'une caméra dédiée de 54 CCDs. Durant les automnes 2008 et 2009, BOSS a réutilisé la caméra du SDSS pour compléter la photométrie du pôle Sud galactique (SGC), obtenant une zone continue d'environ  $3.100 \text{ deg}^2$ . L'image finale, couvrant  $14.555 \text{ deg}^2$ , est visible en figure 3.3. Elle contient plus de 932 millions d'objets détectés individuellement. L'ensemble des données a été rendu public lors de la huitième publication de données (Data Release 8) et est disponible via le site internet du SDSS<sup>2</sup>. Les principales caractéristiques techniques de la caméra utilisée pour réaliser la photométrie sont décrites en section 3.4.2.

La prise de données a été réalisée sous conditions photométriques<sup>3</sup>, lorsque la Lune était sous l'horizon et lorsque le *seeing* dans la bande photométrique  $r$  était meilleur que  $2''$ . La technique utilisée est celle du *time-delay and integrate* (TDI) *drift scanning* : le télescope est laissé fixe, pointant vers une certaine direction, tandis que le ciel défile sous l'effet de la rotation de la Terre. La caméra intègre la lumière d'un objet, de son entrée dans le plan focal, jusqu'à sa sortie. Ce mode de fonctionnement permet de minimiser le temps de lecture des CCDs (qui serait comparable

2. <http://www.sdss3.org/dr8/>

3. Les conditions photométriques correspondent à un ciel sans nuage, pour lequel l'extinction atmosphérique dépend simplement de la colonne d'air.

au temps d'exposition pour un fonctionnement en « photographie standard ») ainsi que le temps de pointage du télescope. Il permet d'atteindre une efficacité d'observation dépassant les 90% du temps disponible. Le temps d'exposition effectif pour chaque objet est de 55 s dans chacune des bandes photométriques. Le champ de vue du télescope est si large ( $3^\circ$  de diamètre) qu'il est nécessaire d'effectuer le TDI *drift scanning* le long de grands cercles.

### 3.3.2 Phase spectroscopique

Les mesures spectroscopiques nécessitent des temps d'intégration substantiellement plus longs que les mesures photométriques. Il n'est donc pas envisageable de produire les spectres de l'intégralité des objets détectés en photométrie. La première phase du relevé spectroscopique consiste à sélectionner l'ensemble des cibles dont on souhaite faire le spectre. Une fois les cibles sélectionnées, l'étape dite du « pavage » permet d'optimiser les observations en minimisant le temps nécessaire pour achever le relevé. Une fois ces deux étapes réalisées, il est possible de démarrer les observations. Cette section présente chacune des étapes du relevé spectroscopique.

#### Sélection des objets cibles

La sélection des quasars a été effectuée sur une région couvrant  $7.578 \text{ deg}^2$  du NGC et  $2.663 \text{ deg}^2$  du SGC, pour un total d'environ  $10.200 \text{ deg}^2$ , bien inférieur aux  $14.555 \text{ deg}^2$  de photométrie disponibles. La zone considérée a été réduite afin de se limiter aux régions hors du plan galactique, où la sélection des objets cibles est à la fois plus complète et plus uniforme. Le manque d'information dans la bande photométrique  $u$  dans certaines régions ainsi que la volonté d'achever le relevé dans le temps imparti ont également amené à diminuer la zone dédiée à la spectroscopie. Notons que grâce à une météo très favorable, nous sommes aujourd'hui en avance par rapport au programme ; le relevé devrait être achevé au bout de 87,5% du temps initialement prévu. L'utilisation des 12,5% restant (correspondant au premier semestre 2014) est actuellement débattu par la collaboration.

Pour obtenir la sensibilité souhaitée sur l'échelle BAO grâce aux forêts Lyman- $\alpha$  des quasars à grands décalages spectraux, il est nécessaire d'atteindre une densité moyenne de  $15 \text{ deg}^{-2}$  quasars correspondant à une magnitude limite  $g < 22,0$  (McDonald and Eisenstein, 2007; McQuinn and White, 2011). Cette sensibilité montre une dépendance linéaire à la densité de quasars observés pour un signal sur bruit fixe. L'enjeu de la sélection est donc de maximiser la densité de quasars observés. Différentes méthodes ont été développées pour sélectionner les quasars par leur photométrie, la majorité d'entre elles étudiant la position des objets dans des diagrammes couleur-couleur. Ross et al. (2012) ont mené une étude détaillée de la sélection des quasars pour BOSS. Le chapitre 4 présente une méthode de sélection des quasars par variabilité qui représente l'un des travaux de cette thèse. La sélection des quasars est divisée en plusieurs échantillons, les principaux étant :

- Un échantillon dénommé *CORE* d'objets sélectionnés de manière uniforme sur l'essentiel du relevé. La méthode de sélection de l'échantillon *CORE* a été fixée environ un an après le début du relevé et n'a pas été modifiée depuis. Le nombre de cibles *CORE* est variable en fonction de la latitude galactique et peut aller de 15 à  $30 \text{ deg}^{-2}$ .
- Un échantillon appelé *BONUS* d'objets sélectionnés de manière non-uniforme pour maximiser la densité de quasars. La méthode de sélection de l'échantillon *BONUS* varie en fonction des optimisations des différentes méthodes et de la disponibilité de données complémentaires telles que des données photométriques venant de relevés différents. Le nombre de cibles de l'échantillon *BONUS* est le complémentaire de l'échantillon *CORE* pour atteindre

40  $\text{deg}^{-2}$  cibles.

- Un échantillon *KNOWN* de quasars à grands décalages spectraux déjà connus. La majorité de ces quasars ont été découverts par SDSS et SDSS-II. La plus large bande passante des spectrographes de BOSS, notamment du côté bleu où se situe la forêt Lyman- $\alpha$ , ainsi que leur meilleure efficacité de détection, justifient la ré-observation de ces quasars. Cette ré-observation permet également d'assurer l'uniformité des spectres de quasar sur l'ensemble du relevé.

Avec cette stratégie et un budget de 40  $\text{deg}^{-2}$  cibles, BOSS observe de 15  $\text{deg}^{-2}$  à 18  $\text{deg}^{-2}$  quasars à  $z > 2.15$ , en fonction de la quantité d'étoile et de la disponibilité de données complémentaires. La méthode de sélection par variabilité, présentée au chapitre 4, appliquée dans une région de 220  $\text{deg}^2$  ayant servi à la recherche de supernovae et disposant donc de données photométriques multi-époques, permet d'atteindre une densité moyenne de 24  $\text{deg}^{-2}$  quasars.

### « Pavage » du ciel

La prise de données spectroscopiques s'effectue à l'aide de 1000 fibres optiques, insérées dans une plaque rigide positionnée au plan focal du télescope, couvrant un champ de vue de 7  $\text{deg}^2$  (voir section 3.4). Les plaques sont percées antérieurement aux observations, aux endroits où doivent être positionnées les fibres optiques. Ce protocole impose de découper l'intégralité des 10.000  $\text{deg}^2$  de ciel qui constituent le relevé en plaques de 3 degrés de diamètre, et d'attribuer à chaque plaque les cibles appartenant à leur champ de vue. Cette procédure, dénommée « pavage » (ou *tiling* en anglais), est entièrement décrite par Blanton et al. (2003). Son objectif est de maximiser le nombre de cibles se voyant assigner une fibre, tout en minimisant le nombre de plaques nécessaires pour compléter l'intégralité du relevé. Le pavage n'est pas effectué en une seule fois sur l'intégralité des 10.000  $\text{deg}^2$ , mais réalisé indépendamment sur 31 morceaux, dénommés de BOSS1 à BOSS31. De cette façon le pavage peut prendre en compte les modifications d'algorithmes de sélection des objets cibles même une fois le relevé débuté.

Le pavage d'une bande commence par la distribution uniforme des plaques sur la zone à couvrir. Cette distribution est ensuite perturbée pour prendre en compte les variations de densité des cibles. Sur les 1.000 fibres que comporte une plaque, 80 sont positionnées à des endroits où le relevé photométrique n'a pas détecté de sources. Ces fibres dites « du ciel » permettent de mesurer la contribution du fond du ciel au bruit des données. 20 fibres sont également attribuées à l'observation d'étoiles de type F de la séquence principale permettant l'étalonnage du rendement du système en fonction de la longueur d'onde. Sur les 900 fibres restantes, 5 sont attribuées à des objets également observés par d'autres plaques, permettant le contrôle de la reproductibilité des résultats spectroscopiques. Les 895 autres fibres se répartissent approximativement de 160 à 200 pour les cibles quasars, de 560 à 630 pour les cibles galaxies, et de 20 à 90 pour les cibles de programmes *ancillary*. La nécessité de faire se recouvrir les plaques, pour éviter la présence de trous dans la couverture du ciel, limite l'efficacité du pavage. Cette efficacité est définie comme la fraction des 895 fibres attribuées à des cibles quasars, galaxies ou *ancillary* uniques. Elle est en moyenne de 0,927. Les fibres restantes sont attribuées à des cibles de priorités inférieures telles que des objets déjà observés.

### Préparation des plaques et observations

Une fois le pavage d'une bande fixé, les plaques sont percées à l'Université de Washington, plusieurs mois avant les observations. Pour déterminer la position des trous, le temps sidéral propice à l'observation de chaque plaque est estimé. La position de chaque fibre est alors déterminée en tenant compte de la réfraction atmosphérique. Cette réfraction dépendant de la longueur

d'onde, la position de la fibre ne sera pas la même suivant que l'on souhaite optimiser la partie bleue, ou la partie rouge du spectre. Pour les quasars, la position des trous est centrée sur la lumière à 4.000 Å pour maximiser le signal sur bruit dans la forêt Lyman- $\alpha$ . Pour les galaxies, elle est optimisée pour 5.400 Å. Le plan focal à 4.000 Å diffère de celui à 5.400 Å de 0 à 300  $\mu\text{m}$  en fonction de la distance à l'axe optique (centre de la plaque). Pour corriger cet effet, des oeillets adhérents sont collés à l'arrière de la plaque pour les cibles quasars, permettant de positionner les fibres optiques légèrement en retrait du plan focal à 5.400 Å. En plus des 1.000 trous correspondant aux 1.000 cibles mentionnées précédemment, 16 trous sont percés pour permettre le pointage et le suivi pendant l'observation. Des trous sont également percés aux emplacements des étoiles brillantes pour éviter les réflexions sur la surface de la plaque.

Une fois les plaques percées, elles sont envoyées sur le site de l'observatoire, où chaque jour le personnel prépare les observations qui se tiendront la nuit. Les plaques devant être observées pendant la nuit sont positionnées sur les cartouches le jour même. Les 1000 fibres sont alors branchées une par une à la main. La préparation d'une cartouche dure environ 45 minutes et requiert deux personnes. Neuf cartouches peuvent être branchées simultanément.

Les observations débutent quand le Soleil est à environ 12° sous l'horizon. La première cartouche devant être observée est alors montée au plan focal du télescope. L'observation d'une cartouche débute par la mesure des lampes à arc, utile pour l'étalonnage des spectres, puis se poursuit avec 15 minutes de prise de données scientifiques. Au bout des 15 minutes, un programme simple de réduction des données réalise l'étalonnage et l'extraction des spectres à une dimension. Cette extraction simplifiée a pour but d'estimer le signal sur bruit des spectres. L'observation d'une cartouche est considérée terminée lorsque tous les spectres de galaxies ont un signal sur bruit au carré  $(S/B)^2 \geq 20$  pour la caméra rouge et un  $(S/B)^2 \geq 10$  pour la caméra bleue. Le critère utilise le signal sur bruit au carré car il dépend linéairement du temps d'exposition. Il a été défini de sorte à maximiser la capacité du programme complet de réduction des données à déterminer le décalage spectral des galaxies, tout en minimisant le temps d'exposition pour achever le relevé à temps. A la fin des 15 minutes d'exposition, si le critère n'est pas vérifié, la même plaque est observée pendant 15 minutes additionnelles. Si le critère est vérifié, le personnel procède à la mise en place de la prochaine cartouche. A la fin d'une nuit d'observation, si la dernière plaque n'est pas terminée, celle-ci n'est pas débranchée et est ré-observée la nuit suivante. Il n'y a pas de critère sur le signal sur bruit des spectres de quasars car il a été suggéré (McDonald and Eisenstein, 2007; Font-Ribera et al., 2012) que le gain de sensibilité sur l'échelle BAO venant d'une couverture plus large du ciel, dépasse celui venant de l'amélioration du signal sur bruit des spectres. Cette affirmation est aujourd'hui remise en question pour la préparation du prochain relevé eBOSS.

### 3.4 Caractéristiques techniques de BOSS

Comme décrit dans Gunn et al. (2006), l'exécution d'un relevé photométrique multi-bandes, couvrant une large partie du ciel observable et associé à un relevé spectroscopique d'une telle envergure, a nécessité un certain nombre d'innovations techniques :

- Un télescope possédant un très grand champ de vue, un plan focal très peu distordu, une orientation très précise ainsi que la capacité d'alterner entre photométrie et spectroscopie. Les principales caractéristiques optiques du télescope conçu pour ces objectifs sont décrites en section 3.4.1.
- Une caméra couvrant l'intégralité du champ de vue du télescope et permettant l'observation simultanée dans différentes bandes photométriques avec une précision suffisante pour la calibration astrométrique. La caméra du SDSS est brièvement étudiée en section 3.4.2.

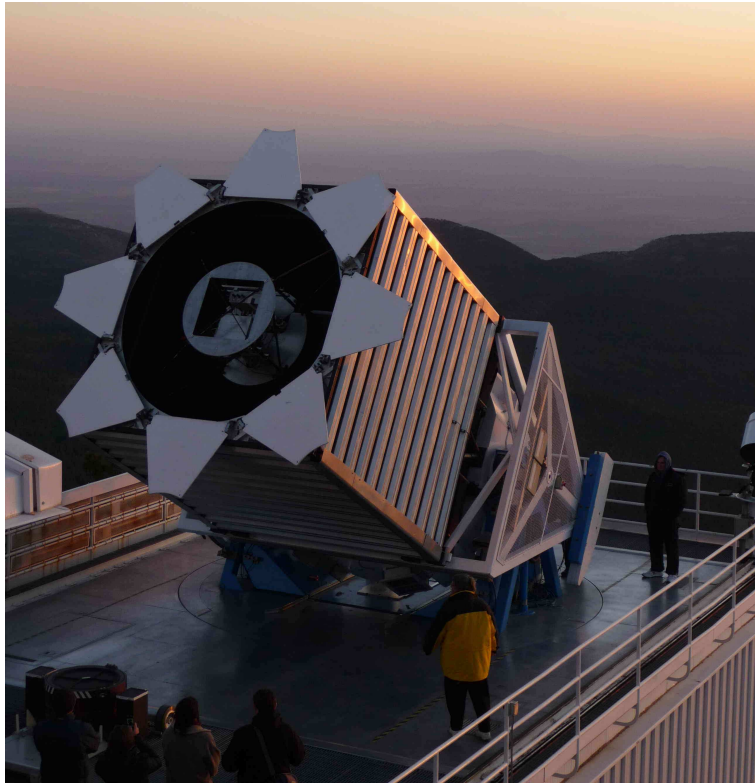


FIGURE 3.4: Photographie du télescope de 2,5 m de diamètre utilisé par le SDSS.

- Un système de spectrographes capables d’obtenir simultanément les spectres de plusieurs centaines d’objets avec une large couverture en longueur d’onde. Les spectrographes de BOSS sont étudiés en détails en section 3.4.3.
- Un système d’acquisition permettant le stockage des données ainsi qu’un contrôle en temps réel de leur qualité.
- Un système de traitement des données capable de rapidement calibrer les observations et d’identifier les objets à la fois pour la photométrie et pour la spectroscopie. La procédure de réduction des données spectroscopiques est détaillée en section 3.5.

### 3.4.1 Le télescope

Pour remplir les objectifs du SDSS, le télescope utilisé doit détenir un grand champ de vue, un plan focal présentant extrêmement peu de distorsion, un pointage très précis ainsi qu’une capacité à alterner entre photométrie et spectroscopie. C’est dans cette optique qu’un télescope de 2,5 m, possédant un champ de vue de  $3^\circ$ , a été conçu puis construit à l’Apache Point Observatory (APO). Cet observatoire est situé dans l’état du Nouveau Mexique aux Etats-Unis, à 2.800 mètres d’altitude. Une photo du télescope est visible en figure 3.4. Cette section présente les principales caractéristiques optiques du télescope. Une étude complète de l’optique ainsi que de la mécanique a été menée par Gunn et al. (2006).

Le télescope du SDSS est de type Ritchey-Chrétien, ce qui signifie que les deux miroirs (primaire et secondaire) sont hyperboliques. Sa conception diffère des Ritchey-Chrétien classiques par l’ajout de deux correcteurs optiques. Le premier est un correcteur Gascoigne, permettant de



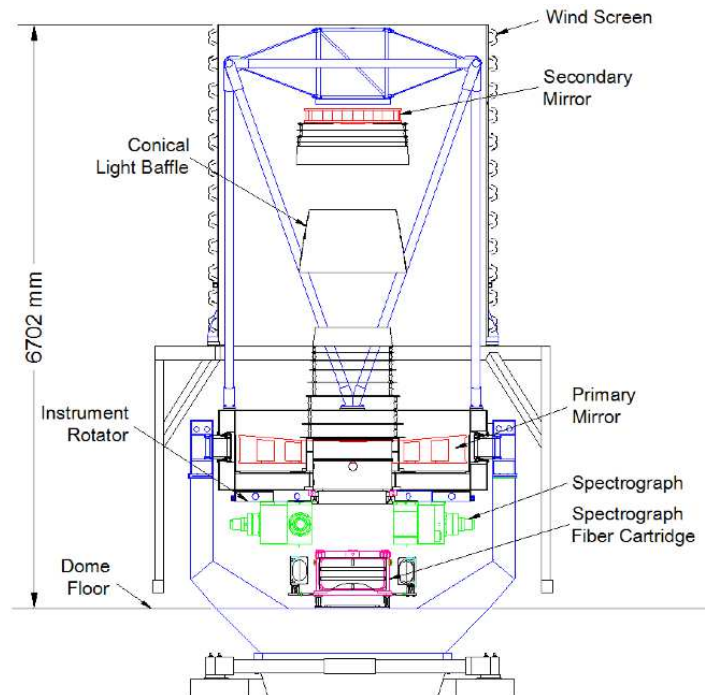


FIGURE 3.5: Schéma du télescope du SDSS. Les miroirs primaire et secondaire sont représentés en rouge. Egaleme nt visibles sont les spectrographes, en vert, et une cartouche (en anglais, cartridge), en rose (voir section 3.4.3).

réduire l'astigmatisme du système optique. Le second consiste en une paire de correcteurs interchangeables suivant que le télescope soit en prise de données photométriques ou spectroscopiques. Un schéma du télescope est visible en figure 3.5.

### Les miroirs primaire et secondaire

Le miroir primaire est un miroir de 2,5 m de diamètre et de rapport focale sur diamètre  $f/2,25$ . Il est percé en son centre d'un trou de 1,17 m de diamètre. Ce trou est surmonté d'une paroi conique d'un diamètre maximal légèrement supérieur (1,20 m) empêchant les lumières parasites de se diriger vers le plan focal du télescope.

Le miroir secondaire a un diamètre de 1,08 m. C'est un miroir sphérique, de rayon de courbure  $7.331 \pm 6$  mm. Il est entouré d'une paroi conique de 1,30 m de diamètre au maximum, le protégeant des lumières parasites. L'ensemble miroir secondaire plus parois obstrue 27% du flux de lumière arrivant sur le télescope.

L'ensemble des miroirs primaire et secondaire forment un système optique de rapport focale sur diamètre  $f/5,0$  et de  $3^\circ$  de champ de vue. Le miroir secondaire étant placé à seulement 3,6 m du miroir primaire, le plan focal se situe 0,76 m en retrait du miroir primaire, permettant un accès facile pour les instruments de mesure.

### Les correcteurs

Le premier correcteur est de type Gascoigne. Il constitue la dernière pièce d'optique commune aux configurations photométrique et spectroscopique, c'est pourquoi il est parfois désigné comme

correcteur « commun ». Son rôle est de diminuer significativement l'astigmatisme, caractéristique des télescopes de type Ritchey-Chrétien. C'est une lentille de 802 mm de diamètre d'environ 12 mm d'épaisseur.

Le second correcteur, ou correcteur « final » diffère suivant que le télescope soit utilisé en photométrie ou en spectroscopie. Il est donc optimisé suivant les deux cas de figure. Le correcteur photométrique est brièvement étudié dans la section 3.4.2, dédiée à la caméra, tandis que le correcteur spectroscopique est décrit dans la section 3.4.3 dédiée aux spectrographes de BOSS.

### 3.4.2 La caméra

La caméra utilisée par BOSS pour sa première phase de photométrie est la même que celle utilisée lors des deux premières générations du SDSS. Elle a été optimisée pour fonctionner en TDI *drift scanning* et couvrir au mieux le champ de vue de 3° du télescope. Une description détaillée de cette caméra a été réalisée par Gunn et al. (1998).

La première pièce de la caméra est un correcteur optique ayant pour but de corriger la distorsion du plan focal induite par l'optique du télescope. Il s'agit d'une lentille de 45 mm d'épaisseur au centre, possédant une face d'entrée fortement asphérique et une face de sortie plate. En plus de son rôle de correcteur optique, cette lentille est un élément structurel de la caméra sur laquelle les détecteurs sont montés. Elle joue un rôle primordial dans le maintien des propriétés mécaniques de la caméra, influant directement sur la qualité des images et de l'astrométrie.

La caméra est ensuite composée de 6 colonnes de 5 capteurs CCDs (un pour chaque bande photométrique  $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$ ) de  $2.048 \times 2.048$  pixels. L'espace restant au plan focal est utilisé pour positionner 24 capteurs CCDs de plus petite taille ( $2.048 \times 400$  pixels), 22 d'entre eux servant à l'astrométrie et 2 à la mise au point. Ces 24 capteurs sont répartis de part et d'autre des CCDs destinés à l'imagerie, comme visible en figure 3.6. La direction du TDI *drift scanning* est vers le haut sur cette figure, si bien qu'une étoile entrant dans le plan focal rencontre premièrement un CCD d'astrométrie, puis successivement les CCDs de photométrie  $r$ ,  $i$ ,  $u$ ,  $z$  et  $g$ , et enfin un deuxième CCD d'astrométrie.

Les CCDs destinés à la photométrie sont tous identiques mais surmontés de filtres permettant de découper leur gamme de sensibilité (allant de la coupure atmosphérique dans l'UV située aux alentours de 3.000 Å jusqu'à la limite de sensibilité des détecteurs silicium à 11.000 Å) en 5 bandes photométriques  $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$ . Les gammes de longueur d'onde définies par ces filtres ainsi que les efficacités quantiques<sup>4</sup> du système complet (optique, filtre et CCD) dans chaque bande sont représentées en figure 3.7. La très faible efficacité dans la bande  $u$  justifie son positionnement au centre du plan focal. Les magnitudes limites de détection sont définies pour un signal sur bruit  $S/B = 5$ . Pour une étoile, elles valent approximativement  $u = 22,1$  mag,  $g = 23,2$  mag,  $r = 23,1$  mag,  $i = 22,5$  mag et  $z = 20,8$  mag.

### 3.4.3 Les spectrographes

BOSS est constitué de deux spectrographes identiques. La conception de ces spectrographes hérite largement des technologies utilisées pour ceux des deux premières générations du SDSS, leur caractéristique principale étant qu'ils sont alimentés par un ensemble de fibres optiques, positionnées au plan focal du télescope. Toutefois, pour atteindre les objectifs de BOSS, différentes améliorations ont été nécessaires, telle que l'augmentation du nombre de fibres par exposition ou de l'efficacité optique. Cette section résume les principales caractéristiques des spectrographes

4. L'efficacité quantique est définie comme le rapport entre le nombre d'électrons collectés et le nombre de photons incident sur le système.

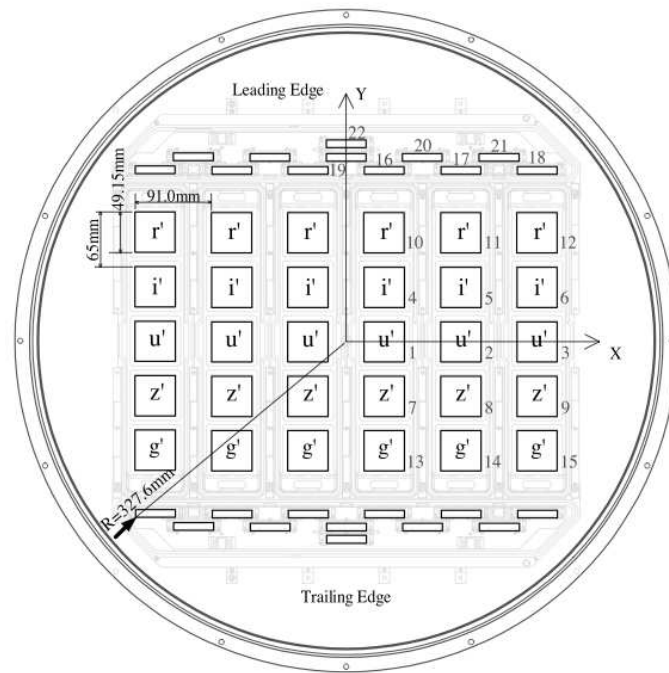


FIGURE 3.6: Disposition du plan focal de la caméra utilisée pour BOSS ainsi que pour SDSS et SDSS-II. Les 6 colonnes de 5 capteurs CCDs dédiés à l'imagerie sont visibles au centre, tandis que les 24 CCDs destinés à l'astrométrie et à la mise au point se situent de part et d'autre du plan focal. Le TDI *drift scanning* s'effectue du haut (*leading edge*) vers le bas (*trailing edge*) si bien qu'une étoile traversant le plan focal rencontre en premier lieu un capteur *r* et en dernier lieu un capteur *g*.  
 Source : Gunn et al. (1998)

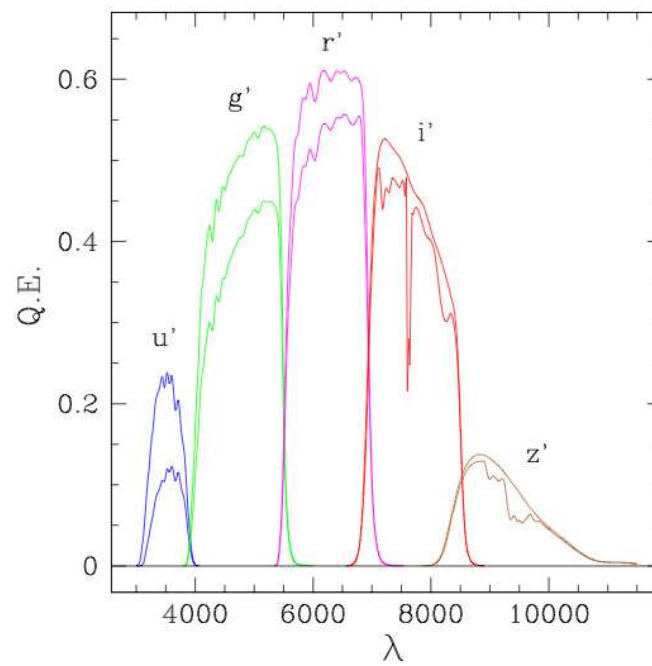


FIGURE 3.7: Efficacité quantique du système dans chaque bande photométrique. Les courbes supérieures prennent en compte la transmission des filtres, l'efficacité quantique des CCDs ainsi que les pertes liées à l'optique. Les courbes inférieures prennent également en compte l'extinction atmosphérique. *Source : Gunn et al. (1998)*

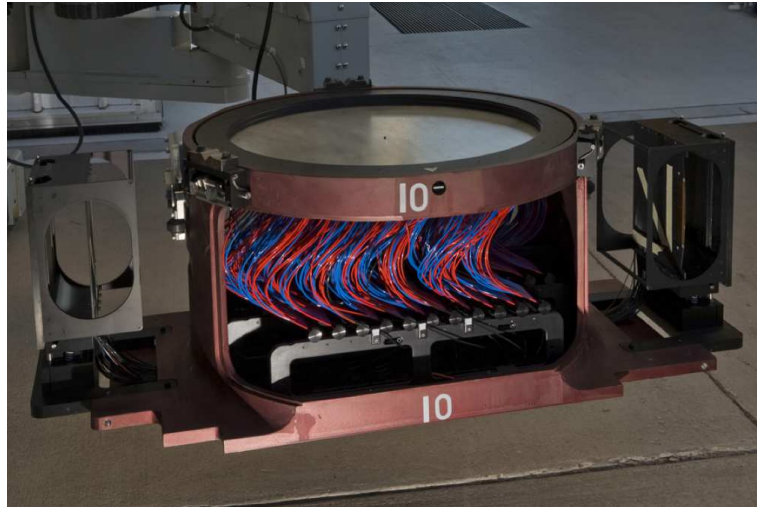


FIGURE 3.8: Photographie d'une cartouche prête à être montée au plan focal du télescope. Les fibres optiques (câbles rouges et bleus) sont branchées sur la plaque, visible au sommet de la cartouche. L'autre extrémité des fibres est divisée en deux faisceaux (500 fibres par faisceaux), dirigés chacun vers un côté de la cartouche. Ces deux faisceaux, visibles à l'intérieur des blocs noirs situés de part et d'autre de la cartouche, sont destinés à être introduit dans chacun des spectrographes. La cartouche sera placée à l'arrière du télescope, la plaque se situant au plan focal. *Source : Smee et al. (2012)*

conçus pour BOSS. Une étude détaillée des spectrographes des deux premières générations du SDSS ainsi que des améliorations apportées pour BOSS a été menée par Smee et al. (2012).

### Les cartouches

Pour obtenir les spectres des objets désirés, une plaque en aluminium est positionnée au plan focal du télescope. Cette plaque mesure 3,2 mm d'épaisseur et 0,813 m de diamètre pour 4,3 kg. Le télescope transforme l'ascension droite et la déclinaison ( $\alpha, \delta$ ) définissant la position d'un objet dans le ciel par rapport à l'axe du télescope, en position ( $X, Y$ ) dans le plan focal. La plaque est donc percée, avant son positionnement dans le plan focal, aux endroits où viendront converger les rayons lumineux issus des objets à spectrer. Des fibres optiques sont ensuite insérées dans ces trous pour rediriger la lumière des objets vers les grismes (combinaison d'un réseau de diffraction et d'un prisme) VPH (pour l'anglais *volume holographic phase*), puis vers les CCDs. L'ensemble plaque plus fibres optiques est maintenu par une structure en aluminium, le tout formant une « cartouche », dont un exemple est visible en figure 3.8. Une cartouche peut être montée sur le télescope par un seul opérateur en quelques minutes.

### Fibres optiques

Dans le but d'améliorer le rendement du relevé, le nombre de fibres par plaque a été augmenté de 640 à 1000 lors du passage de SDSS-II à BOSS. Le diamètre des fibres a également été diminué, passant de  $180 \mu\text{m}$  (3" sur le ciel) à  $120 \mu\text{m}$  (2" sur le ciel). La diminution de la taille des fibres était nécessaire pour réutiliser l'essentielle de l'optique des spectrographes du SDSS-II tout en augmentant le nombre de fibres par plaque. Elle résulte également en l'augmentation du signal sur

bruit des objets à grands décalages spectraux, en réduisant la contribution du fond du ciel. Chaque fibre optique collecte la lumière au plan focal du télescope sous forme d'un cône d'ouverture numérique<sup>5</sup> 0,1 (correspondant au f/5 du télescope) et ressort sous forme d'un cône légèrement plus large d'ouverture numérique 0,125 (f/4). Cette dégradation du rapport focale sur diamètre se produit lors de la propagation de la lumière le long de la fibre optique. Le collimateur, située à la sortie des fibres optiques, est prévu pour recevoir un cône d'ouverture numérique 0,125. Tout rayon émit par les fibres en dehors de ce cône est perdu. Il est donc nécessaire de maîtriser l'élargissement du cône à la sortie des fibres pour maximiser le rendement. Pour ce faire, les spectrographes sont montés directement sur le télescope (comme on peut le voir sur la figure 3.5), limitant ainsi les déplacements relatifs entre les deux extrémités d'une fibre et par la même occasion les contraintes mécaniques qui pourraient dégrader le rapport focale. Cette configuration permet également de minimiser la taille des fibres, chacune d'elle mesurant  $1.830 \pm 25$  mm.

Les 1.000 fibres sont divisées en deux jeux de 500 fibres regroupées en faisceaux et dirigées vers deux fentes situées chacune à une extrémité de la cartouche. Lors du positionnement de la cartouche à l'arrière du télescope, chacune des fentes est introduite dans l'un des deux spectrographes.

### Le système optique et les CCDs

Le système optique des spectrographes de BOSS est quasiment identique au système original du SDSS. La bande passante a toutefois été étendue, passant de  $3.900 \lesssim \lambda \lesssim 9.000 \text{ \AA}$  à  $3.600 \lesssim \lambda \lesssim 10.000 \text{ \AA}$ . L'extension vers le bleu permet d'augmenter la partie détectée de la forêt Lyman- $\alpha$  des quasars ayant un décalage spectral inférieur à  $2,5^6$ . L'augmentation du côté rouge est motivée par l'amélioration de la détection de discontinuités dans les spectres de galaxies. Ces améliorations vers le bleu et le rouge, ont été rendues possibles par les progrès des technologies CCDs, mais également par l'amélioration de la transmission optique des optiques et notamment des grismes VPH.

Un schéma du système optique est visible en figure 3.9. La lumière pénètre dans les spectrographes par une fente située à l'une des extrémités de la cartouche, dans laquelle sont insérées 500 fibres optiques. Dans un premier temps, la lumière issue des fibres rencontre le collimateur qui la réfléchit vers un miroir dichroïque (ou semi-réfléchissant), sous forme d'un faisceau parallèle de 160 mm de diamètre. Les longueurs d'onde inférieures à 640 nm sont alors réfléchies par le miroir dichroïque vers le canal bleu, tandis que les longueurs d'onde supérieures à 640 nm sont transmises vers le canal rouge. Dans chaque canal, la lumière est ensuite dispersée par un grisme VPH. La densité des grismes est de 400 lignes/mm pour le canal rouge est 520 lignes/mm pour le canal bleu.

Les caméras des canaux bleus et rouges sont situées aux sorties des grismes. Elles sont constituées d'une succession de lentilles (un singulet, un triplet puis un doublet) puis de deux aplatisseurs de champ montés directement devant les capteurs CCDs. Ces capteurs CCDs sont constitués de  $4k \times 4k$  pixels de  $15 \mu\text{m}$  de côté. Le parcours des rayons lumineux à l'intérieur des caméras est représenté en figure 3.10.

5. Rappelons que l'ouverture numérique est définie par  $O.N. = n \cdot \sin i$  où  $n$  est l'indice du milieu d'observation ( $n_{air}$  dans notre cas) et  $i$  est l'angle formé par l'axe optique du système (la fibre optique) et le rayon le plus écarté de l'axe optique.

6. En considérant que la forêt Lyman- $\alpha$  est exploitable jusqu'à environ  $10 \text{ \AA}$  de la raie Lyman- $\beta$ , soit jusqu'à  $1.035 \text{ \AA}$ , la coupure  $\lambda \gtrsim 3.600 \text{ \AA}$  permet de détecter entièrement la forêt des quasars ayant un décalage spectral supérieur à  $z \approx 3.600/1.035 - 1 \approx 2,48$ . Tout quasar ayant un décalage spectral inférieur aura une forêt tronquée.

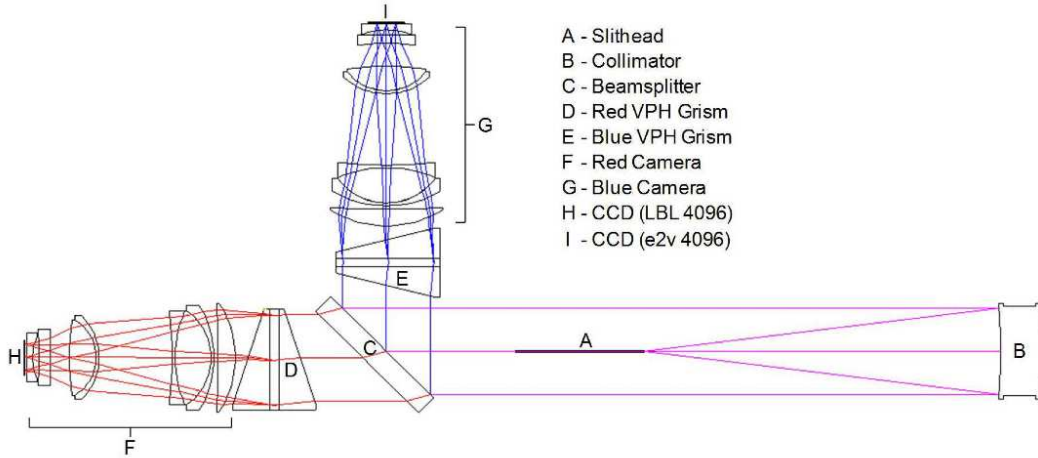


FIGURE 3.9: Schéma optique des spectrographes. La lumière pénètre dans le système par la fente (A) maintenant l'extrémité du faisceau de 500 fibres optiques (slithead). Elle est alors dirigée vers le collimateur (B), qui la renvoie dans la direction opposée sous forme d'un faisceau parallèle. Le miroir dichroïque (C) (en anglais, beamsplitter) sépare le faisceau lumineux en une composante rouge transmise (vers la gauche sur le schéma) et une composante bleue réfléchie (vers le haut). Chaque composante est alors dispersée par un grisme (D et E) puis est focalisée par les caméras (F et G) sur les capteurs CCDs (H et I). *Source : Smees et al. (2012)*

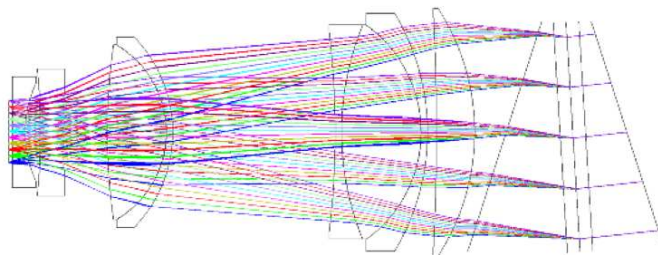


FIGURE 3.10: Parcours des rayons lumineux diffractés par le grisme (situé à droite de l'image) puis focalisés sur le CCD par le système optique de la caméra. *Source : Smees et al. (2012)*

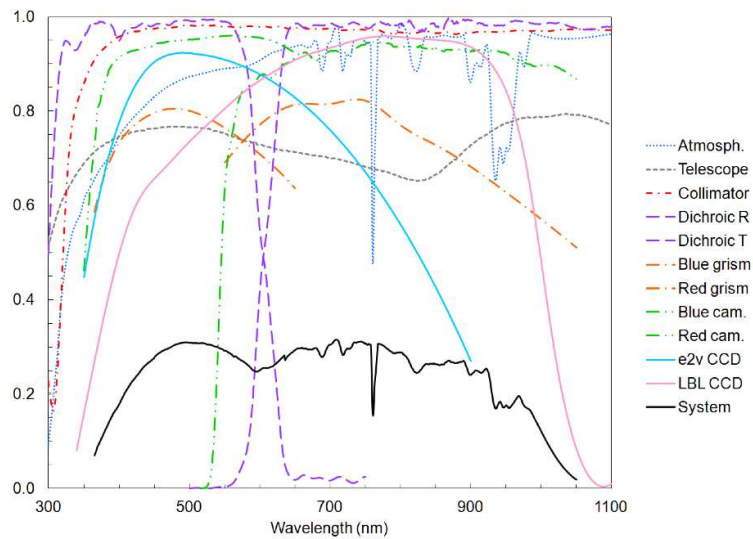


FIGURE 3.11: Efficacités optiques prévisionnelles pour chaque partie du système, ainsi que pour le système complet. *Source : Smee et al. (2012)*

### Performances Optiques

L'efficacité optique du système a été prédit (Smee et al., 2012) en fonction de la longueur d'onde considérée en prenant en compte la contribution de chaque composante, de l'extinction atmosphérique jusqu'à l'efficacité des capteurs CCDs. Le résultat des prédictions est visible en figure 3.11. L'efficacité maximale prédite pour le système complet est d'environ 30% dans le rouge, comme dans le bleu. Cette prédiction peut être comparée avec les rendements mesurés pour les deux spectrographes de BOSS, visibles en figure 3.12. Ces mesures sont effectuées à l'aide de 84 étoiles standards. Pour chacune d'elle, le rendement est défini comme le rapport du flux mesuré sur le flux arrivant en dehors de l'atmosphère. Le rendement de chaque spectrographe est alors obtenu en moyennant les rendements individuels. Les rendements des deux spectrographes sont très comparables mais légèrement inférieur aux prédictions, atteignant environ 26% dans le bleu et 28% dans le rouge. Ils représentent toutefois une amélioration substantielle par rapport aux rendements des spectrographes du SDSS, également visibles en figure 3.12.

La résolution spectrale est mesurée avant l'observation de chaque plaque, à l'aide de lampes à arc dédiées à la calibration. Premièrement, les lampes à arc sont imagées par les spectrographes. Secondement, chaque raie d'émission visible dans les spectres obtenus, est ajustée par une gaussienne de largeur  $\sigma_\lambda$ . Les valeurs de  $\sigma_\lambda$  ainsi calculées sont ajustées par un polynôme d'ordre quatre pour obtenir  $\sigma_\lambda$  en fonction de la longueur d'onde, sur toute la plage couverte par les spectrographes. Le pouvoir de résolution est alors défini par  $R = \frac{\lambda}{2,35 \times \sigma_\lambda}$ , où  $2,35 \times \sigma_\lambda$  représente la largeur à mi-hauteur de la raie à la longueur d'onde  $\lambda$ . Les pouvoirs de résolution moyennés sur 100 plaques observées par BOSS sont représentés en figure 3.13 pour les deux spectrographes. Le décrochage visible autour de 6.200 Å est lié à la séparation faisceau en deux gammes de longueurs d'onde.

Le pouvoir de résolution des spectrographes ne dépend pas uniquement de la longueur d'onde, Il dépend également de la position dans le plan focal du CCD. Cette dépendance est illustrée en figure 3.14 où  $\sigma_\lambda$  est représenté en unité de pixel, en fonction de la position sur les différents CCDs. Pour les caméras bleues, il est clair que la qualité optique diminue sur les bords du champ, le pouvoir de résolution pouvant y être jusqu'à deux fois inférieur à celui aux centres des capteurs.



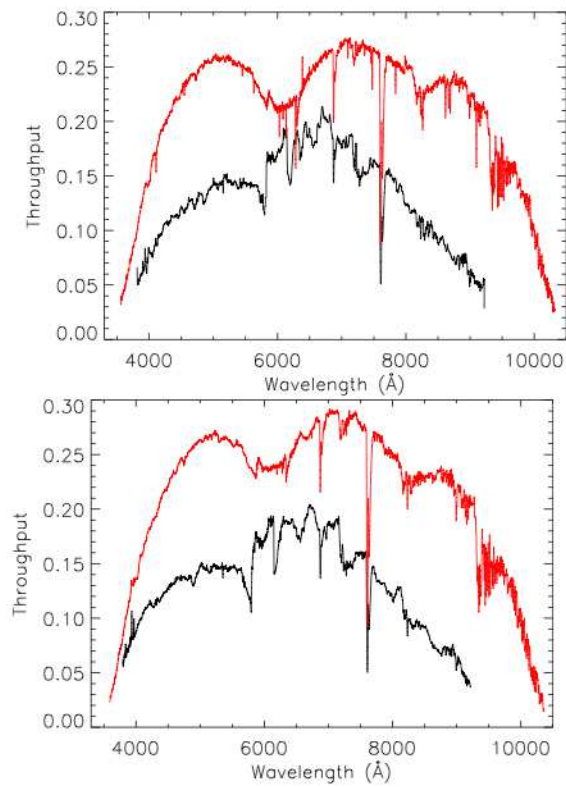


FIGURE 3.12: Efficacités optiques mesurées pour les deux spectrographes de BOSS (courbes rouges), comparées aux rendements des spectrographes du SDSS (courbes noires). *Source : Smee et al. (2012)*

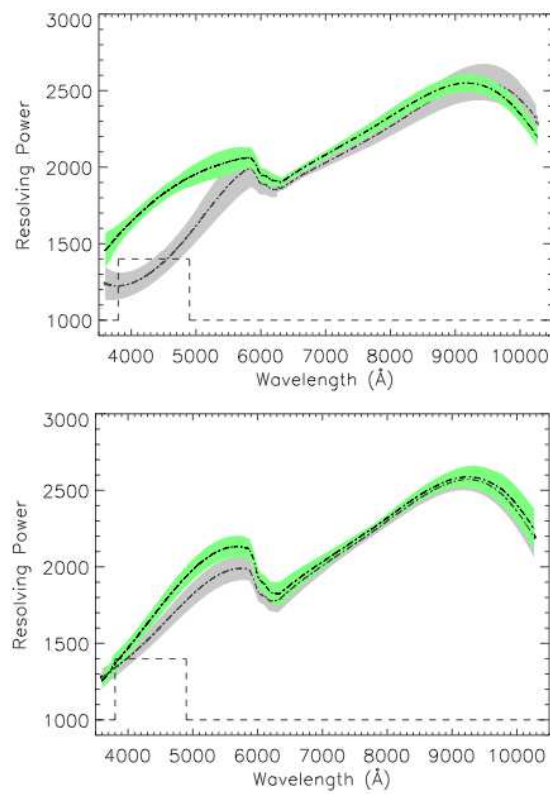


FIGURE 3.13: Pouvoirs de résolution requis (courbes en pointillés) et mesurés (courbes grises et vertes) pour les deux spectrographes de BOSS. Les courbes vertes correspondent aux pouvoirs de résolution moyennés sur 100 plaques ainsi que les zones à 68% de confiance pour la fibre centrale. Les courbes grises correspondent aux pouvoirs de résolution moyennés sur 100 plaques ainsi que les zones à 68% de confiance pour une fibre proche du bord du spectrographe. *Source* : *Smee et al. (2012)*

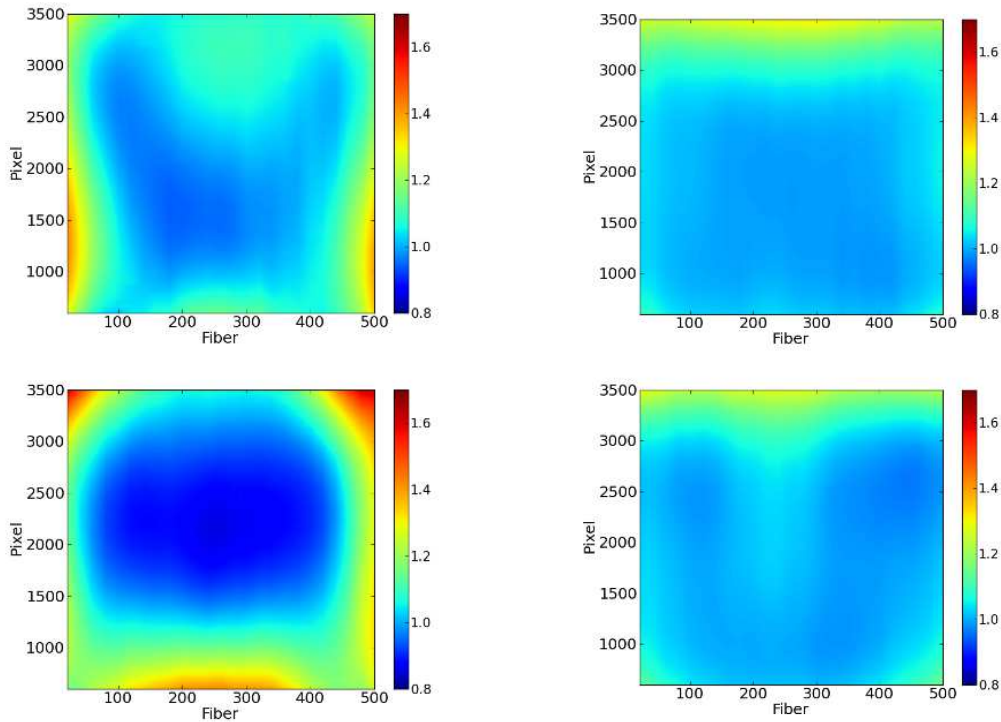


FIGURE 3.14: Largeur  $\sigma_\lambda$  d'une raie en unité de pixel, en fonction de la position dans le plan focal du CCD, pour chacun des quatre CCDs des spectrographes. Les deux figures de gauche correspondent aux CCDs bleus tandis que les figures de droites correspondent aux CCDs rouges. *Source : Smeed et al. (2012)*

Les caméras rouges paraissent plus homogènes avec une résolution quasiment constante pour des longueurs d'ondes  $\lambda < 9.500 \text{ \AA}$ .

### 3.5 Réduction des données

La réduction des données est effectuée par un programme travaillant par plaque. Il est développé de sorte à extraire, calibrer, additionner, classifier et évaluer les décalages spectraux des 1.000 objets en incluant toutes les expositions d'une plaque. Ce programme, brièvement décrit en 3.5.1, est capable de distinguer avec une bonne efficacité les quasars, les galaxies et les étoiles. Il a été complété d'une inspection visuelle de l'intégralité des spectres de quasars décrite en 3.5.2.

#### 3.5.1 Réduction automatique des données

La première étape traite les expositions individuelles. Elle consiste à transformer les images en deux dimensions des CCDs en différents spectres à une dimension. Ces spectres sont alors étalonnés à l'aide des données des lampes à arc. La variance par pixel est déterminée à l'aide du bruit de lecture ainsi que du nombre de photons enregistrés dans chaque pixel. La variance inverse est ensuite multipliée par un facteur tenant compte des défauts connus des CCDs. Les pixels affectés par des rayons cosmiques sont identifiés et masqués. Enfin, le flux est étalonné à l'aide de modèles ajustés aux spectres des étoiles standards observées à cet effet.

La seconde étape consiste à additionner les expositions individuelles pour former un spectre composite pour chaque fibre. Chaque spectre regroupe les données des caméras bleue et rouge, couvrant ainsi l'intégralité de la gamme en longueur d'onde de l'instrument, allant de 3.600 Å à plus de 10.000 Å. L'échelle des spectres est modifiée de sorte à être linéaire en  $\ln(\lambda)$ . La variance du spectre est estimée en propageant les variances des poses individuelles, les covariances étant négligées. Les spectres des objets ayant été observés à plusieurs reprises sont produits séparément, puis comparés ; Le spectre possédant le meilleur signal sur bruit est désigné en tant que spectre « primaire <sup>7</sup> ».

La dernière étape consiste à ajuster les spectres obtenus par des modèles d'étoiles, de galaxies et de quasars. Cette procédure, décrite dans Aihara et al. (2011) permet de déterminer le type de l'objet ainsi que son décalage spectral.

Lors de la neuvième publication de données du SDSS (DR9), la collaboration BOSS a rendu publique les spectres additionnés de 493.845 galaxies, 93.003 quasars et 82.645 étoiles, (hors spectres hérités des catalogues de SDSS et SDSS-II) ainsi que toutes les données dont elle dispose (expositions individuelles, variances des pixels, vecteurs d'étalonnage, etc.). Les données sont disponibles sur le site du SDSS <sup>8</sup>.

### 3.5.2 Inspection visuelle

Les quasars présentent des structures spectrales variées, comme les BAL et DLA décrits en section 2.4.3, rendant difficile l'automatisation de leur classification. Pour palier le problème, des membres du groupe de participation français du SDSS ont inspecté visuellement l'intégralité des spectres de quasars du DR9. Cette inspection visuelle permet une identification fiable du type de l'objet ainsi qu'une détermination de son décalage spectral. Une description complète de cette inspection est donnée dans Paris et al. (2012). 180.268 spectres ont été inspectés visuellement pour l'échantillon DR9. Les spectres inspectés correspondent aux cibles quasars des échantillons *CORE* et *BONUS*, aux cibles des projets *ancillary* destinés à l'observation des quasars ainsi qu'à une fraction des spectres de galaxie classifiés comme quasars par le programme automatique de réduction des données.

L'inspection visuelle a permis d'estimer l'efficacité de la procédure de classification automatique des spectres. Elle a notamment montré que l'échantillon d'étoiles obtenu par classification automatique est extrêmement pur. Seuls 0,14% des objets classés de manière automatique comme étoiles sont en fait des quasars. L'échantillon de quasars dont les décalages spectraux vérifient  $z > 2,15$  (correspondant aux quasars utilisés pour l'analyse BAO) présente une plus forte contamination avec 8% des objets classés automatiquement comme quasars s'avérant être des étoiles. Cette contamination est proche de 50% pour les quasars à plus bas décalages spectraux ( $z < 2,15$ ).

Grâce à l'inspection visuelle, il est également possible de quantifier l'efficacité de la réduction automatique des données à déterminer le décalage spectral des objets identifiés comme quasars. Pour DR9, 0,3% des décalages spectraux mesurés par la procédure automatique ont une erreur  $\Delta z > 0,1$ . De telles erreurs surviennent majoritairement lorsque le décalage spectral du quasar est inférieur à 2,0 et que la raie d'émission Lyman- $\alpha$  du quasar n'est pas observable.

L'inspection visuelle permet enfin de repérer les caractéristiques spectrales propres à chaque spectre de quasar telles que la présence de BAL ou de DLA. Les DLA repérés dans les spectres sont caractérisés à l'aide d'une procédure automatique (Noterdaeme et al., 2009) déterminant la colonne de densité du système. L'indice de « balnicité » des spectres présentant un BAL est également déterminé par une procédure automatique.

7. Seuls les spectres primaires ont été considérés pour l'analyse BAO présentée en section 5

8. <http://www.sdss3.org/dr9/>



## Chapitre 4

# Sélection des quasars

*« I have called this principle, by which each slight variation, if useful, is preserved, by the term of Natural Selection, in order to mark its relation to man's power of selection. »*

—C. Darwin, On The Origin of Species

---

Ce chapitre adresse l'une des principales problématiques des relevés spectroscopiques : la sélection des objets cibles. Contrairement à la sélection des galaxies, qui a été développée dans le but d'être le plus homogène possible sur l'ensemble de la zone du ciel couverte par BOSS, la sélection des quasars se veut être la plus complète possible sans soucis d'homogénéité. Cette liberté nous a permis de développer une sélection basée sur la variabilité des quasars et de l'appliquer, au cours du relevé, aux objets pour lesquels nous disposons d'une photométrie multi-époque.

---

### 4.1 Sélection par couleurs

La technique usuelle de sélection des objets cibles (quasars potentiels) consiste à appliquer un jeu de coupures dans différents diagrammes couleur - couleur. Une couleur est définie comme la différence de magnitude dans deux bandes différentes. Dans le cas de BOSS, nous disposons pour chaque objet de sa magnitude dans cinq bandes : à savoir les bandes  $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$  (voir section 3.4.2). A partir de ces cinq mesures, il est possible de construire une base de quatre couleurs indépendantes « primaires », par exemple  $(u - g)$ ,  $(g - r)$ ,  $(r - i)$  et  $(i - z)$ . Toute autre couleur émanant alors de la combinaison linéaire de ces couleurs « primaires ».

Les propriétés intrinsèques aux différents objets astrophysiques (étoiles, galaxies, quasars, *etc.*) conduisent ces objets à occuper préférentiellement différentes parties des diagrammes couleur - couleur. Pensons en particulier au pic Lyman- $\alpha$  marqué, présent dans les spectres de quasars (voir section 2.4.1) mais absent des spectres d'étoiles. Un jeu de coupures savamment choisies dans l'espace des quatre couleurs permet alors, avec plus ou moins de réussite, de les séparer.

## 4.2 Sélection par variabilité

Contrairement à la majorité des étoiles qui ont un flux constant dans le temps, les quasars possèdent généralement un flux variable. Les travaux de Sesar et al. (2007) ont montré qu'au moins 90% des quasars présentent des variations dans leur flux d'au moins 0.03 magnitude sur une période de plusieurs années. D'après d'autres travaux (Vanden Berk et al., 2004) les changements de luminosité d'un quasar sont de l'ordre de 10% sur des échelles de temps comprises entre un mois et un an. Le mécanisme à l'origine de la variabilité des quasars n'est pas connu avec certitude. Il est fortement probable qu'il soit lié aux mécanismes d'accrétion par le trou noir super massif au centre du quasar. D'autres phénomènes tels que des collisions d'étoiles, des supernovae ou des effets de microlentille ont également été proposés. Cette section présente la méthode de sélection des quasars cibles que nous avons mis au point au cours de ma thèse. Ces travaux ont été publiés dans la revue *Astronomy & Astrophysics* (Palanque-Delabrouille et al., 2011)

### 4.2.1 La région du *Stripe 82*

Nous menons notre étude sur une partie du ciel appelée *Stripe 82*. Cette bande, de  $220 \text{ deg}^2$ , est définie par les coordonnées équatoriales  $-43^\circ < \alpha_{J2000} < 45^\circ$  et  $-1,25^\circ < \delta_{J2000} < 1,25^\circ$ . Entre les années 2000 et 2005, SDSS a produit la photométrie de cette région de une à trois fois par an. SDSS-II a ensuite couvert cette région de 10 à 20 fois par an de 2005 à 2008, dans le cadre d'un projet de recherche de supernova (Frieman et al., 2008). Avec un échantillonnage moyen de 53 époques par objet sur une période allant de 5 à 10 ans (Abazajian et al., 2009), le *Stripe 82* est parfaitement adapté au développement d'un algorithme de sélection par variabilité.

### 4.2.2 Echantillons tests

Pour mettre en place et tester nos algorithmes de sélection reposant sur la variabilité des objets, nous avons défini deux échantillons de contrôle, un d'étoiles et un de quasars. Ils seront également employés pour l'entraînement du réseau de neurones (voir 4.2.6). L'échantillon de quasars consiste en une liste de 13.328 objets, confirmés comme quasars par spectroscopie. Cette liste est issue de la fusion du catalogue de quasars 2dF (Croom et al., 2004), du catalogue de galaxies brillantes rouges et quasars 2dF-SDSS (Croom et al., 2009), de la base de données spectroscopique SDSS-DR7 (Abazajian et al., 2009), du catalogue de quasars SDSS-DR7 (Schneider et al., 2010) et de la première année d'observation de BOSS. Ces quasars ont un décalage spectral compris dans l'intervalle  $0.05 \leq z \leq 5.0$  (voir fig 4.1) et une magnitude dans la bande photométrique  $g$  appartenant à l'intervalle  $18 \leq g \leq 23$ , après correction de l'extinction galactique.

L'échantillon d'étoiles comporte 2.697 objets initialement marqués comme quasars potentiels par la sélection en couleurs, puis identifiés comme étoiles après spectroscopie par BOSS. Ces étoiles, occupant la même position dans les diagrammes couleur-couleur que les quasars, correspondent exactement au type d'objets que nous souhaitons discriminer par la méthode de variabilité. L'utilisation de ce catalogue d'étoiles représente donc une approche conservative.

Les courbes de lumière des objets de chacun des catalogues ont été construites grâce aux données collectées par SDSS à différentes époques. Les figures 4.2 et 4.3 présentent des exemples typiques de courbes de lumière obtenues pour des étoiles et pour des quasars. L'augmentation de la cadence d'échantillonnage après la date (exprimée en jour Julien modifié) 53.500 correspond au début de la recherche systématique de supernovae par SDSS-II.

Les étoiles et les quasars présentent le même échantillonnage temporel, typique du *Stripe 82*. Le nombre d'époques (c'est-à-dire le nombre de points de mesure pour une bande photométrique donnée) varie entre 1 et 140, avec une moyenne de 53 et un écart type de 20. L'écart en temps

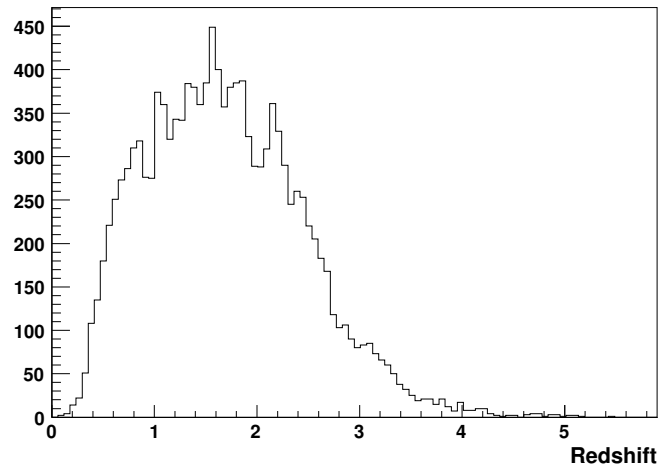


FIGURE 4.1: Distribution des décalages spectraux de l'échantillon de quasars.

entre la première et la dernière mesure est compris entre 8 et 10 ans pour 74% de l'échantillon, entre 5 et 7 ans pour 24% des objets et est au plus de 4 ans pour les 2% restants. Seuls les objets possédant un minimum de 4 observations ont été considérés pour cette étude indépendamment de l'écart temporel entre les observations. Les 13.063 quasars et 2.609 étoiles vérifiant ce critère possèdent tous une courbe de lumière s'étalant sur un minimum de deux ans consécutifs.

### 4.2.3 Nettoyage des courbes de lumière

La première étape consiste à nettoyer les courbes de lumière de points aberrants qui pourraient substantiellement altérer le résultat d'un algorithme de sélection des quasars par variabilité. Nous utilisons pour ce faire une procédure en deux temps, indépendante de l'origine de la déviance (qu'elle soit technique ou photométrique). La première étape consiste en l'application d'un filtre médian à trois points sur les courbes de lumière, pour chaque bande photométrique et pour chaque objet. Le filtre médian à trois points remplace chaque point  $x_i$  de la courbe de lumière par la valeur médiane de l'ensemble  $\{x_{i-1}, x_i, x_{i+1}\}$ .<sup>1</sup> Les courbes de lumière sont ensuite nettoyées de tout point déviant significativement d'un ajustement de la courbe de lumière par un polynôme de degré cinq. Pour ne pas retirer un trop grand nombre de points photométriques dans le cas de courbes de lumière intrinsèquement bruitées ou variables, le seuil de déviance, initialisé à  $5\sigma$ , est itérativement augmenté jusqu'à ce que le nombre de points à supprimer ne dépasse pas 10%.

### 4.2.4 Quantification de la variabilité

Le but de notre étude est le développement d'un algorithme basé sur la variabilité, capable de sélectionner les quasars de la région du Stripe 82, tout en rejetant un maximum d'étoiles. Pour ce faire nous définissons deux jeux de variables discriminantes. Le premier regroupe les  $\chi^2$  obtenus en ajustant, pour chaque bande photométrique  $u$ ,  $g$ ,  $r$ ,  $i$  et  $z$ , la courbe de lumière par son flux moyen. Les objets non-variables ont alors des  $\chi^2$  réduits proches de l'unité, ce qui est le cas pour

1. Bien que la fréquence des mesures de SDSS-I soit bien inférieure à celle de SDSS-II, le filtre médian est appliqué sans distinction à l'intégralité de la courbe de lumière.



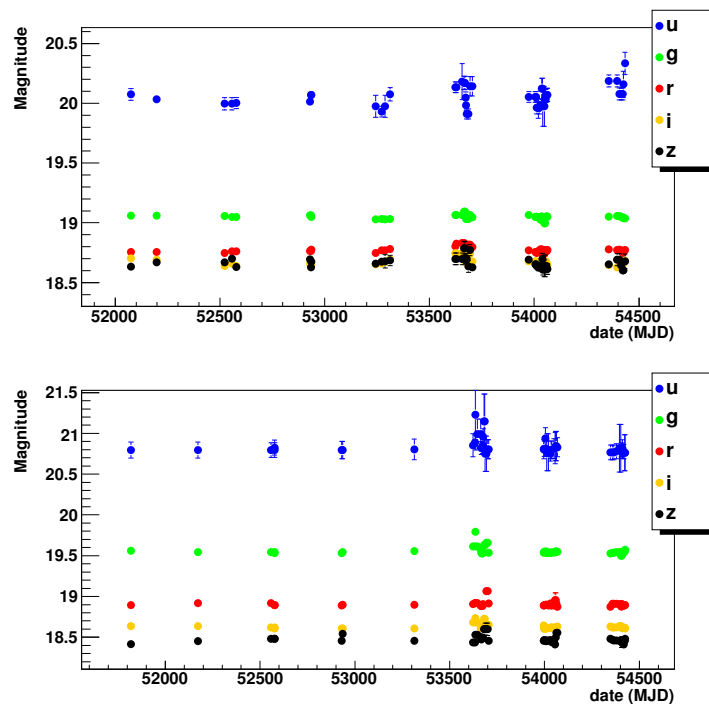


FIGURE 4.2: Exemples de courbe de lumière (après nettoyage tel qu'expliqué en section 4.2.3) dans les cinq bandes photométriques pour deux étoiles du Stripe 82.

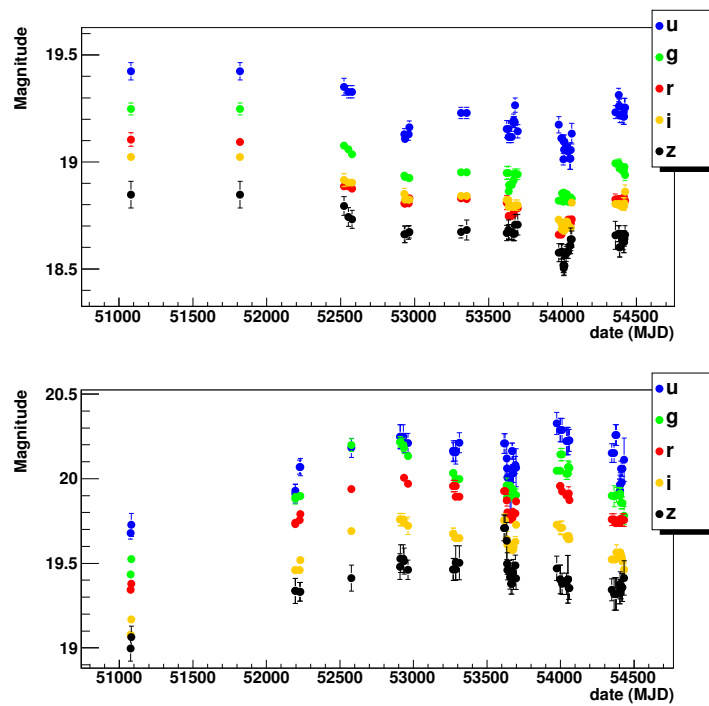


FIGURE 4.3: Exemples de courbe de lumière (après nettoyage tel qu'expliqué en section 4.2.3) dans les cinq bandes photométriques pour deux quasars du Stripe 82.

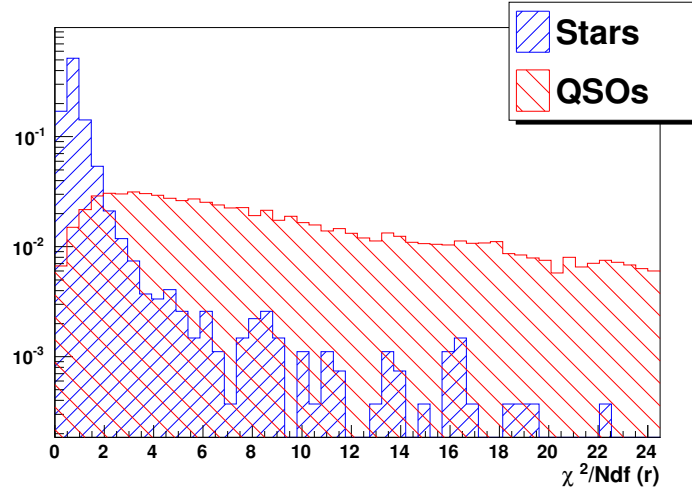


FIGURE 4.4: Distributions normalisées des  $\chi^2$  réduits obtenus en ajustant les courbes de lumière par une constante pour la bande  $r$ , pour l'échantillon d'étoiles (bleu) et celui de quasars (rouge). Les quasars présentent des valeurs de  $\chi^2$  réduits plus grandes que les étoiles ce qui confirme qu'ils sont moins bien ajustés par une constante et donc plus variables que les étoiles.

la majorité des étoiles. Les quasars, à l'inverse, tendent à avoir des plus grandes valeurs de  $\chi^2$  réduits. Ceci est illustré en figure 4.4 pour la bande spectroscopique  $r$ .

Ce jeu de cinq variables nous permet de distinguer les objets variables des étoiles non-variables, mais ne sépare pas les quasars des étoiles variables. La grande majorité de ces étoiles variables sont rejetées de l'échantillon par la coupure  $18 < g$ . Pour affiner la sélection, nous introduisons un nouveau jeu de paramètres tirés de la « fonction de structure » de la courbe de lumière. Cette fonction de structure a été introduite par de Vries et al. (2005) et MacLeod et al. (2008) puis utilisée par Schmidt et al. (2010) pour la sélection des quasars. Elle rend compte de la variabilité d'une courbe de lumière en quantifiant le changement d'amplitude  $\Delta m_{ij}$  en fonction de l'écart en temps  $\Delta t_{ij}$  pour deux observations aux époques  $i$  et  $j$ . Elle est donnée par :

$$\mathcal{V}(\Delta t_{ij}) = |\Delta m_{ij}| - \sqrt{\sigma_i^2 + \sigma_j^2} \quad (4.1)$$

où  $\sigma$  désigne l'erreur sur la mesure de la magnitude. Pour chaque bande photométrique, cette fonction de structure peut-être modélisée par la loi de puissance  $A (\Delta t)^\gamma$ ,  $A$  et  $\gamma$  étant les paramètres de la modélisation,  $\gamma$  étant défini strictement positif. Ce modèle illustre que pour deux observations, l'écart type de la distribution de l'écart en magnitude tend à augmenter avec l'écart en temps. Pour ajuster les paramètres  $A$  et  $\gamma$  nous définissons la fonction de vraisemblance :

$$\mathcal{L}(A, \gamma) = \prod_{j>i} \mathcal{L}_{ij} \quad (4.2)$$

où, pour chaque paire d'observations, nous supposons la fonction de densité de probabilité (f.d.p) des valeurs de  $\Delta m$  gaussienne :

$$\mathcal{L}_{ij} = \frac{1}{\sqrt{2\pi\sigma^2(\Delta m)}} \exp\left(-\frac{\Delta m_{ij}^2}{2\sigma^2(\Delta m)}\right) \quad (4.3)$$

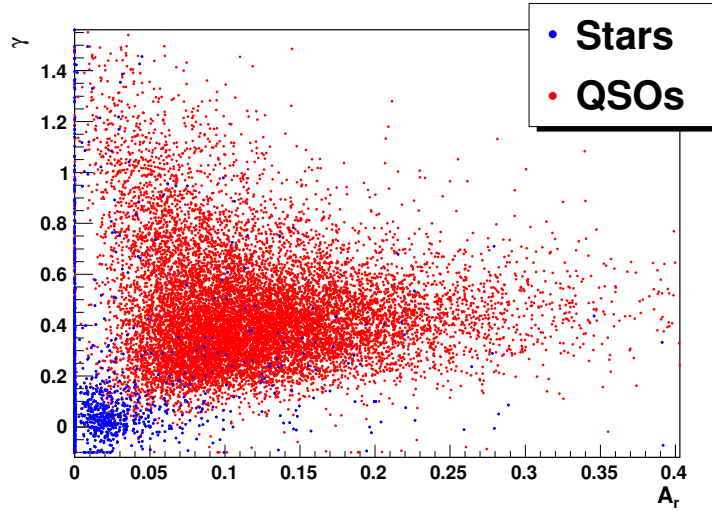


FIGURE 4.5: Paramètres  $\gamma$  et  $A_r$  de la fonction de structure pour l'échantillon d'étoiles (en bleu) et celui de quasars (en rouge). Un paramètre  $A$  élevé correspond à des fluctuations en amplitude importantes. Un paramètre  $\gamma$  élevé correspond à une augmentation des fluctuations en amplitude avec le temps.

La variabilité de l'objet, décrite par une loi de puissance, est naturellement introduite dans notre modèle via la définition de la variance  $\sigma(\Delta m)$  de la f.d.p :

$$\sigma^2(\Delta m) = [A(\Delta t_{ij})^\gamma]^2 + (\sigma_i^2 + \sigma_j^2) \quad (4.4)$$

Les valeurs des paramètres  $A$  et  $\gamma$  sont ensuite obtenues par maximisation de la fonction de vraisemblance. Les résultats pour les échantillons d'étoiles et de quasars sont représentés en figure 4.5. La figure 4.6 présente un exemple de fonction de structure obtenue pour un quasar typique.

#### 4.2.5 Choix des paramètres

Les quasars présentent peu de flux dans la bande  $u$  à cause de la forte absorption par l'hydrogène du milieu intergalactique pour des longueurs d'ondes au repos inférieures à la limite de la série de Lyman située à 91.2 nm (cette gamme de longueurs d'onde est aussi appelée continu de Lyman). De plus le fond du ciel, la profondeur optique de l'atmosphère ainsi que la bande passante de l'optique du télescope rendent les bandes  $u$  et  $z$  plus bruitées que les autres. Elles ont donc un pouvoir discriminant bien plus faible que les bandes  $g$ ,  $r$  et  $i$ . Les  $\chi^2$  étant des quantités robustes, nous pouvons tout de même les utiliser pour les cinq bandes. Ce n'est pas le cas des paramètres des fonctions de structure qui résultent d'un ajustement non-linéaire. Nous nous sommes donc restreints à l'utilisation des fonctions de structures pour les trois bandes photométriques  $g$ ,  $r$  et  $i$ . Nous avons également remarqué que le paramètre  $\gamma$  de l'ajustement est quasiment indépendant de la bande photométrique étudiée. Par ailleurs, les amplitudes obtenues pour les différentes bandes sont fortement corrélées mais non identiques. A titre d'exemple, l'amplitude de la bande  $g$  est en moyenne supérieure à l'amplitude de la bande  $r$  par 0,04. Pour réduire l'incertitude sur les paramètres des ajustements, nous nous restreignons à quatre paramètres : un  $\gamma$  commun aux trois bandes et trois amplitudes ( $A_g$ ,  $A_r$  et  $A_i$ ). Notons que nous obtenons une très bonne corrélation entre les valeurs des amplitudes obtenues en ajustant un  $\gamma$  commun aux trois bandes et en ajustant

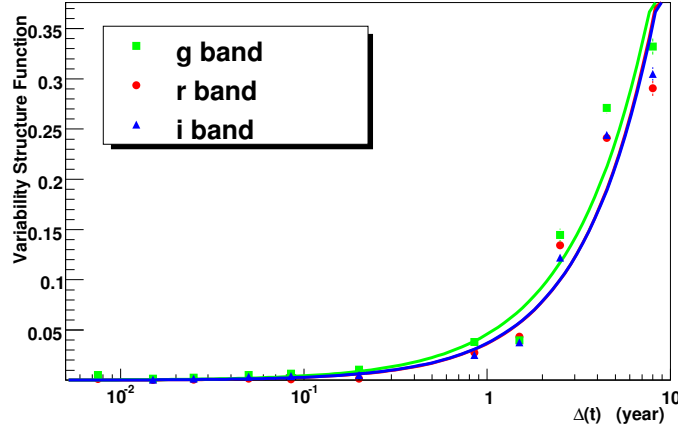


FIGURE 4.6: Exemple de fonction de structure  $\mathcal{V}(\Delta t)$  obtenue pour un quasar typique dans les trois bandes photométriques g, r et i. Les courbes correspondent aux ajustements des points par une loi de puissance de la forme  $A(\Delta t)^\gamma$

un  $\gamma$  indépendant par bande. En combinant les paramètres obtenus pour les courbes de lumière de chaque bande photométrique, par la méthode du  $\chi^2$  et par la fonction de structure, nous obtenons un jeu de 9 variables par objet.

#### 4.2.6 Sélection des quasars par réseau de neurones

Pour différencier les quasars des étoiles, en se basant sur les 9 variables disponibles par objet définies en section 4.2.5, nous avons mis en place un réseau de neurones (Bishop, 1995)<sup>2</sup>. Un réseau de neurones est constitué d'un ensemble de neurones disposés en différentes couches. La figure 4.7 présente l'architecture du réseau de neurones utilisé pour notre étude. Celui-ci est constitué de 4 couches (en anglais : *layers*), avec  $N_l$  neurones dans la couche  $l$ .

La variable d'entrée de chaque neurone de la première couche est l'un des 9 paramètres définissant un objet. Nous imposons donc  $N_1 = 9$ . La variable d'entrée d'un neurone situé dans l'une des couches suivantes ( $l = 2, 3, 4$ ) est une combinaison linéaire des sorties  $x_j^{l-1}$  des  $N_{l-1}$  neurones de la couche précédente. L'entrée  $y_j^l$  du neurone  $j$  de la couche  $l$  est donnée par :

$$y_j^l = \sum_{i=1}^{N_{l-1}} w_{ij}^l x_i^{l-1} + \theta_j^l \quad l \geq 2 \quad (4.5)$$

faisant apparaître le « poids »  $w_{ij}^l$  ainsi que le « décalage »  $\theta_j^l$ . La sortie  $x_j^l$  du neurone est, quant à elle, définie par la fonction non-linéaire :

$$x_j^l = \frac{1}{1 + \exp(-y_j^l)} \quad 2 \leq l \leq 3 \quad (4.6)$$

La quatrième et dernière couche ne possède qu'un neurone dont la sortie  $y_{NN} \equiv y_1^4$  reflète la probabilité que l'objet soit un quasar, d'après les  $N_1$  paramètres de variabilité de l'objet mis en entrée.

2. Nous avons utilisé le paquet C++ TMultiLayerPerceptron, faisant partie de la bibliothèque ROOT (<http://root.cern.ch>)

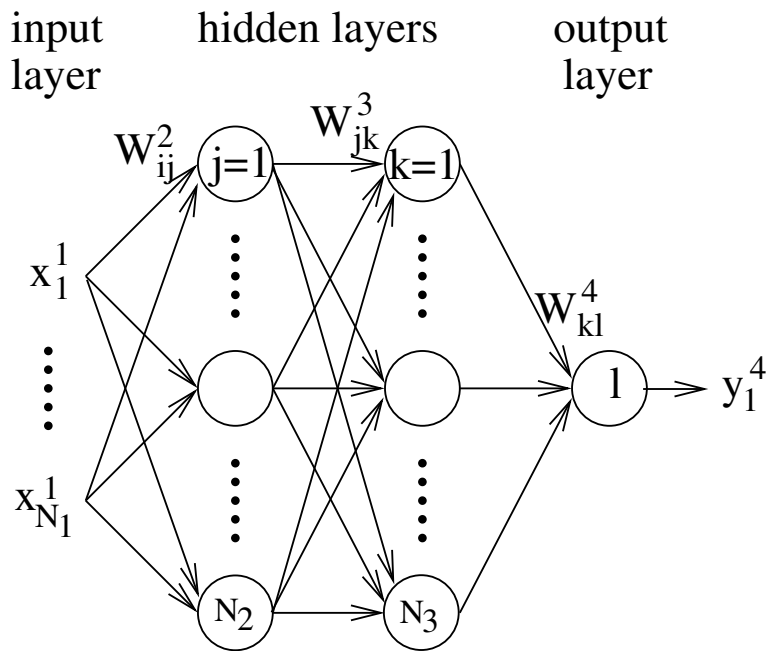


FIGURE 4.7: Représentation schématique du réseau de neurones utilisé, comportant ici  $N_1$  variables d'entrée, deux couches cachées, et un neurone de sortie.

Certains aspects du réseau de neurones, tels que le nombre de couches ou le nombre de neurones par couche, sont plus ou moins arbitraires. Ils ont donc été choisis par expérience et pour des raisons de simplicité. Par contre, les poids et décalages apparaissant dans la formule d'entrée des neurones de l'équation 4.5 doivent être optimisés de sorte que la sortie  $y_{NN}$  du réseau reflète correctement la probabilité que l'objet d'entrée soit un quasar. Pour ce faire, les échantillons d'étoiles et de quasars sont eux mêmes séparés en deux sous-échantillons. L'un des deux sous-échantillons de quasars ainsi que l'un des sous-échantillons d'étoiles sont utilisés pour optimiser les poids et décalages, de sorte que le réseau soit capable de séparer convenablement les quasars des étoiles. Cette phase est appelée entraînement du réseau. L'efficacité du réseau est ensuite testée sur les sous-échantillons de quasars et d'étoiles restants, appelés échantillons tests. Cette procédure permet de s'assurer que le réseau de neurones n'apprenne pas « par coeur » à séparer les échantillons d'entraînement. Le résultat du réseau de neurones sur les échantillons de quasars et d'étoiles tests est donné en figure 4.8.

Sur les 13.063 objets de l'échantillon de quasars, seuls 383 (3%) ne sont pas reconnus comme quasars par le réseau de neurones. Une inspection visuelle de leur courbe de lumière confirme qu'ils ne montrent pas de signe de variabilité que ce soit à court ou à long terme. Une perte d'environ 3% des quasars est donc à attendre d'un algorithme sélectionnant les quasars par leur variabilité. Cette perte atteint environ 5% pour le sous échantillon de 3.571 quasars à grands décalages spectraux ( $z > 2.15$ ). La hausse est due à la moins bonne précision des mesures photométriques de ces objets ainsi qu'au plus faible écart en temps dans le référentiel au repos des quasars à grands décalages spectraux.

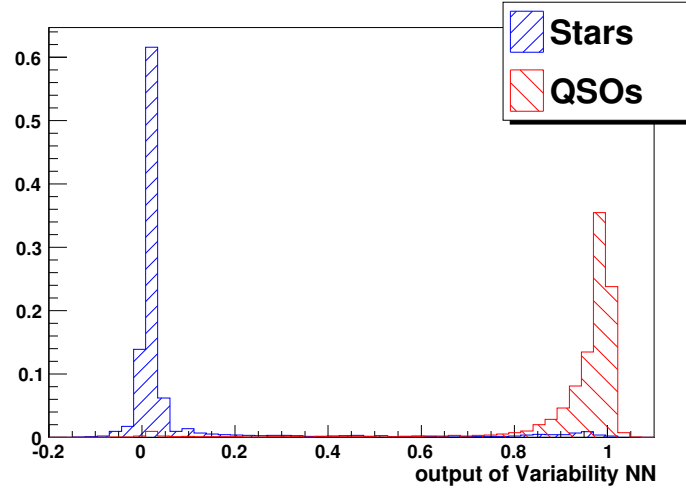


FIGURE 4.8: Sortie du réseau de neurones pour l'échantillon de quasars (en rouge) et celui d'étoiles (en bleu). 97% des quasars ont  $y_{NN} > 0.5$ . Les histogrammes sont normalisés.

#### 4.2.7 Performance de la sélection

Pour quantifier la performance de notre sélection nous introduisons la complétude  $C$  ainsi que la pureté  $P$  :

$$C = \frac{\text{Nombre de quasars sélectionnés}}{\text{Nombre total de quasars confirmés}} \quad (4.7)$$

$$P = \frac{\text{Nombre de quasars sélectionnés}}{\text{Nombre total d'objets sélectionnés}} \quad (4.8)$$

Nous définissons également la réjection stellaire :

$$R = 1 - \frac{\text{Nombre d'étoiles sélectionnées}}{\text{Nombre total d'étoiles dans l'échantillon}} \quad (4.9)$$

La figure 4.9 montre les performances de notre sélection pour différentes gammes de décalages spectraux. Il apparaît notamment que pour une réjection stellaire identique, la complétude des quasars diminue à grand décalage spectral. Cette dépendance est illustrée en figure 4.10 qui présente la complétude des quasars en fonction du décalage spectral. Pour un seuil lâche sur la sortie du réseau de neurones ( $y_{NN} > 0.5$ ), la sélection par variabilité atteint une complétude élevée pour tout décalage spectral. Au contraire, pour un seuil élevé ( $y_{NN} > 0.95$ ), la complétude dépend fortement du décalage spectral. Toutefois, contrairement aux sélections en couleurs, notre sélection par variabilité dépend de manière monotone du décalage spectral et ne présente donc pas de minimum local. Comme il a été mentionné en section 4.2.6, cette dépendance vient à la fois de la baisse de la qualité des points photométriques, ainsi que du plus faible écart en temps dans le référentiel propre des quasars à grands décalages spectraux.

La pureté est plus délicate à estimer car elle dépend d'un échantillon de référence. Les échantillons utilisés pour entraîner le réseau de neurones sont des sous-échantillons de la population d'objets cibles. Ils n'incluent pas, par exemple, les objets sélectionnés par leur variabilité mais non par leurs couleurs. La pureté sera donc quantifiée dans les sections 4.3.1 et 4.3.2 qui traitent des deux cas pour lesquels la sélection par variabilité a été appliquée aux données.

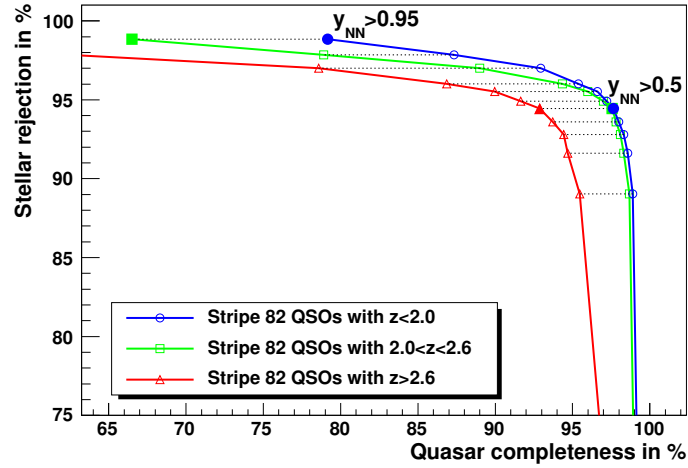


FIGURE 4.9: Réjection stellaire en fonction de la complétude des quasars pour trois gammes de décalages spectraux. Les marqueurs remplis sont obtenus pour  $y_{NN} > 0.5$  et  $y_{NN} > 0.95$  correspondant aux valeurs utilisées pour les sélections définies en sections 4.3.1 et 4.3.2.

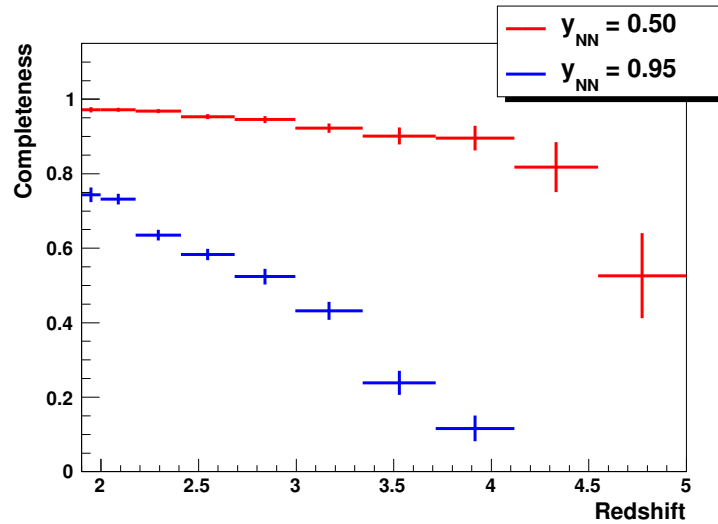


FIGURE 4.10: Complétude en fonction du décalage spectral pour deux seuils de sortie du réseau de neurones ( $y_{NN} > 0.5$  et  $y_{NN} > 0.95$ ) correspondant aux seuils appliqués en sections 4.3.1 et 4.3.2.



### 4.3 Application à la région du Stripe 82

BOSS a pour but d'observer des quasars à grands décalages spectraux ( $z > 2.15$ ) avec une densité d'environ  $20 \text{ deg}^{-2}$  de ciel. Pour cela,  $40 \text{ deg}^{-2}$  fibres optiques sont dédiées à l'observation des quasars. Avec un tel budget, une sélection en couleur typique atteint une densité de quasars à grands décalages spectraux de  $10 - 15 \text{ deg}^{-2}$ . Dans ce contexte, la sélection des cibles par variabilité peut significativement améliorer la pureté de la liste des candidats quasars, permettant d'atteindre l'objectif de  $20 \text{ deg}^{-2}$  quasars, sans augmenter le nombre de fibres optiques dédiées à leur détection. La « sélection principale », décrite en section 4.3.1, vise à remplir cet objectif.

Une autre approche consiste en la recherche de quasars additionnels, systématiquement manqués par les sélections en couleurs usuelles car possédant des couleurs atypiques, mais qui pourraient être sélectionnés par leur variabilité. Nous pouvons nous attendre à ce qu'une telle sélection soit moins biaisée en décalage spectral qu'une sélection par couleurs, contribuant ainsi à améliorer nos connaissances sur la population de quasars à grands décalages spectraux. La section 4.3.2, décrivant la sélection d'objets présentant une variabilité extrême, est dédiée à cette étude.

#### 4.3.1 Sélection principale

Le but de cette sélection est d'obtenir une liste d'environ  $35 \text{ deg}^{-2}$  cibles possédant une pureté en quasars élevée. Pour ce faire, nous commençons par appliquer des coupures en couleurs lâches. Bien que la couleur des quasars varie au cours du temps (leurs fonctions de structure possèdent différentes amplitudes  $A$  dans différentes bandes photométriques mais un unique paramètre  $\gamma$ ), cette variation sur une décennie est faible (en moyenne de l'ordre de 0.1 magnitude). Nous avons donc additionné les différentes observations d'un objet donné pour en améliorer la photométrie et donc les mesures de couleurs. Les coupures en couleurs appliquées sont les suivantes :

- Sortie d'un réseau de neurones basé sur les couleurs  $> 0.2$  pour supprimer les objets situés loin de la zone de population des quasars dans l'espace des couleurs (comme décrit dans Yèche et al. (2010) et Ross et al. (2012)).
- $(u - g) > 0.15$  pour favoriser la sélection des quasars à grands décalages spectraux ( $z > 2.15$ ) par rapport aux quasars à faibles décalages spectraux. Cette coupure ne rejette que 1% des quasars à grands décalages spectraux déjà connus.

Ce jeu de coupure nous permet de définir une liste d'environ  $70 \text{ deg}^{-2}$  objets encore dominée par les étoiles par au moins un facteur 2. La complétude des quasars est estimée à 85% ce qui correspond à une limite supérieure pour la complétude de notre « sélection principale ».

En imposant  $y_{NN} > 0.5$  puis en enlevant les quasars à faibles décalages spectraux déjà connus, nous obtenons une sélection de 7.586 objets avec une densité moyenne de  $34.5 \text{ deg}^{-2}$ . Cette sélection correspond à notre « sélection principale ». Des raisons techniques liées à la stratégie d'observation (telles que la distance minimale à laquelle deux fibres du spectrographe peuvent être disposées ou les règles de priorité entre cibles de diverses catégories : quasars, galaxies... Voir Blanton et al. (2003)) réduisent notre sélection à une densité moyenne de cibles observées de  $31.1 \text{ deg}^{-2}$ . Comme le montre la figure 4.9, un tel seuil sur la sortie du réseau de neurones entraîne une complétude des quasars de 95%. Combinée avec la complétude de 85% liée aux coupures en couleurs, nous pouvons espérer une complétude des quasars d'environ 81%.

Pour comparaison, il est possible de remplacer la sélection par variabilité par des coupures en couleurs plus fortes afin d'obtenir une nouvelle liste contenant également 7586 objets. Cette sélection, uniquement basée sur les couleurs et notre sélection principale contiennent 73% des objets en commun.

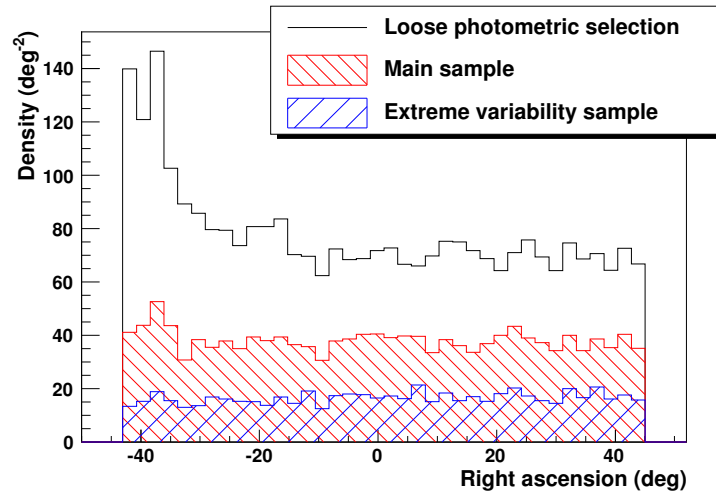


FIGURE 4.11: Distributions en ascension droite des objets de la sélection principale après pré-sélection en couleurs (histogramme noir), après sélection par variabilité (histogramme rouge) et des objets sélectionnés par variabilité extrême (histogramme bleu).

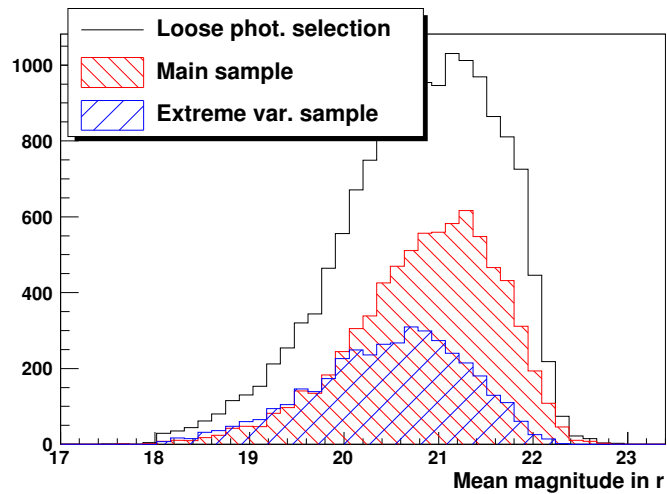


FIGURE 4.12: Distributions des magnitudes dans la bande photométrique  $r$  des objets de la sélection principale après pré-sélection en couleurs (histogramme noir), après sélection par variabilité (histogramme rouge) et des objets sélectionnés par variabilité extrême (histogramme bleu).

La densité d'étoiles dans le ciel n'est pas isotrope : elle augmente fortement lorsque l'on se rapproche du plan galactique. Nous pouvons donc nous attendre à ce qu'une sélection fortement contaminée par les étoiles ne soit également pas isotrope. La figure 4.11 montre que la densité d'objets passant les coupures en couleurs lâches dépend fortement de l'ascension droite. Notamment, elle croit fortement lorsque l'on se rapproche du plan galactique (pour une ascension droite  $\alpha_{J2000} \approx -43$  deg). Une fois la sélection par variabilité effectuée, la densité d'objets cibles devient isotropique, renforçant l'idée qu'il s'agit bien d'objets extragalactiques.

La figure 4.12 présente la distribution de la magnitude dans la bande photométrique  $r$  des objets de la sélection principale, avant et après sélection par variabilité. La chute du nombre d'objets pour  $r > 21$  vient de la pré-sélection en couleurs. Il est notable que la forme de la distribution ne change pas avant et après coupure sur la sortie du réseau de neurones, en accord avec la figure 4.10, d'après laquelle la condition  $y_{NN} > 0.5$  n'introduit pas de dépendance en décalage spectral (et donc en magnitude).

### 4.3.2 Variabilité extrême

En sélectionnant les objets présentant une variabilité extrême mais rejetés par les coupures en couleurs, nous avons l'opportunité de sonder une nouvelle région de l'espace des couleurs, très largement peuplée par les étoiles. Les figures 4.5 et 4.8 montrent l'efficacité de notre algorithme de variabilité pour discriminer les quasars des étoiles, rendant une telle sélection envisageable. Il existe tout de même un risque d'obtenir un échantillon dominé par les étoiles variables.

La région du Stripe 82 contient plusieurs millions d'objets ponctuels. Le calcul des paramètres de variabilité pour un tel nombre d'objets demande d'importantes ressources en espace disque et en temps de calcul. Pour simplifier l'étude, nous avons effectué de très lâches pré-sélections à fin d'obtenir environ  $1.000 \text{ deg}^{-2}$  objets. Les critères de la présélection sont les suivants :

- $i > 18$  et  $(u - g) > 0.4$  pour limiter la contribution des quasars à bas décalages spectraux.
- $g < 22.3$  pour assurer la qualité photométrique des spectres obtenus.
- $(g - i) < 2.2$  pour exclure les étoiles de type M.
- $c_1 < 1.5$  et  $c_3 < 0$  pour exclure une région de l'espace des couleurs éloignée des populations de quasars mais fortement peuplée par les étoiles.<sup>3</sup>

Ces coupures réduisent le nombre d'objets par un facteur dix. Nous obtenons ainsi 235.000 objets répartis sur les  $220 \text{ deg}^{-2}$  du Stripe 82. Elles rejettent également 9% des quasars à grands décalages spectraux déjà connus, sans biais apparent en magnitude.

Pour sélectionner les objets les plus variables, nous imposons la condition  $y_{NN} > 0.95$ . Nous obtenons ainsi une liste de 4.360 cibles, soit une densité moyenne d'environ  $20 \text{ deg}^{-2}$  objets, que nous appelons notre échantillon de variabilité extrême.

Comme pour la sélection principale, le fait que la distribution en ascension droite des objets de notre échantillon de variabilité extrême soit plate (voir figure 4.11) indique qu'il est dominé par des objets extragalactiques. Comme on le voit sur la figure 4.12, l'efficacité de cette sélection décroît lorsque la magnitude augmente. Ce résultat est en accord avec celui de la figure 4.10 qui montre la forte dépendance en décalage spectral d'une sélection imposant  $y_{NN} > 0.95$ .

Une fois encore, pour des raisons techniques liées à la stratégie d'observation, seuls  $\sim 15 \text{ deg}^{-2}$  objets pourraient être observés. 78% de ces objets se trouvent également dans la sélection principale de la section 4.3.1. Les 748 objets restants constituent ce que nous appelons « l'échantillon d'extrême variabilité seule ». Nous avons mis en place un programme « ancillary » dédié à l'observation de cet échantillon, correspondant à une densité de cibles de  $3.4 \text{ deg}^{-2}$ . Ce programme a

3. Les couleurs  $c_1$  et  $c_3$  sont définies dans Fan (1999) par  $c_1 = 0.95(u - g) + 0.31(g - r) + 0.11(r - i)$  et  $c_3 = -0.39(u - g) + 0.79(g - r) + 0.47(r - i)$ .

été accepté ; les résultats sont présentés dans la section suivante.

### 4.3.3 Resultats des sélections

La totalité des cibles des échantillons décrits précédemment a été observée et les données réduites par la chaîne d'acquisition de BOSS. Les spectres ont ensuite été inspectés visuellement (voir section 3.5.2). L'intégralité des spectres obtenus font partie de la 9<sup>ème</sup> publication de données du SDSS (Ahn et al., 2012) et sont disponibles via le site internet du relevé<sup>4</sup>.

Les résultats des observations sont résumés dans le tableau 4.1. Au total, 5.270 quasars à grands décalages spectraux ont été confirmés dans la région du Stripe 82. Parmi eux, 4.900 appartiennent à la sélection principale dont 2.280 font également partie de l'échantillon de variabilité extrême. 370 n'appartiennent qu'à l'échantillon de variabilité extrême. 2.770 quasars, soit environ la moitié, n'étaient pas identifiés comme tel avant cette étude. 90% des quasars à grand décalages spectraux déjà connus ont été retrouvés par leur variabilité. 92% des objets observés sont des quasars, la contamination par d'autres objets est donc de 8%. Cette pureté élevée est en accord avec la distribution plate de la figure 4.11, qui indiquait une faible contamination par les étoiles.

Echantillon	Densité de cibles	Tous quasars		$z > 2.15$		
		Dens.	$P(\%)$	Dens.	$P(\%)$	$C(\%)$
Sélection principale	31.1	29.0	93	22.3	72	84
Var. Extrême	15.1	14.6	96	12.1	80	45
Var. Extrême seule	3.4	2.9	86	1.7	49	6
Total	34.5	31.9	92	24.0	69	90

Echantillon	$2.15 < z < 3.0$			$z > 3.0$		
	Dens.	$P(\%)$	$C(\%)$	Dens.	$P(\%)$	$C(\%)$
Sélection principale	18.1	58	86	4.2	14	76
Var. extrême	10.4	68	49	1.7	11	31
Var. extrême seule	1.4	41	7	0.3	8	5
Total	19.5	56	92	4.5	13	81

TABLE 4.1: Densité, pureté  $P$  et complétude  $C$  des quasars sélectionnés par variabilité. Les densités sont les moyennes en  $\text{deg}^{-2}$  sur une surface de  $220 \text{ deg}^2$ . La pureté est définie comme le rapport entre la densité de quasars dans un échantillon donné et la densité de cibles. La complétude inclut tous les quasars identifiés à grands décalages spectraux, quelle que soit leur sélection d'origine (couleurs, variabilité, émission radio, etc.). La colonne « cibles » correspond à tous les candidats. « Tous quasars » correspond à tous les quasars, quel que soit leur décalage spectral. « Var. extrême » contient à la fois l'échantillon de variabilité extrême et les candidats de la sélection principale qui vérifient  $y_{NN} > 0.95$ . « Var. extrême seule » correspond aux candidats de l'échantillon de variabilité extrême rejetés de la sélection principale par les coupures en couleurs.

La sélection principale atteint une pureté de 93% pour l'ensemble des quasars, et de 72% pour les quasars à grands décalages spectraux. La figure 4.13 présente les distributions en décalages spectraux des quasars confirmés pour les différents échantillons. Les quasars à grands décalages spectraux ( $z > 2.15$ ) sont clairement favorisés. Notons que 86% des objets de l'échantillon de variabilité extrême seule, tous rejetés par les coupures en couleurs, sont des quasars. De plus, la

4. <http://www.sdss3.org/>

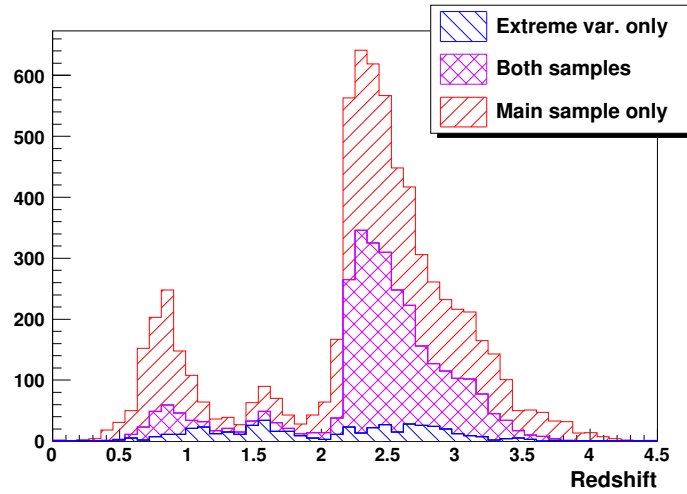


FIGURE 4.13: Distributions des décalages spectraux des quasars confirmés. Chaque quasar n'apparaît que dans un seul histogramme (variabilité extrême seule, sélection principale seule ou intersection des deux échantillons). L'échantillon de variabilité extrême est donc la somme des distributions bleue et mauve, tandis que la sélection principale est la somme des distributions mauve et rouge.

moitié d'entre eux sont à grands décalages spectraux. L'emplacement de ces quasars dans les diagrammes couleur-couleur, visible en figure 4.14, est bel et bien proche des populations d'étoiles. Pour autant, il ne s'agit pas d'une nouvelle population de quasars car leurs caractéristiques (couleur et variabilité) sont typiques des quasars connus. Bien que l'étude de la complétude de l'échantillon de variabilité extrême seule ne soit pas pertinente, du fait du peu de fibres attribuées, il est notable qu'avec  $\sim 3 \text{ deg}^{-2}$  cibles, cet échantillon augmente à lui seul la complétude de 6%. La complétude des quasars à grands décalages spectraux pour la sélection principale est de 84% soit légèrement supérieure aux 81% prédits en section 4.3.1.

Une séparation en deux gammes de décalages spectraux ( $2.15 < z < 3$  et  $z > 3$ ) montre, comme attendu, que la sélection est plus efficace à bas décalage spectral. Cette dépendance est particulièrement visible pour l'échantillon de variabilité extrême dont la complétude passe de 49% pour la gamme inférieure à 31% pour la gamme supérieure. La pureté passe respectivement de 68% à 11%.

#### 4.3.4 Comparaison avec une sélection en couleurs

Nous définissons deux sélections en couleurs. La première est une sélection utilisant une unique pose photométrique par objet. La seconde est une sélection utilisant la photométrie moyennée sur toutes les poses dont nous disposons. La seconde sélection utilisant plus d'informations, l'on peut s'attendre à ce qu'elle soit plus complète. Pour ces deux sélections en couleurs, nous produisons une liste de  $34.5 \text{ deg}^{-2}$  objets, comme nous l'avons fait pour l'ensemble de la sélection par variabilité (sélection principale et variabilité extrême). Nous comparons alors le nombre de quasars retrouvés par chacune des méthodes parmi l'ensemble des quasars connus appartenant au Stripe 82. Le nombre de quasars trouvés en fonction du décalage spectral pour les trois sélections apparaît en figure 4.15. La méthode par variabilité est clairement plus performante, avec un gain allant jusqu'à un facteur 2, et le nombre de quasars sélectionnés décroît linéairement avec

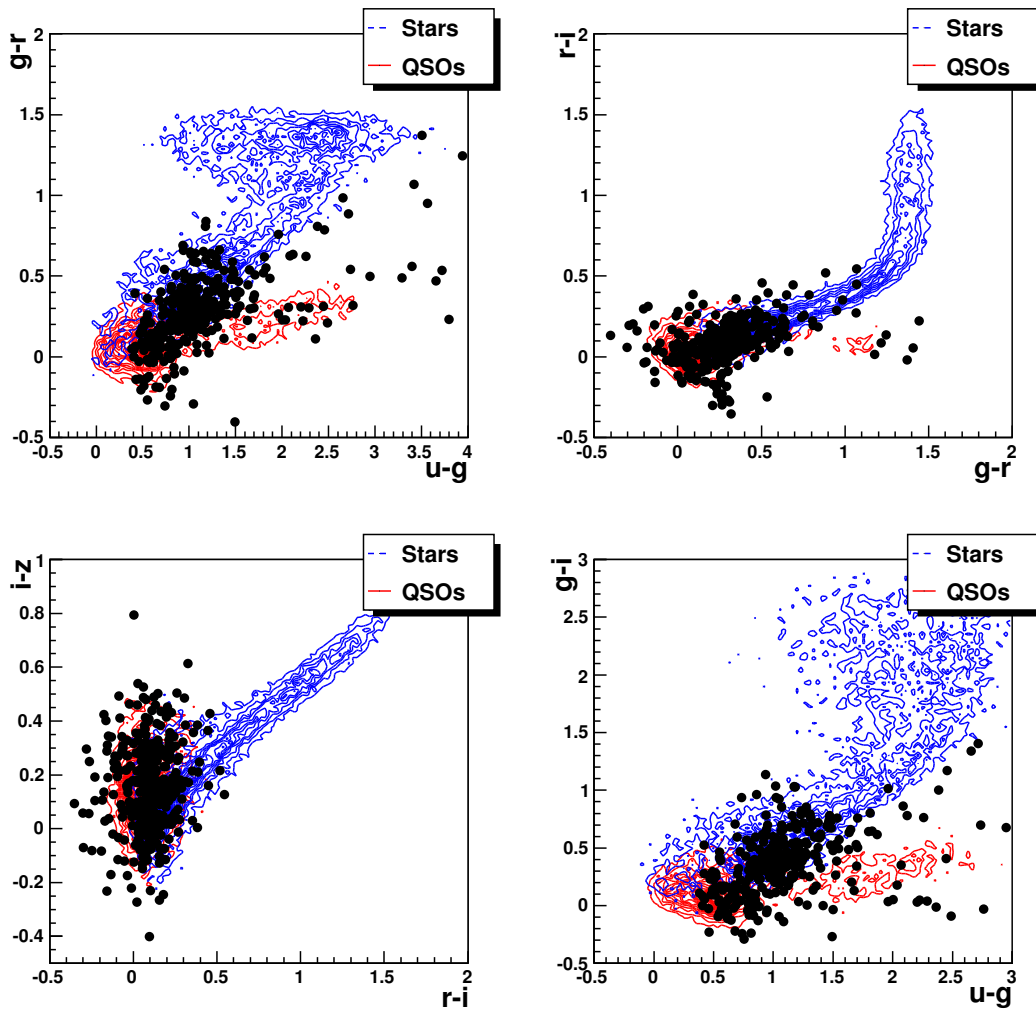


FIGURE 4.14: Diagrammes couleur-couleur indiquant les populations d'étoiles (en bleu), de quasars (en rouge) ainsi que la position des 370 quasars à grands décalages spectraux rejetés par les coupures en couleurs mais sélectionnés par la méthode de variabilité extrême (points noirs).

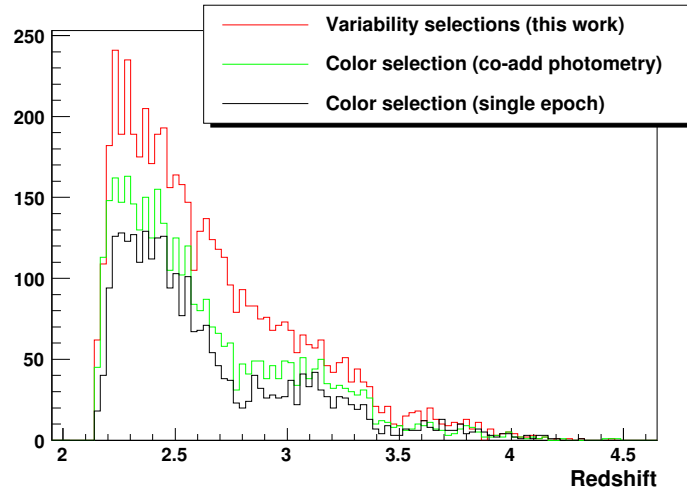


FIGURE 4.15: Comparaison des distributions des décalages spectraux des quasars obtenues par notre méthode de variabilité (rouge), par une sélection en couleurs sur une époque photométrique (noir) et par une sélection en couleurs sur la moyenne des poses photométriques (vert).

le décalage spectral. A contrario, le biais de sélection introduit par les méthodes basées sur les couleurs est visible autour de  $z = 2.8$ .

## 4.4 Prévisions pour les relevés futurs

Nous avons montré l'efficacité de la sélection des quasars par variabilité appliquée aux  $220 \text{ deg}^{-2}$  la région du Stripe 82. En section 4.5 nous étendrons la sélection par variabilité à l'ensemble du relevé de BOSS. Le manque de points photométriques dans les courbes de lumière des objets en dehors du Stripe 82 rendra toutefois cette sélection bien moins efficace que celle présentée précédemment. Dans cette section nous étudions la possibilité de combiner les données photométriques de différents relevés pour mettre en place une sélection par variabilité des quasars. Une telle sélection pourra s'avérer déterminante pour les futurs relevés tels que BigBOSS<sup>5</sup> ou LSST (LSST Science Collaboration et al., 2009). Deux relevés, The Palomar Transient Factory (PTF) et Pan-STARRS 1, observent actuellement la même partie du ciel que SDSS. Ils couvriront les  $10.000 \text{ deg}^{-2}$  à plusieurs reprises, sur une période de 3 à 5 ans. Nous étudions l'efficacité d'une sélection basée sur les données photométriques du SDSS combinées à l'un ou l'autre de ces relevés.

### 4.4.1 Extrapolation pour le Palomar Transient Factory

PTF prend des données dans la bande photométrique  $r$  depuis décembre 2008 à raison d'une mesure toutes les 5 nuits (Rau et al., 2009). Ces données peuvent être additionnées pour obtenir 4 images par an possédant une bonne qualité photométrique. SDSS ne couvrant qu'une seule fois le ciel complet, la combinaison des catalogues, une fois la prise de données de PTF terminée, consistera essentiellement en un point photométrique du SDSS (utile pour augmenter le bras de levier temporel) et de 20 points de PTF (4 points par année sur 5 ans).

5. bigboss.lbl.gov

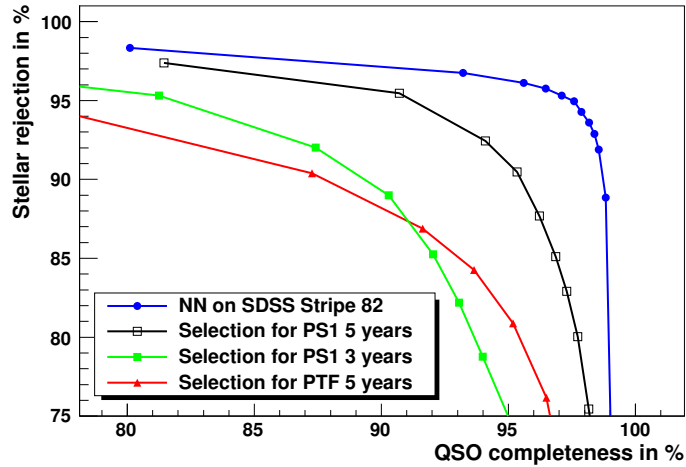


FIGURE 4.16: Réjections stellaires en fonction de la complétude des quasars obtenues par variabilité sur le Stripe 82 (ronds bleus), prédites pour Pan-STARRS 3 et 5 ans (carrés verts et noirs) et prédites pour PTF 5 ans (triangles rouges). Les marqueurs correspondent à différentes coupures sur la valeur de la sortie du réseau de neurones  $y_{NN}$ .

Pour simuler un tel catalogue nous utilisons les courbes de lumière des objets de la région du Stripe 82 mesurées par SDSS. Pour chaque courbe de lumière, 4 points par an, régulièrement répartis au cours du temps, ont été sélectionnés parmi les 5 dernières années de prise de données du SDSS afin de simuler les points PTF. L'unique point du SDSS est sélectionné parmi les mesures précédentes, en s'assurant de conserver un écart temporel d'au moins 2 ans pour assurer un bras de levier réaliste.

Seules les courbes de lumière synthétiques présentant les 21 points photométriques (1 pour SDSS et 20 pour PTF) sont considérées. Nous obtenons ainsi 2.248 courbes de lumière d'étoiles et 11.456 de quasars soit respectivement 83% et 86% des échantillons définis en section 4.2.2.

PTF observant uniquement dans la bande photométrique  $r$ , nous nous limitons aux paramètres de variabilité  $r$ ,  $A_r$  et  $\chi_r^2$  définis en section 4.2.4. Un réseau de neurones (voir section 4.2.6) utilisant ces 3 paramètres a été entraîné. La réjection stellaire obtenue en fonction de la complétude des quasars apparaît en figure 4.16 (triangles rouges). Il est par exemple possible d'obtenir une complétude des quasars proche de 88% pour une réjection stellaire de 90%. Pour comparaison, les résultats obtenus en utilisant les données complètes du Stripe 82 sont également tracés (ronds bleus).

#### 4.4.2 Extrapolation pour Pan-STARRS1

Pan-STARRS 1 (PS1) a démarré ses observations en mars 2009, observant l'intégralité du ciel visible 3 fois pendant la période sombre de chaque cycle lunaire. La première partie du projet devait durer 3 ans, elle est actuellement en phase terminale. Nous simulons les observations de PS1 suivant le même procédé que pour PTF, à la différence que PS1 dispose d'observations dans 5 bandes photométriques :  $g$ ,  $r$ ,  $i$ ,  $z$  et  $y$ . Pour établir notre sélection nous nous limitons aux 4 bandes communes avec SDSS. Nous simulons les courbes de lumière pour deux durées possibles pour PS1, 3 et 5 ans. La qualité photométrique des mesures de PS1 correspond environ à 4 points photométriques du SDSS par an. Nous ne gardons donc que les courbes possédant l'intégralité des



13 points photométriques (1 pour SDSS et 12 pour PS1) pour un relevé de 3 ans et 21 points pour un relevé de 5 ans. Nous pouvons alors calculer 8 paramètres de variabilité par objet : les 4  $\chi^2$ , les amplitudes des fonctions de structure  $A_g, A_r, A_i$  et un unique  $\gamma$ .

Les performances d'un réseau de neurones basé sur ces paramètres sont représentées en figure 4.16. Alors que les taux de réjection stellaire et de complétudes des quasars sont similaires à PTF pour un relevé 3 ans, l'extension à un relevé de 5 ans améliorerait considérablement les performances. Il est ainsi possible d'atteindre une réjection stellaire de 95% pour une complétude des quasars de près de 92%.

L'absence du point photométrique de SDSS entraîne une diminution de la complétude des quasars de l'ordre de 3%, sans avoir d'influence sur la réjection stellaire puisque la plupart des étoiles ont une courbe de lumière plate, quelle que soit la période couverte.

## 4.5 Application à l'ensemble du relevé

32% de l'ensemble des objets observés en photométrie par SDSS ont été observés de deux à neuf fois sur une période d'au moins un mois. Cette multi-observation d'environ un tiers des objets résulte de la stratégie d'observation appliquée lors des phases de photométrie (voir section 3.3). Pour couvrir la totalité des  $14.555 \text{ deg}^2$  de ciel observés en photométrie avec le champ de vue sphérique de  $7 \text{ deg}^2$ , SDSS a utilisé la méthode du « drift scanning » consistant à laisser le télescope fixe, utilisant la rotation de la Terre pour couvrir une bande photométrique. Pour observer les régions du pôle Sud galactique et du pôle Nord galactique de manière connexe, il est alors nécessaire de scanner de nombreuses bandes photométriques dont une partie au moins se recouvre. Nous obtenons ainsi plusieurs points photométriques pour un même objet. Les bandes, résultant de cette stratégie d'observation, sont en partie visibles sur l'image finale du ciel produite par le SDSS (figure 3.3). Nous utilisons cette information multi-époque dans le but de sélectionner les quasars en utilisant leur variabilité intrinsèque sur l'ensemble du relevé.

### 4.5.1 Échantillons

Pour mener cette étude nous nous basons sur les mêmes échantillons de quasars et d'étoiles que ceux décrits en section 4.2.2. Pour chaque objet de ces échantillons, la courbe de lumière est dégradée de manière à reproduire l'échantillonnage typique du relevé complet (hors Stripe 82). Pour ce faire, un certain nombre d'observations est sélectionné de manière aléatoire. Les figures 4.17 et 4.18 présentent les nombres d'observations ainsi que l'écart en temps entre la première et la dernière observation, pour les courbes de lumière simulées, ainsi que pour le relevé complet (hors Stripe 82).

### 4.5.2 Tests de variabilité et stratégie d'observation

Le peu d'information contenu dans les courbes de lumière simulées (peu de points photométriques sur des échelles de temps souvent courtes) proscrit l'application de la méthode décrite précédemment (voir section 4.2.4) à l'ensemble du relevé. Il n'est notamment pas possible d'ajuster les deux paramètres  $A$  et  $\gamma$  de la fonction de structure à une courbe de lumière ne possédant que deux observations. Pour simplifier la procédure, permettant son application lorsque peu de données sont disponibles, nous ne considérons que les trois bandes photométriques  $g, r$  et  $i$  (les plus discriminantes d'après l'étude menée sur le Stripe 82). Après avoir calculé les fonctions de structure telles que définies en équation 4.1, nous ajustons celles-ci par une loi de puissance de la

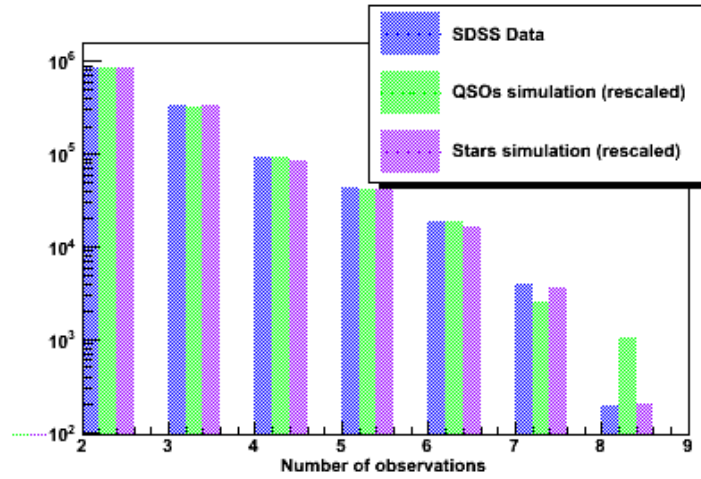


FIGURE 4.17: Nombre de courbes de lumière en fonction du nombre d'observations pour le relevé complet (hors Stripe 82), ainsi que pour les échantillons de quasars et d'étoiles (normalisés au relevé complet.)

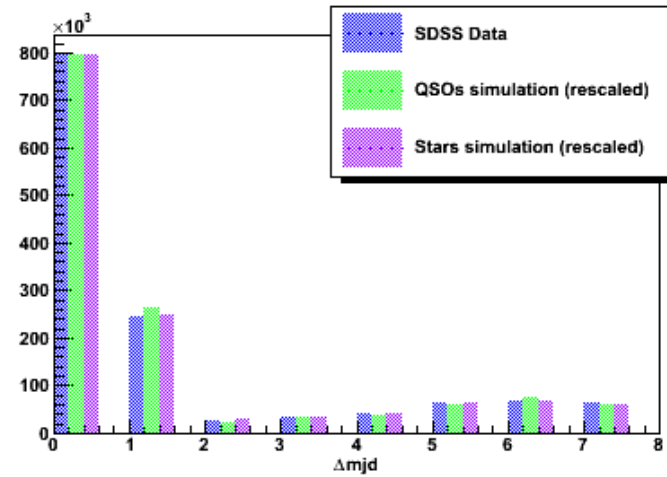


FIGURE 4.18: Nombre de courbes de lumière en fonction de l'écart en temps entre la première et la dernière observation ( $\Delta mjd$  en année) pour le relevé complet (hors Stripe 82), ainsi que pour les échantillons de quasars et d'étoiles (normalisés au relevé complet.)

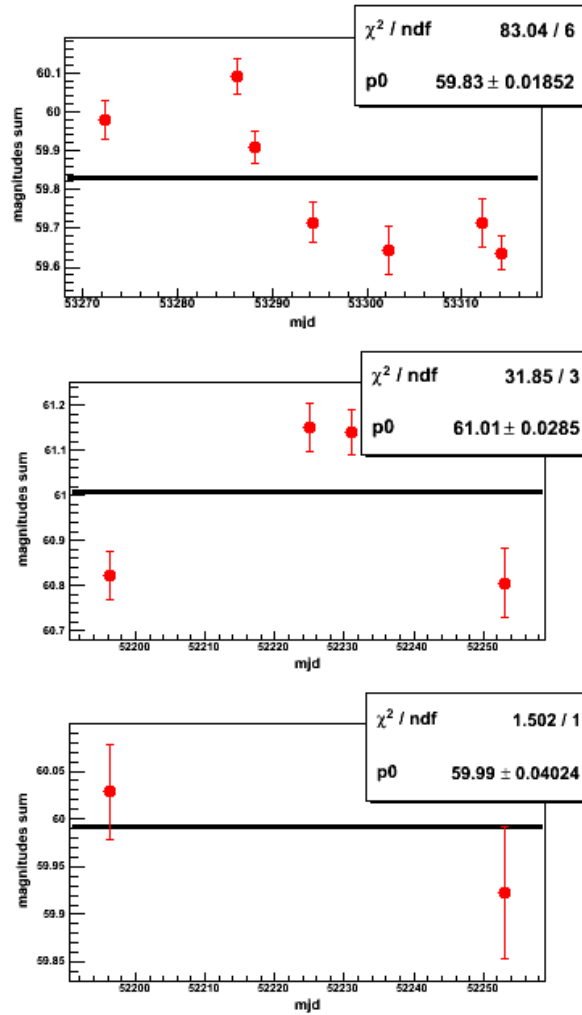


FIGURE 4.19: Courbes de lumière obtenues en sommant les données photométriques dans les bandes  $g$ ,  $r$  et  $i$  et ajustées par une constante  $p_0$  pour trois objets tirés aléatoirement de l'échantillon de quasars.

forme  $A(\Delta t)^\gamma$  où  $\gamma$  est fixé à sa valeur moyenne sur l'échantillon de quasars (non dégradé) et  $A$  est le seul paramètre libre, commun aux trois bandes photométriques.

Comme pour la sélection sur le Stripe 82, les  $\chi^2$  résultant de l'ajustement des courbes de lumière par une constante sont également considérés. Dans le but d'améliorer la photométrie des objets, nous moyennons, pour chaque observation, les magnitudes dans les trois différentes bandes photométriques produisant une unique courbe de lumière par objet. C'est cette courbe de lumière moyenne qui est ensuite ajustée par une constante, donnant un unique  $\chi^2$  par objet. Les quasars présentant une forte variabilité, on s'attend à ce que leurs  $\chi^2$  réduits soient supérieurs à un, tandis que ceux des étoiles, plus stables, devraient être proche de un. Trois exemples d'ajustement de courbes de lumière sont donnés en figure 4.19 pour des quasars et en figure 4.20 pour des étoiles. Ces six objets ont été sélectionnés aléatoirement. Les distributions des  $\chi^2$  pour les échantillons d'étoiles et de quasars sont représentées en figure 4.21

Disposant du paramètre  $A$  et du  $\chi^2$  pour chaque objet de chaque échantillon, il est possible d'effectuer une sélection en rejetant tous les objets possédant un des deux paramètres en-dessous

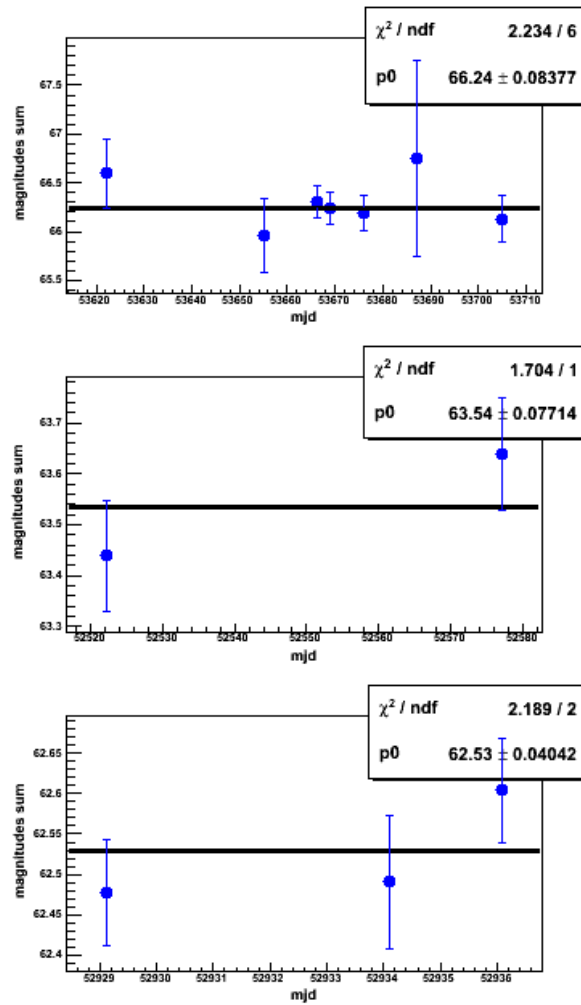


FIGURE 4.20: Courbes de lumière obtenues en sommant les données photométriques dans les bandes  $g$ ,  $r$  et  $i$  et ajustées par une constante  $p_0$  pour trois objets tirés aléatoirement de l'échantillon d'étoiles.

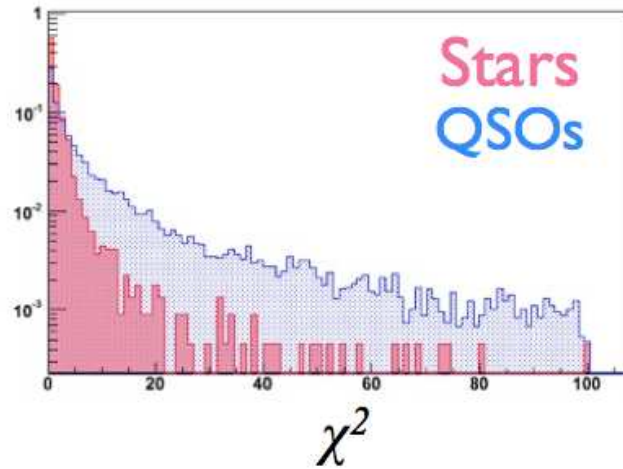


FIGURE 4.21: Distributions des  $\chi^2$  obtenus en ajustant les courbes de lumière moyennes pour l'échantillon d'étoiles et celui de quasars.

d'un certain seuil. Plus ce seuil est élevé, plus la probabilité de sélectionner des quasars est grande mais plus la complétude des quasars est faible. Ce comportement est illustré en figure 4.22 où la réjection stellaire est tracée en fonction de la complétude des quasars pour différentes coupures sur la valeur des  $\chi^2$  et sur le paramètre  $A$  de la fonction de structure. Quel que soit le paramètre sur lequel est effectué la coupure, une forte réjection stellaire entraîne une très faible complétude des quasars. Les performances obtenues sont bien en deçà de celles atteintes sur la région très favorable du Stripe 82. De ce fait la sélection par variabilité des quasars sur l'ensemble du relevé n'est pas compétitive avec la sélection par couleurs.

Il est toutefois possible de sélectionner une faible quantité d'objets présentant une variabilité extrême comme nous l'avons fait sur le Stripe 82. L'idée est de sélectionner les objets les plus variables qui ne passent pas les coupures photométriques. L'amplitude de la fonction de structure  $A$  et le  $\chi^2$  étant très fortement corrélés, je ne considère que le  $\chi^2$  qui permet une meilleure réjection stellaire pour des coupures strictes.

Pour limiter le nombre d'objets à traiter, des coupures en couleurs lâches, identiques à celles utilisées pour l'échantillon de variabilité extrême du Stripe 82 (voir section 4.3.2) sont appliquées à l'ensemble du relevé. Suite aux coupures en  $\chi^2$ , une densité moyenne sur l'ensemble du relevé de  $6.5 \text{ deg}^{-2}$  cibles est attendue, passant à  $5 \text{ deg}^{-2}$  après avoir enlevé les quasars connus ou sélectionnés par couleurs. Nous appelons le lot ainsi sélectionné « échantillon de variabilité SDSS ». Cette densité peut varier fortement en fonction de la région considérée. Du fait de la stratégie d'observation, les régions proches du plan galactique possèdent plus d'objets observés à plusieurs reprises. Elles sont donc plus favorables pour cette sélection. De plus, les cibles de l'échantillon de variabilité SDSS sont attribuées aux fibres restantes lors du troisième tour du « pavage » des bandes spectroscopiques, c'est-à-dire une fois que les galaxies et les quasars des listes principales et secondaires ont été attribués, renforçant les différences entre bandes spectroscopiques.

### 4.5.3 Résultats

Les résultats suivants sont obtenus pour la bande spectroscopique BOSS 19, favorable à cette sélection car proche du plan galactique. 945 objets de l'échantillon de variabilité SDSS ont été

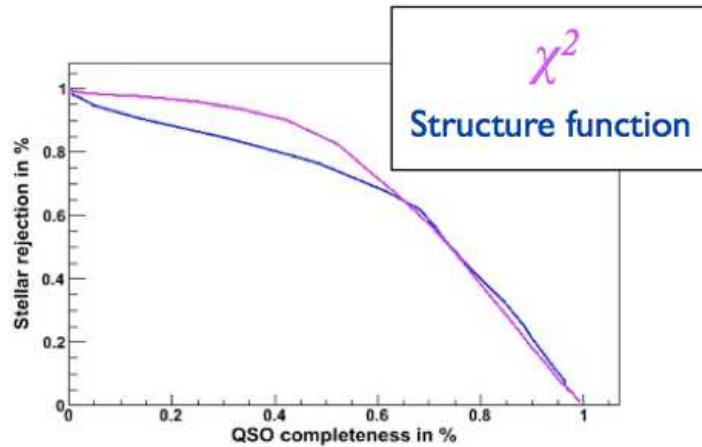


FIGURE 4.22: Réjection stellaire en fonction de la complétude des quasars pour différentes coupures sur le  $\chi^2$  de l'ajustement de la courbe de lumière par une constante et sur le paramètre A de la fonction de structure.

observés sur les  $229 \text{ deg}^2$  qui constituent la bande 19, soit une densité moyenne de  $4.1 \text{ deg}^{-2}$  cibles. Parmi ces objets, la moitié, soit  $\sim 2 \text{ deg}^{-2}$ , sont des quasars et  $\sim 0.8 \text{ deg}^{-2}$  sont des quasars à grands décalages spectraux. Il est remarquable que cette sélection atteigne une pureté de 50% avec si peu d'information photométrique. Compte tenu de son efficacité, cette sélection est à présent implémentée de façon systématique sur l'ensemble du relevé BOSS, pour compléter les sélections standards.



## Chapitre 5

# Fonction de corrélation des forêts Lyman- $\alpha$

*« Dire des idioties, de nos jours où tout le monde réfléchit profondément, c'est le seul moyen de prouver qu'on a une pensée libre et indépendante. »*

—B. Vian, Le Goûter des généraux

---

Ce chapitre présente le principal travail de cette thèse, à savoir le développement d'une analyse dédiée à la recherche du pic BAO dans la fonction de corrélation des forêts Lyman- $\alpha$  des quasars. Au commencement de cette thèse, il n'était pas évident que cette analyse déboucherait sur une détection. Grâce à la qualité exceptionnelle des données de BOSS, due au design des instruments de mesure ainsi qu'à l'effort important fourni pour la réduction des données, mais également grâce à une météo particulièrement clémente nous ayant permis d'accumuler les données plus rapidement que prévu, ce chapitre s'étend jusqu'à tirer des contraintes cosmologiques de la position du pic BAO.

---

### 5.1 Echantillon de données

Le travail présenté ici est une mise à jour de l'analyse des données DR9 ayant donné lieu à la publication Busca et al. (2013). Cette publication, retranscrite en annexe B, présente deux méthodes d'analyse, appelées méthode 1 et méthode 2, similaires en de nombreux points excepté la détermination du flux moyen des quasars que nous aborderons en section 5.2.2. Le développement de la méthode 1 constitue une grande partie de ce travail de thèse. Nous présentons ici les résultats obtenus par cette méthode appliquée à une version préliminaire de la dixième publication de données du SDSS, appelée Data Release 10 alpha (DR10 $\alpha$ ). Nous utilisons cette version préliminaire car la version finale n'était pas disponible au moment où nous avons mis à jour cette analyse. Cette version ne possède que 0,5% de quasars en moins que la version finale de DR10. Nous nous attendons donc à ce que les résultats restent inchangés lors du passage de DR10 $\alpha$  à DR10 final. Les données spectroscopiques utilisées couvrent 6.161 deg<sup>2</sup> soit environ deux tiers de



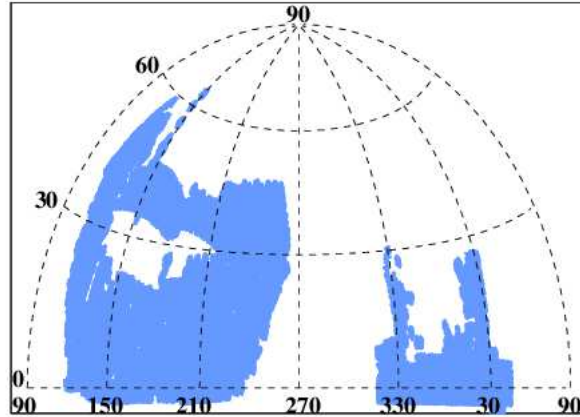


FIGURE 5.1: Projection Hammer-Aitoff en coordonnées équatoriales des données de DR10. Les observations couvrent  $6.161 \text{ deg}^2$ , soit environ deux tiers du relevé final.

la surface du relevé complet. Une projection Hammer-Aitoff en coordonnées équatoriales de cette couverture est visible en figure 5.1. La réduction des données a été effectuée comme décrit en 3.5. L'inspection visuelle a confirmé 165.957 appartenant à DR10 $\alpha$ .

### 5.1.1 Sélection de l'échantillon de quasar

Pour l'analyse de la fonction de corrélation des forêts Lyman- $\alpha$ , nous ne retenons que les quasars dont le décalage spectral  $z_q$  vérifie  $2,1 \leq z_q \leq 3,5$ . La borne supérieure vient du très faible nombre de quasars ayant un  $z_q$  supérieur à 3,5 (voir figure 5.3). Comme nous le verrons en section 5.2.2, le calcul de la fonction de corrélation des forêts Lyman- $\alpha$  passe par une estimation du flux du quasar, avant absorption par le milieu intergalactique. Cette estimation est réalisée pour chaque quasar, par ajustement d'une fonction aux données nécessitant un certain nombre de points de mesure dans la forêt Lyman- $\alpha$ . La borne inférieure en décalage spectral permet de ne retenir que les spectres ayant un morceau de forêt suffisamment important pour que cet ajustement soit fiable. Elle dérive de la coupure dans le bleu des spectrographes aux alentours de  $3.600 \text{ \AA}$  (voir figure 3.13) qui correspond à une coupure sur le décalage spectral des absorbeurs (c'est-à-dire des nuages d'hydrogène neutre produisant l'absorption) :  $z_{abs} > 1,96$ .

L'inspection visuelle a confirmé 115.024 quasars vérifiant nos coupures en décalage spectral. La figure 5.2 montre le signal sur bruit par pixel, moyenné sur la forêt de ces quasars. Le signal sur bruit typique  $S/B \sim 3$  reflète le temps d'exposition relativement court attribué à chaque plaque ( $\sim 1\text{h}$ ) dans le but d'achever le relevé à temps. Parmi les 115.024 quasars, 13.122 ont été identifiés visuellement comme possédant au moins un système DLA et 12.162 comme étant des quasars de type BAL. Ces nombres sont résumés en table 5.1. Comme décrit en section 2.4.3, les spectres des quasars de type BAL présentent une très forte absorption proche du pic Lyman- $\alpha$  qui n'est pas associée à l'hydrogène neutre du MIG. Cette absorption se situe généralement à plus faible longueur d'onde, donc dans la forêt Lyman- $\alpha$ . Des BAL peuvent également affecter les raies d'émission des métaux et notamment celles appartenant à la région de la forêt Lyman- $\alpha$  telles que la raie FeII à  $1.070,95 \text{ \AA}$  ou la raie FeIII à  $1.123,17 \text{ \AA}$ . Pour éviter d'introduire des corrélations

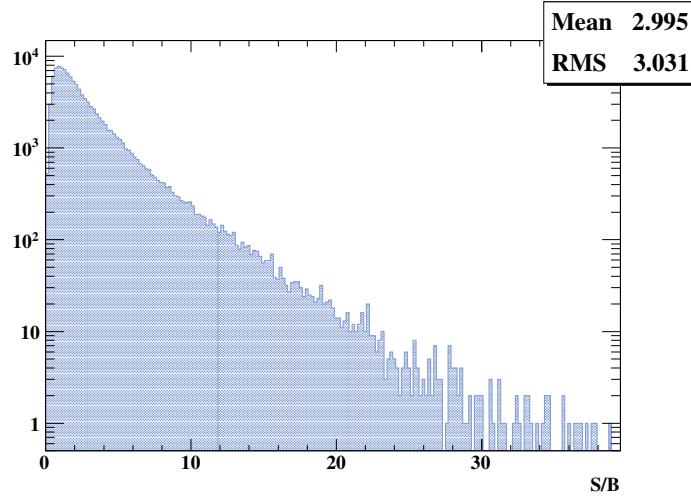


FIGURE 5.2: Distribution du signal sur bruit S/B des pixels moyenné sur la forêt des 115.024 quasars vérifiant nos coupures en décalage spectral.

	DR10 $\alpha$
Total	165.957
$2.1 \leq z \leq 3.5$	115.024
BAL	12.162
DLA	13.122
BAL $\cap$ DLA	740
Coupures qualités	773
Echantillon final	89.322

TABLE 5.1: Résumé du nombre de quasars pour la version préliminaire de DR10 utilisée pour cette analyse (DR10 $\alpha$ ).

non désirées liées à la présence de ces raies d'absorption, les quasars de types BAL sont retirés de notre échantillon. Nous retirons également les quasars présentant un système DLA, car ces systèmes peuvent avoir un impact sur le flux du quasar dans la forêt Lyman- $\alpha$  pouvant s'étendre sur plusieurs dizaines d'angström autour de leur position centrale. Notons que nous effectuons ces coupures d'après l'inspection visuelle mais qu'il existe également des catalogues automatiques de recherche des systèmes DLA (Noterdaeme et al., 2009).

Parmi les quasars restant, 773 sont supprimés suivant des critères de qualité (signal sur bruit trop faible ou nombre de points de mesure dans la forêt insuffisant); Nous obtenons donc un échantillon final de 94.561 quasars (environ le double de l'échantillon utilisé pour l'analyse DR9). La distribution en décalage spectral des quasars appartenant à notre échantillon final est visible en figure 5.3.

### 5.1.2 Définition de la forêt Lyman- $\alpha$

La forêt Lyman- $\alpha$  est encadrée par les raies d'émissions Lyman- $\alpha$  et Lyman- $\beta$  respectivement à  $\lambda_{Ly-\alpha} = 1.215 \text{ \AA}$  et  $\lambda_{Ly-\beta} = 1.025 \text{ \AA}$ . Pour la mesure du flux transmis, nous choisissons l'intervalle de longueurs d'onde dans le référentiel du quasar :

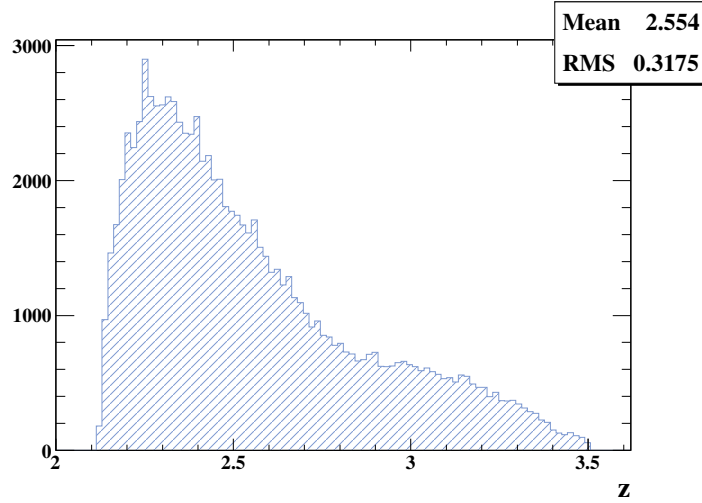


FIGURE 5.3: Distribution des décalages spectraux des quasars utilisés pour l'analyse.

$$1.040 < \lambda_{rq} < 1.200 \text{ \AA} \quad (5.1)$$

Cet intervalle est légèrement plus large que celui choisi pour l'analyse de DR9 :  $1.045 < \lambda_{rq} < 1.180 \text{ \AA}$  (Busca et al., 2013). Il est toutefois plus restreint que celui utilisé par Slosar et al. (2013) :  $1.030 < \lambda_{rq} < 1.210 \text{ \AA}$ . L'augmentation par rapport à l'étude menée sur DR9 permet de gagner en statistique. Le gain est de 12% sur le nombre de pixels utilisés pour le calcul de la fonction de corrélation. L'élargissement de la forêt est également justifié par la figure 5.4 montrant la variance du flux mesuré en fonction de la longueur d'onde dans le référentiel du quasar. Lorsque l'on est loin des raies d'émissions Lyman- $\alpha$  et Lyman- $\beta$ , cette variance est dominée par le bruit de mesure ainsi que la variance intrinsèque de la forêt Lyman- $\alpha$ . Elle devient dominée par la variance de l'intensité des raies Lyman- $\alpha$  et Lyman- $\beta$  d'un quasar à l'autre lorsque l'on s'approche de ces raies. Comme nous le verrons en section 5.2.2, les différences d'intensité des raies d'émissions Lyman- $\alpha$  et Lyman- $\beta$  d'un quasar à l'autre ne sont pas prises en compte pour l'estimation du flux moyen des quasars. Dans la région choisie, elle est inférieure à la somme de la variance intrinsèque de la forêt et du bruit de mesure.

### 5.1.3 Pré-traitement des données

Les spectres subissent un traitement préliminaire au calcul de la fonction de corrélation. Ce pré-traitement s'effectue en trois étapes :

- Pour chaque spectre, l'intégrale du flux compris entre  $1.275 \text{ \AA}$  et  $1.285 \text{ \AA}$  dans le référentiel du quasar est normalisée à 1. Cette région a été choisie pour la normalisation car elle est dépourvue de raies d'émission.
- Un masque visant à retirer les longueurs d'ondes contaminées par des raies d'émission ou d'absorption est appliqué à chaque spectre. Les régions masquées sont données en table 5.2. Parmi elles figurent les raies d'absorptions H et K produites par le calcium ionisé. Ces raies résultent certainement d'une combinaison d'absorption par le halo de la Voie Lactée, le milieu interstellaire ainsi que l'environnement proche du Soleil. Parmi les raies masquées

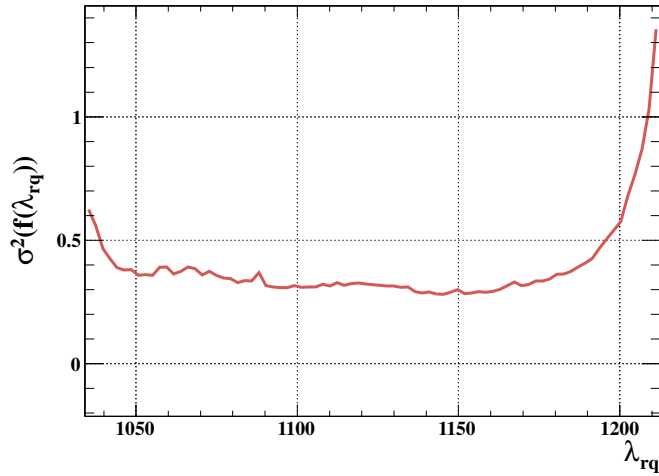


FIGURE 5.4: Variance du flux en fonction de la longueur d'onde dans le référentiel du quasar  $\lambda_{rq}$ , calculée sur l'échantillon final de quasar.

Longueurs d'onde masquées	Raie correspondante
$3.932 < \lambda < 3.937 \text{ \AA}$	Ca (H)
$3.966 < \lambda < 3.972 \text{ \AA}$	Ca (K)
$4.042 < \lambda < 4.050 \text{ \AA}$	Hg
$4.357 < \lambda < 4.362 \text{ \AA}$	Hg
$5.458 < \lambda < 5.464 \text{ \AA}$	Hg
$5.574 < \lambda < 5.582 \text{ \AA}$	OI
$5.682 < \lambda < 5.695 \text{ \AA}$	?
$5.886 < \lambda < 5.900 \text{ \AA}$	?

TABLE 5.2: Liste des longueurs d'ondes masquées car contaminées par des raies d'absorptions ou d'émissions. Parmi elles figurent les raies d'absorptions H et K du calcium ionisé ainsi que les raies d'émissions du mercure.

figurent également les raies d'émissions du mercure produites par les éclairages des villes situées à proximité de l'observatoire.

- Les pixels de chaque spectre sont regroupés afin de former des « super-pixels » de  $3 \text{ \AA}$  de large en moyenne. Ces  $3 \text{ \AA}$  correspondent environ à  $3 \text{ Mpc}$  pour un décalage spectral de  $2,3$ . Dans la suite de ce chapitre, quand nous mentionnerons les pixels des spectres, nous ferons référence à ces super-pixels. Le regroupement des pixels permet de diminuer le nombre total de paires de pixels et donc d'accélérer le calcul de la fonction de corrélation.

## 5.2 Mesure de la fonction de corrélation

La mesure de la fonction de corrélation des galaxies nécessite l'utilisation d'estimateurs tels que l'estimateur de Landy-Szalay, faisant appel à des échantillons aléatoires. Ces échantillons aléatoires permettent de corriger les inhomogénéités du relevé. Cette difficulté est liée au fait que

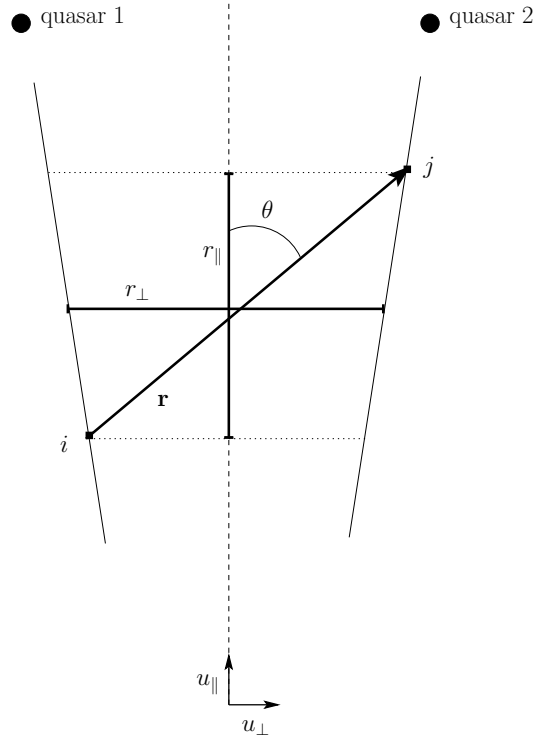


FIGURE 5.5: Schéma représentant deux lignes de visée de deux quasars. La paire de pixels  $(i, j)$  est séparée par le vecteur  $\mathbf{r}$ , faisant un angle  $\theta$  avec la ligne de visée moyenne (ligne en pointillés). Le vecteur  $\mathbf{r}$  peut être séparé en une composante  $r_{\parallel}$  sur le vecteur unitaire  $\mathbf{u}_{\parallel}$  et une composante  $r_{\perp}$  sur le vecteur unitaire  $\mathbf{u}_{\perp}$ .

les objets que l'on cherche à observer, les galaxies, sont également les traceurs de la matière. Un relevé inhomogène peut donc entraîner un biais de la fonction de corrélation. Dans le cas des forêts Lyman- $\alpha$ , la position des objets que l'on souhaite observer, les quasars, est (au premier ordre au moins) décorrélée de la densité d'hydrogène que l'on mesure le long de leur ligne de visée. Il est donc possible de revenir à un estimateur plus simple de la fonction de corrélation.

### 5.2.1 Estimateur

La fonction de corrélation peut être estimée par une simple procédure à deux étapes. La première consiste à attribuer à chaque pixel de la forêt de chaque quasar  $q$ , une estimation du flux moyen attendu à la longueur d'onde du pixel  $\lambda$ , c'est-à-dire le flux  $\langle f_q(\lambda) \rangle$  que l'on observerait s'il n'y avait pas de fluctuation dans la distribution d'hydrogène neutre du MIG. Une fois cette estimation obtenue, elle est comparée au flux mesuré  $f_q(\lambda)$  à cette même longueur d'onde, définissant le « champ de delta » :

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{\langle f_q(\lambda) \rangle} - 1 \quad (5.2)$$

La fonction de corrélation est alors estimée par la somme pondérée du produit des paires de  $\delta$  :

$$\hat{\xi}_A = \frac{\sum_{i,j \in A} w_{ij} \delta_i \delta_j}{\sum_{i,j \in A} w_{ij}} \quad (5.3)$$

où  $w_{ij}$  est le poids associé à la paire de pixels  $\delta_i \delta_j$  (déterminé en section 5.2.4) et où les indices  $i$  et  $j$  font référence à une mesure de flux à une longueur d'onde donnée, sur une ligne de visée d'un quasar donné. La somme sur les indices  $(i, j)$  est définie comme la somme sur toutes les paires de pixels de toutes les paires de quasar appartenant à  $A$ , définissant une région dans l'espace de séparation des pixels que nous définissons explicitement plus bas. Elle ne tient pas compte des paires de pixels provenant de la ligne de visée d'un unique quasar afin d'éviter d'introduire des erreurs corrélées entre les  $\delta_i$  et  $\delta_j$  dues à une mauvaise estimation du flux moyen  $\langle f_q(\lambda) \rangle$ .

Si le vecteur  $\mathbf{r}$  représente la séparation entre deux pixels, la fonction de corrélation peut être estimée en deux dimensions dans les coordonnées  $(r, \mu)$  définies par :

$$r = |\mathbf{r}| \quad \text{et} \quad \mu = \frac{\mathbf{r} \cdot \mathbf{u}_{\parallel}}{r} = \cos\theta \quad (5.4)$$

où  $\mathbf{u}_{\parallel}$  et  $\mathbf{u}_{\perp}$  sont les vecteurs unitaires respectivement le long et transversalement à la bissectrice des lignes de visée, et  $\theta$  est l'angle formé par  $\mathbf{u}_{\parallel}$  et  $\mathbf{r}$  (voir figure 5.5). La région  $A$  est alors définie par les gammes  $r_{\min} < r < r_{\max}$  et  $\mu_{\min} < \mu < \mu_{\max}$ .

Le vecteur  $\mathbf{r}$  peut également être divisé en une composante radiale, parallèle à la ligne de visée, et une composante transverse, perpendiculaire à la ligne de visée :

$$r_{\parallel} = \mathbf{r} \cdot \mathbf{u}_{\parallel} \quad \text{et} \quad r_{\perp} = \mathbf{r} \cdot \mathbf{u}_{\perp} \quad (5.5)$$

La fonction de corrélation peut alors être étudiée en deux dimensions dans les coordonnées  $(r_{\parallel}, r_{\perp})$ .

Dans tous les cas, les coordonnées observées pour chaque pixel sont l'ascension droite (ra), la déclinaison (dec) et le décalage spectral ( $z$ ). Le passage des coordonnées  $(ra, dec, z)$  aux coordonnées  $(r, \mu)$  ou  $(r_{\parallel}, r_{\perp})$  nécessite l'utilisation d'une cosmologie dite fiducielle. Nous prenons comme cosmologie fiducielle le modèle  $\Lambda$ CDM avec les valeurs des paramètres suivant :

$$(\Omega_m, \Omega_{\Lambda}, \Omega_b h^2, h, \sigma_8, n_s) = (0.27, 0.73, 0.0227, 0.7, 0.8, 0.97) \quad (5.6)$$

Seuls  $\Omega_m$  et  $\Omega_{\Lambda}$  sont nécessaires au calcul des distances jusque  $z = 3,5$ . Les autres paramètres sont toutefois nécessaires pour la génération de simulations, comme nous le verrons en section 5.4.

## 5.2.2 Estimation du flux moyen attendu

Pour un quasar  $q$ , le flux moyen attendu  $\langle f_q(\lambda) \rangle$  est le produit du flux non-absorbé du quasar  $C_q(\lambda)$ , également appelé « continu » du quasar, et de la fraction de flux transmise moyenne  $\bar{F}(z)$  au décalage spectral  $z_{abs}$  de l'absorbeur HI :

$$\langle f_q(\lambda) \rangle = C_q(\lambda) \bar{F}(z_{abs}) \quad (5.7)$$

où le décalage spectral  $z_{abs}$  est relié à la longueur d'onde  $\lambda$  par la relation  $z_{abs} = \lambda/\lambda_{Ly\alpha} - 1$ . Notre méthode d'estimation du flux moyen suppose que le continu  $C_q$  de chaque quasar est, en première approximation, proportionnel à un continu universel  $C_u$ . Ce continu universel est uniquement fonction de la longueur d'onde dans le référentiel du quasar  $\lambda_{rq} = \lambda/(1 + z_q)$ , où  $z_q$  est le décalage spectral du quasar. La fraction de flux transmise moyenne  $\bar{F}(z_{abs})$  varie quant à elle avec

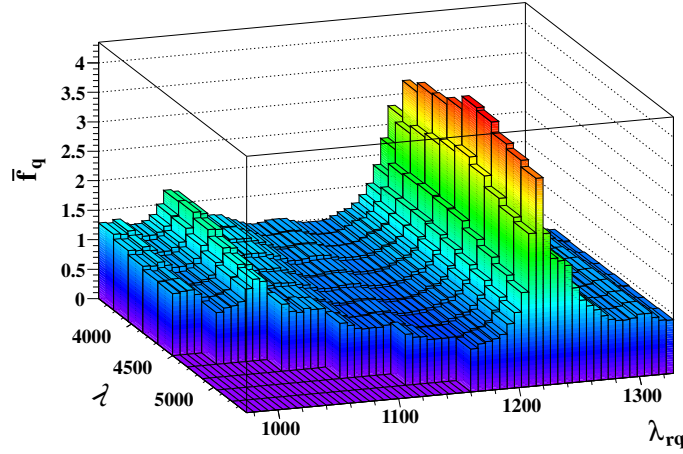


FIGURE 5.6: Moyenne pondérée du flux de chaque quasar en fonction de la longueur d'onde observée  $\lambda$  et de la longueur d'onde dans le référentiel du quasar  $\lambda_{rq}$ .

le décalage spectral de l'absorbeur. Il est possible d'estimer le produit  $C_u(\lambda_{rq})\bar{F}(\lambda)$  en moyennant le flux mesuré pour chaque pixel de chaque ligne de visée en fonction de la longueur d'onde mesurée  $\lambda$  et de la longueur d'onde dans le référentiel du quasar  $\lambda_{rq}$  :

$$\bar{f}(\lambda, \lambda_{rq}) = \frac{\sum_q w_q(\lambda) f_q(\lambda)}{\sum_q w_q(\lambda)} \quad (5.8)$$

où les poids  $w_q(\lambda)$  sont déterminés comme expliqué en 5.2.4. La moyenne  $\bar{f}(\lambda, \lambda_{rq})$  obtenue sur l'échantillon de quasars sélectionnés est présentée en figure 5.6. Sur cette figure, le pic Lyman- $\alpha$  est clairement visible en  $\lambda_{rq} = 1.215 \text{ \AA}$ . Le pic Lyman- $\beta$  est également visible en  $\lambda_{rq} = 1.025 \text{ \AA}$ . Le fait qu'il n'y ait pas d'information pour  $\lambda \gtrsim 4.500$  et  $\lambda_{rq} \lesssim 1.150$  vient de la coupure sur le décalage spectral des quasars  $z_q \leq 3,5$ .

Le produit  $C_q\bar{F}$  est ensuite estimé en ajustant à la forêt de chaque quasar la loi de puissance suivante :

$$C_q\bar{F} = a_q \left( \frac{\lambda}{\bar{\lambda}} \right)^{b_q} \bar{f}(\lambda, \lambda_{rq}) \quad (5.9)$$

où  $a_q$  et  $b_q$  sont les paramètres de l'ajustement et où  $\bar{\lambda}$  est la longueur d'onde observée moyenne de la forêt du quasar. Le paramètre  $a_q$  est appelé « normalisation » tandis que le paramètre  $b_q$  est appelé « paramètre de déformation ». Le rôle de ce dernier est de prendre en compte la diversité des indices spectraux des différents spectres de quasar, mais également un éventuel mauvais étalonnage spectroscopique de ces spectres. Un exemple d'ajustement pour l'un des spectres de quasar de l'échantillon est visible en figure 5.7. La figure 5.8 montre quant à elle les distributions des paramètres  $a_q$  et  $b_q$  pour l'ensemble de l'échantillon de quasars considéré. Les spectres ayant  $a_q < 0.1$  sont sortis de l'échantillon. Cette coupure concerne quelques centaines de spectres, parmi les plus bruyants, dont certains possèdent même un paramètre de normalisation négatif.

Notre procédure d'estimation du continu produit un effet indésirable mais inévitable lorsque l'estimation se base sur le flux dans la forêt du quasar. L'utilisation des données de la forêt pour

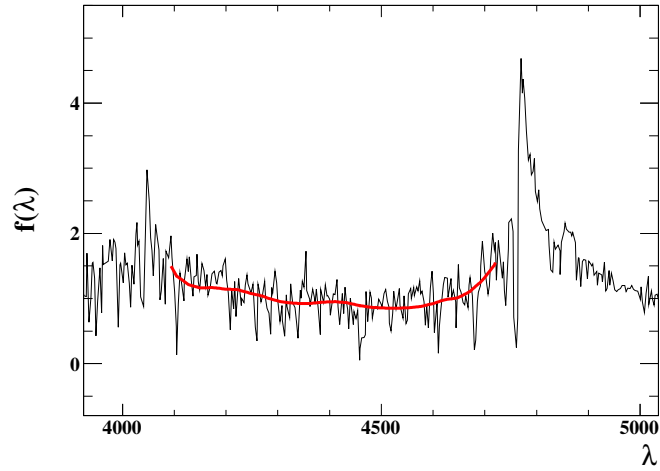


FIGURE 5.7: Exemple d'estimation du flux moyen attendu (ligne rouge) pour un spectre de quasar de l'échantillon (ligne noire). Les lignes d'émission Lyman- $\alpha$  et Lyman- $\beta$  sont clairement visibles. La forêt utilisée est définie par  $1.040 \text{ \AA} < \lambda_{rq} < 1.200 \text{ \AA}$ .

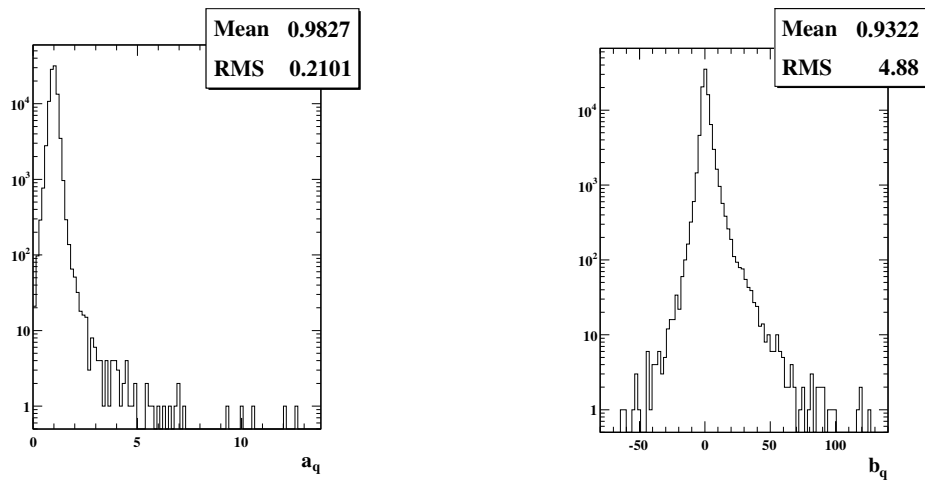


FIGURE 5.8: Histogrammes des paramètres de normalisation (à gauche) et de déformation (à droite) pour l'ensemble de l'échantillon de quasars considéré.



ajuster le continu force l'absorption moyenne de chaque quasar à être proche de celle de l'échantillon de quasar complet (mais pas rigoureusement égale puisque nous ajustons à la fois la normalisation  $a_q$  et la déformation  $b_q$ ). Cette approche crée une corrélation négative entre pixels d'un même quasar aux grandes échelles. Cette corrélation négative n'entre pas directement dans notre calcul de la fonction de corrélation puisque nous ne tenons pas compte des paires de pixels venant tous deux de la même ligne de visée d'un quasar. Elle entre tout de même au second ordre en introduisant une corrélation non physique, négative, entre les paires de pixels  $\delta_i^1 \delta_j^2$  et  $\delta_i^1 \delta_k^2$ , où  $\delta_i^1$  est un pixel de la ligne de visée d'un quasar et  $\delta_j^2$  et  $\delta_k^2$  sont deux pixels de la ligne de visée d'un second quasar. Heureusement, comme nous le verrons par l'étude de simulations en section 5.4, cet effet dépend de manière monotone de la séparation  $r$  et peut être simplement paramétrisé lors de l'ajustement de la fonction de corrélation sans avoir d'effet sur la détection et la mesure de la position du pic BAO.

Il est possible de contourner ce problème en estimant le continu du quasar dans la forêt à partir du flux en dehors de la forêt. Lee et al. (2012) ont développé une méthode basée sur la technique de décomposition en composantes principales utilisant le flux des quasars à plus grandes longueurs d'onde que le pic Lyman- $\alpha$ . Le principal problème d'une telle méthode est qu'elle introduit des systématiques liées à l'étalonnage relatif des CCDs bleus et rouges des spectrographes.

### 5.2.3 Champ de $\delta$

La figure 5.9 montre la moyenne des  $\delta_q(\lambda)$  obtenus en fonction de la longueur d'onde observée. La moyenne fluctue autour de zéro avec des déviations de l'ordre de 2% et pouvant aller jusque 5%. Plusieurs résidus de lignes d'émission et/ou d'absorption sont visibles, notamment les raies de la série de Balmer, dont les plus fortes sont la raie H $\gamma$  à 4.340 Å puis la raie H $\delta$  à 4.100 Å. Ces raies de Balmer sont des artefacts créés lors de l'étalonnage en flux des spectres. Pour la conversion du comptage CCD en flux, cet étalonnage utilise des spectres d'étoiles standards de type F, dont le flux est connu. Un mauvais masquage de la série de Balmer dans les spectres de ces étoiles standards est responsable des artefacts visibles dans les spectres de quasars. L'étalonnage étant réalisé lors de l'observation de chaque plaque et pour chaque spectrographe, ces artefacts sont constants pour la moitié des spectres d'une plaque.

Sans correction, les artefacts de la figure 5.9 produiraient des corrélations non-désirées, avec un effet dominant pour les paires de pixels de distance radiale  $r_{\parallel}$  nulle. Nous avons fait une correction globale en soustrayant la moyenne  $\langle \delta \rangle(\lambda)$  à chaque mesure individuelle  $\delta_q(\lambda)$ . Cette soustraction est justifiée dans la limite où la variance des artefacts est faible d'une « demi-plaque » à l'autre. L'étude des vecteurs d'étalonnage<sup>1</sup> montre que l'écart type des artefacts de Balmer d'une « demi-plaque » à l'autre est inférieur à 20% de la moyenne des artefacts, ce qui justifie notre correction globale.

La figure 5.10 montre la distribution du décalage spectral des paires de  $\delta$ . Le décalage spectral moyen de 2,38 est le décalage spectral auquel nous mesurons la fonction de corrélation et donc déterminons les contraintes cosmologiques.

### 5.2.4 Détermination des poids

McQuinn and White (2011) présentent une discussion sur la pondération optimale pour la fonction de corrélation de données Lyman- $\alpha$ . Dans cette analyse, nous choisissons les poids  $w_{ij}$  de sorte à approximativement minimiser l'erreur sur  $\hat{\xi}_A$  défini par l'équation 5.3. Insistons sur

1. Un vecteur d'étalonnage contient le spectre d'une étoile standard de type F dont le flux est connu et permet donc d'étalonner les spectres des autres objets mesurés par le spectrographe.

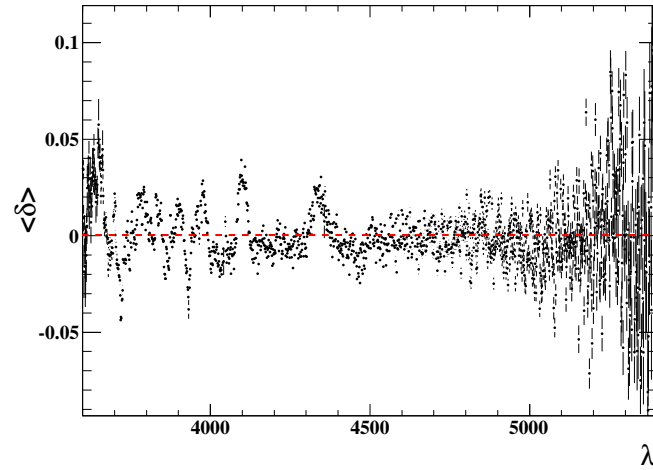


FIGURE 5.9: Moyenne des  $\delta_q$  en fonction de la longueur d'onde observée. Les raies de la série de Balmer sont visibles et notamment les raies H $\gamma$  à 4.340 Å et H $\delta$  à 4.100 Å.

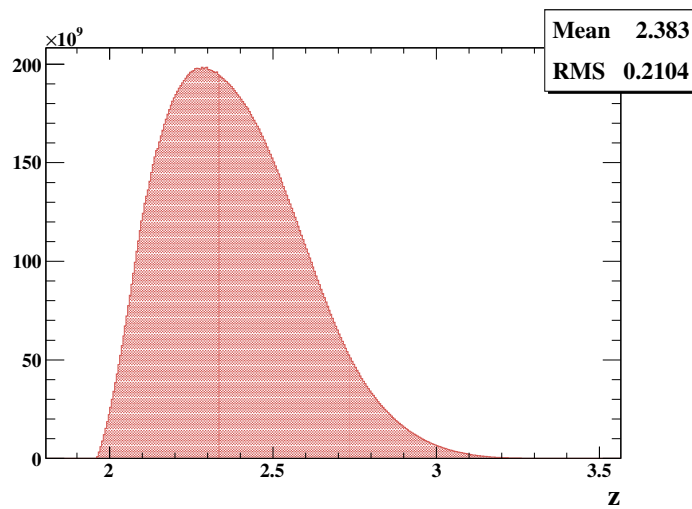


FIGURE 5.10: Distribution pondérée du décalage spectral de l'ensemble des paires de  $\delta$ . Le décalage spectral moyen des paires est de 2,38.

le fait que l'utilisation d'un poids sous-optimal réduit la précision de mesure de la fonction de corrélation, ce qui peut diminuer le niveau de confiance de la détection des BAO, mais n'introduit pas de biais sur la fonction de corrélation et a fortiori sur la position reconstruite du pic acoustique. Dans l'approximation de pixels décorrélés, la variance de  $\hat{\xi}_A$  est :

$$\text{var}(\hat{\xi}_A) = \frac{\sum_{i,j \in A} w_{ij}^2 \xi_{ii} \xi_{jj}}{\left( \sum_{i,j \in A} w_{ij} \right)^2} \quad (5.10)$$

où  $\xi_{ii} = \langle \delta_i^2 \rangle$  est la variance du pixel  $i$ , dépendant du bruit d'observation, mais également de la variance intrinsèque de la forêt Lyman- $\alpha$ . Le signal sur bruit de la fonction de corrélation est alors :

$$\left( \frac{S}{B} \right)^2 = \frac{\langle \hat{\xi}_A \rangle^2}{\text{var}(\hat{\xi}_A)} = \frac{\left( \sum_{i,j \in A} w_{ij} \xi_{ij} \right)^2}{\sum_{i,j \in A} w_{ij}^2 \xi_{ii} \xi_{jj}} \quad (5.11)$$

Du fait de l'évolution en décalage spectral de la variance intrinsèque et de la fraction de flux transmise moyenne, les  $\xi_{ij}$  dépendent également de ce décalage spectral. Nous modélisons cette dépendance par l'évolution en décalage spectral de la fonction de corrélation à une dimension mesurée par McDonald et al. (2006) :

$$\xi_{ij}(z) = \left( \frac{1+z_i}{1+z_0} \right)^{\frac{\gamma}{2}} \left( \frac{1+z_j}{1+z_0} \right)^{\frac{\gamma}{2}} \xi_{ij}(z_0) \quad (5.12)$$

où  $\gamma \approx 3.8$ . En optimisant le signal sur bruit nous obtenons :

$$w_{ij} \propto \frac{(1+z_i)^{\frac{\gamma}{2}} (1+z_j)^{\frac{\gamma}{2}}}{\xi_{ii}^2 \xi_{jj}^2} \quad (5.13)$$

Pour utiliser cette expression il est nécessaire d'estimer la variance  $\xi_{ii}$ . Pour se faire nous supposons qu'elle peut être vue comme la somme d'un terme de bruit d'observation, dont la chaîne d'analyse donne une estimation  $\sigma_{pipeline}^2$  et d'un terme de variance intrinsèque à la forêt  $\sigma_{LSS}^2$  :

$$\xi_{ii}^2 = \frac{\sigma_{pipeline,i}^2}{\eta(z_i)} + \sigma_{LSS}^2(z_i) \quad (5.14)$$

où  $z_i = \lambda_i / \lambda_{Ly\alpha} - 1$  est le décalage spectral de l'absorbeur et  $\eta$  est un terme permettant de corriger un éventuel biais dans l'estimation du bruit par la chaîne d'analyse. Le terme  $\sigma_{LSS}^2$  représente la variance physique de la densité d'hydrogène à un décalage spectral donné. Elle permet de définir un poids maximum à attribuer aux pixels, dans le cas d'une mesure sans bruit. Les données sont organisées en différentes classes de décalage spectral et de  $\sigma_{pipeline}^2$ , comme visible en figure 5.11. Pour chaque classe,  $\xi_{ii}^2$  est estimé par la variance des  $\delta_i = \delta(\lambda_i)$ . Les valeurs de  $\sigma_{LSS}^2(z)$  et de  $\eta(z)$  sont alors déterminées en ajustant une droite aux données pour le décalage spectral considéré. Les poids de l'équation 5.8 sont définis par :

$$w_q(\lambda_i) \propto \frac{(\lambda_i / \lambda_{Ly\alpha})^{\frac{\gamma}{2}}}{\xi_i^2} \quad (5.15)$$

Les ajustements sont visibles en figure 5.11. Le premier graphique montre que la variance mesurée dépend linéairement de  $\sigma_{pipeline}^2$  avec un coefficient de proportionnalité égal à  $1/\eta$  et une ordonnée à l'origine égale à  $\sigma_{LSS}^2$ . Les deux graphiques suivant montrent l'évolution de  $\sigma_{LSS}^2$  et de  $\eta$  avec le décalage spectral. Ces évolutions sont ajustées par des polynômes du second degré. Pour  $z > 3$ , il n'y a plus suffisamment de paires pour estimer correctement  $\eta(z)$ . Ce dernier est alors estimé par extrapolation du polynôme ajusté à plus bas décalage spectral puis fixé à 1 lorsque le polynôme dépasse 1. La correction  $\eta(z)$  trouvée est en bon accord avec l'étude similaire, mais plus poussée de la correction du bruit estimé par le *pipeline*, menée pour l'analyse du spectre de puissance à une dimension des forêts Lyman- $\alpha$  (Palanque-Delabrouille, 2013, in prep.).

### 5.2.5 $\hat{\xi}(r_{\parallel}, r_{\perp})$

La figure 5.12 présente la fonction de corrélation obtenue avec les données DR10 $\alpha$  en deux dimensions, dans les coordonnées  $(r_{\parallel}, r_{\perp})$ . L'anisotropie produite par les distortions en *redshift* est clairement visible. Ces distortions en *redshift* ont déjà été détectées à plus de  $5\sigma$  dans les données DR9 par Slosar et al. (2011). Bien que ce travail n'ait pas pour but de mesurer de manière précise le paramètre  $\beta$  (défini en équation 2.31) quantifiant l'anisotropie, nous souhaitons insister sur son interprétation physique : Il s'agit de l'effondrement des sur-densités d'hydrogène sur elles-mêmes, étape préliminaire à la formation des structures.

### 5.2.6 $\hat{\xi}(r, \mu)$

Nous calculons la fonction de corrélation  $\hat{\xi}(r, \mu)$  en classes de  $r$  de  $4 h^{-1}$  Mpc de large (centrées en 2 ; 6 ; ... ; 198  $h^{-1}$  Mpc) et en classes de  $\mu$  de 0,02 (centrées en 0,01 ; 0,03 ; ... ; 0,99). La figure 5.13 présente la fonction de corrélation  $\hat{\xi}(r, \mu)$  obtenue moyennée en trois classes de  $\mu$ . Le pic acoustique est principalement visible autour de  $r \sim 105 h^{-1}$  Mpc pour  $0,8 < \mu < 1$ , soit pour un angle inférieur à 37 degrés de la ligne de visée. Il est également visible pour  $0,5 < \mu < 0,8$ . Les courbes correspondent à un ajustement pour le modèle  $\Lambda$ CDM d'après la procédure décrite en section 5.5.3 Le fait que le pic acoustique soit principalement visible à grand  $\mu$  nous laisse penser que nous serons plus sensible au paramètre de Hubble  $H(z)$  qu'à la distance angulaire  $D_A(z)$ .

## 5.3 Etude en multipôles

Cette section présente la décomposition de la fonction de corrélation à deux dimensions en harmoniques sphériques, suivant la décomposition standard proposée par Hamilton (1992). Elle détaille également le calcul de la matrice de covariance des multipôles obtenus.

### 5.3.1 Détermination des multipôles

La décomposition en harmoniques sphériques consiste à projeter la fonction de corrélation à deux dimensions  $\hat{\xi}(r, \mu)$  sur la suite de polynômes de Legendre  $P_l(\mu)$ , où  $l \in \mathbb{N}_+$  et :

$$\begin{aligned}
 P_0(\mu) &= 1 \\
 P_1(\mu) &= \mu \\
 P_2(\mu) &= \frac{3\mu^2 - 1}{2} \\
 P_3(\mu) &= \frac{5\mu^3 - 3\mu}{2} \\
 P_4(\mu) &= \frac{35\mu^4 - 30\mu^2 + 3}{8} \\
 &\dots
 \end{aligned} \tag{5.16}$$

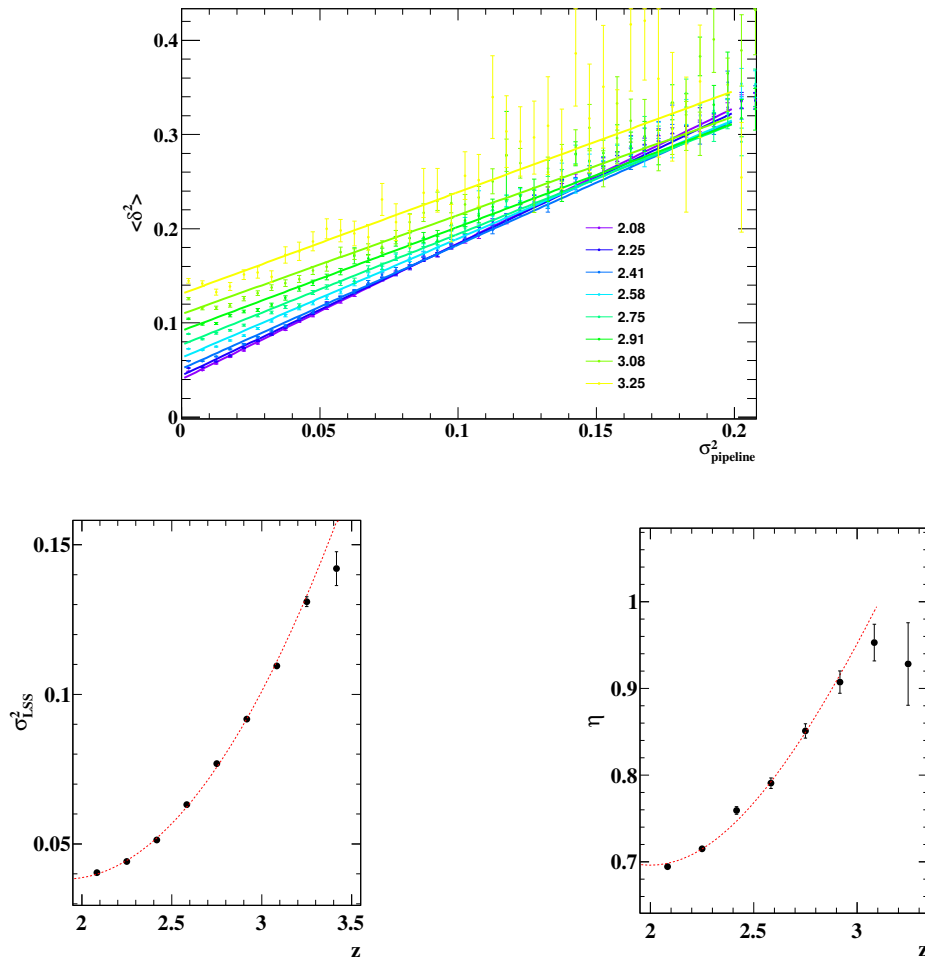


FIGURE 5.11: En haut : variance des  $\delta$  en fonction de la variance estimée par la chaîne d'analyse pour différentes gammes de décalage spectral. Les points de mesure de chaque gamme de décalage spectral sont ajustés d'après l'équation 5.14. En bas : évolution des paramètres de l'ajustement (la variance intrinsèque  $\sigma_{LSS}^2$  à gauche et le facteur de correction  $\eta$  à droite) en fonction du décalage spectral.

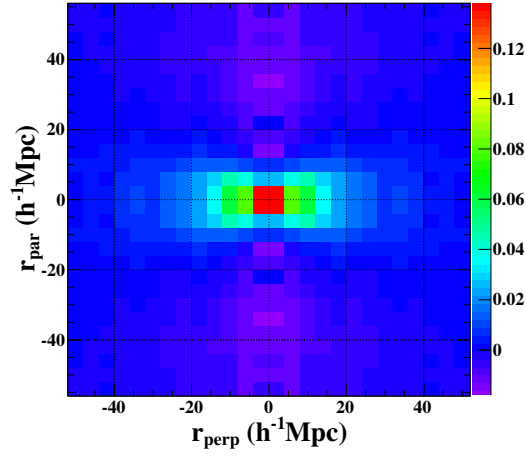


FIGURE 5.12: Fonction de corrélation  $\hat{\xi}$  en fonction de la séparation perpendiculaire à la ligne de visée  $r_{perp}$  et de la séparation le long de la ligne de visée  $r_{par}$ . L'anisotropie produite par les distortions de *redshift* est clairement visible. Les parties du graphique à  $r_{perp} < 0$  et/ou  $r_{par} < 0$  sont obtenues en répliquant la partie à  $r_{perp} > 0$  et  $r_{par} > 0$ .

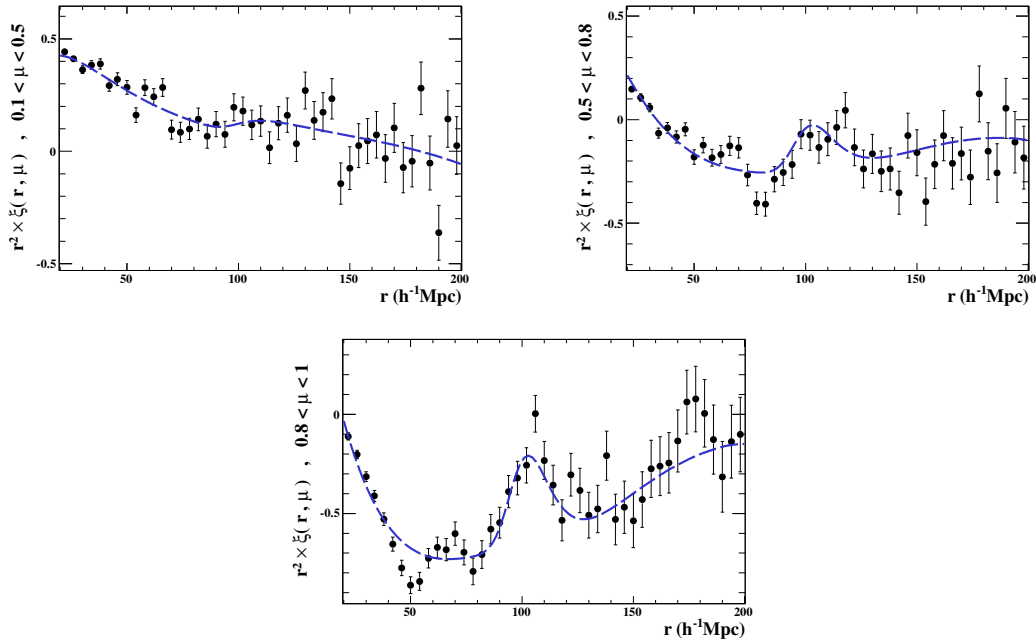


FIGURE 5.13: Moyenne de la fonction de corrélation sur trois intervalles en  $\mu$  :  $0,1 < \mu < 0,5$  (en haut à gauche),  $0,5 < \mu < 0,8$  (en haut à droite) et  $0,8 < \mu < 1$  (en bas). Les courbes résultent de l'ajustement isotrope décrit en section 5.5.3. Le pic acoustique est principalement visible pour  $\mu > 0,8$  mais apparaît également pour  $0,5 < \mu < 0,8$ .

Les polynômes de Legendre sont orthogonaux entre eux :

$$\langle P_i, P_j \rangle = \int_{-1}^1 P_i(\mu) P_j(\mu) d\mu = 0 \quad (5.17)$$

Le pôle  $\xi_l(r)$  est donc défini comme la projection de  $\hat{\xi}(r, \mu)$  sur le polynôme de degré  $l$  (à un facteur près correspondant à la norme du polynôme de Legendre).  $\hat{\xi}(r, \mu)$  est une fonction paire par rapport à  $\mu$  (nous n'utilisons d'ailleurs pas le signe de  $\mu$  dans notre analyse), sa projection sur les polynômes de degré impair est donc nulle. La décomposition de  $\hat{\xi}(r, \mu)$  en multipôles peut donc s'écrire :

$$\hat{\xi}(r, \mu) = \sum_{l=0,2,4,\dots} \xi_l(r) P_l(\mu) \quad (5.18)$$

Kaiser (1987) a montré que seuls les monopôle ( $l = 0$ ), quadrupôle ( $l = 2$ ) et hexadécapôle ( $l = 4$ ) sont générés par les distorsions en *redshift* dans l'approximation linéaire. Tout terme d'ordre supérieur ( $l \geq 6$ ) est donc le résultat de la méthode d'analyse ou d'effets non-linéaires. Nous négligeons ces termes dans la suite. Nous négligeons également le terme  $l = 4$  à la fois faible et mal contraint (une détermination de l'hexadécapôle est visible en figure 5.16, étant parfaitement compatible avec 0, il n'est plus ajusté par la suite) si bien que la décomposition se simplifie sous la forme :

$$\hat{\xi}(r, \mu) = \left[ \xi_0(r) - \frac{1}{2} \xi_2(r) \right] + \left[ \frac{3}{2} \xi_2(r) \right] \mu^2 \quad (5.19)$$

Les monopôles ( $l = 0$ ) et quadrupôles ( $l = 2$ ) peuvent donc être obtenus pour chaque valeur de  $r$ , en ajustant  $\hat{\xi}(r, \mu^2)$  par une droite. Cet ajustement est effectué par une simple minimisation de  $\chi^2$ , en utilisant la variance définie en équation 5.10 et en négligeant les corrélations entre les points de mesure. Cette méthode rend la détermination de  $\xi_0(r)$  et  $\xi_2(r)$  légèrement sous-optimale. Nous excluons de l'ajustement la partie  $\mu < 0.1$  (transverse à la ligne de visée) pour éviter un biais résiduel dû à la soustraction du fond du ciel, identique pour chaque spectre d'une « demie-plaque », ou à la présence de raies résiduelles, telles que les raies de Balmer. Cette coupure n'enlève que peu de statistiques. L'un de ces ajustements, pour  $r = 34 h^{-1} \text{Mpc}$ , est visible en figure 5.14.

La figure 5.15 montre les monopôle et quadrupôle obtenus. Le pic attendu est visible à la fois dans le monopôle et le quadrupôle. Cette figure montre également la combinaison  $\xi_0 + 0.3\xi_2$  qui, du fait de l'anti-corrélation du monopôle et du quadrupôle (voir section 5.3.2), est une quantité mieux contrainte. Le pic est encore plus évident sur cette combinaison.

### 5.3.2 Estimation de la covariance

Les différents points de mesure de  $\hat{\xi}(r, \mu)$  sont corrélés entre eux car les  $\delta_i$  intervenant dans leur calcul (voir équation 5.3) ne sont eux mêmes pas indépendants. La corrélation des  $\delta_i$  provient à la fois des structures à grandes échelles de l'Univers (c'est ce que nous souhaitons mesurer) et de corrélations induites par des effets instrumentaux ou par l'estimation du continu. Pour calculer le niveau de confiance de la détection du pic, ainsi que pour en tirer des contraintes cosmologiques, il est nécessaire d'estimer la matrice de covariance.

Plutôt que d'estimer la matrice de covariance de  $\hat{\xi}(r, \mu)$  de dimension  $2.500 \times 2.500$  (la fonction de corrélation est estimée sur 50 classes de  $\mu$  et 50 classes de  $r$  donc 2.500 classes au total), nous estimons directement la matrice de covariance du monopôle et du quadrupôle de dimension  $100 \times 100$  (50 classes en  $r$  pour le monopôle et 50 pour le quadrupôle) dont le calcul est plus aisé. Nous utilisons deux techniques standards d'estimation de la covariance, basées sur le découpage du relevé en  $N$  sous échantillons en fonction de la position sur le ciel. La première technique

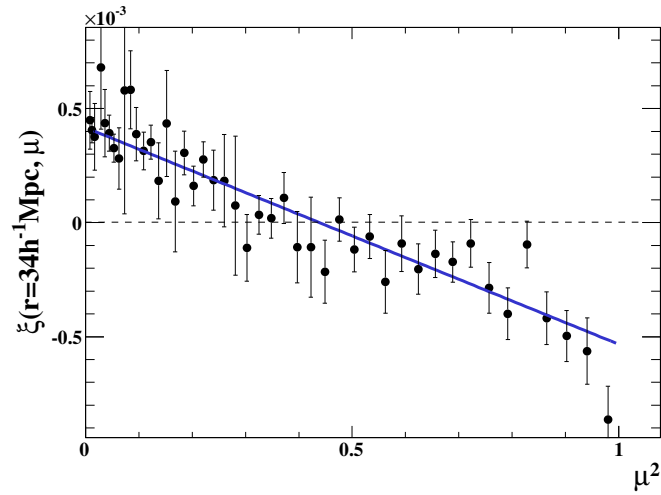


FIGURE 5.14:  $\hat{\xi}(r, \mu)$  en fonction de  $\mu^2$  pour  $r = 34 h^{-1} \text{Mpc}$ . La droite bleue est ajustée par minimisation du  $\chi^2$ . Les points de mesure en  $\mu^2 < 0.01$  ne sont pas pris en compte pour l'ajustement.

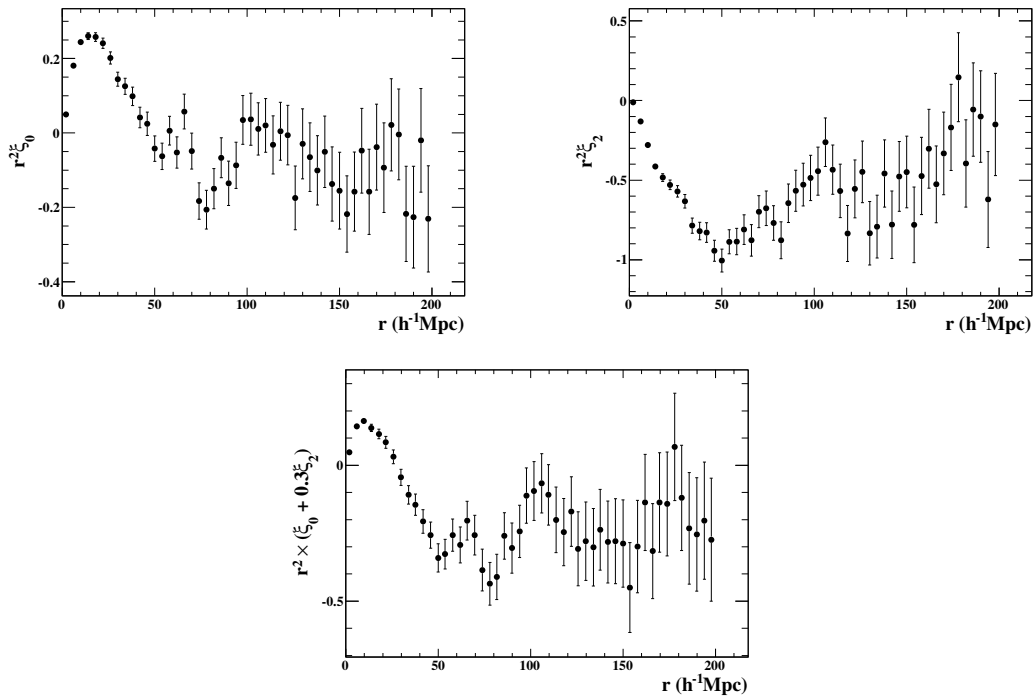


FIGURE 5.15: Monopôle (en haut à gauche) et quadrupôle (en haut à droite) de la fonction de corrélation  $\hat{\xi}(r, \mu)$ . Les erreurs correspondent à la racine carré des éléments diagonaux de la matrice de covariance obtenue par la technique du *bootstrap* (voir section 5.3.2). Le graphique du bas montre la combinaison  $\xi_0 + 0.3\xi_2$  qui, du fait de l'anti-corrélation du monopôle et du quadrupôle, est une quantité mieux contrainte.



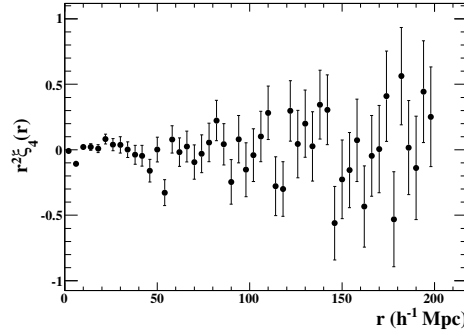


FIGURE 5.16: Estimation de l'hexadécapôle de la fonction de corrélation. N'étant pas contraint, il n'est plus ajusté par la suite.

consiste à estimer un monopôle et un quadrupôle par échantillon puis de calculer directement la covariance de ces  $N$  monopôles et quadrupôles. La deuxième technique, appelée *bootstrap* (Efron and Gong, 1983), consiste à remplacer les  $N$  sous-échantillons formant l'intégralité des données, par un jeu de  $N$  sous-échantillons tirés aléatoirement parmi les données et avec remplacement. Ce nouvel échantillon aléatoire est appelé échantillon de *bootstrap*. La covariance des données est alors estimée par la covariance d'un grand nombre d'échantillons de *bootstrap*. Ces deux techniques sont décrites en détails dans la suite de cette section. Nous commençons par introduire les sous-échantillons sur lesquels elles sont toutes les deux basées.

### Définitions des sous-échantillons

Les sous-échantillons sont définis par les plaques utilisées pour maintenir les fibres dans le plan focal. Pour éviter que les sous-échantillons soient trop corrélés, chaque paire de quasars n'appartient qu'à un seul sous-échantillon. Dans le cas où les deux quasars d'une paire appartiennent à une plaque différente, la paire est associée à la plaque du quasar possédant l'ascension droite la plus faible. Ce découpage par plaque permet de définir  $N_s = 1.576$  sous-échantillons, soit autant que de plaques observées. La surface couverte par les sous-échantillons est variable mais vaut en moyenne  $4 \text{ deg}^2$ .

### Variance des sous-échantillons

Pour cette première estimation de la covariance du monopôle et du quadrupôle, nous calculons pour chaque sous-échantillon  $s$  la fonction de corrélation  $\hat{\xi}_s(r, \mu)$  d'après la formule 5.3. Il est alors possible d'estimer pour chaque sous-échantillon  $s$  un monopôle  $\xi_{0,s}$  et un quadrupôle  $\xi_{2,s}$ . La covariance entre les différents multipôles des différents sous-échantillons est alors donnée par :

$$\text{cov}[\xi_l^s(r), \xi_{l'}^s(r')] = \frac{\sum_{i=0}^{N_s} \sqrt{w_i(r)w_i(r')} [\xi_l^{s_i}(r) - \bar{\xi}_l^s(r)] [\xi_{l'}^{s_i}(r') - \bar{\xi}_{l'}^s(r')]}{\sum_{i=0}^{N_s} \sqrt{w_i(r)w_i(r')}} \quad (5.20)$$

où  $(l, l') \in (0, 2)^2$  et l'exposant  $s$  rappelle qu'il s'agit de multipôles estimés sur des sous-échantillons. Le poids  $w_i(r)$  du sous-échantillon  $i$  est proportionnel au nombre de paires de  $\delta$  utilisées pour l'es-

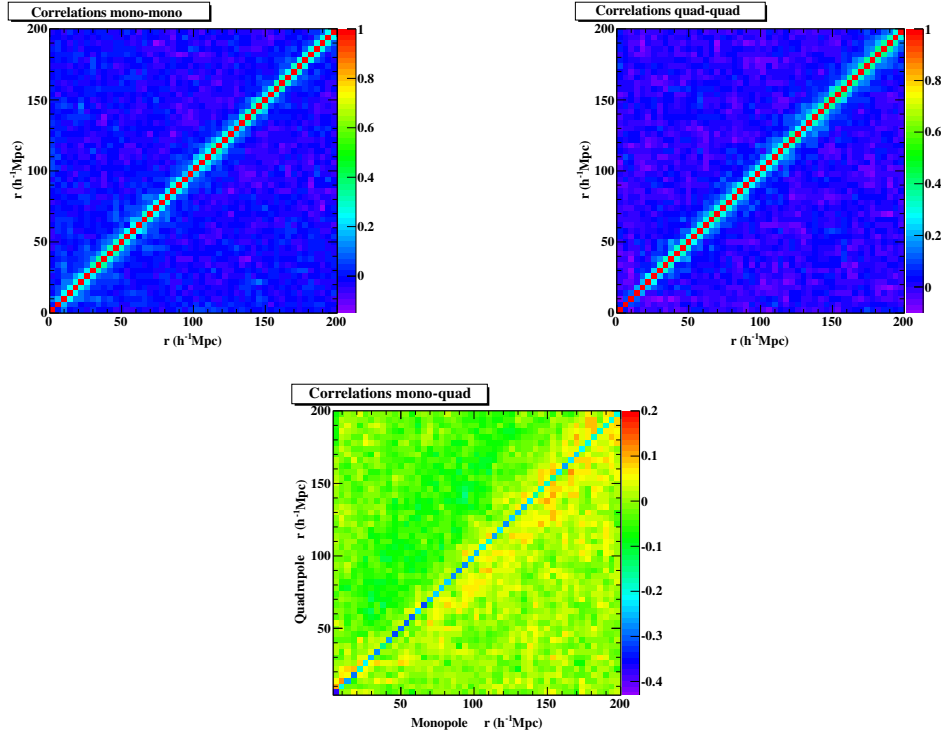


FIGURE 5.17: Matrices de corrélation du monopôle et du quadrupôle calculées par la variance des sous-échantillons. Les éléments des matrices monopôle-monopôle et quadrupôle-quadrupôle sont normalisés par la variance  $\text{Corr}_{ij} = C_{ij} / \sqrt{C_{ii}C_{jj}}$ , où  $C_{ij}$  représente la matrice de covariance. Les éléments de la matrice monopôle-quadrupôle sont normalisés par le produit des variances du monopôle et du quadrupôle :  $\text{Corr}_{ij} = C_{ij}^{mq} / \sqrt{C_{ii}^m C_{jj}^q}$ . Les premiers éléments hors diagonale de la matrice monopôle-monopôle sont de l'ordre de 25% tandis que ceux de la matrice quadrupôle-quadrupôle sont de l'ordre de 30%.

estimation de  $\xi_l^{s_i}(r)$  et :

$$\bar{\xi}_l^s(r) = \frac{\sum_{i=0}^{N_s} w_i(r) \xi_l^{s_i}(r)}{\sum_{i=0}^{N_s} w_i(r)} \quad (5.21)$$

L'estimateur de la covariance du monopôle et quadrupôle de l'échantillon complet est alors obtenu par normalisation de la covariance des sous-échantillons :

$$\text{cov}[\xi_l(r), \xi_{l'}(r')] = \frac{\text{cov}[\xi_l^s(r), \xi_{l'}^s(r')]}{N_s - 1} \quad (5.22)$$

Les matrices de covariance obtenues avec cet estimateur sont visibles en figure 5.17. La faiblesse de cet estimateur est qu'il nécessite l'estimation d'un monopôle et d'un quadrupôle pour chaque sous-échantillon. Le nombre de quasar appartenant à un sous-échantillon pouvant être dans certains cas de quelques dizaines, voir moins, la qualité de l'estimation du monopôle et

du quadrupôle n'est alors pas garantie (notamment à bas  $r$ , où l'estimation nécessite des paires de quasars proches).

### **Bootstrap**

Afin d'obtenir une estimation plus robuste de la covariance du monopôle et du quadrupôle de l'échantillon complet, nous avons utilisé une autre approche basée sur la technique de *bootstrap*. L'idée est de tirer aléatoirement et avec remplacement  $N_s$  sous-échantillons parmi les  $N_s$  disponibles. Un tel tirage est appelé échantillon de *bootstrap*. Chaque échantillon de *bootstrap* contient approximativement le même nombre de lignes de visée de quasar que l'échantillon de données complet. La fonction de corrélation de l'échantillon de *bootstrap*  $\hat{\xi}_{BS}(r, \mu)$  est obtenue en moyennant les fonctions de corrélation de chaque sous-échantillon d'après une formule équivalente à l'équation 5.21.  $\hat{\xi}_{BS}(r, \mu)$  est alors séparée en un monopôle  $\xi_0^{BS}(r)$  et un quadrupôle  $\xi_2^{BS}(r)$  d'après la procédure décrite en section 5.3.1. La covariance de l'échantillon de données est ensuite estimée en calculant la covariance de  $N_{BS} = 10.000$  échantillons de *bootstrap* :

$$\text{cov}[\xi_l(r), \xi_{l'}(r')] = \frac{\sum_{i=0}^{N_{BS}} \sqrt{w_i(r)w_i(r')} [\xi_{l,BS}^i(r) - \overline{\xi_{l,BS}}(r)] [\xi_{l',BS}^i(r') - \overline{\xi_{l',BS}}(r')]}{\sum_{i=0}^{N_{BS}} \sqrt{w_i(r)w_i(r')}} \quad (5.23)$$

où  $w_i(r)$  est l'inverse de l'erreur au carré de  $\xi_{l,BS}^i(r)$  obtenue par la procédure d'ajustement :

$$w_i(r) = \frac{1}{\sigma^2(\xi_{l,BS}^i(r))} \quad (5.24)$$

Les matrices de covariances obtenues par cette technique sont visibles en figure 5.18. Elles sont en très bon accord avec les matrices estimées par la variance des sous-échantillons ce qui donne confiance dans les deux estimations. La figure 5.19 montre la différence entre les matrices de covariance monopôles-monopôles obtenues par les deux méthodes. Les désaccords sont en moyenne de 2% et au maximum de 8%. Notons tout de même que nos sous-échantillons sont partiellement corrélés (puisque un quasar peut entrer dans plusieurs échantillons, via des paires de quasars différentes) tandis que les deux méthodes d'estimation de la covariance que nous utilisons supposent des sous-échantillons décorrélés. Il se peut donc que nos estimations de la covariance soient légèrement sous-estimées.

La technique du *bootstrap* permet de calculer la covariance sur un plus grand nombre d'échantillons que le calcul direct de la variance sur les sous-échantillons. Elle permet également une évaluation plus fiable des multipôles, car basée sur un plus grand nombre de lignes de visée. Nous utiliserons donc la covariance obtenue par cette technique pour la détermination du niveau de confiance de détection ainsi que pour les ajustements cosmologiques.

## **5.4 Comparaison aux simulations**

Certaines études visant à mesurer la fonction de corrélation des galaxies (voir par exemple Anderson et al. (2012)) estiment la matrice de covariance par des simulations. Il est alors crucial d'avoir des simulations réalistes, ce qui, pour l'analyse Lyman- $\alpha$ , nécessiterait d'avoir recours à des simulations hydrodynamiques avec un volume suffisamment grand pour couvrir l'intégralité

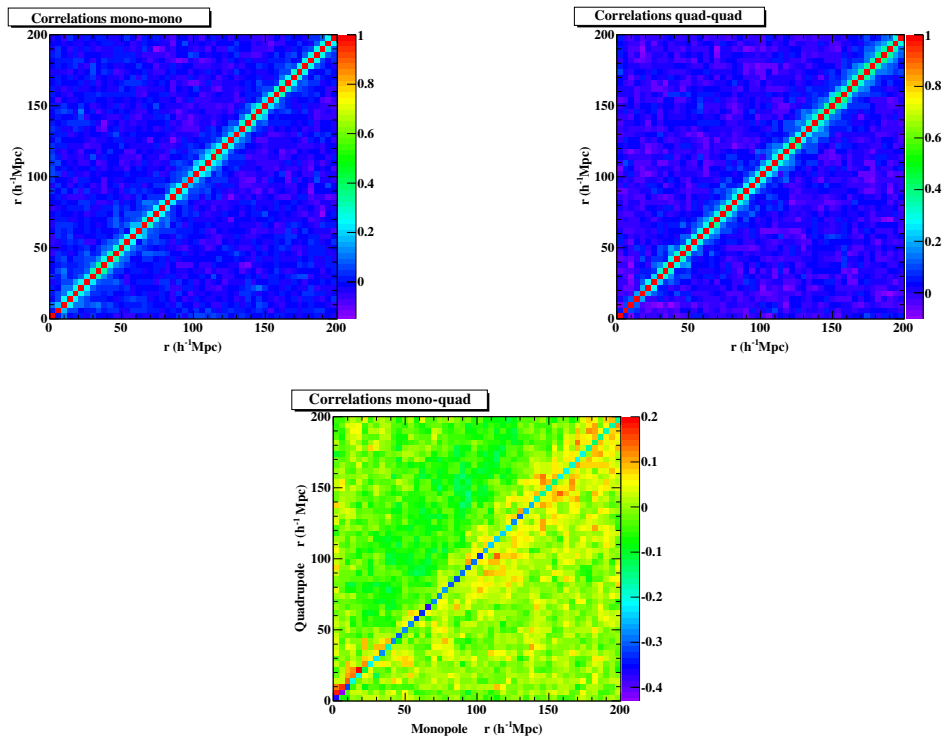


FIGURE 5.18: Matrices de corrélation du monopôle et du quadrupôle obtenues par la technique du *bootstrap*. Les résultats sont très proches de ceux obtenus par calcul direct de la variance des sous-échantillons

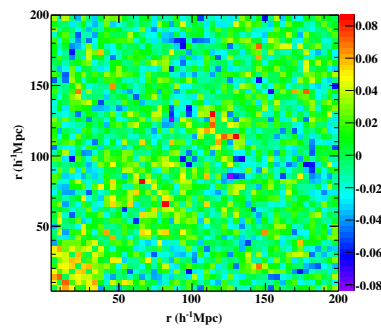


FIGURE 5.19: Différence des matrices de corrélation monopôle-monopôle estimées par la variance des sous-échantillons et par la technique du *bootstrap*. L'écart est en moyenne de 2% et au maximum de 8%.

du relevé (plusieurs Gpc<sup>3</sup>) mais également une résolution suffisante pour distinguer la longueur de Jeans du gaz (quelques dizaines de kpc). Avec les technologies actuelles, il n'est pas possible de produire une telle simulation, c'est pourquoi nous estimons la matrice de covariance directement des données.

Ces dernières années, différentes méthodes ont toutefois été proposées pour générer des simulations simplifiées de relevés Lyman- $\alpha$  en appliquant une transformation non-linéaire à un champ gaussien (Le Goff et al., 2011; Font-Ribera et al., 2012). Pour cette étude, nous disposons de 25 simulations générées d'après la procédure décrite par Font-Ribera et al. (2012). Nous utilisons ces simulations essentiellement pour développer ainsi que pour tester l'analyse. Elles nous permettent également de contrôler la cohérence des résultats sur les données.

#### 5.4.1 Production des simulations

Nous décrivons ici les grandes étapes de la production des simulations. Pour une description détaillée nous renvoyons le lecteur à la publication de Font-Ribera et al. (2012). Nos simulations possèdent la même géométrie que les données DR10 dans le sens où les positions en ascension droite et en déclinaison des quasars, ainsi que leurs décalages spectraux, sont identiques à ceux des données. Les spectres non-absorbés des quasars sont générés à partir des vecteurs propres d'une décomposition en composantes principales de 50 spectres à hautes résolutions obtenus par le télescope spatial Hubble (HST) (Suzuki et al., 2005). Les amplitudes de chaque vecteur propre de la décomposition sont tirées aléatoirement de distributions gaussiennes dont les variances sont données par Suzuki (2006). Une fois générés, les spectres sont normalisés de sorte à reproduire la magnitude dans la bande  $g$  des spectres des données.

Font-Ribera et al. (2012) fournissent ensuite un algorithme permettant de générer les spectres de fraction de flux transmise  $F(\lambda)$  en partant d'un champ aléatoire gaussien  $g(\lambda)$ . Plus précisément, ils donnent une recette pour choisir les paramètres  $a$  et  $b$  ainsi que le spectre de puissance  $P_g(k)$  du champ  $g(\lambda)$  de sorte que la transformation  $F(\lambda) = \exp[-a \exp(bg(\lambda))]$  donne la valeur moyenne de  $F(\lambda)$  ainsi que le spectre de puissance désiré. Le spectre de puissance recherché est donné par (McDonald, 2003) :

$$P_F(k, \mu_k) = b^2(1 + \beta\mu^2)^2 P_L(k, \mu_k) D(k, \mu_k) \quad (5.25)$$

où  $\mu_k$  est le cosinus de l'angle formé par la ligne de visée et le vecteur d'onde  $\mathbf{k}$ ,  $b$  est le biais de  $F(\lambda)$ ,  $\beta$  le paramètre de distortion en *redshift* et  $P_L(k)$  est le spectre de puissance linéaire du modèle  $\Lambda$ CDM avec les paramètres fiduciels donnés en équation 5.6.  $D(k, \mu_k)$  est un terme paramétrisant les non-linéarités du spectre de puissance, tendant vers 1 à petit  $k$ . Le biais  $b$  dépend du décalage spectral d'après la formule de McDonald et al. (2006) :

$$b = 0,14 \left( \frac{1+z}{3,25} \right)^{1,9} \quad (5.26)$$

tandis que les distortions en *redshift* sont fixées à la valeur  $\beta = 1,4$ . La dépendance en décalage spectral est introduite en générant des spectres de fraction de flux transmise pour douze décalages spectraux différents. Pour chaque pixel de longueur d'onde  $\lambda$ ,  $F(\lambda)$  est alors obtenu par interpolation entre les spectres à différents décalages spectraux en tenant compte du décalage spectral du pixel.

Le spectre final du quasar est alors obtenu en multipliant le continu non-absorbé par  $F(\lambda)$  puis par convolution avec la résolution du spectrographe. La dernière étape consiste à ajouter le bruit en tenant compte de l'efficacité optique des spectrographes de BOSS et de mesures du fond du ciel, comme décrit par Le Goff et al. (2011), puis à regrouper les pixels des spectres pour qu'ils aient la taille des pixels des spectres des données.

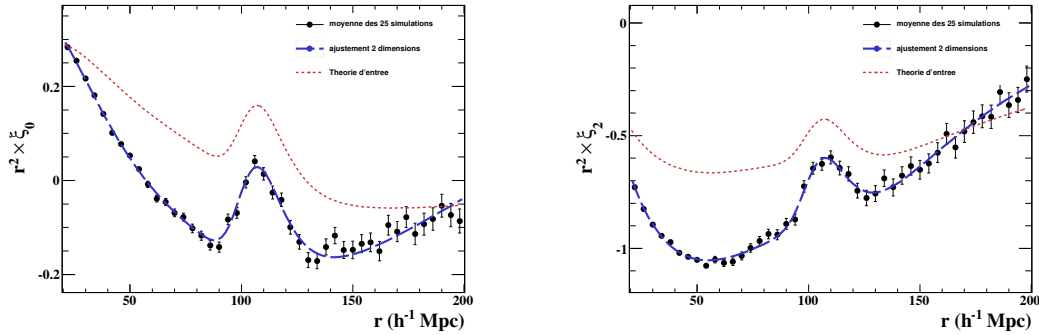


FIGURE 5.20: Moyennes des monopôles et quadrupôles obtenus sur les 25 simulations. Les courbes bleues correspondent à un ajustement à deux dimensions tel que décrit en section 5.5.4. Les courbes rouges correspondent aux multipôles théoriques utilisés pour générer les simulations.

#### 5.4.2 Analyse des simulations

Nous analysons les 25 simulations de la même façon que nous analysons les données à l'exception des BAL, qui ne sont pas inclus dans les simulations, et des systèmes DLA, puisque nous ne disposons pas d'inspection visuelle pour les simulations. Par simplicité, nous avons décidé d'ignorer la présence des systèmes DLA dans les simulations. Cette décision est temporaire, à terme nous souhaitons utiliser une détection automatique à la fois sur les données et sur les simulations. Cette procédure devrait être en place pour la prochaine publication sur le catalogue de données DR11.

La figure 5.20 montre le monopôle et le quadrupôle obtenus en moyennant les résultats sur les 25 simulations. Également présentés sur cette figure se trouvent les ajustements obtenus d'après la procédure décrite en 5.5.4 et les monopôle et quadrupôle théoriques, utilisés pour générer les simulations. Comme décrit en section 5.2.2, notre procédure d'estimation du flux moyen entraîne une distorsion de la fonction de corrélation mesurée par rapport à la fonction de corrélation réelle. Cette distorsion est clairement visible en figure 5.20 en comparant les monopôles et quadrupôles calculés par rapport aux monopôles et quadrupôles théoriques. Comme attendu, le monopôle calculé devient négatif plus tôt, aux alentours de  $60 h^{-1}\text{Mpc}$ , que le monopôle théorique qui reste positif jusque  $r \sim 130 h^{-1}\text{Mpc}$ . Cette distorsion n'a toutefois pas d'effet apparent sur la position du pic acoustique. Cette affirmation qualitative sera confirmée de manière quantitative en section 5.5.2 lors de l'ajustement des multipôles obtenus sur les simulations.

La figure 5.21 compare les multipôles obtenus sur les données et la moyenne des multipôles calculés sur les simulations. Cette figure montre que les simulations ne sont pas en parfait accord avec les données, en particulier pour le monopôle à bas  $r$ . Ce désaccord peut avoir plusieurs origines ; parmi elles figurent :

- Le traitement des systèmes DLA n'est pas identique pour les simulations et les données. Nous savons de plus que la distribution des systèmes DLA dans les simulations n'est pas représentative des données.
- Le biais  $b_\delta$  ainsi que sont évolution en décalage spectral et le paramètre des distorsions en *redshift* choisis pour nos simulations ne sont sans doute pas optimaux.

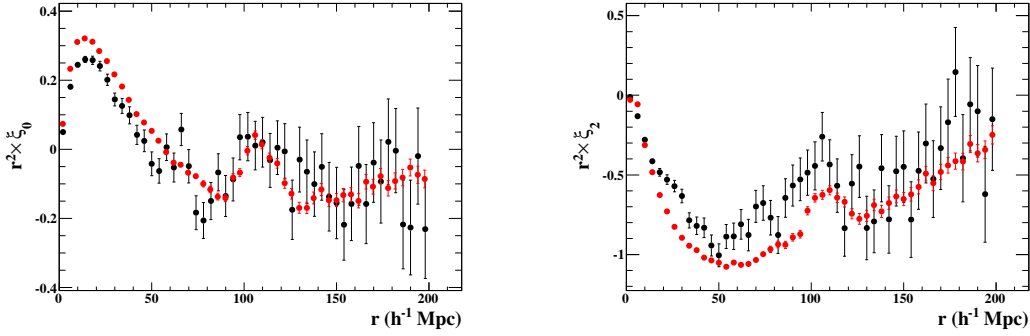


FIGURE 5.21: Moyennes des monopôles et quadrupôles obtenus sur les 25 simulations (en rouge) comparées aux monopôle et quadrupôle calculés sur les données. Les simulations et les données montrent un désaccord, principalement à bas  $r$ .

## 5.5 Détection et mesure de la position du pic

Cette section est dédiée à l'estimation du niveau de confiance de la détection du pic acoustique, ainsi qu'à la mesure de sa position. Pour déterminer le niveau de confiance de la détection, il est nécessaire de comparer la qualité de l'ajustement d'une fonction théorique « avec un pic » à celui d'une fonction « sans pic »<sup>2</sup>. Une manière standard de procéder serait de choisir une statistique test, telle que le  $\chi^2$ , de calculer les distributions de cette estimateur de la qualité de l'ajustement sur un grand nombre de simulations sans pic et avec pic, puis de vérifier l'accord entre les données et ces distributions. Puisque la production de simulations requiert d'importantes ressources informatiques, et que seules 25 simulations de DR10 étaient disponibles au moment de cette analyse, nous choisissons de simplement comparer les  $\chi^2$  d'un ajustement avec pic et d'un ajustement sans pic. L'ajustement avec pic est également celui qui permet la mesure de la position du pic BAO.

### 5.5.1 Paramétrisation de l'ajustement

La position du pic acoustique dans le monopôle  $\xi_0(r)$  et le quadrupôle  $\xi_2(r)$  des données est déterminée par deux cosmologies : la cosmologie « réelle » et la cosmologie « fiducielle ». La vraie cosmologie est celle utilisée par la nature pour déterminer la valeur de l'échelle BAO  $r_s(z_{drag})$  (simplement  $r_s$  dans la suite). C'est également celle qui permet de transformer la séparation physique des absorbeurs Lyman- $\alpha$  en angles sur le ciel et en différences de décalages spectraux :

$$\theta_{BAO} = \frac{r_s}{(1+z)D_A(z)} \quad , \quad \Delta z = r_s H(z) \quad (5.27)$$

De notre côté, nous utilisons une cosmologie fiducielle, le modèle  $\Lambda$ CDM, avec le jeu de paramètres définis en équation 5.6. Cette cosmologie fiducielle nous permet de transformer les angles et décalages spectraux des absorbeurs en distances comobiles, aux décalages spectraux en question. Si la cosmologie fiducielle est la vraie cosmologie, le pic reconstruit sera à la position  $r_{s, fid}$  prédit par la cosmologie fiducielle. Dans le cas contraire, le rapport entre la position reconstruite du pic et la position prédite par le modèle fiduciel permet de contraindre les paramètres de la vraie cosmologie.

2. Les appellations « avec pic » et « sans pic » sont abusives puisque, comme visible en figure 5.22, les BAO ont un effet sur le quadrupôle et l'hexadécapôle de la fonction de corrélation qui n'a pas la forme d'un pic.

L'utilisation d'une cosmologie fiducielle incorrecte (donc de mauvaises valeurs de  $D_A$ ,  $H$  et  $r_s$ ) entraîne donc un déplacement du pic acoustique dans la fonction de corrélation. Ce déplacement peut avoir lieu de manière différente dans la direction transverse et dans la direction radiale. Nous définissons les paramètres  $\alpha_t$  et  $\alpha_p$  comme le rapport entre la position reconstruite du pic et la position prédite par la cosmologie fiducielle, respectivement transversalement et le long de la ligne de visée<sup>3</sup> :

$$\alpha_t = \frac{[D_A/r_s]}{[D_A/r_s]_{fid}} \quad , \quad \alpha_p = \frac{[Hr_s]_{fid}}{[Hr_s]} \quad (5.28)$$

où l'indice *fid* désigne le modèle fiduciel.

Suivant la méthode proposée par Xu et al. (2013), nous utilisons une fonction  $\tilde{\xi}_l(r)$  pour l'ajustement du monopôle et du quadrupôle dont la position du pic dépend de  $\alpha_t$  et  $\alpha_p$  :

$$\tilde{\xi}_l(r) = \hat{\xi}_l(r, \alpha_t, \alpha_p, b, \beta) + A_l(r) \quad , \quad l = 0, 2 \quad (5.29)$$

où les fonctions  $A_l(r)$  sont définies en équation 5.39. Les multipôles  $\hat{\xi}_l$  sont obtenus par projection de la fonction de corrélation théorique  $\hat{\xi}_F$  sur les polynômes de Legendre :

$$\hat{\xi}_0(r) = \int_{-1}^1 \hat{\xi}_F(r', \mu') d\mu \quad (5.30)$$

et

$$\hat{\xi}_2(r) = \frac{1}{5} \int_{-1}^1 \frac{3\mu'^2 - 1}{2} \hat{\xi}_F(r', \mu') d\mu \quad (5.31)$$

où, suivant la démarche de Xu et al. (2013), nous avons transformé les paramètres  $\alpha_t$  et  $\alpha_p$  en  $\alpha$  et  $\epsilon$  :

$$\begin{cases} \alpha_p &= \alpha(1 + \epsilon)^2 \\ \alpha_t &= \frac{\alpha}{1 + \epsilon} \end{cases} \quad (5.32)$$

et introduit les paramètres  $r'$  et  $\mu'$  définis par :

$$r' = \alpha r \sqrt{\frac{1 - \mu^2}{(1 + \epsilon)^2} + \mu^2(1 + \epsilon)^4} \quad (5.33)$$

et

$$\mu' = \frac{\mu}{\sqrt{\frac{1 - \mu^2}{(1 + \epsilon)^6} + \mu^2}} \quad (5.34)$$

Le facteur  $1/5$  de l'expression 5.31 est la normalisation du polynôme de Legendre d'ordre 2. La fonction de corrélation  $\hat{\xi}_F$  est quant à elle la transformée de Fourier du spectre de puissance  $P_F(\mathbf{k})$  des  $\delta_q$  défini en équation 5.25. Hamilton (1992) a montré qu'elle est donnée par :

$$\hat{\xi}_F(r, \mu) = \sum_{l=0,2,4} b^2 C_l(\beta) \xi_{l,F} P_l(\mu) \quad (5.35)$$

où  $P_l(\mu)$  sont les polynômes de Legendre définis en équation 5.16 et les  $C_l(\beta)$  sont des polynômes du second degré en  $\beta$  :

$$\begin{aligned} C_0 &= 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 \\ C_2 &= \frac{4}{3}\beta + \frac{4}{7}\beta^2 \\ C_4 &= \frac{8}{35}\beta^2 \end{aligned} \quad (5.36)$$

3. Comme nous le verrons en section 5.7.1, l'application de ces formules aux données suppose que l'analyse ne biaise pas la position reconstruite du pic acoustique, ce qui est vérifié en section 5.5.2



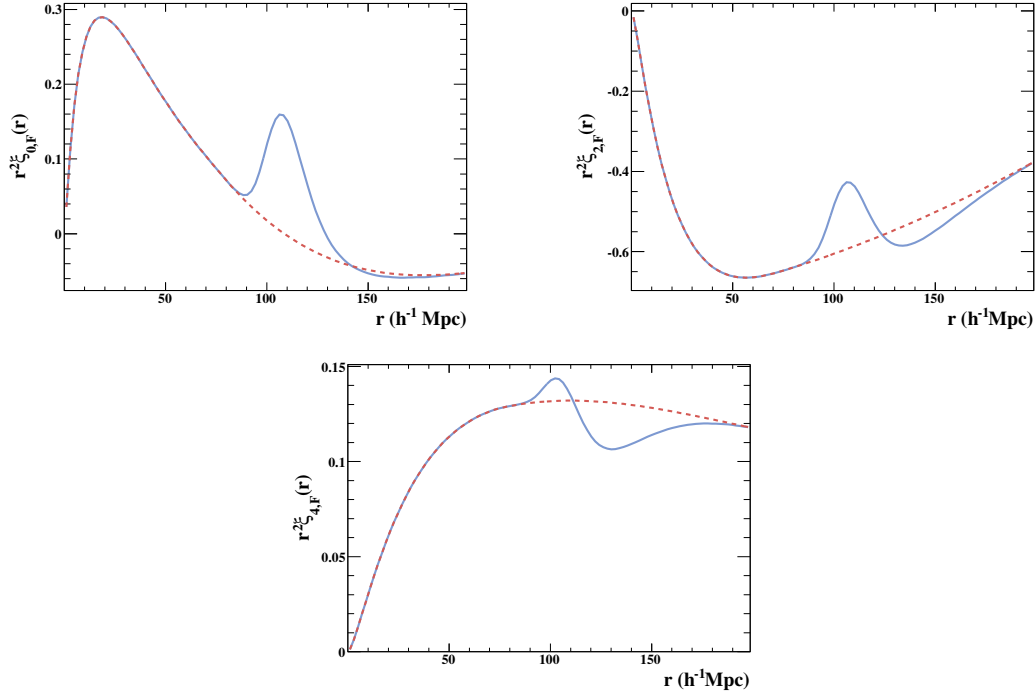


FIGURE 5.22: En bleu : multipôles  $\xi_{l,F}$  calculés à partir du spectre de puissance  $P_F(\mathbf{k})$  de l'équation 5.25 et pour  $b = 0,14$  et  $\beta = 1,4$ . En rouge : multipôles « sans pic »  $\xi_{l,F}^{sp}$  obtenus en ajustant une spline d'ordre 3 aux  $\xi_{l,F}$  en masquant la région comprise entre 80 et 200  $h^{-1}\text{Mpc}$ .

Les multipôles  $\xi_{l,F}$  résultent de la transformation de Fourier en trois dimensions. Ils sont donnés par :

$$\xi_{l,F} = (2\pi)^3 \int P_F(k) j_l(k.r) d^3 k \quad (5.37)$$

où les  $j_l$  sont les fonctions de Bessel  $j_l(x)$  définies par :

$$\begin{aligned} x \cdot j_0(x) &= \sin(x) \\ x^3 \cdot j_2(x) &= x^2 \sin(x) - 3 \sin(x) + 3 \cos(x) \\ x^5 \cdot j_4(x) &= x^4 \sin(x) - 45x^2 \sin(x) + 10x^3 \cos(x) - 105x \cos(x) \end{aligned} \quad (5.38)$$

La figure 5.22 montre les multipôles  $\xi_{l,F}$  calculés à partir du spectre de puissance  $P_F(\mathbf{k})$  de l'équation 5.25. Nous définissons également des multipôles « sans pic »  $\xi_{l,F}^{sp}$  en ajustant une spline d'ordre 3 aux  $\xi_{l,F}$  en masquant la région comprise entre 80 et 200  $h^{-1}\text{Mpc}$ . Ils seront utilisés pour les ajustements sans pic.

Les fonctions  $A_l(r)$  de l'équation 5.29 sont de la forme :

$$A_l(r) = \frac{a_l}{r^2} + \frac{b_l}{r} + c_l + \frac{d_l}{\sqrt{r}} \quad (5.39)$$

où  $a_l$ ,  $b_l$ ,  $c_l$  et  $d_l$  sont des paramètres de nuisance de l'ajustement. Ces fonctions permettent de prendre en compte la distorsion de la fonction de corrélation introduite par la méthode d'obtention

du flux moyen. Elles permettent également de s'affranchir du fait que la forme globale de la fonction de corrélation varie de manière non triviale lorsque l'on varie les paramètres de la cosmologie fiducielle. En d'autres termes, ces fonctions permettent de faire varier librement la forme globale de l'ajustement et de s'assurer que les paramètres  $\alpha_t$  et  $\alpha_p$  ne sont contraints que par la position du pic dans les données.

### 5.5.2 Ajustement des simulations

Nous commençons par valider sur les simulations notre méthode d'estimation des multipôles ainsi que notre procédure d'ajustement. Pour ce faire nous procédons à deux types d'ajustements. Le premier est un ajustement isotrope, pour lequel  $\alpha_t = \alpha_p = \alpha_{iso}$ . L'ajustement comporte alors 79 degrés de liberté. Le deuxième est un ajustement à 2 dimensions, c'est-à-dire pour lequel  $\alpha_t$  et  $\alpha_p$  peuvent varier indépendamment. Dans tous les cas, nous limitons notre ajustement aux points situés à  $r > 20 h^{-1}\text{Mpc}$  pour ne pas tirer de contraintes de la position du pic de  $r^2\xi_0$  à  $r \sim 15 h^{-1}\text{Mpc}$ . La figure 5.23 présente les monopôles et quadrupôles obtenus ainsi que les meilleurs ajustements<sup>4</sup> isotropes et à deux dimensions pour 5 réalisations des simulations choisies aléatoirement. Dans la plupart des cas, les ajustements isotropes et à deux dimensions sont très semblables. Les figures 5.24 et 5.25 présentent quant à elles les distributions des  $\chi^2$  obtenus pour chacun de ces ajustements. Que ce soit pour les ajustements isotropes ou à deux dimensions, les distributions obtenues sont en bon accord avec des distributions de  $\chi^2$  à environ 80 degrés de liberté.

La figure 5.26 présente les valeurs de  $\alpha_{iso}$  ajustées pour les différentes simulations. La valeur moyenne est  $\alpha_{iso} = 1,0044 \pm 0,0036$ . La valeur de l'ajustement de la moyenne des 25 simulations est également représentée, elle vaut  $\alpha_{iso} = 1,0049 \pm 0,0039$ . A ce niveau de précision, de l'ordre de quatre pour mille, il n'y a donc pas de biais de la position du pic acoustique introduit par le calcul de la fonction de corrélation ou la procédure d'ajustement isotrope. Une simulation sur les 25 est à plus de deux écarts standards du modèle fiduciel ce qui est le nombre attendu.

La figure 5.27 présente quant à elle les valeurs des ajustements à deux dimensions pour 24 des 25 simulations ainsi que pour la moyenne des 25 simulations. La simulation manquante correspond à la numéro 23 pour laquelle la détection du pic acoustique est trop faible pour contraindre l'ajustement à deux dimensions. Les valeurs obtenues pour la moyenne des 25 simulations sont  $\alpha_p = 1,0038 \pm 0,0073$  et  $\alpha_t = 1,0072^{+0,0140}_{-0,0134}$ . Une fois encore les valeurs retrouvées sont compatibles avec le modèle fiduciel et il n'y a pas d'indication de biais introduit par la méthode de calcul de la fonction de corrélation ou par la procédure d'ajustement.

### 5.5.3 Données : ajustement isotrope et niveau de confiance de la détection

Nous commençons par fixer  $\alpha_{iso}$  à 1 pour vérifier que nous trouvons des valeurs raisonnables du  $\chi^2$  et des paramètres de l'ajustement pour le modèle fiduciel. Le  $\chi^2$  de cet ajustement est de 93,24 pour 80 degrés de liberté (ces chiffres sont résumés en table 5.3). Les paramètres  $b_\delta$  et  $\beta$  sont largement dégénérés car à la fois le monopôle et le quadrupôle ont des amplitudes proportionnelles à  $b_\delta^2$  et à des polynômes en  $\beta$ . Le produit  $b_\delta(1+\beta)$  est, quant à lui, mieux déterminé. Nous trouvons  $b_\delta(1+\beta) = 0,451^{+0,050}_{-0,057}$ , ce qui est sensiblement plus élevé que la valeur mesurée par Slosar et al. (2011) à  $40 h^{-1}\text{Mpc}$  :  $b_\delta(1+\beta) = 0,336 \pm 0,012$ . Les plus grandes barres d'erreurs de notre ajustement résultent de la souplesse introduite par les termes  $A_l(r)$ .

4. L'intégralité des ajustements de ce travail sont réalisés grâce à la bibliothèque C++ MINUIT. Pour plus d'information, voir le site internet <http://seal.web.cern.ch/seal/work-packages/mathlibs/minuit/>.

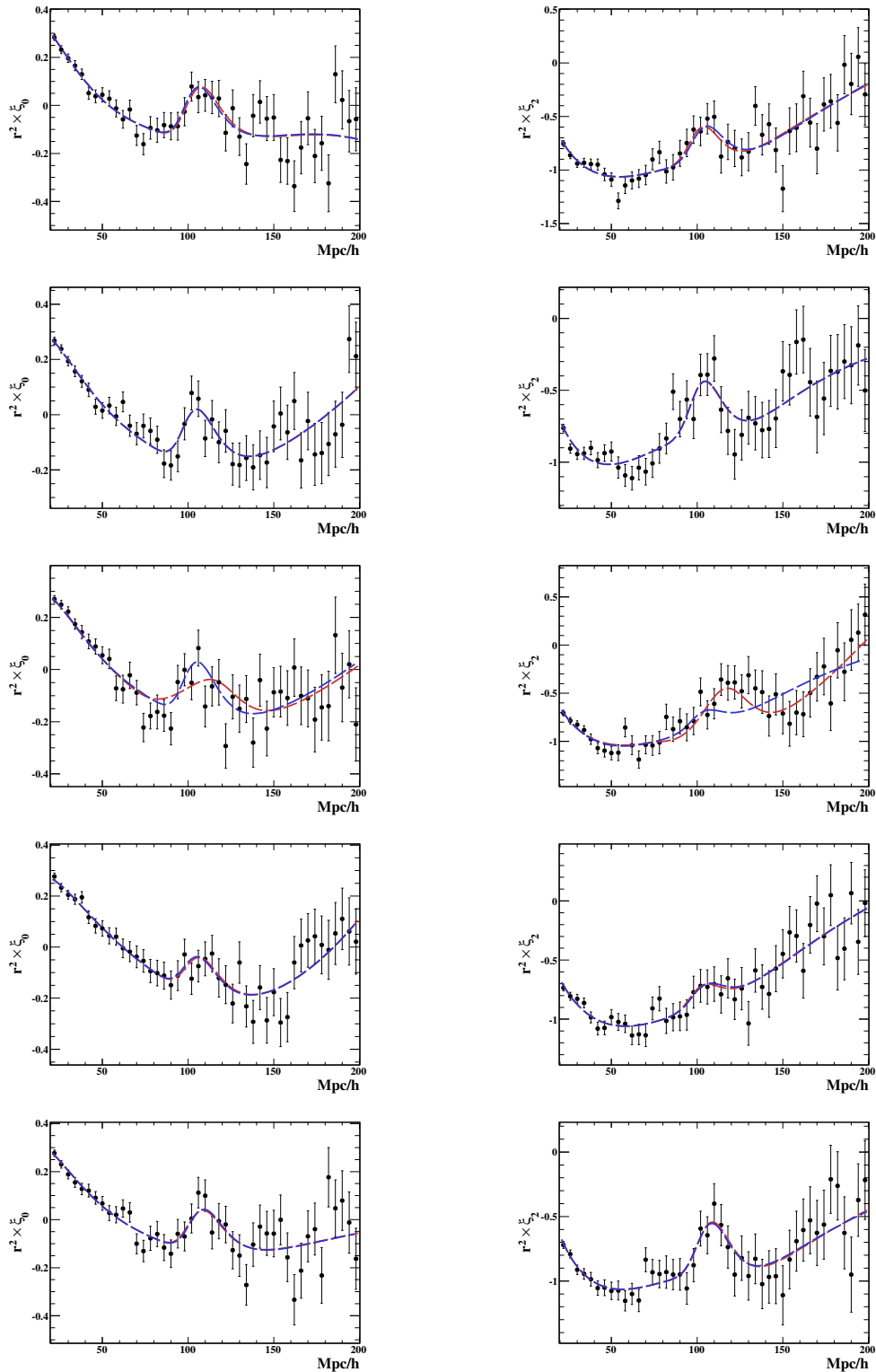


FIGURE 5.23: Monopôles et quadrupôles ainsi que les ajustements obtenus sur 5 simulations choisies aléatoirement parmi les 25 disponibles (les réalisations représentées correspondent aux numéros 003, 007, 014, 019 et 022). Les ajustements isotropes sont en bleu, les ajustements en deux dimensions sont en rouge.

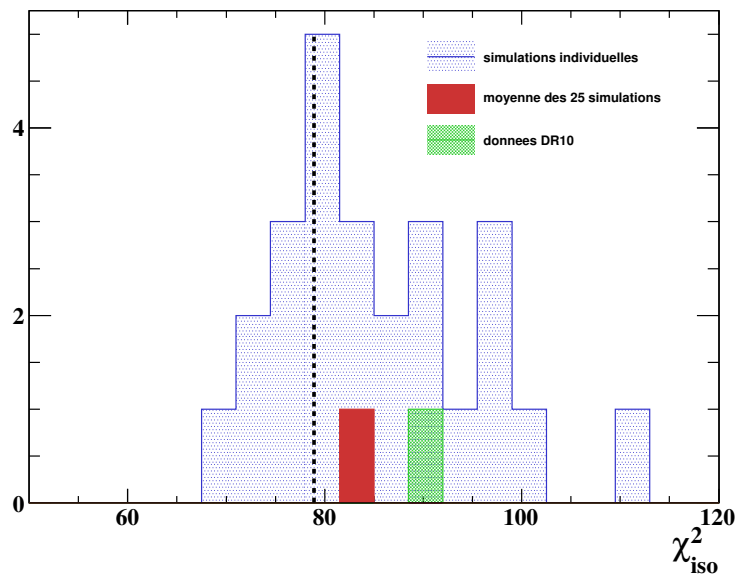


FIGURE 5.24: Distributions des  $\chi^2$  des ajustements isotropes pour les 25 simulations (en bleu). La ligne verticale noire correspond aux nombres de degrés de liberté de l'ajustement (79). Egalement représentés sont les  $\chi^2$  des ajustements de la moyenne des simulations (en rouge) et des données (en vert).

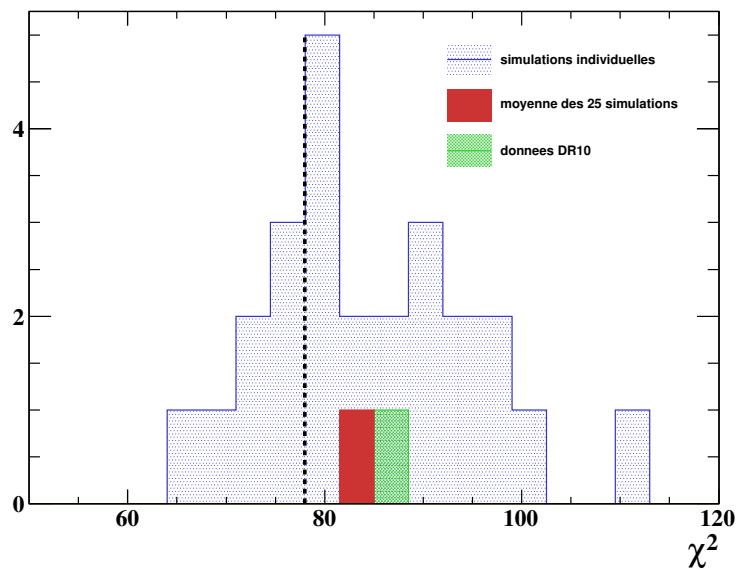


FIGURE 5.25: Distributions des  $\chi^2$  des ajustements à deux dimensions pour les 25 simulations (en bleu). La ligne verticale noire correspond aux nombres de degrés de liberté de l'ajustement (78). Egalement représentés sont les  $\chi^2$  des ajustements de la moyenne des simulations (en rouge) et des données (en vert).

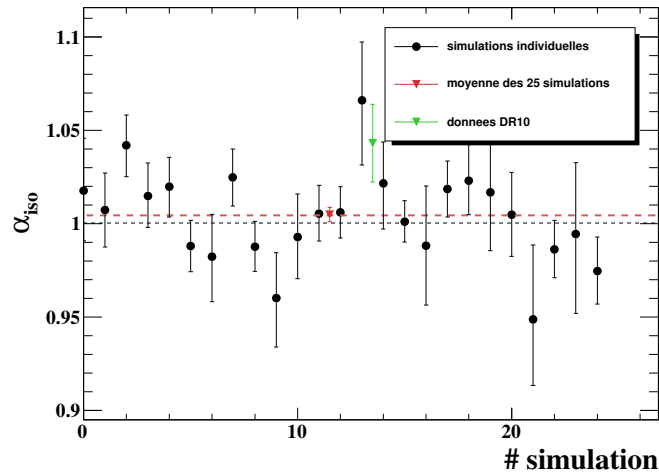


FIGURE 5.26: Valeurs des  $\alpha_{iso}$  mesurées sur les différentes simulations. L'ajustement des  $\alpha_{iso}$  des 25 simulations par une constante  $C$  donne  $C = 1,0044 \pm 0,0036$ . L'ajustement sur la moyenne des 25 simulations est représenté par le triangle rouge ; il est compatible à un  $\sigma$  avec le modèle fiduciel. La valeur de  $\alpha_{iso}$  obtenue sur les données est représentée par le triangle vert.

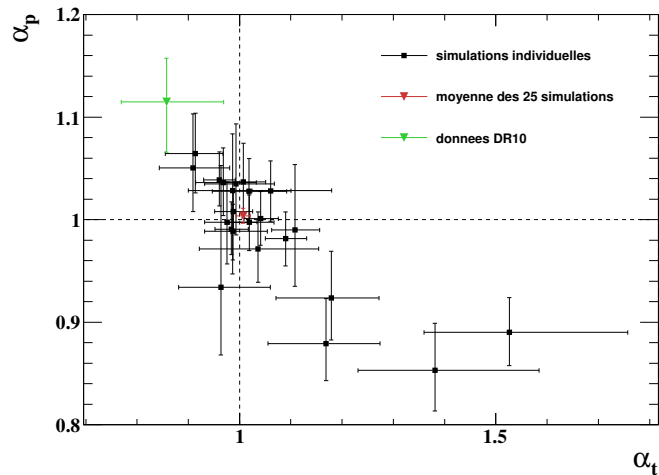


FIGURE 5.27: Valeurs des  $\alpha_t$  et  $\alpha_p$  mesurées sur les différentes simulations. L'ajustement sur la moyenne des 25 simulations est représenté par le triangle rouge ; il est compatible à moins d'un  $\sigma$  avec  $\alpha_t = 1$  et  $\alpha_p = 1$ . Les valeurs obtenues sur les données sont représentées par le triangle vert.

ajustement	fiduciel	isotrope	deux dimensions
$\chi^2_{min}/DDL$	93,24/80	89,15/79	87,70/78
$\alpha_{iso}$	1	$1,043^{+0,021}_{-0,020}$	—
$\alpha_p$	1	—	$1,115^{+0,043}_{-0,049}$
$\alpha_t$	1	—	$0,858^{+0,111}_{-0,088}$
$b(1 + \beta)$	$0,451^{+0,050}_{-0,057}$	$0,479^{+0,065}_{-0,036}$	$0,428^{+0,065}_{-0,067}$

TABLE 5.3: Valeurs des paramètres obtenues pour l'ajustement du modèle fiduciel ( $\alpha_p = \alpha_t = 1$ ), l'ajustement isotrope ( $\alpha_p = \alpha_t$ ) et pour l'ajustement à deux dimensions.

Nous estimons ensuite le niveau de confiance de la détection du pic acoustique dans les multipôles des données. Suivant la démarche d'Anderson et al. (2012), nous calculons le  $\chi^2$  d'un ajustement avec pic et d'un ajustement sans pic en fonction de  $\alpha_{iso}$ . Les deux ajustements obtenus sont représentés en figure 5.29. Le meilleur ajustement pour le modèle avec pic est obtenu pour  $\alpha_{iso} = 1,043$  et a  $\chi^2_{min,ap} = 89,15$  pour 79 degrés de liberté. Comme on peut le voir en figure 5.24, cette valeur est parfaitement compatible avec la distribution de  $\chi^2$  des 25 simulations. Le  $\chi^2$  de l'ajustement sans pic pour cette valeur de  $\alpha_{iso}$  est  $\chi^2_{sp}(1,043) = 106,6$ . Nous définissons le  $\Delta\chi^2$  par :

$$\Delta\chi^2(\alpha_{iso}) = \chi^2(\alpha_{iso}) - \chi^2_{min,ap} \quad (5.40)$$

La figure 5.28 présente les  $\Delta\chi^2(\alpha_{iso})$  pour les deux ajustements. La différence de  $\chi^2$  entre les deux ajustements pour  $\alpha_{iso} = 1,043$  est de 17,5. Rappelons que les deux ajustements utilisent l'équation 5.29, mais que dans un cas cette paramétrisation utilise les  $\xi_{l,F}$  avec pic et dans l'autre cas les  $\xi_{l,F}^{sp}$  sans pic (voir figure 5.22). Le  $\Delta\chi^2$  n'est donc pas rigoureusement égal au carré du niveau de confiance de la détection (il n'y a pas de degré de liberté de différence entre les deux ajustements). La grande différence entre les  $\chi^2$  des deux fits nous assure toutefois de la détection du pic, avec un niveau de confiance de l'ordre de  $4\sigma$ .

Le profil de la courbe  $\chi^2(\alpha_{iso})$  pour l'ajustement avec pic, nous permet d'estimer les erreurs à une déviation standard sur la mesure de la position du pic :  $\alpha_{iso} = 1,043^{+0,021}_{-0,020}$ . Les valeurs de l'ensemble des paramètres physiques sont reproduites en table 5.3. Les erreurs sur les paramètres autres que  $\alpha_{iso}$  sont les erreurs retournées par MINUIT.

Pour vérifier que la valeur mesurée de  $\alpha_{iso}$  n'est pas sensible à la gamme en  $r$  de l'ajustement, nous effectuons différents ajustement sur différentes gammes. Le risque est de contraindre la mesure par la partie décroissante de la fonction de corrélation à bas  $r$ , dont l'évolution en fonction de  $\alpha_{iso}$  n'est pas correctement modélisée par notre ajustement. La table 5.4 résume les valeurs de  $\alpha_{iso}$  obtenues en fonction de la gamme en  $r$  de l'ajustement. Aucune dépendance n'est observée.

#### 5.5.4 Données : ajustement en deux dimensions

Nous libérons enfin la contrainte  $\alpha_p = \alpha_t$  pour ajuster simultanément les deux paramètres. Le  $\chi^2$  de cet ajustement à deux dimensions est de 87,70 pour 78 degrés de liberté. Cette valeur est parfaitement compatible avec la distribution de  $\chi^2$  obtenue pour les 25 simulations, représentée en figure 5.25. Nous trouvons  $\alpha_p = 1,115^{+0,043}_{-0,049}$  à deux écarts standards au-dessus du modèle fiduciel et  $\alpha_t = 0,858^{+0,111}_{-0,088}$  à 1,5 écarts standards en-dessous du modèle fiduciel. Les erreurs sur  $\alpha_p$  sont environ deux fois plus petites que celles sur  $\alpha_t$  ce qui est cohérent avec le fait que le pic soit clairement visible à grand  $\mu$ , mais pas à bas  $\mu$ . Les valeurs des paramètres obtenues sont résumés

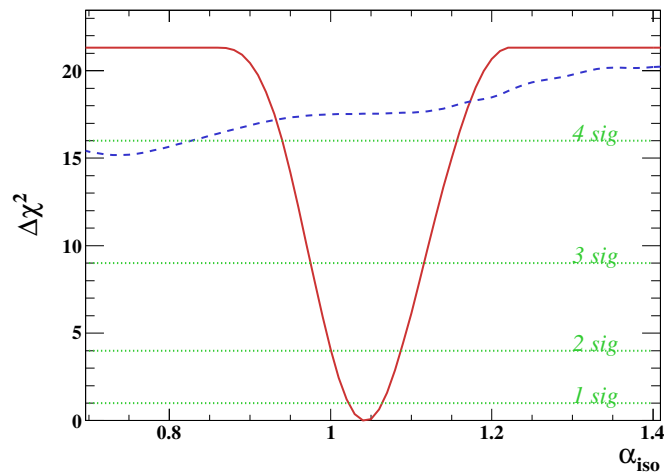


FIGURE 5.28: Evolution du  $\Delta\chi^2$  (défini en équation 5.40) en fonction de  $\alpha_{iso}$  pour un ajustement avec pic (ligne rouge) et un ajustement sans pic (ligne bleue). Le minimum pour l'ajustement avec pic est atteint pour  $\alpha_{iso} = 1,043$ . La différence entre la courbe rouge et la courbe bleue donne une indication du niveau de confiance de la détection.

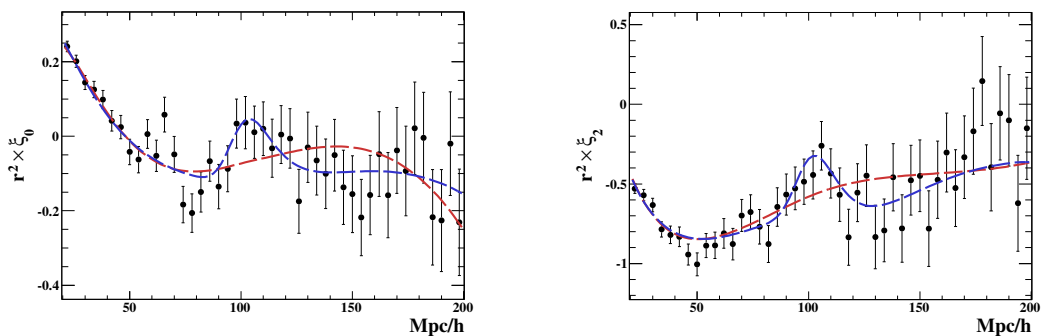


FIGURE 5.29: Ajustement isotrope ( $\alpha_t = \alpha_p$ ) des données DR10 $\alpha$ . La courbe bleue correspond à l'ajustement « avec pic » tandis que la courbe rouge correspond à l'ajustement « sans pic ».

gamme en $r$	$\chi^2/\text{DDL}$	$\alpha_{iso}$
20 – 200	89,15/79	$1,043^{+0,021}_{-0,020}$
30 – 200	83,10/75	$1,045^{+0,020}_{-0,020}$
40 – 200	77,10/69	$1,044^{+0,021}_{-0,020}$
50 – 200	70,30/65	$1,042^{+0,022}_{-0,022}$
20 – 190	85,59/75	$1,042^{+0,021}_{-0,021}$
20 – 180	78,72/69	$1,042^{+0,021}_{-0,021}$
20 – 170	72,81/65	$1,042^{+0,023}_{-0,022}$
50 – 170	55,65/51	$1,038^{+0,023}_{-0,023}$

TABLE 5.4: Meilleurs ajustements des multipôles de la fonction de corrélation obtenue sur DR10 $\alpha$ . La valeur de  $\alpha_{iso}$  n'est pas sensible à la gamme en  $r$  de l'ajustement.

en table 5.3.

La figure 5.30 présente la « nappe de  $\chi^2$  » obtenue en fonction des paramètres  $\alpha_t$  et  $1/\alpha_p$  ainsi que les contours à une et deux déviations standards. Ces contours sont calculés suivant une approche fréquentiste. Le  $\chi^2$  de l'ajustement en deux dimensions est une fonction des 12 paramètres  $(\alpha_t, \alpha_p, b, \beta, A_i)$ , où  $A_i$  représente l'ensemble des 8 paramètres de nuisance. Le minimum  $\chi_0^2$  de la distribution de  $\chi^2$  est obtenu par minimisation de la distribution de  $\chi^2$  en libérant les 12 paramètres. Pour obtenir le niveau de confiance de l'un des paramètres  $\theta_i$ , nous fixons  $\theta_i$  à une valeur que nous incrémentons successivement. Pour chaque valeur de  $\theta_i$ , nous minimisons le  $\chi^2$  avec maintenant 11 paramètres libres. La différence  $\Delta\chi^2(\theta_i)$  entre le nouveau minimum et  $\chi_0^2$  nous permet d'estimer le niveau de confiance  $\text{NC}(\theta_i)$  sur la valeur du paramètre  $\theta_i$ , en supposant les erreurs expérimentales gaussiennes :

$$\text{NC}(\theta_i) = 1 - \int_{\Delta\chi^2(\theta_i)}^{\infty} f_{\chi^2}(t, N_{ddl}) dt \quad (5.41)$$

avec

$$f_{\chi^2}(t, N_{ddl}) = \frac{e^{-t/2} t^{N_{ddl}/2-1}}{\sqrt{2}^{N_{ddl}} \Gamma(N_{ddl}/2)} \quad (5.42)$$

où  $\Gamma$  est la fonction Gamma et le nombre de degrés de liberté  $N_{ddl}$  est égal à 1. Dans le cas de la figure 5.30, nous fixons les paramètres  $\alpha_t$  et  $\alpha_p$  et les incrémentons successivement. Pour chaque couple de valeurs  $(\alpha_t, \alpha_p)$  nous minimisons le  $\chi^2$  avec 10 paramètres libres, produisant la « nappe de  $\chi^2$  » visible en échelle de couleurs. Les niveaux de confiances sont ensuite estimés d'après la formule 5.41 appliquée avec maintenant deux degrés de liberté.

La figure 5.31 montre les meilleurs ajustements pour le monopôle et le quadrupôle. La composition d'un  $\alpha_p$  grand et d'un  $\alpha_t$  petit permet d'élargir le pic du monopôle, en ne modifiant que faiblement sa position.

## 5.6 Recherche de systématiques

La recherche d'effets systématiques est une tâche difficile puisque nous sommes encore largement dominés par les erreurs statistiques. La procédure standard consiste à étudier le comportement du signal lorsque l'on divise l'échantillon de données complet en sous-échantillons, suivant



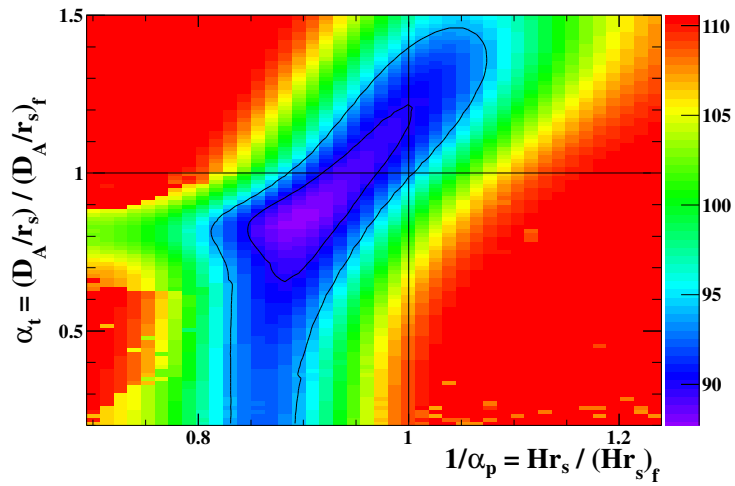


FIGURE 5.30: Nappe de  $\chi^2$  obtenue en fonction de  $\alpha_t$  et  $1/\alpha_p$  ainsi que les contours à 1 et 2  $\sigma$ .

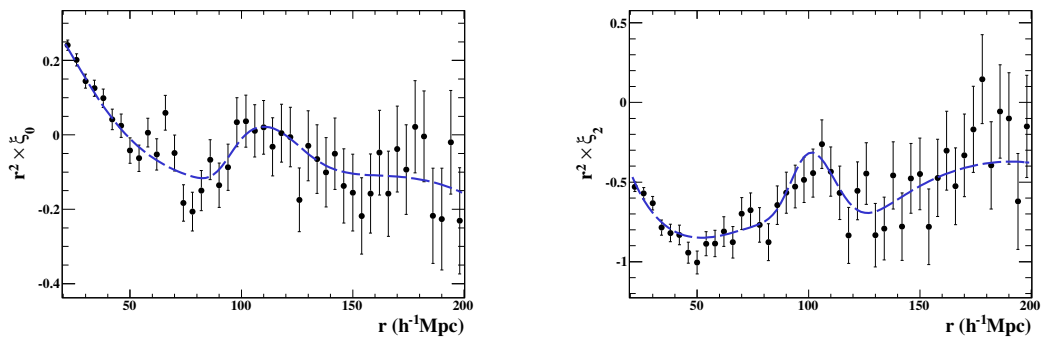


FIGURE 5.31: Ajustement à deux dimensions des multipôles de DR10 $\alpha$ .

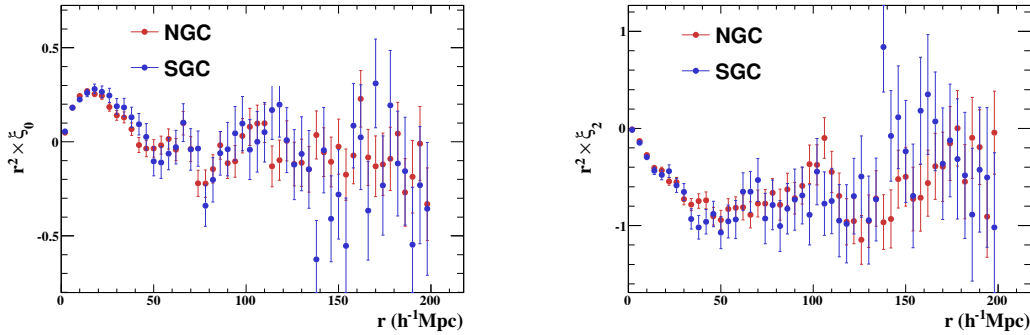


FIGURE 5.32: Monopôles (à gauche) et quadrupôles (à droite) des fonctions de corrélation du pôle Nord galactique (NGC) et du pôle Sud galactique (SGC).

différents critères. Dans notre cas, diviser l'échantillon entraîne une diminution significative du niveau de confiance de la détection, voire la perte du signal BAO. Cette procédure nous permet toutefois de vérifier qu'il n'y a pas d'effet systématique majeur.

### 5.6.1 Nord - Sud

Nous commençons par comparer les multipôles obtenus sur le pôle Nord galactique (NGC) et le pôle Sud galactique (SGC). Ces multipôles sont visibles en figure 5.32. Le SGC contient environ 4 fois moins de quasars que le NGC ce qui justifie que les barres d'erreurs y soient deux fois plus grandes. Il n'y a pas de désaccord apparent entre les multipôles du NGC et du SGC. La figure 5.33 montre le  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  obtenu sur le NGC. Le meilleur ajustement est obtenu pour  $\alpha_{iso} = 1,024_{-0,018}^{+0,018}$ , à un écart standard de la valeur mesurée sur l'échantillon complet. Les erreurs sur  $\alpha_{iso}$  sont légèrement plus faibles que pour l'échantillon de DR10 $\alpha$ , mais le niveau de confiance de la détection l'est également, laissant penser qu'il s'agit d'une fluctuation favorable. Malheureusement, avec seulement 19.000 quasars, le SGC ne permet pas de détecter à lui seul le pic acoustique et de vérifier que sa position est en accord avec celle mesurée sur le NGC.

### 5.6.2 Coupure des raies de Balmer

Dans un second temps nous comparons la fonction de corrélation obtenue en coupant les données contaminées par les résidus des raies de Balmer  $H_\gamma$  et  $H_\delta$ , avec celle calculée sur l'échantillon complet. Les pixels coupés sont ceux dont la longueur d'onde appartient à l'intervalle  $4.080 < \lambda < 4.130$  pour la raie  $H_\delta$  et  $4.310 < \lambda < 4.370$  pour la raie  $H_\gamma$ . Ces coupures retirent environ 8% des pixels. Les multipôles obtenus après coupures sont visibles en figure 5.34. Ils sont très similaires à ceux calculés sur l'échantillon complet. Seuls 5 points sur 100 bougent d'un écart standard. La figure 5.35 montre que le  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  est presque inchangé après coupures des raies  $H_\gamma$  et  $H_\delta$ ; le nouveau minimum étant atteint pour  $\alpha_{iso} = 1,036_{-0,022}^{+0,023}$ .

### 5.6.3 Plaques

Nous séparons enfin les paires de pixels en deux échantillons suivant qu'elles viennent de deux quasars observés par la même plaque, ou par des plaques différentes. Nous obtenons ainsi deux échantillons parfaitement distincts, dans le sens où chaque paire de pixels de l'échantillon complet

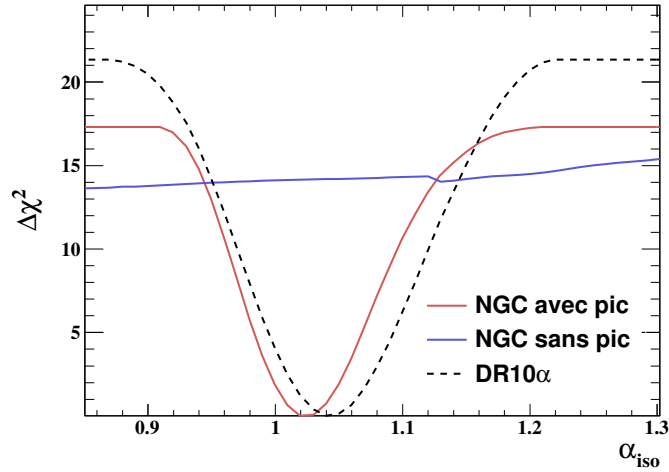


FIGURE 5.33:  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  pour un ajustement avec pic (courbe rouge) et sans pic (courbe bleue) pour les multipôles calculés sur le NGC. La courbe obtenue sur l'échantillon DR10 $\alpha$  complet est également représentée (courbe en pointillés noire) pour comparaison.

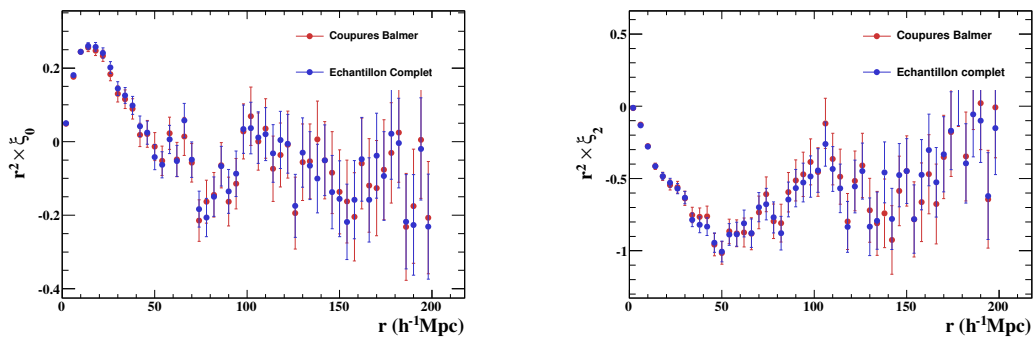


FIGURE 5.34: Monopôles (à gauche) et quadrupôles (à droite) des fonctions de corrélation obtenues avec ou sans coupures des raies de Balmer  $H_\gamma$  et  $H_\delta$ . Les multipôles sans coupures correspondent aux multipôles de la figure 5.29.

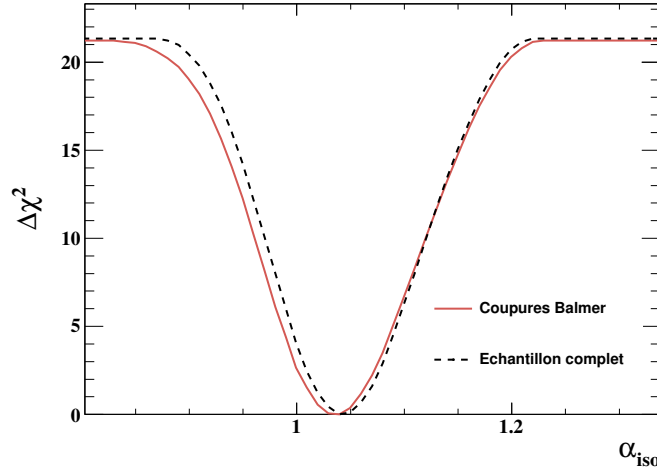


FIGURE 5.35:  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  pour un ajustement de la fonction de corrélation obtenue en coupant les données contaminées par les résidus des raies de Balmer  $H_\gamma$  et  $H_\delta$ . La courbe obtenue sur l'échantillon DR10 $\alpha$  complet est également représentée (courbe en pointillés noire) pour comparaison.

n'appartient qu'à un seul des deux sous échantillons. De nombreuses caractéristiques de la prise de données sont communes par plaque ou demi-plaque (soustraction du fond du ciel, vecteurs d'éta-lonnage, déclinaison lors de l'observation, etc.); cette séparation est donc une piste idéale pour la recherche de systématiques. Les multipôles des fonctions de corrélation calculées sur ces deux sous-échantillons sont visibles en figure 5.36. « L'échantillon de plaques identiques » présente une variance bien plus élevée à grand  $r$ . A l'inverse, les multipôles de « l'échantillon de plaques différentes » possèdent une variance plus élevée à bas  $r$ . Ce comportement est directement lié au nombre de paires entrant en compte dans le calcul des multipôles. La distribution des paires en fonction de la séparation des pixels est présentée, pour les deux échantillons, en figure 5.37. L'échantillon de plaques différentes possède environ dix fois plus de paires pour  $r \sim 200 h^{-1} \text{Mpc}$ , tandis que celui de plaques identiques en possède environ cinq fois plus pour  $r$  proche de 0.

Les  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  pour les deux échantillons sont représentés en figure 5.38. Le meilleur ajustement pour l'échantillon de plaques différentes est obtenu pour  $\alpha_{iso}^{dp} = 0,988^{+0,024}_{-0,023}$ ; celui pour l'échantillon de plaques identiques est en  $\alpha_{iso}^{sp} = 1,053^{+0,027}_{-0,027}$ . La différence entre  $\alpha_{iso}^{dp}$  et  $\alpha_{iso}^{sp}$  vaut  $\alpha_{iso}^{sp} - \alpha_{iso}^{dp} = 0,065 \pm 0,036$ . Elle est compatible avec zéro à moins de deux écarts standards. Cet écart est trop faible pour être considéré comme lié à des effets systématiques.

#### 5.6.4 Conclusion sur la recherche de systématiques

Contrairement à l'étude menée sur DR9, pour laquelle toute division des données en sous échantillons entraînait la perte du signal, nous disposons maintenant de suffisamment de données pour commencer la recherche d'effets systématiques. L'étude menée ici ne permet toutefois pas de mettre en évidence de tels effets. Ce constat implique évidemment que nous sommes encore largement dominé par les erreurs statistiques. Nous nous attendons donc à ce que les erreurs systématiques sur la mesure de  $\alpha_{iso}$  soient inférieures à 2%. L'un des mérites souvent attribué à l'utilisation des BAO pour la mesure de distance est leur faible contamination par des effets

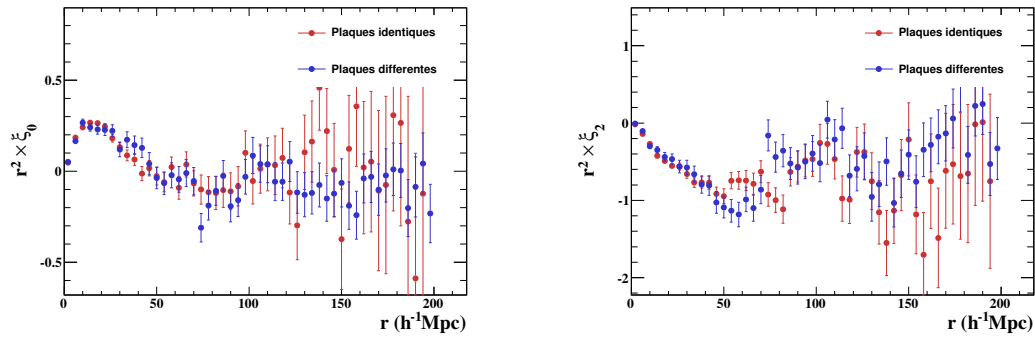


FIGURE 5.36: Multipôles des fonctions de corrélation obtenus en ne considérant que les paires de pixels de quasars observés par la même plaque (points rouges) ou par des plaques différentes (points bleus).

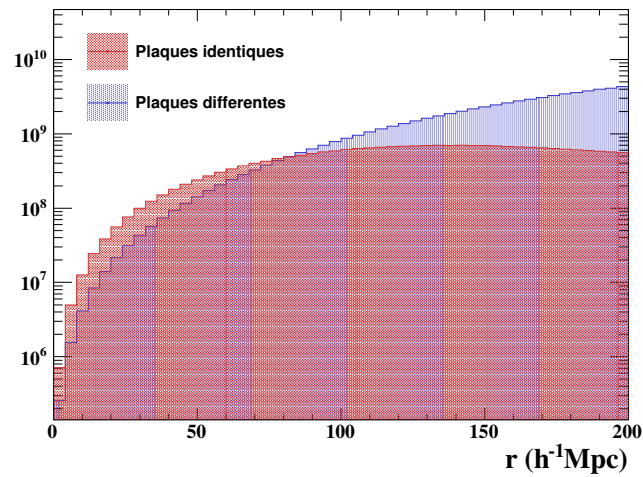


FIGURE 5.37: Distribution des paires de pixels en fonction de la séparation  $r = |\mathbf{r}|$  séparant les pixels. La distribution rouge correspond aux paires de pixels venant de quasars observés par la même plaque, tandis que la distribution bleue correspond aux paires de pixels venant de quasars observés par des plaques différentes.

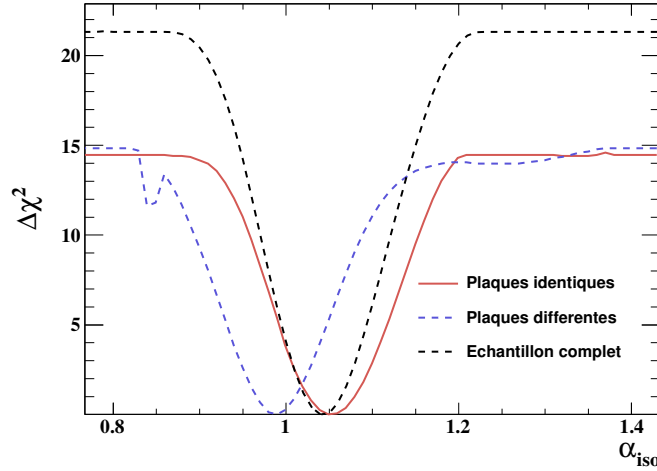


FIGURE 5.38:  $\Delta\chi^2$  en fonction de  $\alpha_{iso}$  pour l'échantillon de « plaques identiques » et l'échantillon de « plaques différentes », comme décrit dans le texte. La courbe obtenue sur l'échantillon DR10 $\alpha$  complet est également représentée (courbe en pointillés noire) pour comparaison.

systématiques. Les analyses basées sur la distribution des galaxies (Anderson et al., 2012; Blake et al., 2011), n'ont à ce jour pas mis en évidence d'effets systématiques sur la mesure de la position reconstruite du pic BAO. Nous commençons à vérifier qu'il en est de même pour les forêts Lyman- $\alpha$ .

## 5.7 Contraintes cosmologiques

Dans cette section, nous obtenons finalement des contraintes cosmologiques de la mesure de la position du pic acoustique présentée en section 5.5. Pour ce faire nous utilisons les formules 5.28 reliant les valeurs de  $\alpha_t$  et  $\alpha_p$  aux fonctions  $D_A(z)$ ,  $H(z)$  et  $r_s(z_{drag})$ . Comme nous allons maintenant le montrer, ces formules supposent que l'analyse n'introduit pas de biais sur la position du pic BAO.

### 5.7.1 Contraintes sur $D_A/r_s$ et $Hr_s$

Pour les simulations, nous connaissons à la fois la position reconstruite du pic, et sa position réelle. Nous pouvons donc simplement définir  $\alpha_t$  et  $\alpha_p$  comme :

$$\alpha_t^{\text{simu}} = \frac{[D_A/r_s]_{\text{rec}}}{[D_A/r_s]_{\text{fid}}} \quad , \quad \alpha_p^{\text{simu}} = \frac{[Hr_s]_{\text{fid}}}{[Hr_s]_{\text{rec}}} \quad (5.43)$$

où l'indice « rec » désigne l'échelle reconstruite et  $r_s = r_s(z_{drag})$  (défini en équation 2.8). Les valeurs moyennes des  $\alpha_p$  et  $\alpha_t$  obtenues sur un grand nombre de simulations permettent donc d'estimer le biais de l'analyse. Nous utilisons rigoureusement la même analyse pour traiter les simulations et les données. Un éventuel biais observé dans les simulations doit donc être présent dans les données. Les valeurs des paramètres  $\alpha$  obtenues sur ces données peuvent donc s'exprimer

sous la forme :

$$\begin{aligned}\alpha_t^{\text{données}} &= \left( \frac{[D_A/r_s]^{\text{données}}}{[D_A/r_s]_{\text{rec}}^{\text{simu}}} \right) \left( \frac{[D_A/r_s]_{\text{rec}}^{\text{simu}}}{[D_A/r_s]_{\text{fidu}}} \right) \\ \alpha_p^{\text{données}} &= \left( \frac{[Hr_s]^{\text{données}}}{[Hr_s]_{\text{rec}}^{\text{simu}}} \right) \left( \frac{[Hr_s]_{\text{rec}}^{\text{simu}}}{[Hr_s]_{\text{fidu}}} \right)\end{aligned}\tag{5.44}$$

Puisque nous n'observons pas de biais sur les simulations, nous simplifions cette expression en considérant :

$$\begin{aligned}[D_A/r_s]_{\text{rec}}^{\text{simu}} &= [D_A/r_s]_{\text{fidu}} \\ [Hr_s]_{\text{rec}}^{\text{simu}} &= [Hr_s]_{\text{fidu}}\end{aligned}\tag{5.45}$$

de sorte que :

$$\begin{aligned}\alpha_t^{\text{données}} &= \left( \frac{[D_A/r_s]^{\text{données}}}{[D_A/r_s]_{\text{fidu}}} \right) \\ \alpha_p^{\text{données}} &= \left( \frac{[Hr_s]^{\text{données}}}{[Hr_s]_{\text{fidu}}} \right)\end{aligned}\tag{5.46}$$

Nous retrouvons ainsi les expressions de l'équation 5.28 permettant de transformer les contraintes sur les paramètres  $\alpha$  en contraintes sur  $D_A/r_s$  et  $Hr_s$ .

## 5.7.2 Contraintes sur les paramètres cosmologiques

La prochaine étape consiste à transformer les contraintes précédentes en contraintes sur les paramètres cosmologiques. Dans le modèle  $\Lambda$ CDM et pour un décalage spectral inférieur à 2,38,  $H(z)$  dépend des paramètres  $H_0$ ,  $\Omega_m$  et  $\Omega_\Lambda$ . La distance angulaire  $D_A(2,38)$  est l'intégrale de 0 à 2,38 de  $1/H(z)$  et dépend donc des mêmes paramètres. La dépendance de l'horizon acoustique  $r_s(z_{\text{drag}})$  est plus subtile. Comme le montre l'équation 2.6,  $r_s(z_{\text{drag}})$  est l'intégral de  $c_s(z)/H(z)$  pour un temps allant de  $t_0 = 0$  à  $t_{\text{drag}}$ . A cette époque,  $H(z)$  dépend essentiellement de la densité de radiation  $\Omega_r$ , de la densité de matière  $\Omega_m$ , et de  $H_0$ . La vitesse du son  $c_s$  dépend quant à elle du rapport de la densité d'énergie des baryons sur celle des photons :  $\rho_b/\rho_\gamma = \Omega_b/[\Omega_\gamma(1+z)]$ . Remarquons que les pré-facteurs  $H_0$  se simplifient dans les expressions de  $D_A/r_s$  et  $Hr_s$ . En résumé,  $D_A/r_s$  et  $Hr_s$  dépendent des paramètres suivant :  $(\Omega_\Lambda, \Omega_m, \Omega_b, \Omega_r, \Omega_\gamma)$ .

Notre mesure est principalement sensible aux paramètres  $\Omega_m$  et  $\Omega_\Lambda$ . Pour transformer les contraintes sur  $\alpha_t$  et  $\alpha_p$  en contraintes sur ces deux paramètres, nous fixons les densités de baryons, photons, et radiation. Nous calculons  $\Omega_\gamma$  à partir de la température du FDC,  $T_{\text{FDC}}$ , en utilisant  $H_0 = 67,3$  (Planck Collaboration et al., 2013b), puis nous obtenons  $\Omega_r$  en postulant 3,04 neutrinos effectifs. Nous calculons enfin  $\Omega_b$  grâce à la mesure de  $\Omega_b h^2$  rapportée par Planck. La figure 5.39 présente les contraintes obtenues pour un modèle  $\Lambda$ CDM ouvert. Egalement représentés sur les graphiques sont les contours obtenus d'après les mesures de  $D_V/r_s$  rapportées par différentes analyses BAO. Les contours obtenus avec notre analyse à  $z = 2,38$  sont orthogonaux aux contours obtenus à plus bas décalage spectral. Ils sont les seuls à nécessiter par eux mêmes la présence d'énergie noire en imposant  $\Omega_\Lambda > 0,5$ .

La figure 5.40 présente les contours obtenus pour un modèle d'univers  $w$ CDM plat, où  $w$  est le rapport de la pression sur la densité de l'énergie noire, comme défini en section 1.3.4. Comme pour la figure 5.39, ces contours incluent la valeur de  $H_0$  de Planck. Nos contours sont les seuls contraintes BAO à imposer  $w < 0$ .

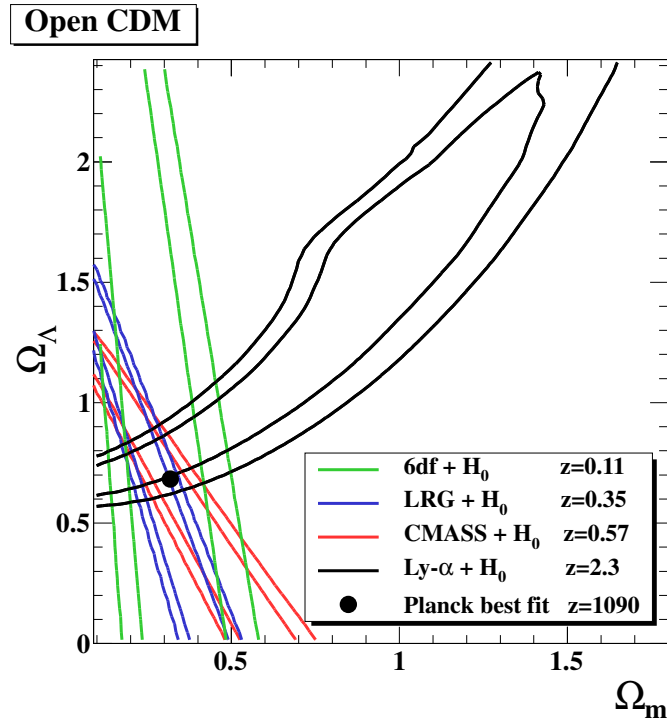


FIGURE 5.39: Contraintes sur la densité de matière  $\Omega_m$  et la densité d'énergie noire  $\Omega_\Lambda$  dans le cadre du modèle  $\Lambda$ CDM ouvert. Les contours représentent les intervalles de confiance à une et deux déviations standards. Les contours noirs sont obtenus à partir des contours de la figure 5.30 et de la mesure de  $H_0$  de Planck (Planck Collaboration et al., 2013b). Les autres contours sont obtenus à partir des mesures de  $D_V/r_s$  d'analyse BAO sur d'autres catalogues à plus bas décalages spectraux, également combinés avec la mesure de  $H_0$  de Planck : 6df à  $z = 0,11$  (Beutler et al., 2011), LRG à  $z = 0,35$  (Percival et al., 2010) et CMASS à  $z = 0,57$  (Anderson et al., 2012). Le point noir correspond au meilleur ajustement des  $C_l$  du FDC (Planck Collaboration et al., 2013b).



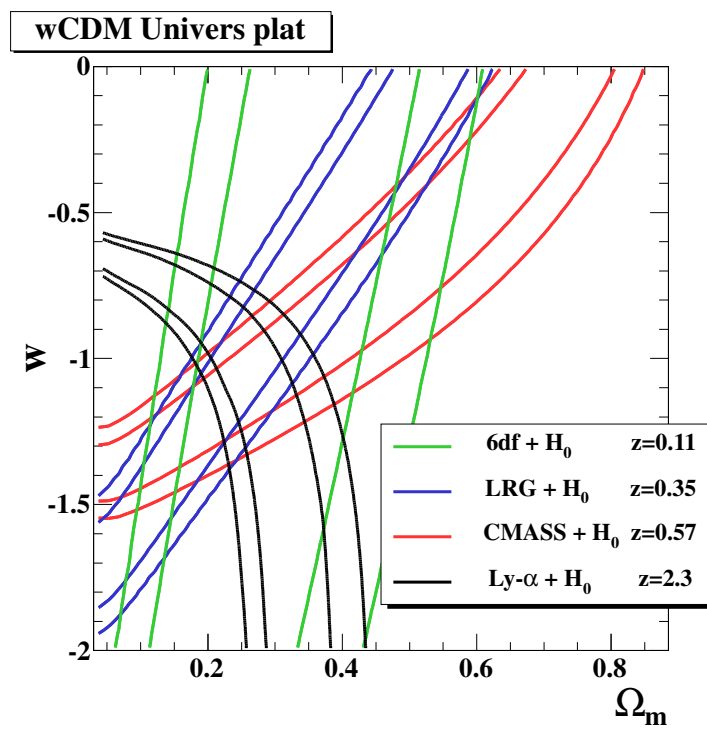


FIGURE 5.40: Contraintes sur la densité de matière  $\Omega_m$  et le rapport de la pression sur la densité de l'énergie noire  $w$  obtenues en imposant  $\Omega_k = 0$ . Voir la légende de la figure 5.39 pour plus d'informations.

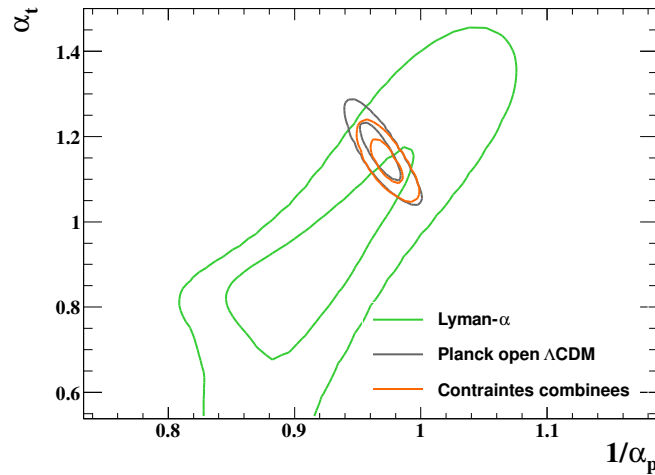


FIGURE 5.41: Contours à un et deux écarts standards dans le plan  $(\alpha_t, 1/\alpha_p)$  pour un décalage spectral  $z = 2,38$ . Les contours verts correspondent à cette analyse et les contours gris aux contraintes imposées par Planck dans le cadre d'un univers  $\Lambda$ CDM ouvert. Les contours oranges sont la combinaison des contraintes individuelles.

### 5.7.3 Contraintes combinées avec Planck

La collaboration Planck met à disposition différentes chaînes de Markov pour différents modèles cosmologiques<sup>5</sup>. Ces chaînes permettent de modéliser la fonction de vraisemblance dans l'espace des paramètres  $(\alpha_t, \alpha_p)$  à un décalage spectral  $z = 2,38$ . Il est ensuite possible d'estimer le niveau de confiance de ces paramètres d'après la démarche décrite en section 5.5.4.

La figure 5.41 montre les contours à un et deux écarts standards obtenus par Planck dans le plan  $(\alpha_t, 1/\alpha_p)$  pour un univers  $\Lambda$ CDM ouvert. C'est contours sont évidemment bien plus petits que ceux obtenus par notre analyse, si bien que la combinaison des deux résultats (en orange sur la figure) n'améliore que peu les contraintes. Les contours combinés sont légèrement décentrés par rapport au modèle fiduciel (basé sur les résultats de WMAP7), pointant vers une plus grande valeur de  $D_A/r_s$  et une plus faible valeur de  $Hr_s$ . La figure 5.42 montre les contraintes obtenues par Planck pour un modèle d'univers «  $w_0w_a$ CDM » plat, c'est à dire pour lequel l'équation d'état de l'énergie noire est déterminée par deux paramètres  $w_0$  et  $w_a$  (voir section 1.3.4). Elles sont obtenues en utilisant la mesure de  $H_0$  de Riess et al. (2011), basée sur l'observation de céphéides et de supernovae proches par le télescope spatial Hubble. Ces contours sont légèrement dégénérés dans la direction  $1/\alpha_p \propto Hr_s$ . La combinaison avec les contraintes de notre analyse lève cette dégénérescence.

Les contours obtenus par Planck pour un modèle d'univers  $\Lambda$ CDM ouvert, et pour un modèle  $w_0w_a$ CDM plat combinés avec une mesure de  $H_0$ , sont en légère tension. Ce résultat n'est pas étonnant puisque les valeurs de  $H_0$  obtenues par Planck (dans le cadre du modèle  $\Lambda$ CDM) et par Riess et al. (2011) sont elles mêmes en tension. Il est intéressant de noter que les contours obtenus par Planck (pour un modèle d'univers  $\Lambda$ CDM ouvert ou dans le cadre d'un modèle  $w_0w_a$ CDM plat), sont orthogonaux aux contours obtenus par notre analyse. Nous pouvons donc nous attendre à ce que la combinaison des résultats Planck et de cette analyse puisse contraindre le modèle

5. <http://pla.esac.esa.int/pla/aio/planckProducts.html>

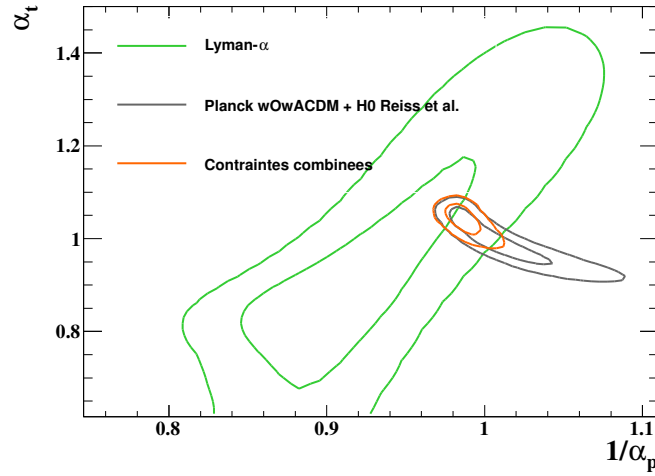


FIGURE 5.42: Contours à un et deux écarts standards dans le plan  $(\alpha_t, 1/\alpha_p)$  pour un décalage spectral  $z = 2,38$ . Les contours verts correspondent à cette analyse et les contours gris aux contraintes imposées par Planck dans le cadre d'un univers  $w_o w_a$ CDM plat. Les contours oranges sont la combinaison des contraintes individuelles.

$w_o w_a$ CDM ouvert, sans avoir recourt à une mesure de  $H_0$ . Malheureusement, la collaboration Planck n'a pas rendu publique de chaîne de Markov pour un tel modèle.

## 5.8 Perspectives

Nous listons ici les améliorations que nous souhaitons apporter à l'analyse dans le but d'en améliorer la mesure de la position du pic BAO et donc les contraintes sur les paramètres cosmologiques :

- Il est possible d'inclure les lignes de visée des quasars présentant un (ou plusieurs) DLA en masquant la région de plus forte absorption et en corrigeant l'impact du DLA sur le reste de la forêt (Slosar et al., 2013). Il n'y a toutefois pas eu d'étude montrant qu'une telle correction n'a pas d'effets indésirables sur la mesure de la fonction de corrélation. Une telle étude nécessite des simulations comprenant des DLA reproduisant fidèlement les données. Bien que nos simulations incorporent des DLA, leur nombre, répartition spatiale, ainsi que leurs profils ne sont pas représentatifs des données. Par ailleurs l'étude nécessite de traiter ces simulations et les données de manière identique. Il est bien évident que nous ne pouvons mettre en place une inspection visuelle sur les simulations. Il est donc nécessaire d'avoir recourt à une détection automatique des DLA (Noterdaeme et al., 2009).
- Il est également possible d'inclure les spectres des quasars de type BAL, en masquant la région de forte absorption (Slosar et al., 2013). Une fois encore, il est nécessaire d'étudier l'impact de l'utilisation de ces spectres sur la mesure de la fonction de corrélation et donc d'avoir recours aux simulations. Nos simulations actuelles ne contiennent pas de spectres de quasars de type BAL.
- Nous travaillons actuellement à la mise en place d'un ajustement s'appliquant directement à la fonction de corrélation en deux dimensions  $\xi(r, \mu)$  (et non aux multipôles  $\xi_0(r)$  et  $\xi_2(r)$ ) (Kirkby et al., 2013). Un tel ajustement nécessite de calculer la matrice de covariance

complète de  $\xi(r, \mu)$  de dimension  $2.500 \times 2.500$ , mais permet de prendre correctement en compte les corrélations le long de la dimension  $\mu$ . Il permet également de mieux modéliser la distorsion de la fonction de corrélation introduite par la méthode d'estimation du flux moyen.

- Pour terminer, l'étude décrite ici est basée sur environ deux tiers du relevé final, qui devrait être achevé d'ici la fin de l'année 2013. L'application de l'analyse au relevé final ne devrait poser aucune difficulté.



# Conclusion

La collaboration SDSS III est un cadre privilégié pour débiter la recherche. Elle regroupe nombre de physiciens parmi les plus influents de la cosmologie actuelle et possède le plus grand catalogue d'objets astrophysiques disponible à ce jour, tant en photométrie qu'en spectroscopie. Les retombées scientifiques des deux premières générations du SDSS ont été majeures ; Nul doute qu'il en sera de même pour la troisième génération.

Mon travail de thèse s'est axé sur deux volets bien différents du Baryon Oscillation Spectroscopic Survey. Dans un premier temps, j'ai participé au développement d'une méthode de sélection des quasars par variabilité. Cette méthode s'est avérée bien plus efficace que les sélections usuelles par couleurs, lorsqu'un nombre important de données photométriques multi-époques sont disponibles. De ce fait, elle a été appliquée comme seule méthode de sélection des quasars sur la région du *Stripe 82*, où de telles données sont disponibles. Elle a également été appliquée à l'ensemble du relevé où environ 30% des objets disposent d'information multi-époque, en complément des sélections standards, à la recherche des quasars les plus variables.

A la lumière de cette étude, il devient clair que la sélection par variabilité est l'un des enjeux des futurs relevés spectroscopiques de quasars. Le croisement des données de différents relevés photométriques est une solution pour une application à court terme de la sélection par variabilité à de grandes régions du ciel. Il représente toutefois un défi du point de vue de l'intercalibration photométrique et astrométrique des relevés.

Le second volet de ce travail concerne la mise en place d'une analyse dédiée au calcul de la fonction d'auto-corrélation des forêts Lyman- $\alpha$  des spectres de quasars de BOSS. Grâce à la qualité exceptionnelle des données, à l'effort important fourni pour leur réduction, ainsi qu'à une météo particulièrement clémente, cette analyse a débouché sur la première détection des oscillations acoustiques de baryons à grand décalage spectral dans la distribution spatiale de l'hydrogène neutre du milieu intergalactique.

Il est remarquable que malgré une mesure à 2% de la position du pic BAO, nous ne puissions mettre en évidence la présence d'effets systématiques sur notre mesure. Cette faible contamination par des effets systématiques est l'un des principaux atouts de l'utilisation des BAO pour contraindre les paramètres cosmologiques. Notre mesure permet notamment de contraindre les densités de matière et d'énergie noire, ainsi que le paramètre de l'équation d'état de l'énergie noire. Il est intéressant de noter que les contraintes obtenues sont orthogonales aux contraintes tirées des mesures BAO à plus bas décalages spectraux.

Outre l'augmentation de la statistique, différentes améliorations vont permettre, à court terme, de renforcer ces contraintes cosmologiques. Parmi les plus importantes figure l'utilisation des DLA et BAL, après corrections des effets indésirables. L'analyse de la fonction de corrélation directement en deux dimensions rend difficile l'estimation de la matrice de covariance, mais permettra de mieux contraindre la distorsion introduite par la méthode d'estimation du flux moyen attendu. La mise en place de telles améliorations est en partie rendue possible par le développement de simulations de plus en plus réalistes.



## Annexe A

# Métrie de Friedmann-Lemaître-Robertson-Walker

---

Dans cette annexe nous cherchons une expression de la métrie commune aux trois types d'espace autorisés par le principe cosmologique : la 3-sphère, l'espace plat et le 3-hyperboloïde.

---

### A.1 Les trois espaces possibles

Considérons la partie spatiale de l'espace-temps, c'est-à-dire les trois dimensions d'espaces. La métrie est l'outil permettant d'associer une distance  $dl$  à une différence de coordonnées. Pour trouver l'une des expressions possible de la métrie de chacun des espaces à 3 dimensions, imaginons un espace euclidien de dimension 4. Il est possible d'attribuer à chaque point de cet espace des coordonnées euclidiennes  $x_i$ , ( $i = 1$  à 4). La distance séparant deux points situés en  $x_i$  et  $x_i + dx_i$  est alors donnée par :

$$dl^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 \quad (\text{A.1})$$

#### A.1.1 3-sphère

Cherchons maintenant l'expression de la distance  $dl$  séparant deux points situés sur une 3-sphère. Dans un espace euclidien à 4 dimensions, une 3-sphère de rayon  $R$  est définie par l'équation :

$$R^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 \quad (\text{A.2})$$

où les  $x_i$  ( $i = 1$  à 4) représentent les coordonnées sur les 4 axes du repère euclidien. Il est possible de définir  $(\chi, \theta, \phi)$  pour passer en coordonnées sphériques de sorte que :

$$\begin{cases} x_1 = R \cos \chi \\ x_2 = R \sin \chi \cos \theta \\ x_3 = R \sin \chi \sin \theta \cos \phi \\ x_4 = R \sin \chi \sin \theta \sin \phi \end{cases} \quad (\text{A.3})$$



La longueur  $dl$  séparant deux points de la 3-sphère situés en  $(\chi, \theta, \phi)$  et  $(\chi + d\chi, \theta + d\theta, \phi + d\phi)$  est alors donnée par l'équation A.1 qui, en coordonnées sphériques et après simplification, donne :

$$dl^2 = R^2 (d\chi^2 + \sin^2\chi (d\theta^2 + \sin^2\theta d\phi^2)) \quad (\text{A.4})$$

### A.1.2 3-hyperboloïde

Cherchons de même l'expression de la distance  $dl$  séparant deux points situés sur un 3-hyperboloïde. Dans un espace euclidien à 4 dimensions, un 3-hyperboloïde de rayon  $R$  est défini par l'équation :

$$R^2 = x_1^2 - x_2^2 - x_3^2 - x_4^2 \quad (\text{A.5})$$

Comme pour la 3-sphère, il est possible de définir  $(\chi, \theta, \phi)$  pour passer en coordonnées sphériques :

$$\begin{cases} x_1 = R \cosh\chi \\ x_2 = R \sinh\chi \cos\theta \\ x_3 = R \sinh\chi \sin\theta \cos\phi \\ x_4 = R \sinh\chi \sin\theta \sin\phi \end{cases} \quad (\text{A.6})$$

La longueur  $dl$  séparant deux points du 3-hyperboloïde situés en  $(\chi, \theta, \phi)$  et  $(\chi + d\chi, \theta + d\theta, \phi + d\phi)$  est alors donnée par l'équation :

$$dl^2 = R^2 (d\chi^2 + \sinh^2\chi (d\theta^2 + \sin^2\theta d\phi^2)) \quad (\text{A.7})$$

la seule différence avec l'équation A.4 étant la présence de  $\sinh^2\chi$  à la place de  $\sin^2\chi$ .

### A.1.3 Espace euclidien

Pour un espace euclidien de dimension 3, la métrique est triviale en coordonnées euclidiennes :

$$dl^2 = dx_1^2 + dx_2^2 + dx_3^2 \quad (\text{A.8})$$

Il est toutefois utile d'utiliser les coordonnées sphériques pour obtenir une expression similaire aux équations A.4 et A.7. Le passage aux coordonnées sphériques pour l'espace euclidien ne nécessite pas de plongement dans un espace de dimension supérieure. Il suffit de définir :

$$\begin{cases} x_1 = R \cos\theta \\ x_2 = R \sin\theta \cos\phi \\ x_3 = R \sin\theta \sin\phi \end{cases} \quad (\text{A.9})$$

ce qui, une fois injecté dans l'équation A.8, permet d'obtenir la relation :

$$dl^2 = dR^2 + R^2 (d\theta^2 + \sin^2\theta d\phi^2) \quad (\text{A.10})$$

Nous obtenons ainsi trois expressions différentes pour l'expression de  $dl$  (et donc trois métriques différentes) pour la 3-sphère, l'espace euclidien, et le 3-hyperboloïde. Il est intéressant de noter que  $R$  est fixe et représente le rayon de courbure pour la 3-sphère et le 3-hyperboloïde, mais est une simple variable pour l'espace euclidien.

### A.1.4 Généralisation de la métrique

Il est possible de choisir une coordonnée temporelle  $t$  telle que l'expression de la métrique de l'espace-temps soit donnée par :

$$d\tau^2 = dt^2 - dl^2 \quad (\text{A.11})$$

où l'expression de  $dl^2$  dépend de la variété de l'univers considéré. Dans le système de coordonnées  $(t, \chi, \theta, \phi)$ , l'expression de la métrique est donc :

$$d\tau^2 = dt^2 - a^2 \left( d\chi^2 + \begin{cases} \sin^2\chi \\ \chi^2 \\ \sinh^2\chi \end{cases} (d\theta^2 + \sin^2\theta d\phi^2) \right) \quad (\text{A.12})$$

où  $a = R$  pour la 3-sphère et le 3-hyperboloïde et où nous avons effectué le changement de variable  $R \rightarrow a\chi$  avec  $a \in \mathbb{R}_+^*$  pour un univers plat.

Dans le but d'obtenir une expression commune aux trois espaces homogènes et isotropes, nous effectuons les changements de variable suivants :

$$r = \begin{cases} \sin\chi & \text{3-sphère} \\ \chi & \text{espace plat} \\ \sinh\chi & \text{3-hyperboloïde} \end{cases} \quad (\text{A.13})$$

Nous obtenons alors une expression commune de la métrique pour les trois espaces :

$$d\tau^2 = dt^2 - a^2 \left( \frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \right) \quad (\text{A.14})$$

où

$$k = \begin{cases} +1 & \text{3-sphère} \\ 0 & \text{espace plat} \\ -1 & \text{3-hyperboloïde} \end{cases} \quad (\text{A.15})$$

et

$$\begin{cases} a = R & \text{3-sphère, 3-hyperboloïde} \\ a \in \mathbb{R}_+^* & \text{espace plat} \end{cases} \quad (\text{A.16})$$

Il s'agit de la métrique de Friedmann-Lemaître-Robertson-Walker.



## **Annexe B**

# **Publication DR9**

Nous reproduisons ici la publication présentant le travail effectué sur les données DR9.

## Baryon Acoustic Oscillations in the Ly $\alpha$ forest of BOSS quasars

Nicolás G. Busca<sup>1</sup>, Timothée Delubac<sup>2</sup>, James Rich<sup>2</sup>, Stephen Bailey<sup>3</sup>, Andreu Font-Ribera<sup>3,27</sup>, David Kirkby<sup>4</sup>, J.-M. Le Goff<sup>2</sup>, Matthew M. Pieri<sup>5</sup>, Anže Slosar<sup>6</sup>, Éric Aubourg<sup>1</sup>, Julian E. Bautista<sup>1</sup>, Dmitry Bizyaev<sup>7</sup>, Michael Blomqvist<sup>4</sup>, Adam S. Bolton<sup>8</sup>, Jo Bovy<sup>9</sup>, Howard Brewington<sup>7</sup>, Arnaud Borde<sup>2</sup>, J. Brinkmann<sup>7</sup>, Bill Carithers<sup>3</sup>, Rupert A.C. Croft<sup>10</sup>, Kyle S. Dawson<sup>8</sup>, Garrett Ebelke<sup>7</sup>, Daniel J. Eisenstein<sup>11</sup>, Jean-Christophe Hamilton<sup>1</sup>, Shirley Ho<sup>10</sup>, David W. Hogg<sup>12</sup>, Klaus Honscheid<sup>13</sup>, Khee-Gan Lee<sup>14</sup>, Britt Lundgren<sup>15</sup>, Elena Malanushenko<sup>7</sup>, Viktor Malanushenko<sup>7</sup>, Daniel Margalá<sup>4</sup>, Claudia Maraston<sup>5</sup>, Kushal Mehta<sup>16</sup>, Jordi Miralda-Escudé<sup>17,18</sup>, Adam D. Myers<sup>19</sup>, Robert C. Nichol<sup>5</sup>, Pasquier Noterdaeme<sup>20</sup>, Matthew D. Olmstead<sup>8</sup>, Daniel Oravetz<sup>7</sup>, Nathalie Palanque-Delabrouille<sup>2</sup>, Kaike Pan<sup>7</sup>, Isabelle Pâris<sup>20,28</sup>, Will J. Percival<sup>5</sup>, Patrick Petitjean<sup>20</sup>, N. A. Roe<sup>3</sup>, Emmanuel Rollinde<sup>20</sup>, Nicholas P. Ross<sup>3</sup>, Graziano Rossi<sup>2</sup>, David J. Schlegel<sup>3</sup>, Donald P. Schneider<sup>21,22</sup>, Alaina Shelden<sup>7</sup>, Erin S. Sheldon<sup>6</sup>, Audrey Simmons<sup>7</sup>, Stephanie Snedden<sup>6</sup>, Jeremy L. Tinker<sup>12</sup>, Matteo Viel<sup>23,24</sup>, Benjamin A. Weaver<sup>12</sup>, David H. Weinberg<sup>25</sup>, Martin White<sup>3</sup>, Christophe Yèche<sup>2</sup>, Donald G. York<sup>26</sup>

<sup>1</sup> APC, Université Paris Diderot-Paris 7, CNRS/IN2P3, CEA, Observatoire de Paris, 10, rueA. Domon & L. Duquet, Paris, France

<sup>2</sup> CEA, Centre de Saclay, IRFU, F-91191 Gif-sur-Yvette, France

<sup>3</sup> Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

<sup>4</sup> Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA

<sup>5</sup> Institute of Cosmology and Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth, PO1 3FX, UK

<sup>6</sup> Bldg 510 Brookhaven National Laboratory, Upton, NY 11973, USA

<sup>7</sup> Apache Point Observatory, P.O. Box 59, Sunspot, NM 88349, USA

<sup>8</sup> Department of Physics and Astronomy, University of Utah, 115 S 1400 E, Salt Lake City, UT 84112, USA

<sup>9</sup> Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA

<sup>10</sup> Bruce and Astrid McWilliams Center for Cosmology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

<sup>11</sup> Harvard-Smithsonian Center for Astrophysics, Harvard University, 60 Garden St., Cambridge MA 02138, USA

<sup>12</sup> Center for Cosmology and Particle Physics, New York University, New York, NY 10003, USA

<sup>13</sup> Department of Physics and Center for Cosmology and Astro-Particle Physics, Ohio State University, Columbus, OH 43210, USA

<sup>14</sup> Max-Planck-Institut für Astronomie, Königstuhl 17, D69117 Heidelberg, Germany

<sup>15</sup> Department of Astronomy, University of Wisconsin, 475 North Charter Street, Madison, WI 53706, USA

<sup>16</sup> Steward Observatory, University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA

<sup>17</sup> Institució Catalana de Recerca i Estudis Avançats, Barcelona, Catalonia

<sup>18</sup> Institut de Ciències del Cosmos, Universitat de Barcelona/IEEC, Barcelona 08028, Catalonia

<sup>19</sup> Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

<sup>20</sup> Université Paris 6 et CNRS, Institut d'Astrophysique de Paris, 98bis Blvd. Arago, 75014 Paris, France

<sup>21</sup> Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA

<sup>22</sup> Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA

<sup>23</sup> INAF, Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34131 Trieste, Italy

<sup>24</sup> INFN/National Institute for Nuclear Physics, Via Valerio 2, I-34127 Trieste, Italy.

<sup>25</sup> Department of Astronomy, Ohio State University, 140 West 18th Avenue, Columbus, OH 43210, USA

<sup>26</sup> Department of Astronomy and Astrophysics and the Enrico Fermi Institute, The University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois, 60615, USA

<sup>27</sup> Institute of Theoretical Physics, University of Zurich, 8057 Zurich, Switzerland

<sup>28</sup> Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile

Received November 13th, 2012; accepted February 4th, 2013

### ABSTRACT

We report a detection of the baryon acoustic oscillation (BAO) feature in the three-dimensional correlation function of the transmitted flux fraction in the Ly $\alpha$  forest of high-redshift quasars. The study uses 48,640 quasars in the redshift range  $2.1 \leq z \leq 3.5$  from the Baryon Oscillation Spectroscopic Survey (BOSS) of the third generation of the Sloan Digital Sky Survey (SDSS-III). At a mean redshift  $z = 2.3$ , we measure the monopole and quadrupole components of the correlation function for separations in the range  $20h^{-1}\text{Mpc} < r < 200h^{-1}\text{Mpc}$ . A peak in the correlation function is seen at a separation equal to  $(1.01 \pm 0.03)$  times the distance expected for the BAO peak within a concordance  $\Lambda\text{CDM}$  cosmology. This first detection of the BAO peak at high redshift, when the universe was strongly matter dominated, results in constraints on the angular diameter distance  $D_A$  and the expansion rate  $H$  at  $z = 2.3$  that, combined with priors on  $H_0$  and the baryon density, require the existence of dark energy. Combined with constraints derived from Cosmic Microwave Background (CMB) observations, this result implies  $H(z = 2.3) = (224 \pm 8)\text{km s}^{-1}\text{Mpc}^{-1}$ , indicating that the time derivative of the cosmological scale parameter  $\dot{a} = H(z = 2.3)/(1 + z)$  is significantly greater than that measured with BAO at  $z \sim 0.5$ . This demonstrates that the expansion was decelerating in the range  $0.7 < z < 2.3$ , as expected from the matter domination during this epoch. Combined with measurements of  $H_0$ , one sees the pattern of deceleration followed by acceleration characteristic of a dark-energy dominated universe.

**Key words.** cosmology, Ly $\alpha$  forest, large scale structure, dark energy

## 1. Introduction

Baryon acoustic oscillations (BAO) in the pre-recombination universe have striking effects on the anisotropies of the Cosmic Microwave Background (CMB) and on the large scale structure (LSS) of matter (Weinberg et al. (2012) and references therein). The BAO effects were first seen in the series of peaks in the CMB angular power spectrum (de Bernardis et al., 2000). Subsequently, the BAO relic at redshift  $z \sim 0.3$  was seen (Eisenstein et al., 2005; Cole et al., 2005) as a peak in the galaxy-galaxy correlation function at a co-moving distance corresponding to the sound horizon at recombination. For the WMAP7 cosmological parameters (Komatsu et al., 2011), the expected comoving scale of the BAO peak is  $r_s = 153$  Mpc, with an uncertainty of  $\approx 1\%$ .

The BAO peak in the correlation function at a redshift  $z$  appears at an angular separation  $\Delta\theta = r_s/(1+z)D_A(z)$  and at a redshift separation  $\Delta z = r_s H(z)/c$ , where  $D_A$  and  $H$  are the angular distances and expansion rates. Measurement of the peak position at any redshift thus constrains the combinations of cosmological parameters that determine  $r_s H$  and  $r_s/D_A$ . While the possibility of measuring both combinations is beginning to be exploited (Chuang & Wang, 2012; Xu et al., 2012), most present measurements have concentrated on the combination  $D_V \equiv [(1+z)^2 D_A^2 cz/H]^2$ , which determines the peak position for an isotropic distribution of galaxy pairs and an isotropic clustering strength. The “BAO Hubble diagram”,  $D_V/r_s$  vs.  $z$ , now includes the Sloan Digital Sky Survey (SDSS) measurement (Eisenstein et al., 2005) updated to the DR7 (Abazajian et al., 2009) sample and combined with 2dF data (Percival et al., 2010), the 6dF point at  $z = 0.1$  (Beutler et al., 2011), the WiggleZ points at ( $0.4 < z < 0.8$ ) (Blake et al., 2011a), and a reanalysis of the SDSS DR7 sample that uses reconstruction (Eisenstein et al., 2007; Padmanabhan et al., 2009) to sharpen the precision of the BAO measurement (Padmanabhan et al., 2012; Mehta et al., 2012). Recently, the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) of SDSS-III (Eisenstein et al., 2011) has added a precise measurement at  $z \sim 0.57$  (Anderson et al., 2012). BOSS has also reported a measurement of  $D_A(z = 0.55)/r_s$  based on the BAO structure in the angular power spectrum of galaxies (Seo et al., 2012).

In this paper, we present an observation of the BAO peak at  $z \sim 2.3$  found in the flux correlation function of the Ly $\alpha$  forest of BOSS quasars. This is the first such observation at a redshift where the expansion dynamics is matter-dominated,  $z > 0.8$ . The possibility of such a measurement was suggested by McDonald (2003) and White (2003) and first studied in detail by McDonald & Eisenstein (2007). While the galaxy BAO measurements are most sensitive to  $D_V \propto D_A^{2/3} H^{-1/3}$ , the Ly $\alpha$  flux transmission is more sensitive to peculiar velocity gradient effects, which enhance redshift distortions and shift our sensitivity to the expansion rate. As we shall show below, the most accurately measured combination from the Ly $\alpha$  forest BAO peak is  $\propto D_A^{0.2} H^{-0.8}$ , and the present BOSS data set allows us to determine its value to a precision of 3.5%. Combining this result with constraints from CMB observations allows us to deduce the value of  $H(z = 2.3)$  accurate to 4%. Comparing our results with measurements of  $H_0$  and of  $H(0.2 < z < 0.8)$  reveals the expected sequence of deceleration and acceleration in models with dark energy.

The last decade has seen increasing use of Ly $\alpha$  absorption to investigate large scale structure. The number of quasars in early studies (Croft et al., 1999; McDonald et al., 2000; Croft et al., 2002; Viel et al., 2004; McDonald et al., 2006) was enough only

to determine the Ly $\alpha$  absorption correlation along individual lines of sight. With the BOSS project the surface density of quasars is sufficient to probe the full three-dimensional distribution of neutral hydrogen. A study using the first 10,000 BOSS quasars was presented by Slosar et al. (2011). This sample provided clear evidence for the expected long-range correlations, including the redshift-space distortions due to the gravitational growth of structure. With the SDSS data release DR9 (Ahn et al., 2012), we now have  $\sim 60,000$  quasars at  $z \sim 2.3$  (Pâris et al., 2012), with a high enough surface density to observe the BAO peak.

The use of Ly $\alpha$  absorption to trace matter has certain interesting differences from the use of galaxies. Galaxy surveys provide a catalog of positions in redshift space that correspond to points of high over-densities. On the other hand, the forest region of a quasar spectrum provides a complete mapping of the absorption over a  $\sim 400 h^{-1}$  Mpc (comoving) range starting about  $100 h^{-1}$  Mpc in front of the quasar (so as to avoid the necessity of modeling the quasar’s Ly $\alpha$  emission line). To the extent that quasar lines of sight are random, a large collection of quasars can provide a nearly unbiased sample of points where the absorption is measured. Cosmological simulations (Cen et al., 1994; Petitjean et al., 1995; Zhang et al., 1995; Hernquist et al., 1996; Miralda-Escudé et al., 1996; Theuns et al., 1998) indicate that most of the Ly $\alpha$  absorption is due to cosmic filamentary structures with overdensities of order one to ten, much lower than the overdensities of virialized halos sampled by galaxies. These simulations have also indicated that, on large scales, the mean Ly $\alpha$  absorption is a linear tracer of the mass overdensity (Croft et al., 1997, 1998; Weinberg et al., 1998; McDonald et al., 2000; McDonald, 2003), implying a relation of the power spectrum of the measured absorption to that of the underlying mass fluctuations. Finally, the forest is observable in a redshift range inaccessible to current large galaxy surveys and where theoretical modeling is less dependent on non-linear effects in cosmological structure formation. These factors combine to make Ly $\alpha$  absorption a promising tracer of mass that is complementary to galaxy tracers.

With Ly $\alpha$  forest measurements along multiple sightlines, one can attempt to reconstruct the underlying 3-dimensional mass density field (Nusser & Haehnelt, 1999; Pichon et al., 2001; Gallerani et al., 2011), from which one can investigate topological characteristics (Caucci et al., 2008) or the power spectrum (Kitaura et al., 2012). However, the BOSS sample is fairly sparse, with a typical transverse sightline separation  $\sim 15 h^{-1}$  Mpc (comoving), and the signal-to-noise ratio in individual spectra is low (see Fig. 2 below), which makes it poorly suited to such reconstruction techniques. In this paper we take the more direct approach of measuring the BAO feature in the correlation function of transmitted flux fraction (the ratio of the observed flux to the flux expected in the absence of absorption).

The use of the Ly $\alpha$  forest is handicapped by the fact that not all fluctuations of the transmitted flux fraction are due to fluctuations of the density of hydrogen. Because the neutral hydrogen density is believed to be determined by photo-ionization equilibrium with the flux of UV photons from stars and quasars, variations of the UV flux may contribute to fluctuations of the transmitted flux (Worseck & Wisotzki, 2006). In addition, hydrogen can be self-shielded to ionizing photons, leading to much higher neutral fractions and damping wings in strong absorption systems (Font-Ribera et al., 2012b). Metals present in the intergalactic medium provide additional absorption superimposed on the Ly $\alpha$  absorption (Pieri et al., 2010). A further complication results from the fact that the flux correlation function uses the

fluctuations of the transmitted flux about its mean value; this requires an estimate of the product of the mean absorption (as a function of absorber redshift) and the unabsorbed flux for individual quasars (Le Goff et al., 2011).

All of these complications are important for a complete understanding of the statistics of the Ly $\alpha$  forest. Fortunately, they are not expected to produce a sharp peak-like feature in the correlation function, so the interpretation of a peak as due to BAO should give robust constraints on the cosmological parameters.

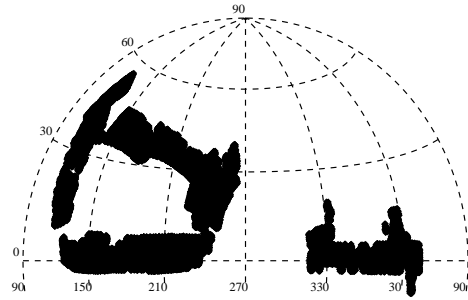
The BOSS collaboration has performed three independent analyses to search for BAO in Ly $\alpha$  forest. This paper presents two of them, both of which aim to analyze the forest using simple procedures at the expense of some loss of sensitivity. The third analysis, with the goal of an optimal measurement of the flux correlation function with a more complex method, is described in a separate publication (Slosar et al., 2013).

Our methods are tested extensively on a set of detailed mock catalogs of the BOSS Ly $\alpha$  forest data set. These mock catalogs, which use the method presented by Font-Ribera et al. (2012a), will be described in detail in a forthcoming public release paper (S. Bailey et al., in preparation). In addition, the BOSS collaboration have also released a fiducial version of the DR9 Ly $\alpha$  forest spectra (Lee et al., 2012b), with various per-object products including masks, continua, and noise correction vectors designed to aid in Ly $\alpha$  forest analysis. While our analyses implement their own sample selection criteria and continuum determination procedures, we have also applied our measurement to this Lee et al. (2012b) sample.

This paper is organized as follows. Section 2 presents the BOSS quasar sample used in this analysis and the procedure used to produce the quasar spectra. Section 3 describes the analysis to measure the correlation function. Section 4 derives the monopole and quadrupole components of the correlation function and determines the significance of the peak observed in these functions at the BAO scale. The cosmological implications of our detection of a BAO peak are discussed in Section 5. Finally, Appendix A provides a brief description of the mock spectra used to test our methodology and Appendix B shows the result of our BAO measurement applied to the BOSS Ly $\alpha$  sample of Lee et al. (2012b).

## 2. The BOSS quasar sample and data reduction

The BOSS project (Dawson et al., 2013) of SDSS-III (Eisenstein et al., 2011) is obtaining the spectra of  $\sim 1.6 \times 10^6$  luminous galaxies and  $\sim 150,000$  quasars. The project uses upgraded versions of the SDSS spectrographs (Smee et al., 2012) mounted on the Sloan 2.5-meter telescope (Gunn et al., 2006) at Apache Point, New Mexico. BOSS galaxy and quasar spectroscopic targets are selected using algorithms based primarily on photometry from the SDSS camera (Gunn et al., 1998; York et al., 2000) in the *ugriz* bands (Fukugita et al., 1996; Smith et al., 2002) reduced and calibrated as described by Stoughton et al. (2002), Pier et al. (2003), and Padmanabhan et al. (2008). Targets are assigned to fibers appropriately positioned in the  $3^\circ$  diameter focal plane according to a specially designed tiling algorithm (Blanton et al., 2003). Fibers are fixed in place by a pierced metal plate drilled for each observed field and fed to one of two spectrographs. Each exposed plate generates 1,000 spectra covering wavelengths of 360 to 1000 nm with a resolving power ranging from 1500 to 3000 (Smee et al., 2012). A median of 631 of these fibers are assigned to galactic targets and 204 to quasar targets. The BOSS spectroscopic targets are observed in dark and gray time,



**Fig. 1.** Hammer-Aitoff projection in equatorial coordinates of the BOSS DR9 footprint. The observations cover  $\sim 3000 \text{ deg}^2$ .

while the bright-time is used by other SDSS-III surveys (see Eisenstein et al. 2011).

The quasar spectroscopy targets are selected from photometric data with a combination of algorithms (Richards et al. 2009; Yeche et al. 2009; Kirkpatrick et al. 2011; Bovy et al. 2011; Palanque-Delabrouille et al. 2011; for a summary, see Ross et al. (2012)). The algorithms use SDSS fluxes and, for SDSS Stripe 82, photometric variability. When available, we also use data from non-optical surveys (Bovy et al., 2012): the GALEX survey (Martin et al., 2005) in the UV; the UKIDSS survey (Lawrence et al., 2007) in the NIR, and the FIRST survey (Becker et al., 1995) in the radio.

The quasar spectroscopy targets are divided into two samples “CORE” and “BONUS”. The CORE sample consists of 20 quasar targets per square degree selected from SDSS photometry with a uniform algorithm, for which the selection efficiency for  $z > 2.1$  quasars is  $\sim 50\%$ . The selection algorithm for the CORE sample (Bovy et al., 2011) was fixed at the end of the first year of the survey, thus making it useful for studies that require a uniform target selection across the sky. The BONUS sample was chosen from a combination of algorithms with the purpose of increasing the density on the sky of observed quasars beyond that of the CORE sample. The combined samples yield a mean density of identified quasars of  $15 \text{ deg}^{-2}$  with a maximum of  $20 \text{ deg}^{-2}$ , mostly in zones where photometric variability, UV, and/or NIR data are available. The combined BONUS plus CORE sample can be used for Ly $\alpha$  BAO studies, which require the highest possible quasar density in a broad sky area but are insensitive to the uniformity of the quasar selection criteria because the structure being mapped is in the foreground of these quasar back-lights.

The data presented here consist of the DR9 data release (Ahn et al., 2012) covering  $\sim 3000 \text{ deg}^2$  of the sky shown in figure 1. These data cover about one-third of the ultimate BOSS footprint.

The data were reduced with the SDSS-III pipeline as described in Bolton et al. (2012). Typically four exposures of 15 minutes were co-added in pixels of wavelength width  $\sim 0.09 \text{ nm}$ . Besides providing flux calibrated spectra, the pipeline provided preliminary object classifications (galaxy, quasar, star) and redshift estimates.

The spectra of all quasar targets were visually inspected, as described in Pâris et al. (2012), to correct for misidentifications or inaccurate redshift determinations and to flag

broad absorption lines (BAL). Damped Ly $\alpha$  troughs are visually flagged, but also identified and characterized automatically (Noterdaeme et al., 2012). The visual inspection of DR9 confirms 60,369 quasars with  $2.1 \leq z_q \leq 3.5$ . In order to simplify the analysis of the Ly $\alpha$  forest, we discarded quasars with visually identified BALs and DLAs, leaving 48,640 quasars.

For the measurement of the flux transmission, we use the rest-frame wavelength interval

$$104.5 < \lambda_{\text{rf}} < 118.0 \text{ nm} . \quad (1)$$

The range is bracketed by the Ly $\alpha$  and Ly $\beta$  emission lines at 121.6 and 102.5 nm. The limits are chosen conservatively to avoid problems of modeling the shapes of the two emission lines and to avoid quasar proximate absorbers. The absorber redshift,  $z = \lambda/\lambda_{\text{Ly}\alpha} - 1$ , is in the range  $1.96 < z < 3.38$ . The lower limit is set by the requirement that the observed wavelength be greater than 360 nm below which the system throughput is less than 10% its peak value. The upper limit comes from the maximum quasar redshift of 3.5, beyond which the BOSS surface density of quasars is not sufficient to be useful. The distribution of absorber redshift is shown in figure 2 (top panel). When given the weights used for the calculation of the correlation function (section 3.3), the absorbers have a mean redshift of  $\langle z \rangle = 2.31$ .

For the determination of the correlation function, we use “analysis pixels” that are the flux average over three adjacent pipeline pixels. Throughout the rest of this paper, “pixel” refers to analysis pixels unless otherwise stated. The effective width of these pixels is  $210 \text{ km s}^{-1}$ , i.e. an observed-wavelength width  $\sim 0.27 \text{ nm} \sim 2 \text{ h}^{-1} \text{ Mpc}$ . The total sample of 48,640 quasars thus provides  $\sim 8 \times 10^6$  measurements of Ly $\alpha$  absorption over a total volume of  $\sim 20 \text{ h}^{-3} \text{ Gpc}^3$ .

Figure 2 (bottom panel) shows the distribution of the signal-to-noise ratio for pixels averaged over the forest region. The relatively modest mean value of 5.17 reflects the exposure times necessary to acquire such a large number of spectra.

In addition to the BOSS spectra, we analyzed 15 sets of mock spectra that were produced by the methods described in appendix A. These spectra do not yet reproduce all of the characteristics of the BOSS sample, but they are nevertheless useful for a qualitative understanding of the shape of the measured correlation function. More importantly, they are useful for understanding the detectability of a BAO-like peak and the precision of the measurement of its position.

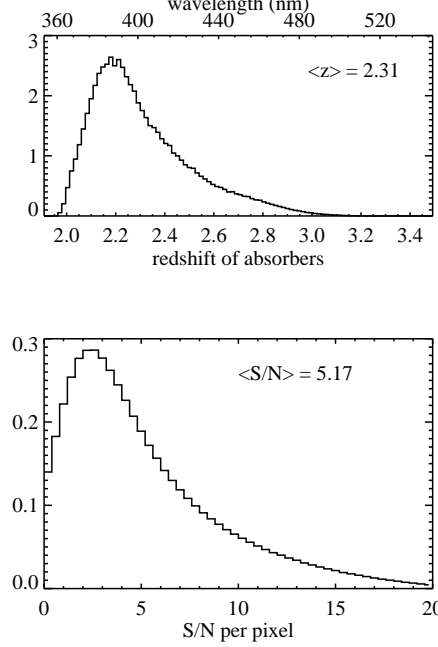
### 3. Measurement of the correlation function

The flux correlation function can be determined through a simple two-step process. In the first step, for each pixel in the forest region (equation 1) of quasar  $q$ , the measured flux  $f_q(\lambda)$  at observed wavelength  $\lambda$  is compared with the mean expected flux,  $C_q(\lambda)\overline{F}(z)$ , thus defining the “delta field”:

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{C_q(\lambda)\overline{F}(z)} - 1 . \quad (2)$$

Here,  $C_q(\lambda)$  is the unabsorbed flux (the so-called “continuum”) and  $\overline{F}(z)$  is the mean transmitted fraction at the HI absorber redshift. The quantities  $\lambda$  and  $z$  in equation 2 are not independent but related via  $z = \lambda/\lambda_{\text{Ly}\alpha} - 1$ .

Figure 3 shows an example of an estimation for  $C_q(\lambda)$  (blue line) and  $C_q\overline{F}$  (red line). Our two methods for estimating  $C_q$  and  $\overline{F}$  are described in sections 3.1 and 3.2.



**Fig. 2.** Top: weighted distribution of absorber redshifts used in the calculation of the correlation function in the distance range  $80 \text{ h}^{-1} \text{ Mpc} < r < 120 \text{ h}^{-1} \text{ Mpc}$ . Bottom: distribution of signal-to-noise ratio for analysis pixels (triplets of pipeline pixels) averaged over the forest region.

In the second step, the correlation function is calculated as a weighted sum of products of the deltas:

$$\hat{\xi}_A = \sum_{i,j \in A} w_{ij} \delta_i \delta_j / \sum_{i,j \in A} w_{ij} , \quad (3)$$

where the  $w_{ij}$  are weights and each  $i$  or  $j$  indexes a measurement on a quasar  $q$  at wavelength  $\lambda$ . The sum over  $(i, j)$  is understood to run over all pairs of pixels of all pairs of quasars within  $A$  defining a region in space of pixel separations,  $\mathbf{r}_i - \mathbf{r}_j$ . The region  $A$  is generally defined by a range  $r_{\text{min}} < r < r_{\text{max}}$  and  $\mu_{\text{min}} < \mu < \mu_{\text{max}}$  with:

$$r = |\mathbf{r}_i - \mathbf{r}_j| \quad \mu = \frac{(\mathbf{r}_i - \mathbf{r}_j)_\parallel}{r} \quad (4)$$

where  $(\mathbf{r}_i - \mathbf{r}_j)_\parallel$  is the component along the line of sight. Separations in observational pixel coordinates  $(\text{ra}, \text{dec}, z)$  are transformed to  $(r, \mu)$  in units of  $\text{h}^{-1} \text{ Mpc}$  by using a  $\Lambda$ CDM fiducial cosmology with matter and vacuum densities of

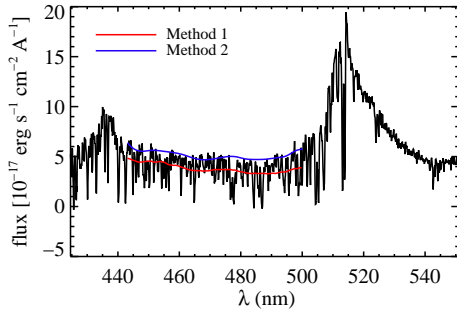
$$(\Omega_M, \Omega_\Lambda) = (0.27, 0.73) . \quad (5)$$

In the sum (3), we exclude pairs of pixels from only one quasar to avoid the correlated errors in  $\delta_i$  and  $\delta_j$  coming from the estimate of  $C_q$ . Note that the weights in eq. 3 are set to zero for pixels flagged by the pipeline as having problems due, e.g., to sky emission lines or cosmic rays.



A procedure for determining  $\xi$  is defined by its method for estimating the expected flux  $C_q \bar{F}$  and by its choice of weights,  $w_{ij}$ . The two methods described here use the same technique to calculate weights but have different approaches to estimate  $C_q \bar{F}$ . We will see that the two methods produce correlation functions that have no significant differences. However, the two independent codes were invaluable for consistency checks throughout the analysis.

The two methods were “blind” to the extent that many of the procedures were defined during tests either with mock data or with the real data in which we masked the region of the peak in the correlation function. Among those aspects fixed in this way were the quasar sample, the continuum determination, the weighting, the extraction of the monopole and quadrupole correlation function and the determination of the peak significance (section 4). This early freezing of procedures resulted in some that are suboptimal but which will be improved in future analyses. We note, however, that the procedures used to extract cosmological information (section 5) were decided on only after de-masking the data.



**Fig. 3.** An example of a BOSS quasar spectrum of redshift 3.239. The red and blue lines cover the forest region used here,  $104.5 < \lambda_{\text{rf}} < 118.0$ . This region is sandwiched between the quasar’s Ly $\beta$  and Ly $\alpha$  emission lines respectively at 435 and 515 nm. The blue line is an estimate of the continuum (unabsorbed flux) by method 2 and the red line is the estimate of the product of the continuum and the mean absorption by method 1.

### 3.1. Continuum fits, method 1

Both methods for estimating the product  $C_q \bar{F}$  assume that  $C_q$  is, to first approximation, proportional to a universal quasar spectrum that is a function of rest-frame wavelength,  $\lambda_{\text{rf}} = \lambda / (1 + z_q)$  (for quasar redshift  $z_q$ ), multiplied by a mean transmission fraction that slowly varies with absorber redshift. Following this assumption, the universal spectrum is found by stacking the appropriately normalized spectra of quasars in our sample, thus averaging out the fluctuating Ly $\alpha$  absorption. The product  $C_q \bar{F}$  for individual quasars is then derived from the universal spectrum by normalizing it to account for the quasar’s mean forest flux and then modifying its slope to account for spectral-index diversity and/or photo-spectroscopic miscalibration.

Method 1 estimates directly the product  $C_q \bar{F}$  in equation 2. An example is given by the red line in figure 3. The estimate is made by modeling each spectrum as

$$C_q \bar{F} = a_q \left( \frac{\lambda}{\langle \lambda \rangle} \right)^{b_q} \bar{f}(\lambda_{\text{rf}}, z) \quad (6)$$

where  $a_q$  is a normalization,  $b_q$  a “deformation parameter”, and  $\langle \lambda \rangle$  is the mean wavelength in the forest for the quasar  $q$  and  $\bar{f}(\lambda_{\text{rf}}, z)$  is the mean normalized flux obtained by stacking spectra in bins of width  $\Delta z = 0.1$ :

$$\bar{f}(\lambda_{\text{rf}}, z) = \sum_q w_q f_q(\lambda) / f_q^{128} / \sum_q w_q. \quad (7)$$

Here  $z$  is the redshift of the absorption line at observed wavelength  $\lambda$  ( $z = \lambda / \lambda_{\text{Ly}\alpha} - 1$ ),  $f_q$  is the observed flux of quasar  $q$  at wavelength  $\lambda$  and  $f_q^{128}$  is the average of the flux of quasar  $q$  for  $127.5 < \lambda_{\text{rf}} < 128.5$  nm. The weight  $w_q(\lambda)$  is given by  $w_q^{-1} = 1 / [\text{ivar}(\lambda) \cdot (f_q^{128})^2] + \sigma_{f_{\text{flux, LSS}}}^2$ . The quantity  $\text{ivar}$  is the pipeline estimate of the inverse flux variance in the pixel corresponding to wavelength  $\lambda$ . The quantity  $\sigma_{f_{\text{flux, LSS}}}^2$  is the contribution to the variance in the flux due to the LSS. We approximate it by its value at the typical redshift of the survey,  $z \sim 2.3$ :  $\sigma_{f_{\text{flux, LSS}}}^2 \sim 0.035$  (section 3.3).

Figure 4 shows the resulting mean  $\delta_i$  as a function of observed wavelength. The mean fluctuates about zero with up to 2% deviations with correlated features that include the H and K lines of singly ionized calcium (presumably originating from some combination of solar neighborhood, interstellar medium and the Milky Way halo absorption) and features related to Balmer lines. These Balmer features are a by-product of imperfect masking of Balmer absorption lines in F-star spectroscopic standards, which are used to produce calibration vectors (in the conversion of CCD counts to flux) for DR9 quasars. Therefore such Balmer artifacts are constant for all fibers in a plate fed to one of the two spectrographs and so they are approximately constant for every ‘half-plate’.

If unsubtracted, the artifacts in figure 4 would lead to spurious correlations, especially between pairs of pixels with separations that are purely transverse to the line of sight. We have made a global correction by subtracting the quantity  $\langle \delta \rangle(\lambda)$  in figure 4 (un-smoothed) from individual measurements of  $\delta$ . This is justified if the variance of the artifacts from half-plate-to-half-plate is sufficiently small, as half-plate-wide deviations from our global correction could, in principle add spurious correlations.

We have investigated this variance both by measuring the Balmer artifacts in the calibration vectors themselves and by studying continuum regions of all available quasars in the DR9 sample. Both studies yield no detection of excess variance arising from these artifacts, but do provide upper limits. The study of the calibration vectors indicate that the square-root of the variance is less than 20% of the mean Balmer artifact deviations and the study of quasar spectra indicate that the square-root of the variance is less than 100% of the mean Balmer artifacts (and less than 50% of the mean calcium line deviations).

We then performed Monte Carlo simulations by adding a random sampling of our measured artifacts to our data to confirm that our global correction is adequate. We found that there is no significant effect on the determination of the BAO peak position, even if the variations are as large as that allowed in our tests.

### 3.2. Continuum fits, method 2

Method 1 would be especially appropriate if the fluxes had a Gaussian distribution about the mean absorbed flux,  $C_q \bar{F}$ . Since this is not the case, we have developed method 2 which explicitly uses the probability distribution function for the transmitted flux fraction  $F$ ,  $P(F, z)$ , where  $0 < F < 1$ . We use the  $P(F, z)$  that results from the log-normal model used to generate mock data (see appendix A).

Using  $P(F, z)$ , we can construct for each BOSS quasar the PDF of the flux in pixel  $i$ ,  $f_i$ , by assuming a continuum  $C_q(\lambda_i)$  and convolving with the pixel noise,  $\sigma_i$ :

$$P_i(f_i, C_q(\lambda_i), z_i) \propto \int_0^1 dF P(F, z_i) \exp\left[-\frac{(C_q F - f_i)^2}{2\sigma_i^2}\right]. \quad (8)$$

The continuum is assumed to be of the form

$$C_q(\lambda) = (a_q + b_q \lambda) \bar{f}(\lambda_{\text{rf}}) \quad (9)$$

where  $\bar{f}(\lambda_{\text{rf}})$  is the mean flux as determined by stacking spectra as follows:

$$\bar{f}(\lambda_{\text{rf}}) = \sum_q w_q(\lambda_{\text{rf}}) [f_q(\lambda_{\text{rf}}) / f_q^{128}] / \sum_q w_q \quad (10)$$

as in equation 7 except that here there is no redshift binning. The parameters  $a_q$  and  $b_q$  are then determined for each quasar by maximizing a likelihood given by

$$L(C_q) = \prod_i P_i[f_i, C_q(\lambda_i)]. \quad (11)$$

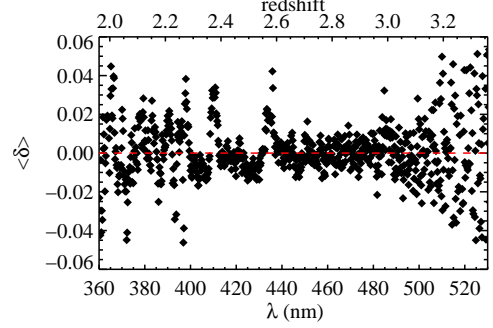
Figure 3 shows the  $C_q(\lambda)$  estimated for a typical quasar (blue line).

The last element necessary to use equation (2) is the mean transmitted flux fraction  $\bar{F}(z)$ . If  $P(F, z)$  derived from the mocks were the true distribution of the transmitted flux fraction, then  $\bar{F}(z)$  could simply be computed from the average of this distribution. Since this is not precisely true, we determine  $\bar{F}(z)$  from the data by requiring that the mean of the delta field vanish for all redshifts. The  $\bar{F}(z)$  we obtain is shown in figure 5. The unphysical wiggles in the derived  $\bar{F}(z)$  are associated with the aforementioned residuals in  $\bar{\delta}(\lambda)$  for method 1 (figure 4).

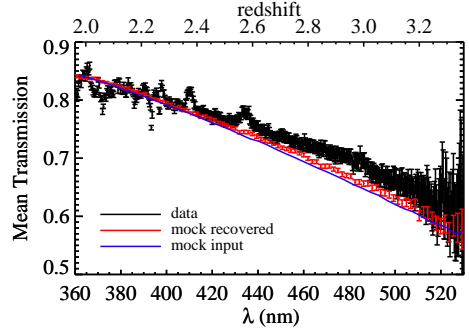
There is one inevitable effect of our two continuum estimation procedures. The use of the forest data in fitting the continuum effectively forces each quasar to have a mean absorption near that of the mean for the entire quasar sample. This approach introduces a spurious negative correlation between pixels on a given quasar even when well separated in wavelength. This negative correlation has no direct effect on our measurement of the flux correlation function because we do not use pixel pairs from the same quasar. However, the physical correlation between absorption on neighboring quasars causes the unphysical negative correlation for individual quasars to generate a negative contribution to the correlation measured with quasar pairs. Fortunately, this distortion is a smooth function of scale so it can be expected to have little effect on the observability or position of the BAO peak. This expectation is confirmed by analysis of the mock spectra (section 5).

### 3.3. Weights

A discussion on the optimal use of weights for the Ly $\alpha$  correlation function is found in McQuinn & White (2011). Here we



**Fig. 4.** The mean of  $\delta(\lambda)$  plotted as a function of observed wavelength (method 1). Systematic offsets from zero are seen at the 2% level. The calcium lines (393.4, 396.8 nm) is present. The features around the hydrogen lines H $\gamma$ ,  $\delta$  and  $\epsilon$  (434.1, 410.2, 397.0 nm) are artifacts from the use of F-stars for the photocalibration of the spectrometer.



**Fig. 5.** Mean transmitted flux fraction as a function of redshift obtained from the continuum fits with method 2. Data are shown in black, mock-000 in red and the input mean transmitted flux fraction in blue.

simply choose the weights  $w_{ij}$  so as to approximately minimize the relative error on  $\hat{\xi}_A$  estimated with equation (3). In the approximation of uncorrelated pixels, the variance of  $\hat{\xi}_A$  is

$$\text{Var}(\hat{\xi}_A) = \frac{\sum_{i,j \in A} w_{ij}^2 \xi_{ii} \xi_{jj}}{\left[\sum_{i,j \in A} w_{ij}\right]^2} \quad \xi_{ii} = \langle \delta_i^2 \rangle \quad (12)$$

where the pixel variance,  $\xi_{ii}$ , includes contributions from both observational noise and LSS. The signal-to-noise ratio is:

$$\left(\frac{S}{N}\right)^2 = \frac{\langle \hat{\xi}_A \rangle^2}{\text{Var}(\hat{\xi}_A)} \approx \frac{\left(\sum_{i,j \in A} \xi_{ij} w_{ij}\right)^2}{\sum_{i,j \in A} \xi_{ii} \xi_{jj} w_{ij}^2}. \quad (13)$$

Because of LSS growth and redshift evolution of the mean absorption, the  $\xi_{ij}$  depend on redshift and we use the measured dependence of the 1d correlation function (McDonald et al., 2006)

$$\xi_{ij}(z) = (1 + z_i)^{\gamma/2} (1 + z_j)^{\gamma/2} \xi_{ij}(z_0) \quad \gamma \sim 3.8. \quad (14)$$

Maximizing the signal-to-noise ratio with respect to  $w_{ij}$  this gives:

$$w_{ij} \propto \frac{(1+z_i)^{\gamma/2}(1+z_j)^{\gamma/2}}{\xi_{ii}^2 \xi_{jj}^2}. \quad (15)$$

For this expression to be used, we require a way of estimating the  $\xi_{ii}$ . We assume that it can be decomposed into a noise term and a LSS term ( $\sigma_{LSS}$ ):

$$\xi_{ii}^2 = \frac{\sigma_{pipeline,i}^2}{\eta(z_i)} + \sigma_{LSS}^2(z_i) \quad z_i = \lambda_i/\lambda_{Ly\alpha} - 1, \quad (16)$$

where  $\sigma_{pipeline,i}^2 = [\text{ivar}(C_q \bar{F})^2]^{-1}$  is the pipeline estimate of the noise-variance of pixel  $i$  and  $\eta$  is a factor that corrects for a possible misestimate of the variance by the pipeline.

We then organize the data in bins of  $\sigma_{pipeline,i}^2$  and redshift. In each such bin, we measure the variance of  $\delta_i$ , which serves as an estimator of  $\xi_{ii}$  for the bin in question. The two functions  $\eta(z)$  and  $\sigma_{LSS}^2(z)$  can then be determined by fitting equation (16).

These fits are shown in figure 6. The top panel shows that the measured inverse variance follows the inverse pipeline variance until saturating at the redshift-dependent LSS variance (shown on the bottom left panel). For  $z > 3$ , there are not enough pixel pairs to determine  $\eta(z)$  and  $\sigma_{LSS}^2(z)$ . In this high redshift range, we assumed  $\eta = 1$  and extrapolated  $\sigma_{LSS}(z)$  with a second-degree polynomial fit to the  $z < 3$  data.

### 3.4. $\xi(r, \mu)$

The procedure described above was used to determine  $\xi(r, \mu)$  through equation 3 in  $r$ -bins of width  $4 \text{ h}^{-1}\text{Mpc}$  (centered at 2.6, ..., 198  $\text{h}^{-1}\text{Mpc}$ ) and in  $\mu$ -bins of width 0.02, (centered at 0.01, 0.03, ... 0.99). The  $50 \times 50 \text{ } r - \mu$  bins have an average of  $6 \times 10^6$  terms in the sum (3) with an average nominal variance of  $\xi$  for individual bins of  $(10^{-4})^2$  as given by (eqn.12).

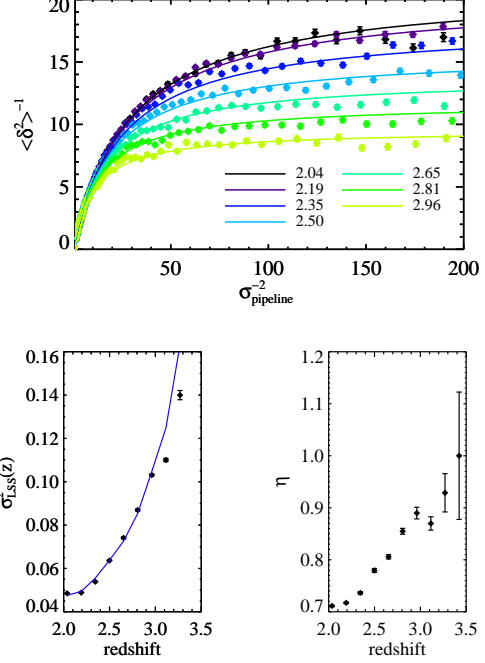
Figure 7 shows an example of  $\xi(r, \mu)$  for the  $r$  bin centered on  $34 \text{ h}^{-1}\text{Mpc}$ . The blue dots are the data and the red dots are the mean of the 15 mocks. The function falls from positive to negative values with increasing  $\mu$ , as expected from redshift distortions. The effect is enhanced by the deformation due to the continuum subtraction.

Figure 8 presents  $\xi(r, \mu)$  averaged over three bins in  $\mu$ . A clear peak at the expected BAO position,  $r_s = 105 \text{ h}^{-1}\text{Mpc}$ , is present in the bin  $0.8 < \mu < 1.0$  corresponding to separation vectors within  $37^\circ$  of the line-of-sight. The curves show the best fits for a  $\Lambda$ CDM correlation function, as described in section 5.

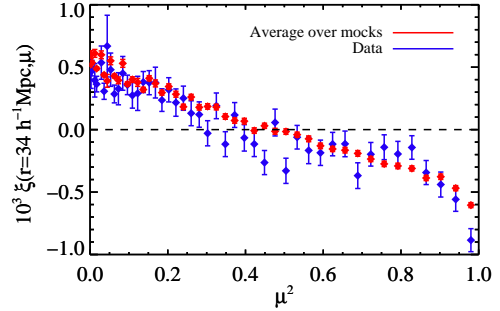
The data were divided into various subsamples to search for systematic errors in  $\xi(r, \mu)$ . For example, searches were made for differences between the northern and southern Galactic cap regions and between higher and lower signal-to-noise ratio quasars. No significant differences were found in the overall shape and amplitude of the correlation function. We also verified that the BAO peak position does not change significantly when wavelength slices of Ly $\alpha$  forest data are eliminated, in particular slices centered on the Balmer features in figure 4. The peak position also does not change significantly if the subtraction of the mean  $\delta$  (figure 4) is suppressed.

## 4. The Monopole and quadrupole

The analysis of the correlation function was performed in the framework of the standard multipole decomposition (Hamilton,



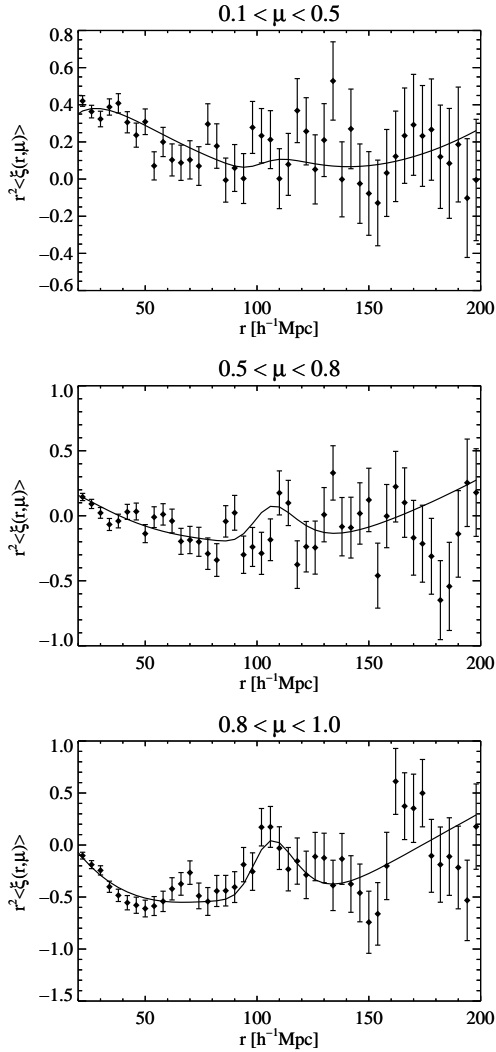
**Fig. 6.** Top panel: Inverse total variance in bins of redshift as a function of the pipeline inverse variance. Bottom panel: Parameters of the fit: the LSS contribution  $\sigma_{LSS}$  (left) and the pipeline correction factor  $\eta$  (right) as a function of redshift. The lines show fits to the data as explained in the text.



**Fig. 7.**  $\xi(r, \mu)$  vs.  $\mu^2$  for the bin centered on  $r = 34 \text{ h}^{-1}\text{Mpc}$ . The red dots are the mean of the 15 mocks and the blue dots are the data.

1992). For each bin in  $r$  we fit a monopole ( $\ell = 0$ ) and quadrupole ( $\ell = 2$ ) to the angular dependence:

$$\hat{\xi}(r, \mu) = \sum_{\ell=0,2} \xi_{\ell}(r) P_{\ell}(\mu) = [\xi_0(r) - \xi_2(r)/2] + [3\xi_2(r)/2]\mu^2 \quad (17)$$



**Fig. 8.**  $\xi(r, \mu)$  averaged over  $0.1 < \mu < 0.5$ ,  $0.5 < \mu < 0.8$  and  $0.8 < \mu < 1$ . The curves give fits (section 5) to the data imposing concordance  $\Lambda$ CDM cosmology. The BAO peak is most clearly present in the data for  $\mu > 0.8$ .

where  $P_\ell$  is the  $\ell$ -Legendre polynomial. We ignore the small and poorly determined  $\ell = 4$  term. This fit is performed using a simple  $\chi^2$  minimization with the nominal variance (equation 12) and ignoring the correlations between bins. This approach makes the fit slightly sub-optimal. (Later, we will correctly take into account correlations between  $r$ -bins of the monopole and quadrupole.) We also exclude from this fit the portion  $\mu < 0.1$  to avoid residual biases due to correlated sky subtraction across quasars; this has a negligible impact on the fits and, at any rate, there is little BAO signal at low  $\mu$ .

Figure 9 displays the monopole and quadrupole signals found by the two methods. The two methods are slightly offset from one another, but the peak structure is very similar. Figure 9 also shows the combination  $\xi_0 + 0.1\xi_2$  which, because of the small monopole-quadrupole anti-correlation (section 4.1), is a better-determined quantity. The peak structure seen in figure 8 is also present in these figures.

Because of the continuum estimation procedure (sections 3.1 and 3.2), we can expect that the monopole and quadrupole shown in figure 9 are deformed with respect to the true monopole and quadrupole. The most important difference is that the measured monopole is negative for  $60 \text{ h}^{-1}\text{Mpc} < r < 100 \text{ h}^{-1}\text{Mpc}$  while the true  $\Lambda$ CDM monopole remains positive for all  $r < 130 \text{ h}^{-1}\text{Mpc}$ . The origin of the deformation in the continuum estimate is demonstrated in appendix A where both the true and estimated continuum can be used to derive the correlation function (figure A.1). As expected, the deformation is a slowly varying function of  $r$  so neither the position of the BAO peak nor its amplitude above the slowly varying part of the correlation function are significantly affected.

#### 4.1. Covariance of the monopole and quadrupole

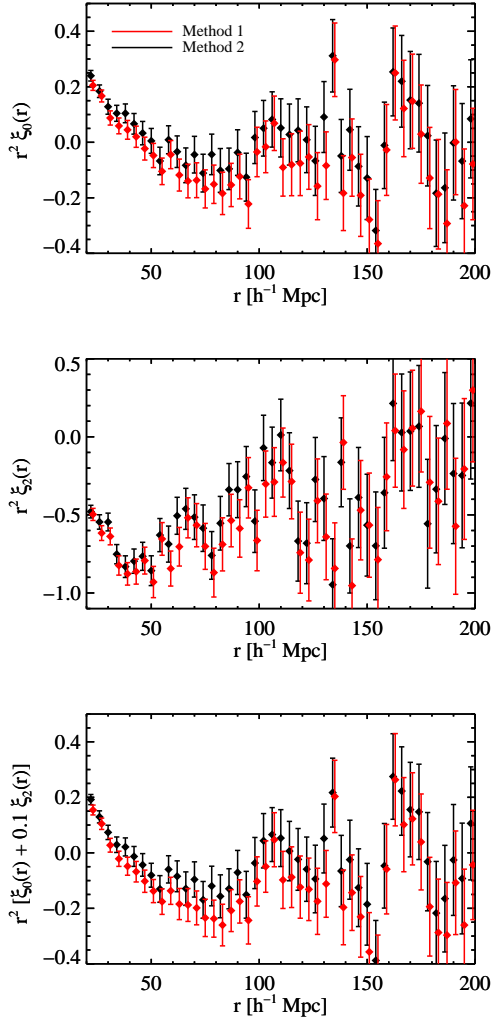
In order to determine the significance of the peak we must estimate the covariance matrix of the monopole and quadrupole. If the fluctuations  $\delta_i$  in equation (3) in different pixels were uncorrelated, the variance of  $\xi_A$  would simply be the weighted products of the fluctuation variances. This yields a result that is  $\sim 30\%$  smaller than the true correlation variance that we compute below. The reason is, of course, that the  $\delta$ -pairs are correlated, either from LSS or from correlations induced by instrumental effects or continuum subtraction; this effect reduces the effective number of pairs and introduces correlations between  $(r, \mu)$  bins.

Rather than determine the full covariance matrix for  $\xi(r, \mu)$ , we determined directly the covariance matrix for  $\xi_0(r)$  and  $\xi_2(r)$  by standard techniques of dividing the full quasar sample into subsamples according to position on the sky. In particular we used the sub-sampling technique described below. We also tried a bootstrap technique (e.g. Efron & Gong, 1983) consisting of substituting the entire set of  $N$  subdivisions of the data by  $N$  of these subdivisions chosen at random (with replacement) to obtain a “bootstrap” sample. The covariances are then measured from the ensemble of bootstrap samples. Both techniques give consistent results.

The adopted covariance matrix for the monopole and quadrupole uses the sub-sampling technique. We divide the data into angular sectors and calculate a correlation function in each sector. Pairs of pixels belonging to different sectors contribute only to the sector of the pixel with lower right ascension. We investigated two different divisions of the sky data: defining 800 (contiguous but disjoint) sectors of similar solid angle, and taking the plates as defining the sectors (this latter version does not lead to disjoint sectors). The two ways of dividing the data lead to similar covariance matrices.

Each sector  $s$  in each division of the data provides a measurement of  $\xi_s(r, \mu)$  that can be used to derive a monopole and quadrupole,  $\hat{\xi}_\ell(r)$ , ( $\ell = 0, 2$ ). The covariance of the whole BOSS sample can then be estimated from the weighted and rescaled covariances for each sector:

$$\begin{aligned} & \sqrt{W(r)W(r')} \text{Cov}[\hat{\xi}_\ell(r), \hat{\xi}_{\ell'}(r')] \\ &= \left\langle \sqrt{W_s(r)W_s(r')} [\hat{\xi}_{\ell s}(r)\hat{\xi}_{\ell' s}(r') - \bar{\xi}_\ell(r)\bar{\xi}_{\ell'}(r')] \right\rangle. \end{aligned} \quad (18)$$



**Fig. 9.** Monopole (upper panel) and quadrupole (middle panel) correlation functions found by method 1 (red) and method 2 (black). The bottom panel shows the combination  $\xi_0 + 0.1\xi_2$  found by method 1 (red) and method 2 (black).

The average denoted by  $\langle \rangle$  is the simple average over sectors, while  $\xi_\ell(r)$  denotes the correlation function measured for the whole BOSS sample. The  $W_s(r)$  are the summed pixel-pair weights for the radial bin  $r$  for the sector  $s$  and  $W(r)$  is the same sum for the whole BOSS sample.

The most important terms in the covariance matrix are the  $r = r'$  terms, i.e. the monopole and quadrupole variances. They are shown in figure 10 as a function of  $r$ . In the figure, they are multiplied by the number  $N$  of pixel pairs in the  $r$ -bin. The product is nearly independent of  $r$ , as expected for a variance nearly equal to the pixel variance divided by  $N$ . For the monopole, the

variances are only about 30% higher than what one would calculate naively assuming uncorrelated pixels and equation (12). Figure 10 also displays the monopole-quadrupole covariance times number of pairs, which also is nearly independent of  $r$ .

Figure 11 displays the monopole-monopole and quadrupole-quadrupole covariances. Nearest-neighbor covariances are of order 20%. Figure 11 also shows monopole-quadrupole covariance.

We used the 15 sets of mock spectra to test our method for calculating the covariance matrix. From the 15 measurements of  $\xi_\ell(r)$  one can calculate the average values of  $\xi_\ell(r)\xi_{\ell'}(r')$  and compare them with those expected from the covariance matrix. Figure 12 shows this comparison for the monopole and quadrupole variance, the monopole and quadrupole covariances between neighboring  $r$ -bins and the monopole-quadrupole covariance. The agreement is satisfactory.

#### 4.2. Detection significance of the BAO peak

In this section, we estimate the significance of our detection of a BAO peak at  $105 \text{ h}^{-1} \text{ Mpc}$ . At the statistical power of the present data, it is clear that the peak significance will depend to some extent on how we treat the so-called “broadband” correlation function on which the peak is superimposed. In particular, the significance will depend strongly on the  $r$ -range over which the correlation functions are fitted. To the extent that the BAO peak is known to be present in the matter correlation function and that the Ly $\alpha$  absorption is known to trace matter, the actual significance is of limited interest for cosmology. Of greater interest is the uncertainty in the derived cosmological parameter constraints (section 5) which will be non-linear reflections of the peak significance derived here.

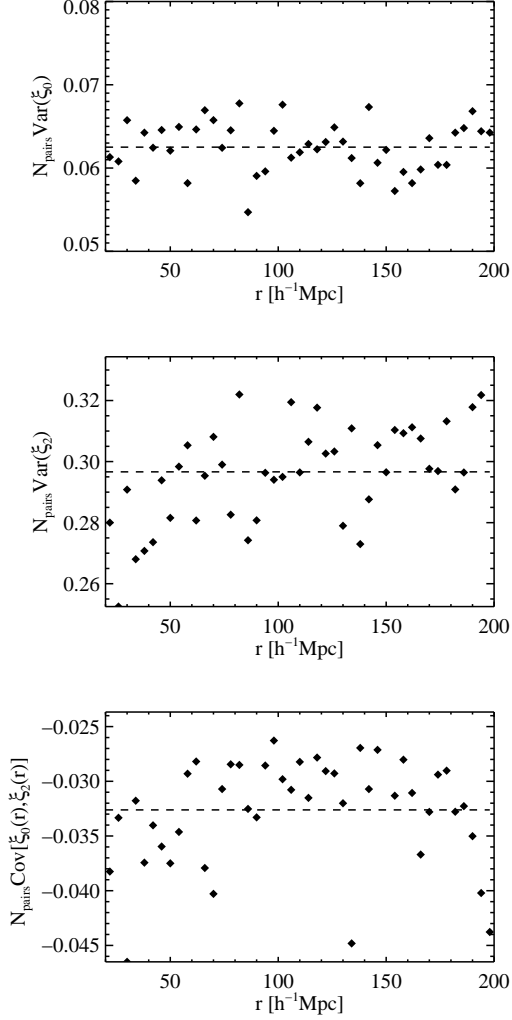
A detection of the BAO peak requires comparing the quality of a fit with no peak (the null hypothesis) to that of a fit with a peak. Typically, this exercise would be performed by choosing a test statistic, such as the  $\chi^2$ , computing the distribution of this quality indicator from a large number of peak-less simulations and looking at the consistency of the data with this distribution. Since our mock data sets are quite computationally expensive and only a handful are available, we chose a different approach.

Our detection approach uses the following expression to fit the observed monopole and quadrupole.

$$\xi_\ell(r) = B_\ell \xi_\ell^{BB}(r) + C_\ell \xi_\ell^{\text{peak}}(r) + A_\ell \xi_\ell^{\text{dist}}(r) \quad (19)$$

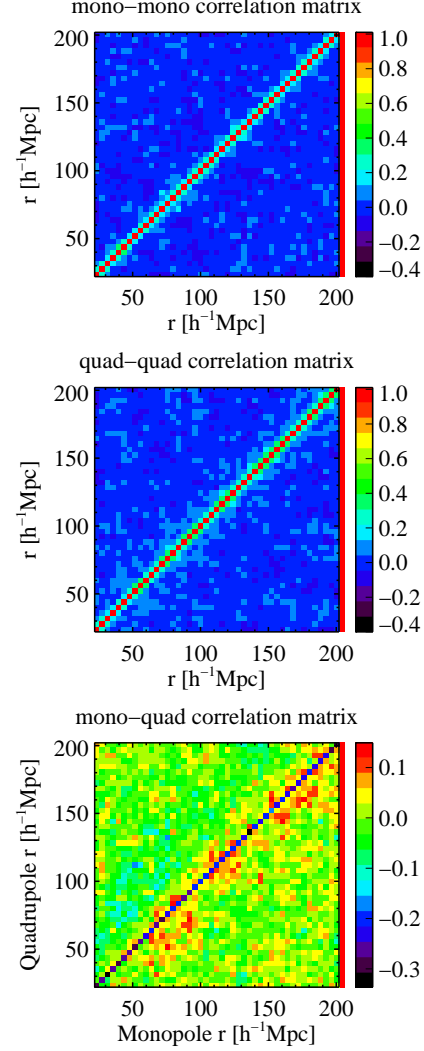
where  $\xi_\ell^{BB}$  is a broadband term to describe the LSS correlation function in the absence of a peak,  $\xi_\ell^{\text{peak}}$  is a peak term, and  $\xi_\ell^{\text{dist}}$  is a “distortion” term used to model the effects of continuum subtraction. The broadband term is derived from the fiducial  $\Lambda$ CDM cosmology defined by the parameters in equation (A.1). It is obtained by fitting the shape of the fiducial correlation function with an 8-node spline function *masking the region of the peak* ( $80 \text{ h}^{-1} \text{ Mpc} < r < 120 \text{ h}^{-1} \text{ Mpc}$ ). The peak term is the difference between the theoretical correlation function and the broadband term. Finally, the distortion term is calculated from simulations, as the difference in the monopole or quadrupole measured using the true continuum and that measured from fitting the continuum as described in appendix A. The three components are shown in figure 13.

Expression (19) contains three parameters each for the monopole and quadrupole (so six in total). We have performed fits leaving all six parameters free and fits where we fix the ratio  $C_2/C_0$  to be equal to its nominal value used to generate our



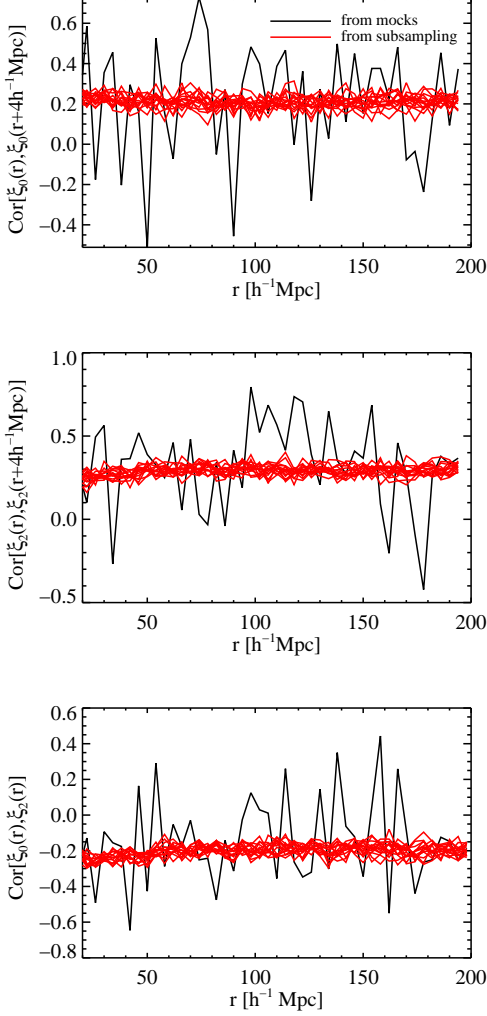
**Fig. 10.** The  $r$ -dependence of the product of the monopole (top) and quadrupole (middle) variances and the number of pairs in the  $r$ -bin. The bottom panel shows the product for the monopole-quadrupole covariance ( $r = r'$  elements). The dotted lines show the means for  $r > 20 \text{ h}^{-1} \text{ Mpc}$ .

mock spectra (the value given by assuming a “redshift distortion parameter”  $\beta = 1.4$ , see appendix A). We define the test statistic as the  $\chi^2$  difference between fitting equation (19) simultaneously to monopole and quadrupole by fixing  $C_0$  to zero (a “peak-less” four or five-parameter fit) and fitting for  $C_0$  (a five or six-parameter fit). In our detection fits we do not fit for the BAO position but fix it to the theoretical prediction. The distribution for this test statistic (“ $\Delta\chi^2_{\text{det}}$ ”) under the null hypothesis is a  $\chi^2$  distribution with one degree of freedom. The significance is then given by  $\sigma = (\Delta\chi^2_{\text{det}})^{1/2}$ .



**Fig. 11.** The monopole and quadrupole covariance matrix. The monopole-monopole and quadrupole-quadrupole elements are normalized to the variance:  $C_{ij}/\sqrt{C_{ii}C_{jj}}$ . The monopole-quadrupole elements are normalized to the mean of the quadrupole and monopole variances. The first off-diagonal elements of the monopole-monopole and quadrupole-quadrupole elements are  $\sim 20\%$  of the diagonal elements. The diagonal elements of the quadrupole-monopole covariance are  $\sim -0.2$  times the geometric mean of the monopole and quadrupole variances.

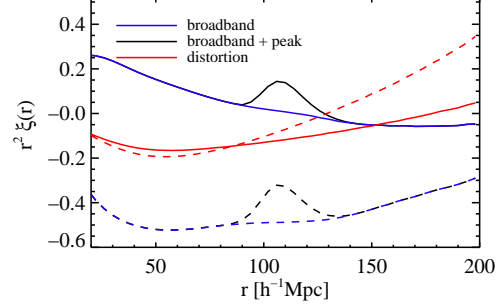
Figure 14 shows the fits to monopole (top panel) and quadrupole (bottom panel) and the corresponding fits with and without peaks and fixing  $\beta = 1.4$ . For method 2, we obtain  $\chi^2/DOF = 93.7/85$  (111.8/86) with (without) a peak, giving  $\Delta\chi^2 = 18.1$  for a detection significance of  $4.2\sigma$ . For method 1, we obtain  $\chi^2/DOF = 93.2/85$  (102.2/86) for a significance



**Fig. 12.** Verification of the off-diagonal elements of the covariance matrix with the 15 sets of mock spectra. The black lines show correlations derived from the dispersion of the 15 measurements and the red lines show the correlations expected from the covariance matrix calculated by sub-sampling. The top and middle panels show the correlation between neighboring bins for monopole and quadrupole respectively. The bottom panel the correlation between monopole and quadrupole measured at the same distance bin.

of  $3.0\sigma$ . Allowing  $\beta$  to be a free parameter gives essentially the same detection significances.

The detection significance of  $\sim 4\sigma$  is typical of that which we found in the 15 sets of mock spectra. For the mocks, the significances ranged from 0 to  $6\sigma$  with a mean of  $3.5\sigma$ .



**Fig. 13.** The fitting functions used for the determination of the peak detection significance:  $r^2 \xi^{bb}(r)$  (blue),  $r^2 [\xi^{peak}(r) + \xi^{bb}(r)]$  (black) and  $r^2 \xi^{dist}(r)$  (red) for the monopole (solid lines) and quadrupole (dashed lines).

Our significance depends strongly on the fitting range. For a lower boundary of the range of  $r_{\min} = 20, 40, 60 h^{-1}\text{Mpc}$  we obtain a significance of  $\sigma = 4.2, 3.2$  and  $2.3$ , respectively (method 2). The reason for this result is illustrated in figure 15, where the results of the fits with and without peaks are compared to data for different values of  $r_{\min} = 20, 40, 60 h^{-1}\text{Mpc}$ . Reducing the fitting range poses less stringent constraints on the distortion and broadband terms, thus allowing some of the peak to be attributed to the broadband. In particular, the statistically insignificant bump in the quadrupole at  $\sim 65 h^{-1}\text{Mpc}$  causes the fitted broadband to increase as  $r_{\min}$  is increased to  $60 h^{-1}\text{Mpc}$ , decreasing the amplitude of the BAO peak. For the monopole, the  $r_{\min} = 60 h^{-1}\text{Mpc}$  fit predicts a positive slope for  $\xi_0(r)$  that decreases the amplitude of the peak but predicts a  $\xi_0(r < 50 h^{-1}\text{Mpc})$  to be much less than what is measured.

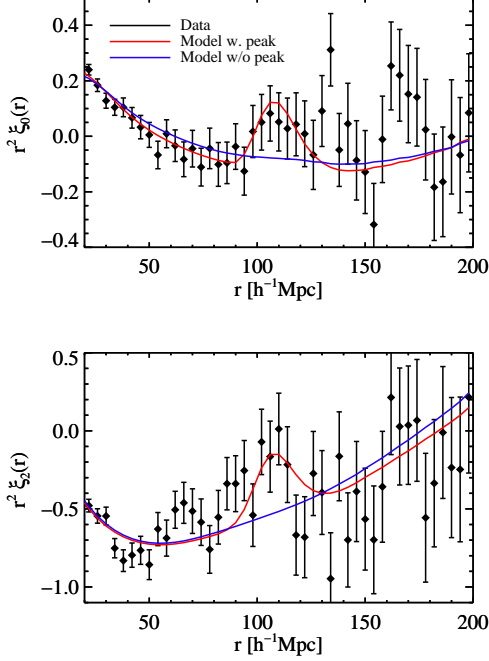
## 5. Cosmology with the BAO peak

The observed position of the BAO peak in  $\xi(r, \mu)$  is determined by two sets of cosmological parameters: the “true” parameters and the “fiducial” parameters. Nature uses the true cosmology to create correlations at the true sound horizon,  $r_s$ . The true cosmology transforms physical separations between Ly $\alpha$  absorbers into angles on the sky and redshift differences:  $\theta_{BAO} = r_s/D_A(z)(1+z)$  and  $\Delta z_{BAO} = r_s H(z)/c$ . We, on the other hand, use a “fiducial” cosmology (defined by equation A.1) to transform angular and redshift differences to local distances at the redshift in question to reconstruct  $\xi(r, \mu)$ . If the fiducial cosmology is the true cosmology, the reconstructed peaks will be at the calculated fiducial sound horizon,  $r_{s,f}$ . Limits on the difference between the fiducial and reconstructed peak position can be used to constrain the differences between the fiducial and true cosmological models.

### 5.1. The peak position

The use of incorrect fiducial  $D_A$ ,  $H$  and  $r_s$  leads to shifts in the BAO peak position in the transverse and radial directions by the multiplicative factors  $\alpha_t$  and  $\alpha_r$ :

$$\alpha_t \equiv \frac{D_A(z)/r_s}{D_{A,f}/r_{s,f}} \quad (20)$$



**Fig. 14.** Monopole and quadrupole fits with a BAO peak (red line) and without a BAO peak (blue line, Method 2). The fitting range is  $r_{\min} < r < 200 \text{ h}^{-1} \text{Mpc}$  with  $r_{\min} = 20 \text{ h}^{-1} \text{Mpc}$ .

$$\alpha_r \equiv \frac{H_f(z)r_{s,f}}{H(z)r_s} \quad (21)$$

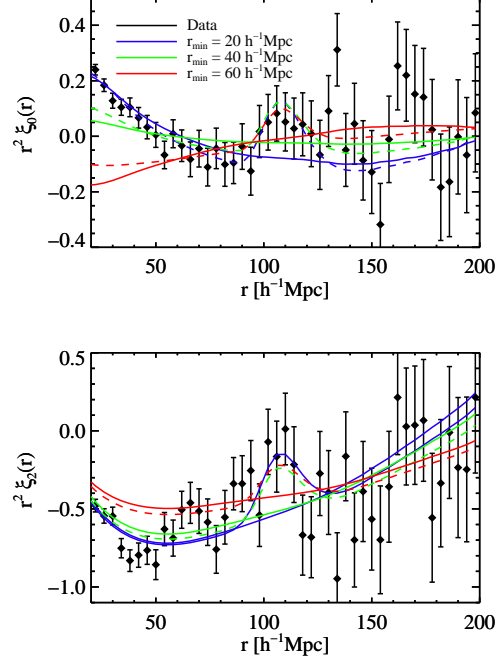
where the subscript  $f$  refers to the fiducial model. Following Xu et al. (2012), we will use a fitting function,  $\tilde{\xi}_\ell(r)$ , for the monopole and quadrupole that follows the expected peak position as a function of  $(\alpha_t, \alpha_r)$ :

$$\tilde{\xi}_\ell(r) = \hat{\xi}_\ell(r, \alpha_t, \alpha_r, b, \beta) + A_\ell(r) \quad \ell = 0, 2. \quad (22)$$

Here, the two functions  $\hat{\xi}_\ell$ , derived from the power spectrum given in appendix A, describe the underlying mass correlation function, the linear bias  $b$  and redshift distortion parameter  $\beta$ , and the movement of the BAO peak for  $\alpha_t, \alpha_r \neq 1$ . The functions  $A_\ell$  take into account distortions, as described below.

For  $\alpha_t = \alpha_r \equiv \alpha_{iso}$ , there is a simple isotropic scaling of the coordinates by  $\alpha_{iso}$  and  $\hat{\xi}_\ell(r)$  is given by  $\hat{\xi}_\ell(r, \alpha_{iso}) = f_\ell(b, \beta)\xi_{\ell,f}(\alpha_{iso}r)$ , where  $\xi_{\ell,f}$  are the fiducial monopole and quadrupole and the normalizations  $f_\ell$  are the functions of the bias and redshift-distortion parameter given by Hamilton (1992). For  $\alpha_t \neq \alpha_r$ , Xu et al. (2012) found an approximate formula for  $\hat{\xi}_\ell(r)$  that was good in the limit  $|\alpha_t - \alpha_r| \ll 1$ . We take the more direct route of numerically expanding  $\xi_{\ell,f}(\alpha_t r_t, \alpha_r r_r)$  in Legendre polynomials,  $P_\ell(\mu)$ , to directly calculate the  $\hat{\xi}_\ell(r, \alpha_t, \alpha_r)$ .

The functions  $A_\ell(r)$  describe broadband distortions due to continuum subtraction and the fact that the broadband correlation function is not expected to change in the same way as the BAO peak position when one deviates from the fiducial model.



**Fig. 15.** Same as figure 14 except for different fitting ranges: blue, green and red curves are for fits with  $r_{\min} = 20 \text{ h}^{-1} \text{Mpc}$ ,  $40 \text{ h}^{-1} \text{Mpc}$  and  $60 \text{ h}^{-1} \text{Mpc}$  respectively (Method 2). The solid lines for fits without a BAO peak and the dashed lines with a peak.

They correspond to the term  $A_\ell \xi_\ell^{\text{dist}}$  in equation (19). We have used two forms to represent  $A_\ell(r)$ :

$$A_\ell^{(1)}(r) = \frac{a_\ell}{r^2} + \frac{b_\ell}{r} + c_\ell \quad (23)$$

and

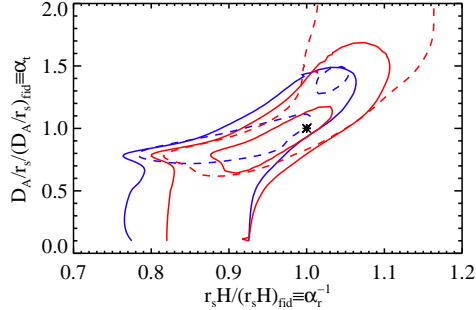
$$A_\ell^{(2)}(r) = \frac{a_\ell}{r^2} + \frac{b_\ell}{r} + c_\ell + \frac{d_\ell}{\sqrt{r}}. \quad (24)$$

The observed monopole and quadrupole can then be fit to (22) with free parameters  $\alpha_t, \alpha_r$ , bias,  $\beta$ , and the nuisance parameters ( $a_\ell, b_\ell, c_\ell$  and  $d_\ell$ ).

We first fixed  $(\alpha_t, \alpha_r) = (1, 1)$  to determine if we find reasonable values of  $(b, \beta)$ . These two parameters are highly degenerate since both the quadrupole and monopole have amplitudes that are proportional to  $b^2$  times polynomials in  $\beta$ . A well-determined combination is  $b(1 + \beta)$ , for which we find a value  $0.38 \pm 0.07$ ; this is in agreement with  $b(1 + \beta) = 0.336 \pm 0.012$  found at  $r \sim 40 \text{ h}^{-1} \text{Mpc}$  by Slosar et al. (2011). The larger error of our fit reflects the substantial freedom we have introduced with our distortion function.

We next freed all parameters to constrain  $(\alpha_t, \alpha_r)$ . The contours for the two methods and two broadband are shown in figure 16 and the  $\chi^2$  for the fiducial and best-fit models are given in table 1. The broadband term in equation (24) fits the data better





**Fig. 16.** The contours for  $(D_A/r_s, r_s H)$  obtained by fitting the monopole and quadrupole to (22). The broadband distortions are eqn. (23, dashed lines) or (24, solid lines). The blue lines are for method 1 and the red lines for method 2. All contours are for  $\Delta\chi^2 = 4$  except for the interior solid red contour which is for  $\Delta\chi^2 = 1$ .

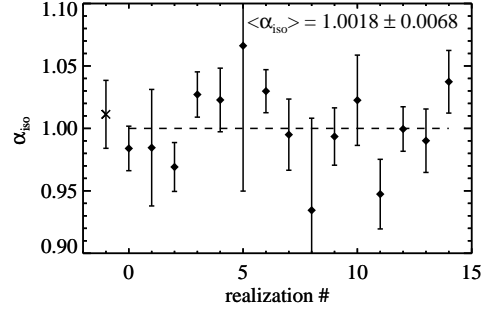
than that in equation (23) both for the fiducial parameters and for the best fit. For broadband in equation (24), the  $\chi^2$  for the fiducial model is acceptable for both methods:  $\chi^2/DOF = 85.0/80$  for method 1 and  $\chi^2/DOF = 71.5/80$  for method 2.

The contours in the figure are elongated along the direction for which the BAO peak position stays approximately fixed at large  $\mu$  (near the radial direction, where the observations are most sensitive). The best constrained combination of  $D_A$  and  $H$  of the form  $(D_A^{\zeta} H^{\zeta-1}/r_s)$  turns out to have  $\zeta \sim 0.2$ . This low value of  $\zeta$  reflects the fact that we are mostly sensitive to the BAO peak in the radial direction. At the one standard-deviation level, the precision on this combination is about 4%. However, even this combination is sensitive to the tails in the contours. A more robust indicator of the statistical accuracy of the peak-position determination comes from fits imposing  $\alpha_t = \alpha_r \equiv \alpha_{iso}$ , as has generally been done in previous BAO studies with the exception of Chuang & Wang (2012) and Xu et al. (2012). This constraint does not correspond to any particular class of cosmological models. It does however eliminate the tails in the contours in a way that is similar to the imposition of outside data sets. The two methods and broadbands give consistent results, as seen in table 1.

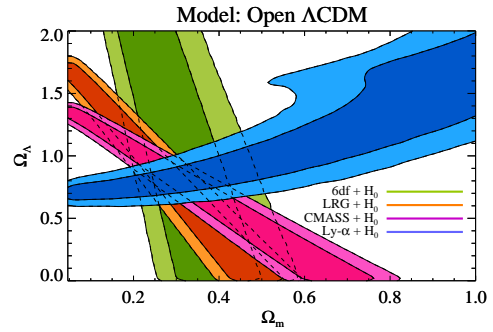
We used the sets of mock spectra to search for biases in our measurement of  $\alpha_{iso}$ . The mean value reconstructed for this quantity on individual mocks is  $1.002 \pm 0.007$ , suggesting that there are no significant biases in the determination of the BAO scale. Figure 17 shows the values and errors for the individual mocks along with that for the data. Both the measured value and its uncertainty for the data is typical of that found for individual sets of mock spectra.

### 5.2. Constraints on cosmological models

Our constraints on  $(D_A/r_s, H r_s)$  can be used to constrain the cosmological parameters. In a  $\Lambda$ CDM cosmology, apart from the pre-factors of  $H_0$  that cancel,  $D_A/r_s$  and  $H r_s$  evaluated at  $z = 2.3$  depend primarily on  $\Omega_M$  through  $r_s$  and on  $\Omega_\Lambda$  which, with  $\Omega_M$ , determines  $D_A$  and  $H$ . The sound horizon also depends on  $H_0$  (required to derive  $\Omega_\gamma$  from  $T_{CMB}$ ), on the effective number of neutrino species  $N_\nu$  (required to derive the radiation density from



**Fig. 17.** The measurements of  $\alpha_{iso}$  ( $= \alpha_t = \alpha_r$ ) for the 15 sets of mock spectra and for the data (realization=-1). The large errors for realization 5 and 8 are due to the very low significance of the BAO peak found on these two sets.



**Fig. 18.** Constraints on the matter and dark-energy density parameters  $(\Omega_M, \Omega_\Lambda)$  assuming a dark-energy pressure-density ratio  $w = -1$ . The blue regions are the one and two standard deviation constraints derived from our contours in figure 16 (method 2, broadband 24) combined with a measurement of  $H_0$  (Riess et al., 2011). Also shown are one and two standard deviation contours from lower redshift measurements of  $D_V/r_s$  (also combined with  $H_0$ ) at  $z = 0.11$  [6dF: Beutler et al. (2011)],  $z = 0.35$  [LRG: Percival et al. (2010)] and  $z = 0.57$  [CMASS: Anderson et al. (2012)]. All constraints use a WMAP7 (Komatsu et al., 2011) prior on the baryon-to-photon ratio  $\eta$  but do not otherwise incorporate CMB results.

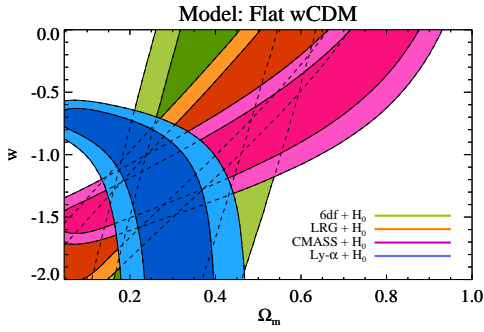
the photon density), and on the baryon-to-photon number ratio,  $\eta$  (required for the speed of sound).

Figure 18 shows the  $\Lambda$ CDM constraints on  $(\Omega_M, \Omega_\Lambda)$  derived from the contours in figure 16 combined with the most recent measurement of  $H_0$  (Riess et al., 2011). We use the contours for method 2 and the broadband of equation 24 which gives better fits to the data than the other method and broadband. The contours also assume  $N_\nu = 3$  and the WMAP7 value of  $\eta$  (Komatsu et al., 2011). Also shown are constraints from BAO measurements of  $D_V/r_s$  (Percival et al., 2010; Anderson et al., 2012; Beutler et al., 2011).

The Ly $\alpha$  contours are nicely orthogonal to the lower redshift  $D_V/r_s$  measurements, reinforcing the requirement of dark en-

**Table 1.** Results with the two methods and two broadband (equations 23 and 24). Columns 2 and 3 give the  $\chi^2$  for the fiducial model and for the model with the minimum  $\chi^2$ . Column 4 gives the best fit for  $\alpha_{iso}$  with the constraint ( $\alpha_r = \alpha_r \equiv \alpha_{iso}$ ). Column 5 gives  $Hr_s/[Hr_s]_{fid}$  with the  $2\sigma$  limits in parentheses. Column 6 gives the  $Hr_s/[Hr_s]_{fid}$  deduced by combining our data with that of WMAP7 (Komatsu et al., 2011) (see section 5.3).

method & broadband	$\chi^2_{fid}/DOF$	$\chi^2_{min}/DOF$	$\alpha_{iso}$	$Hr_s/[Hr_s]_{fid}$	$Hr_s/[Hr_s]_{fid}$ (with WMAP7)
Method 1 (24)	85.0/80	84.6/78	$1.035 \pm 0.035$	$0.876 \pm 0.049$ ( $^{+0.188}_{-0.111}$ )	$0.983 \pm 0.035$
Method 2 (24)	71.5/80	71.4/78	$1.010 \pm 0.025$	$0.954 \pm 0.077$ ( $^{+0.152}_{-0.154}$ )	$1.000 \pm 0.036$
Method 1 (23)	104.3/82	99.9/80	$1.027 \pm 0.031$	$0.869 \pm 0.044$ ( $^{+0.185}_{-0.084}$ )	$0.988 \pm 0.034$
Method 2 (23)	88.4/82	87.7/80	$1.004 \pm 0.024$	$0.994 \pm 0.111$ ( $^{+0.166}_{-0.178}$ )	$1.006 \pm 0.032$



**Fig. 19.** As in figure 18 with constraints on the matter density parameter,  $\Omega_M$ , and dark-energy pressure-density ratio  $w$  assuming  $\Omega_M + \Omega_\Lambda = 1$ .

energy from BAO data. In fact, our measurement is the only BAO measurement that by itself requires dark energy:  $\Omega_\Lambda > 0.5$ . This is because at  $z = 2.3$  the universe is strongly matter dominated and the  $\sqrt{\Omega_M}$  factor in  $H$  partially cancels the  $1/\sqrt{\Omega_M}$  in  $r_s$ , enhancing the importance of the  $\Omega_\Lambda$  dependence of  $H$ .

Figure 19 shows the constraints on  $(\Omega_M, w)$ ; where  $w$  is the dark-energy pressure-density ratio) assuming a flat universe:  $\Omega_k = 0$ . Our result is the only BAO measurement that by itself requires negative  $w$ . Our limit  $w < -0.6$  requires matter domination at  $z = 2.3$ .

$$\frac{\rho_{de}(z = 2.3)}{\rho_m(z = 2.3)} < 0.3 \left( \frac{\Omega_\Lambda/\Omega_M}{0.73/0.27} \right). \quad (25)$$

### 5.3. Constraints on $H(z)$

The contours in figure 16 give the measurements of  $Hr_s$  given in table 1. A measurement of the expansion rate deep in the matter-dominated epoch can be used to demonstrate the deceleration of the expansion at that time. Unfortunately, our data are not yet precise enough to do this. To make a more precise measurement of  $H(z = 2.3)$ , we must add further constraints to eliminate the long tails in figure 16. These tails correspond to models where  $1/H(z = 2.3)$  is increased (resp. decreased) with respect to the fiducial value while  $D_A(z = 2.3)$  is decreased (resp. increased). For flat models, this would imply a change in the mean of  $1/H$  (averaged up to  $z = 2.3$ ) that is opposite to that of the change in  $1/H(z = 2.3)$ , which requires a functional form  $H(z)$  that

strongly differs from the fiducial case. It is possible to construct models with this property by introducing significant non-zero curvature.

Because of the importance of curvature, the tails are eliminated once WMAP7 constraints (Komatsu et al., 2011) are included. This is done in figure 20 within the framework of non-flat models where the dark-energy pressure-density ratio,  $w(z)$ , is determined by two parameters,  $w_0$  and  $w_a$ :  $w(z) = w_0 + w_a z/(1+z)$ . As expected, the WMAP7 results in this framework constrain  $D_A$  and  $1/H$  to migrate in roughly the same direction as one moves away from the fiducial model. Combining WMAP7 constraints with ours gives the values of  $H(z = 2.3)r_s$  given in the last column of table 1. For what follows, we adopt the mean of methods 1 and 2 that use the more flexible broadband of equation (24):

$$\frac{H(z = 2.3)r_s}{[H(z = 2.3)r_s]_{fid}} = 0.992 \pm 0.035. \quad (26)$$

The precision on  $H$  is now sufficient to study the redshift evolution of  $H(z)$ .

The fiducial model has  $r_s = 152.76$  Mpc and  $H(z = 2.3) = 3.23H_0$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . These results produce

$$\frac{H(z = 2.3)r_s}{1+z} = (1.036 \pm 0.036) \times 10^4 \text{ km s}^{-1}. \quad (27)$$

or equivalently

$$\frac{H(z = 2.3)}{1+z} = (67.8 \pm 2.4) \text{ km s}^{-1} \text{ Mpc}^{-1} \left( \frac{152.76 \text{ Mpc}}{r_s} \right). \quad (28)$$

This number can be compared with the measurements of  $H(z)$  at lower redshift shown in table 2 and figure 21. Other than those of  $H_0$ , the measurements that we use can be divided into two classes: those (like ours) that use  $r_s$  as the standard of length and those that use  $c/H_0$  as the standard of length.

The comparison with our measurement is simplest with BAO-based measurements that use  $r_s$  as the standard of length and therefore measure  $H(z)r_s$  (as is done here). The first attempt at such a measurement was made by Gaztañaga et al. (2009), a result debated in subsequent papers by Miralda-Escudé (2009), Yoo & Miralda-Escudé (2010), Kazin et al. (2010), and Cabré & Gaztañaga (2011). Here, we use four more recent measurements. Chuang & Wang (2012) and Xu et al. (2012) studied the SDSS DR7 LRG sample and decomposed the BAO peak into radial and angular components, thus extracting directly  $Hr_s$  and  $D_A/r_s$ . Blake et al. (2012) and Reid et al. (2012) took a more indirect route. They first used the angle-averaged peak position to derive  $D_V(z)/r_s = ((1+z)^2 D_A^2 cz H^{-1})/r_s$ . They then studied

**Table 2.** Recent measurements of  $H(z)/(1+z)$ . The BAO-based measurements use  $r_s = 152.76$  Mpc as the standard of length and are shown as the filled circles in figure 21. The quoted uncertainties in  $H(z)$  do not include uncertainties in  $r_s$  which are expected to be negligible,  $\approx 1\%$  (Komatsu et al., 2011). The measurements of Blake et al. (2011b) use supernova data and therefore measure  $H(z)$  relative to  $H_0$ . We quote the results they obtain without assuming a flat universe and plot them as the open green circles in figure 21 assuming  $h = 0.7$ .

$z$	$H(z)/(1+z)$ km s $^{-1}$ Mpc $^{-1}$	method	reference
2.3	$66.5 \pm 7.4$	BAO	this work
2.3	$67.8 \pm 2.4$	BAO+WMAP7	this work
0.35	$60.8 \pm 3.6$	BAO	Chuang & Wang (2012)
0.35	$62.5 \pm 5.2$	BAO	Xu et al. (2012)
0.57	$58.8 \pm 2.9$	BAO + AP	Reid et al. (2012)
0.44	$57.4 \pm 5.4$	BAO + AP	Blake et al. (2012)
0.60	$54.9 \pm 3.8$		
0.73	$56.2 \pm 4.0$		
0.2	$(1.11 \pm 0.17)H_0$	AP + SN	Blake et al. (2011b)
0.4	$(0.83 \pm 0.13)H_0$		
0.6	$(0.81 \pm 0.08)H_0$		
0.8	$(0.83 \pm 0.10)H_0$		
0	$73.8 \pm 2.5$		Riess et al. (2011)

the Alcock-Paczynski effect on the broadband galaxy correlation function to determine  $D_A(z)H(z)$ . Combining the two measurements yielded  $H(z)r_s$ .

It is evident from comparing our  $H(z)$  measurement (filled red circle in figure 21) to the other BAO-based measurements (other filled circles) that  $H(z)/(1+z)$  decreases between  $z = 2.3$  and  $z = 0.35 - 0.8$ . To demonstrate deceleration quantitatively, we fit the eight BAO-based values of  $H(z)$  in table 2 to the  $\Lambda$ CDM form  $H(z) = H_0(\Omega_\Lambda + \Omega_M(1+z)^3 + (1 - \Omega_\Lambda - \Omega_M)(1+z)^2)^{1/2}$ . Marginalizing over  $\Omega_\Lambda$  and  $H_0$  we find

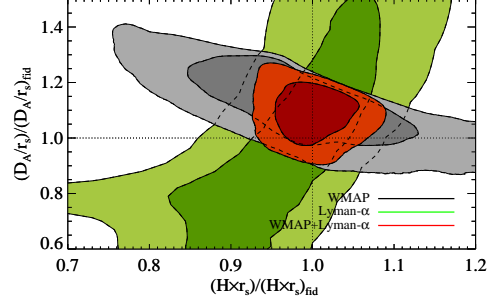
$$\frac{[H(z)/(1+z)]_{z=2.3}}{[H(z)/(1+z)]_{z=0.5}} = 1.17 \pm 0.05, \quad (29)$$

clearly indicating deceleration between  $z = 2.3$  and  $z = 0.5$ . This measurement is in good agreement with the fiducial value of 1.146. We emphasize that this result is independent of  $r_s$ , assuming only that the BAO-peak position is redshift-independent in comoving coordinates. The result also does not assume spatial flatness.

To map the expansion rate over the full range  $0 < z < 2.3$ , we must adopt the fiducial value of  $r_s$  and compare the resulting  $H(z)$  with  $H_0$  and with other BAO-free measurements. Besides the  $H_0$  measurement of Riess et al. (2011), we use the WiggleZ analysis combining their Alcock-Paczynski data with distant supernova data from the Union-2 compilation (Amanullah et al., 2010). The supernova analysis does not use the poorly known mean SNIa luminosity, so the SNIa Hubble diagram gives the luminosity distance in units of  $H_0^{-1}$ ,  $D_L(z)H_0$ . Combining this result with the Alcock-Paczynski measurement of  $D_A(z)H(z)$  yields  $H(z)/H_0$ . The values are given in table 2.

We fit all the data in table 2 (filled and open circles in figure 21) to the  $\Lambda$ CDM form of  $H(z)$ . This yields an estimate of the redshift of minimum  $H(z)/(1+z)$

$$z_{d-a} = 0.82 \pm 0.08 \quad (30)$$



**Fig. 20.** Constraints on  $(D_A/r_s, r_s H)_{z=2.3}$  within the framework of  $\Lambda$ CDM models. The green contours are our  $1\sigma$  and  $2\sigma$  constraints using method 2 and broadband W4. The gray contours are the  $1\sigma$  and  $2\sigma$  constraints from WMAP7 (Komatsu et al., 2011). The red contours show the combined constraints.

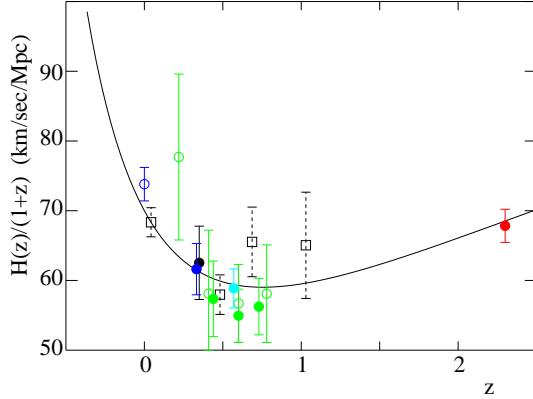
which compares well with the fiducial value:  $z_{d-a} = (2\Omega_\Lambda/\Omega_M)^{1/3} - 1 = 0.755$ .

In this analysis, we have not used two other sources of information on  $H(z)$  at high redshift. The first use high-redshift type Ia supernovae to probe the era where the universe transitions from deceleration to acceleration (e.g., Riess et al. (2004, 2007)). The data of Riess et al. (2007) (plotted as the open squares in figure 21) yielded useful measurements up to  $z \sim 1.1$ . However, this data yields constraints on  $H(z)$  that are weaker than those of BAO-based methods because of the need to differentiate the distance-redshift relation. Moreover, these inferences of  $H(z)$  assume spatial flatness. Fitting the SNe data to a model with an evolving deceleration parameter  $q(z) = q_0 + (dq/dz)_0 z$  and assuming flatness, Riess et al. (2007) and Riess et al. (2004) were able to demonstrate that  $(dq/dz)_0 > 0$ , i.e. a negative 3rd-derivative of  $a(t)$ . However, we point out that in a more general  $q(z)$  model, the demonstration that  $dq/dz > 0$  at low redshift is not equivalent to a demonstration that  $\ddot{a}$  becomes negative in the past.

Another approach to determining  $H(z)$  uses the evolution of stellar populations as a clock to infer  $dt/dz$  (Stern et al., 2010; Moresco et al., 2012). This method yields results that are consistent with  $\Lambda$ CDM expectations, but the uncertainties (statistical and systematic) are larger than those of the determinations in Table 2, so we have not plotted them in figure 21.

## 6. Conclusions

In this paper, we have presented the first observation of the BAO peak using the Ly $\alpha$  forest. It represents both the first BAO detection deep in the matter dominated epoch and the first to use a tracer of mass that is not galactic. The results are consistent with concordance  $\Lambda$ CDM, and require, by themselves, the existence of dark energy. Combined with CMB constraints, we deduce the expansion rate at  $z = 2.3$  and demonstrate directly the sequence of deceleration and acceleration expected in dark-energy dominated cosmologies. These results have been confirmed with higher precision by Slosar et al. (2013) using the same underlying DR9 data set but more aggressive data cuts and a more nearly optimal statistical method.



**Fig. 21.** Measurements of  $H(z)/(1+z)$  vs  $z$  demonstrating the acceleration of the expansion for  $z < 0.8$  and deceleration for  $z > 0.8$ . The BAO-based measurements are the filled circles: [this work: red], [Xu et al. (2012): black] [Chuang & Wang (2012): blue], [Reid et al. (2012), cyan], and [Blake et al. (2012): green]. The open green circles are from WiggleZ (Blake et al., 2011b) Alcock-Paczynski data combined with supernova data yielding  $H(z)/H_0$  (without the flatness assumption) plotted here assuming  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . The open blue circle is the  $H_0$  measurement of Riess et al. (2011). The open black squares with dashed error bars show the results of Riess et al. (2007) which were derived by differentiating the SNIa Hubble diagram and assuming spatial flatness. (For visual clarity, the Riess et al. (2007) point at  $z = 0.43$  has been shifted to  $z = 0.48$ .) The line is the  $\Lambda$ CDM prediction for  $(h, \Omega_M, \Omega_\Lambda) = (0.7, 0.27, 0.73)$ .

BOSS continues to acquire data and will eventually produce a quasar sample three times larger than DR9. We can thus expect improved precision in our measurements of distances and expansion rates, leading to improved constraints on cosmological parameters. The Ly $\alpha$  forest may well be the most practical method for obtaining precise  $D_A(z)$  and  $H(z)$  measurements at  $z > 2$ , thanks to the large number of independent density measurements per quasar. It is reassuring that the first sample large enough to yield a detection of BAO produces a signal in good agreement with expectations. In the context of BAO dark energy constraints, high redshift measurements are especially valuable for breaking the degeneracy between curvature and the equation of state history. More generally, however, by probing an epoch largely inaccessible to other methods, BAO in the Ly $\alpha$  forest have the potential to reveal surprises, which could provide critical insights into the origin of cosmic acceleration.

*Acknowledgements.* We thank Carlos Allende Prieto, Ashley Ross and Uros Seljak for stimulating discussions and Adam Riess for providing the data points of Riess et al. (2007) in figure 21.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

The French Participation Group of SDSS-III was supported by the Agence Nationale de la Recherche under contract ANR-08-BLAN-0222.

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. [PIIF-GA-2011-301665].

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory,

University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## Appendix A: Mock quasar spectra

We have produced mock spectra in order to tune the analysis procedure and to study statistical uncertainties and systematic effects in the measured correlation function.

In some galaxy clustering studies (e.g. Anderson et al. (2012)) the covariance matrix of the measured correlation function is obtained from mock data sets. In this case, it is crucial to have very realistic mocks with the right statistics.

In order to do so, we would need to generate several realizations of hydrodynamical simulations, with a large enough box to cover the whole survey (several  $\text{Gpc}^3$ ) and at the same time have a good enough resolution to resolve the Jeans mass of the gas (tenths of kpc). This type of simulations are not possible to generate with current technology, but luckily in this study the covariance matrix is obtained from the data itself, and the mock data sets are only used to test our analysis and to study possible systematic effects.

In the last few years there have been several methods proposed to generate simplified mock Lyman- $\alpha$  surveys by combining Gaussian fields and nonlinear transformations of the field (Le Goff et al., 2011; Greig et al., 2011; Font-Ribera et al., 2012a). In this study we used a set of mocks generated using the process described in Font-Ribera et al. (2012a), the same method used in the first publication of the Lyman- $\alpha$  correlation function from BOSS (Slosar et al., 2011).

The mock quasars were generated at the angular positions and redshifts of the BOSS quasars. The unabsorbed spectra (continua) of the quasars were generated using the Principal Component Analysis (PCA) eigenspectra of Suzuki et al. (2005), with amplitudes for each eigenspectrum randomly drawn from Gaussian distributions with sigma equal to the corresponding eigenvalues as published in Suzuki (2006) table 1. A detailed description will be provided by Bailey et al. (in preparation), accompanying a public release of the mock catalogs.

We generated the field of transmitted flux fraction,  $F$ , that have a  $\Lambda$ CDM power spectrum with the fiducial parameters

$$(\Omega_M, \Omega_\Lambda, \Omega_b h^2, h, \sigma_8, n_s)_{\text{fid}} = (0.27, 0.73, 0.0227, 0.7, 0.8, 0.97) \quad (\text{A.1})$$

where  $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . These values produce a fiducial sound horizon of

$$r_{s,\text{fid}} = 152.76 \text{ Mpc} . \quad (\text{A.2})$$

Here, we use the parametrized fitting formula introduced by McDonald (2003) to fit the results of the power spectrum from several numerical simulations,

$$P_F(k, \mu_k) = b_\delta^2 (1 + \beta \mu_k^2)^2 P_L(k) D_F(k, \mu_k) , \quad (\text{A.3})$$

where  $\mu_k = k_{\parallel}/k$  is the cosine of the angle between  $\mathbf{k}$  and the line of sight,  $b_\delta$  is the density bias parameter,  $\beta$  is the redshift distortion parameter,  $P_L(k)$  is the linear matter power spectrum,

and  $D_F(k, \mu_k)$  is a non-linear term that approaches unity at small  $k$ . This form of  $P_F$  is the expected one at small  $k$  in linear theory, and provides a good fit to the 3-D Ly $\alpha$  observations reported in Slosar et al. (2011). We do not generate a density and a velocity field, but directly create the Ly $\alpha$  forest absorption field instead, with the redshift distortions being directly introduced in the input power spectrum model of equation (A.3), with the parameter  $\beta$  that measures the strength of the redshift distortion.

To model the evolution of the forest with redshift,  $b_\delta$  varies with redshift according to  $b_\delta = 0.14[(1+z)/3.25]^{1.9}$  (McDonald et al., 2006). The redshift distortion parameter is given a fixed value of  $\beta_F = 1.4$ . The non-linear correction factor  $D(k, \mu_k)$  is taken from McDonald (2003). The flux field was constructed by generating Gaussian random fields  $g$  with an appropriately chosen power spectrum (Font-Ribera et al., 2012a) to which the log-normal transformation  $F = \exp(-ae^{v\delta})$  is applied (Coles & Jones, 1991; Bi et al., 1992; Gnedin & Hui, 1996). Here  $a$  and  $v$  are free parameters chosen to reproduce the flux variance and mean transmitted flux fraction (McDonald et al., 2006).

DLA's were added to the spectra according to the procedure described in Font-Ribera et al. (2012b).

Finally, the spectra were modified to include the effects of the BOSS spectrograph point spread function (PSF), readout noise, photon noise, and flux miscalibration.

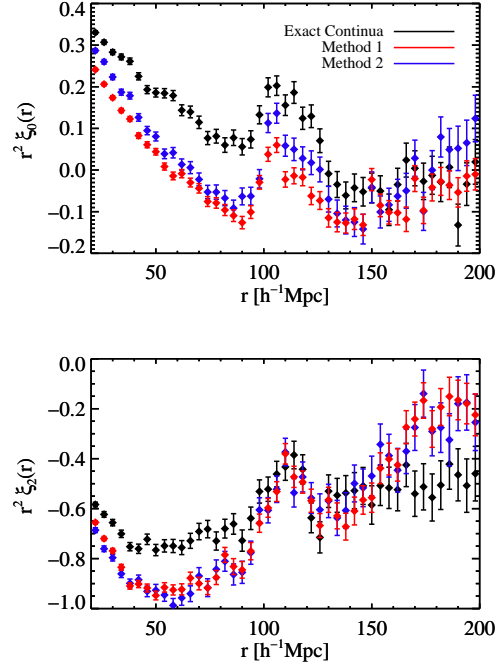
Fifteen independent realizations of the BOSS data were produced and analyzed with the same procedures as those for the real data.

We used the mock spectra to understand how our analysis procedure modifies the correlation function. Figure A.1 shows the average over 15 mocks of the reconstructed quadrupole and monopole using methods 1 and 2 (sections 3.1 and 3.2) and that reconstructed with the true continuum. The monopole and quadrupole for the two methods have a general shape that follows that found with the true continuum including the position of the BAO peak. However, both methods produce a monopole that becomes negative for  $60 \text{ h}^{-1}\text{Mpc} < r < 100 \text{ h}^{-1}\text{Mpc}$  while the true monopole remains positive for all  $r < 130 \text{ h}^{-1}\text{Mpc}$ . As discussed in section 3.2, this result is due to the continuum estimation of the two methods which introduced negative correlations. For both methods, however, the BAO peak remains visible with a deviation above the ‘‘broadband’’ correlation function that is hardly affected by the distortion.

Figure A.2 presents  $\xi_0(r)$  and  $\xi_2(r)$  found with the data, along with the mean of 15 mocks. The figure demonstrates that our mocks do not perfectly reproduce the data. In particular, for  $r < 80 \text{ h}^{-1}\text{Mpc}$ , the monopole is underestimated and the quadrupole overestimated. Since we use only peak positions to extract cosmological constraints, we only use the mocks qualitatively to search for possible systematic problems in extracting the peak position.

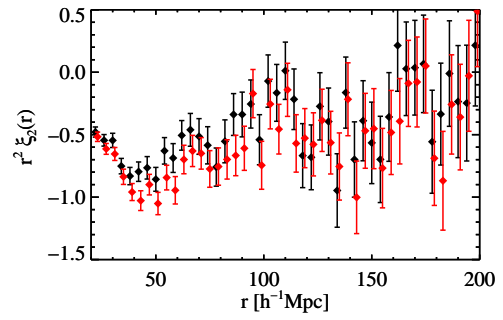
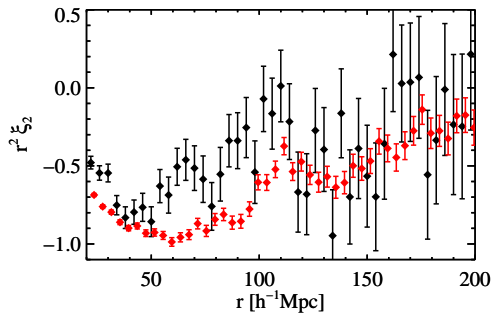
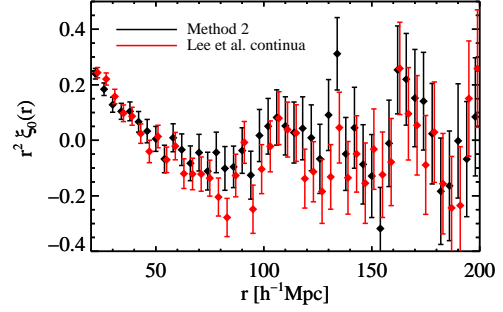
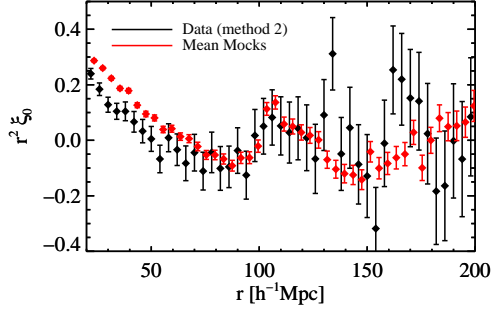
## Appendix B: Results for a Fiducial BOSS Ly $\alpha$ Forest Sample

The spectra analyzed here are all available through SDSS DR9 (Ahn et al., 2012), and the DR9 quasar catalog is described by P aris et al. (2012). A Ly $\alpha$  forest analysis requires many detailed choices about data selection and continuum determination. To aid community analyses and comparison of results from different groups, Lee et al. (2012b) has presented a fiducial BOSS Ly $\alpha$  forest sample that uses constrained PCA continuum determination (Lee et al., 2012a) and reasonable choices of masks



**Fig. A.1.** The effect of the continuum estimation procedure on the correlation function found with the mock spectra. The black dots are the average of monopole and quadrupole obtained with the 15 sets using the exact continua. The blue (red) dots show those obtained with the continuum estimation of method 1 (method 2) as described in section 3.1 (3.2).

for DLAs, BALs, and data reduction artifacts. Both the data selection and the continuum determination differ from those used here. Figure B.1 compares the Method 2 correlation function from this paper’s analysis to that obtained by applying the Method 2 weights and correlation measurement code directly to the continuum-normalized spectra of the fiducial Lee et al. sample. The good agreement in this figure, together with the good agreement between our Method 1 and Method 2 results, demonstrates the robustness of the BAO measurement, and the more general correlation function measurement, to the BOSS DR9 data.



**Fig. A.2.** Comparison of the correlation function for the mock spectra and that for the data. The red dots show the mean of the 15 sets of mock spectra and the black dots show the data.

**Fig. B.1.** Comparison of the monopole and quadrupole correlation functions for the sample used here (black dots) and for the sample and continua of Lee et al. (2012b) (red dots).

## References

- Abazajian K.N., J.K. Adelman-McCarthy, M.S. Agüeros et al. 2009, *ApJS*, 182, 543
- Ahn, C.P., R. Alexandroff, C. Allende Prieto et al. 2012, *ApJS*, 203, 21
- Alcock, C. & B. Paczynski 1979, *Nature*, 281, 358
- Amanullah, R. et al. 2010, *ApJ*, 716, 712
- Anderson, L., E. Aubourg, S. Bailey et al. 2012, *MNRAS*, 427, 3435
- Bailey, S. et al. 2012, in preparation
- Becker, R.H., R.L. White & D.J. Helfand 1995, *ApJ*, 450, 559B
- Beutler F. et al. 2011, *MNRAS*, 416, 3017
- Bi, H.G., G. Börner & Y.Chu 1992, *A&A*, 266, 1
- Blake C. et al. 2011, *MNRAS*, 415, 2892
- Blake C. et al. 2011, *MNRAS*, 418, 1725
- Blake C. et al. 2012, *MNRAS*, 425, 405
- Blanton, M.R., H. Lin, R. H. Luptonn et al. 2003, *AJ*, 125, 2276
- Bolton A. et al. 2012, *AJ*, 144, 144
- Bovy, J. et al. 2011, *ApJ*, 729, 141
- Bovy, J. et al. 2012, *ApJ*, 749, 41
- Cabré, A. & E. Gaztañaga 2011, *MNRAS*, 412, L98
- Cauci, S., S. Colombi, C. Pichon et al. 2008, *MNRAS*, 386, 211
- Cen R., Miralda-Escudé J., Ostriker J. P., Rauch M. 1994, *ApJ*, 437, L9
- Chuang, Chia-Hsun & Yun Wang 2012, *MNRAS*, 426, 226
- Cole, S. et al. 2005, *MNRAS*, 362, 505
- Coles, P. & B. Jones. 1991, *MNRAS*, 248, 1
- Croft, R. et al. 1997, *ApJ*, 488, 532
- Croft, R. et al. 1998, *ApJ*, 495, 44
- Croft, R.A.C., et al. 1999, *ApJ*, 520, 1
- Croft, R.A.C., D. Weinberg, M. Bolte, et al. 2002, *ApJ*, 581, 20
- Dawson, K., D. Schlegel, C. Ahn et al. 2013, *AJ*, 145, 10
- de Bernardis, P., P. A. R. Ade, J. J. Bock, et al. 2000, *Nature*, 404, 955
- Eisenstein D. J., Zehavi, I., Hogg, D.W. et al. 2005, *ApJ*, 633, 560
- Eisenstein, D.J. 2007, *ApJ*, 664, 675
- Eisenstein, D.J. et al. 2011, *AJ*, 142, 72
- Efron, B. & G. Gong 1983, *The American Statistician*, 37, 36
- Font-Ribera, A., P. McDonald & J. Miralda Escudé 2012, *J. Cosmology Astropart. Phys.*, 01, 001
- Font-Ribera, A. & Miralda-Escudé, J. 2012, *J. Cosmology Astropart. Phys.*, 07, 028
- Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, *AJ*, 111, 1748
- Gallerani, S., F. S. Kitaura & A. Ferrara 2011, *MNRAS*, 413L, 6
- Gaztañaga, E., A. Cabré & L. Hui(2009), *MNRAS*, 399, 1663
- Gnedin, N. & L. Hui 1996, *ApJ*, 472 L73
- Greig, B., J. Bolton, J. Wytthe & B. Stuart 2011, *MNRAS*, 418, 1980
- Gunn, J.E., et al. 1998, *AJ*, 116, 3040
- Gunn, J.E., et al. 2006, *AJ*, 131, 2332
- Hamilton, A.J.S. 1992, *ApJ*, 385L, 5
- Hernquist L., Katz N., Weinberg D. H., Miralda-Escudé J. 1996, *ApJ*, 457, L51
- Kazin, E. A., M. R. Blanton, R. Scoccimarro et al. 2010, *ApJ*, 719, 1032
- Kirkpatrick, J.A., D. J. Schlegel, N. P. Ross, et al., 2011, *AJ*, 743, 125
- Kitaura, F.-S., S. Gallerani & A. Ferrara 2012, *MNRAS*, 420, 61
- Komaatsu, E. et al. 2011, *ApJS*, 192, 18
- Lawrence, A., S.J. Warren, O. Almaini et al. 2007, *MNRAS*, 379, 1599
- Le Goff, J. M. et al. 2011, *A&A*, 534, 135
- Lee, K.-G., Suzuki, N., & Spergel, D. N. 2012a, *AJ*, 143, 51
- Lee, K.-G., Bailey, S., Bartsch, L. E., et al. 2012b, arXiv:1211.5146
- Martin, D.C., J. Fanson, D. Schiminovich et al. 2005, *ApJ*, 619, L1
- McDonald, P. et al. 2000, *ApJ*, 543, 1
- McDonald, P. 2003, *ApJ*, 585, 34
- McDonald, P., U. Seljak, D.J. Schlegel et al. 2006, *ApJS*, 163, 80
- McDonald, P. & D. Eisenstein 2007, *Phys. Rev. D*, 76, 063009
- McQuinn, M. & M. White 2011, *MNRAS*, 415, 2257
- Mehta, K.T., A. Cuesta, X. Xu et al. 2012, *MNRAS*, 427, 2168
- Miralda-Escudé J., Cen R., Ostriker J. P., Rauch M. 1996, *ApJ*, 471, 582
- Miralda-Escudé J. (2009), arXiv:0901.1219
- Moresco, M., A. Cimatti, R. Jimenez et al. 2012, *J. Cosmology Astropart. Phys.*, 006, 1208
- Noterdaeme, P., P. Petitjean, W.C. Carithers et al. 2012, *A&A*, 547, L1
- Nusser, A. & M. Haehnelt 1999, *MNRAS*, 303, 179
- Padmanabhan, N., Schlegel, D. J., Finkbeiner, D. P., et al. 2008, *ApJ*, 674, 1217
- Padmanabhan, N. et al. 2009, *Phys. Rev. D*, 79, 3523
- Padmanabhan, N., X. Xu, D.J. Eisenstein et al. 2012, *MNRAS*, 427, 2132
- Palanque-Delabrouille, N., C. Yèche, A. D. Myers et al. 2011, *A&A*, 530, 122
- Pâris, I., Petitjean, P., Rollinde, et al. 2011, *A&A*, 530, 50
- Pâris, I., P. Petitjean, E. Aubourg et al. 2012, *A&A*, 548, 66
- Percival, W.J., B.A. Reid, D.J. Eisenstein et al. 2010, *MNRAS*, 401, 2148
- Petitjean P., Mückert J. P., Kates R. E. 1995, *A&A*, 295, L9
- Pichon, C., J. L. Vergely, E. Rollinde et al. 2001, *MNRAS*, 326, 597
- Pier, J. R., Munn, J. A., Hindsley, R. B., et al. 2003, *AJ*, 125, 1559
- Pieri, M. M. et al. 2010, *ApJ*, 724, L69
- Reid, B.A., L. Samushia, M. White et al. 2012, *MNRAS*, 426, 2719
- Richards, G.T. et al. 2009, *ApJS*, 180, 67
- Riess, A.G., L.-G. Strolger, J. Tonry et al. 2004, *ApJ*, 607, 665
- Riess, A.G., L.-G. Strolger, S. Casertano et al. 2007, *ApJ*, 659, 98
- Riess, A.G., L. Macri, S. Casertano et al. 2011, *ApJ*, 730, 119
- Ross, N.P. et al. 2012, *ApJS* 199, 3
- Seo, H.-J., S. Ho, M. White et al. 2012, *ApJ*, 761, 13
- Slosar, A. et al. 2011, *J. Cosmology Astropart. Phys.*, 09, 001
- Slosar, A. et al. (2013), arXiv:1301.3459
- Smee, S., J.E. Gunn, A. Uomoto et al. 2012, arXiv:1208.2233
- Smith, J. A., Tucker, D. L., Kent, S., et al. 2002, *AJ*, 123, 2121
- Stern, D., et al. 2010, *J. Cosmology Astropart. Phys.*, 02, 008
- Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, *AJ*, 123, 485
- Suzuki, N., D. Tytler, D. Kirkman et al. 2005, *ApJ*, 618, 592
- Suzuki, N. 2006, *ApJS*, 163, 110
- Theuns T., Leonard A., Efstathiou, G., Pearce F.R., Thomas, P.A. 1998, *MNRAS*, 301, 478
- Viel, M., M.G. Haehnelt & V. Springel 2004, *MNRAS*, 354, 684
- Weinberg, D.H. et al. 1998, *ASPC*, 148, 21
- Weinberg, D.H., J. J. Mortonson, D. J. Eisenstein et al. 2012, arXiv:1201.2434
- White, M. 2003, *The Davis Meeting On Cosmic Inflation*, 2003 March 22-25, Davis CA., p.18
- Worseck, G. & L. Wisotzki 2006, *A&A*, 450, 495; A. Lidz et al. (2010) *ApJ* 718, 199
- Xu, X., Cuesta, A.J., N. Padmanabhan et al. 2012, arXiv:1206.6732
- Yèche, C., P. Petitjean, J. Rich, et al. 2009, *A&A*, 523, A14
- Yoo, J. & J. Miralda-Escudé 2010, *Phys. Rev. D*, 82, 3527
- York, D. G. *et al.* (SDSS Collaboration) 2000, *AJ*, 120, 1579
- Zhang, Y., Anninos P., Norman M. L. 1995, *ApJ*, 453, L57

# Liste des tableaux

2.1	Valeurs des paramètres cosmologiques mesurés par Planck . . . . .	35
2.2	Valeurs du biais $b$ et de $\beta$ pour différents traceurs . . . . .	38
4.1	Efficacités des différentes sélections de quasars basées sur la variabilité . . . . .	83
5.1	Nombre de quasars de l'échantillon de l'analyse BAO . . . . .	97
5.2	Longueurs d'onde masquées pour l'analyse BAO . . . . .	99
5.3	Valeurs des paramètres des différents ajustements des données . . . . .	125
5.4	Dépendance des paramètres de l'ajustement en fonction de la zone ajustée . . . . .	127





# Table des figures

1.1	3 types d'espaces homogènes et isotropes . . . . .	12
1.2	Diagramme de Hubble des SNIa . . . . .	22
1.3	Utilisation d'une règle standard transverse à la ligne de visée . . . . .	23
1.4	Comparaison des distances comobile, de luminosité et angulaire . . . . .	24
1.5	Utilisation d'une règle standard le long de la ligne de visée . . . . .	25
2.1	Propagation d'une sur-densité adiabatique dans l'Univers primordial . . . . .	30
2.2	$C_l$ du fond diffus cosmologique obtenus par Planck . . . . .	34
2.3	Spectre de puissance, fluctuations de densité et fluctuations de potentiel gravitationnel, au temps présent . . . . .	37
2.4	Fonction de corrélation des galaxies de l'échantillon CMASS . . . . .	40
2.5	Spectre de puissance des galaxies de l'échantillon CMASS . . . . .	40
2.6	Représentation schématique d'un AGN . . . . .	42
2.7	Spectre d'un quasar typique de BOSS . . . . .	44
2.8	Spectre d'un quasar présentant un système DLA . . . . .	44
2.9	Spectre d'un quasar présentant des BAL . . . . .	45
3.1	Déroulement prévu du SDSS-III . . . . .	48
3.2	Prévisions de sensibilité sur la distance volumique $D_V$ . . . . .	50
3.3	Image des 14.555 deg <sup>2</sup> de ciel observés en photométrie par SDSS . . . . .	51
3.4	Photographie du télescope du SDSS . . . . .	55
3.5	Schéma du télescope du SDSS . . . . .	56
3.6	Disposition du plan focal de la caméra de BOSS . . . . .	58
3.7	Efficacité quantique du système télescope + caméra . . . . .	59
3.8	Photographie d'une cartouche de BOSS . . . . .	60
3.9	Schéma optique des spectrographes . . . . .	62
3.10	Parcours des rayons lumineux après dispersion par le grisme . . . . .	62
3.11	Efficacités optiques prévisionnelles pour chaque partie du système télescope + spectrographe . . . . .	63
3.12	Efficacités optiques mesurées pour les deux spectrographes . . . . .	64
3.13	Pouvoirs de résolution requis et mesurés des deux spectrographes . . . . .	65
3.14	Largeur d'une raie en fonction de la position dans le plan focal des CCD des spectrographes . . . . .	66
4.1	Distribution des décalages spectraux de l'échantillon de quasars de la sélection par variabilité . . . . .	71
4.2	Exemples de courbe de lumière pour deux étoiles . . . . .	72
4.3	Exemples de courbe de lumière pour deux quasars . . . . .	73

4.4	Distributions des $\chi^2$ réduits obtenus en ajustant les courbes de lumière par une constante . . . . .	74
4.5	Paramètres $\gamma$ et $A_r$ de la fonction de structure . . . . .	75
4.6	Exemple de fonction de structure $\mathcal{V}(\Delta t)$ pour un quasar typique . . . . .	76
4.7	Représentation schématique du réseau de neurones . . . . .	77
4.8	Sortie du réseau de neurones . . . . .	78
4.9	Réjection stellaire en fonction de la complétude des quasars . . . . .	79
4.10	Complétude en fonction du décalage spectral . . . . .	79
4.11	Distributions en ascension droite des objets de la sélection . . . . .	81
4.12	Distributions des magnitudes dans la bande photométrique $r$ des objets de la sélection . . . . .	81
4.13	Distributions des décalages spectraux des quasars confirmés . . . . .	84
4.14	Diagrammes couleur-couleur indiquant les populations d'étoiles et de quasars . . . . .	85
4.15	Distributions des décalages spectraux des quasars obtenues par différentes méthodes . . . . .	86
4.16	Prédiction de la réjection stellaire en fonction de la complétude des quasars en utilisant différents relevés . . . . .	87
4.17	Nombre de courbes de lumière en fonction du nombre d'observations pour le relevé complet . . . . .	89
4.18	Nombre de courbes de lumière en fonction de l'écart en temps entre la première et la dernière observation pour le relevé complet . . . . .	89
4.19	Courbes de lumière pour trois objets de l'échantillon de quasars . . . . .	90
4.20	Courbes de lumière pour trois objets de l'échantillon d'étoiles . . . . .	91
4.21	Distributions des $\chi^2$ pour les échantillons d'étoiles et de quasars . . . . .	92
4.22	Réjection stellaire en fonction de la complétude des quasars pour différentes coupures . . . . .	93
5.1	Projection Hammer-Aitoff des données de DR10 . . . . .	96
5.2	Distribution du signal sur bruit des spectres . . . . .	97
5.3	Distribution des décalages spectraux des quasars . . . . .	98
5.4	Variance du flux en fonction de la longueur d'onde . . . . .	99
5.5	Schéma de deux lignes de visée de quasars . . . . .	100
5.6	Moyenne du flux des quasars en fonction de la longueur d'onde observée et de la longueur d'onde dans le référentiel du quasar . . . . .	102
5.7	Exemple d'estimation du flux moyen attendu . . . . .	103
5.8	Histogrammes des paramètres de normalisation et de déformation . . . . .	103
5.9	Moyenne des $\delta_q$ en fonction de la longueur d'onde observée . . . . .	105
5.10	Distribution du décalage spectral des paires de $\delta$ . . . . .	105
5.11	Variance des $\delta$ en fonction de la variance estimée par la chaîne d'analyse . . . . .	108
5.12	Fonction de corrélation en fonction de $r_{perp}$ et $r_{par}$ . . . . .	109
5.13	Moyenne de la fonction de corrélation sur trois intervalles en $\mu$ . . . . .	109
5.14	Fonction de corrélation en fonction de $\mu^2$ pour $r = 34 h^{-1}\text{Mpc}$ . . . . .	111
5.15	Monopôle et quadrupôle de la fonction de corrélation . . . . .	111
5.16	Hexadécapôle de la fonction de corrélation . . . . .	112
5.17	Matrices de corrélation calculées par la variance des sous-échantillons . . . . .	113
5.18	Matrices de corrélation calculées par la technique du <i>bootstrap</i> . . . . .	115
5.19	Différence des matrices de corrélation estimées par la variance des sous-échantillons et par la technique du <i>bootstrap</i> . . . . .	115
5.20	Moyennes des monopôles et quadrupôles des 25 simulations . . . . .	117
5.21	Comparaison des monopôles et quadrupôles des données et des simulations . . . . .	118

5.22	Multipôles avec et sans pics utilisés pour les ajustements . . . . .	120
5.23	Ajustement des multipôles de 5 simulations . . . . .	122
5.24	Distributions des $\chi^2$ des ajustements isotropes . . . . .	123
5.25	Distributions des $\chi^2$ des ajustements à deux dimensions . . . . .	123
5.26	Valeurs des $\alpha_{iso}$ . . . . .	124
5.27	Valeurs des $\alpha_t$ et $\alpha_p$ . . . . .	124
5.28	Evolution du $\Delta\chi^2$ en fonction de $\alpha_{iso}$ . . . . .	126
5.29	Ajustement isotrope des données . . . . .	126
5.30	Nappe de $\chi^2$ en fonction de $\alpha_t$ et $1/\alpha_p$ . . . . .	128
5.31	Ajustement à deux dimensions des données . . . . .	128
5.32	Comparaison des multipôles du NGC et du SGC . . . . .	129
5.33	$\Delta\chi^2$ en fonction de $\alpha_{iso}$ pour le NGC . . . . .	130
5.34	Comparaison des multipôles avec et sans coupure des raies de Balmer . . . . .	130
5.35	$\Delta\chi^2$ en fonction de $\alpha_{iso}$ avec et sans coupure des raies de Balmer . . . . .	131
5.36	Comparaison des multipôles des échantillons de plaques identiques et plaques différentes . . . . .	132
5.37	Distribution des paires de pixels en fonction de la séparation entre pixels pour les échantillons de plaques identiques et plaques différentes . . . . .	132
5.38	$\Delta\chi^2$ en fonction de $\alpha_{iso}$ pour les échantillons de plaques identiques et de plaques différentes . . . . .	133
5.39	Contraintes sur $\Omega_m$ et $\Omega_\Lambda$ . . . . .	135
5.40	Contraintes sur $\Omega_m$ et $w$ . . . . .	136
5.41	Contraintes combinées avec Planck pour un univers $\Lambda$ CDM ouvert . . . . .	137
5.42	Contraintes combinées avec Planck pour un univers $w_o w_a$ CDM ouvert . . . . .	138



# Bibliographie

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A. and et al. (2009). The Seventh Data Release of the Sloan Digital Sky Survey. *ApJS* 182, 543–558.

Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anderson, S. F., Anderton, T., Andrews, B. H., Aubourg, É., Bailey, S., Balbinot, E., Barnes, R. and et al. (2012). The Ninth Data Release of the Sloan Digital Sky Survey : First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. *ApJS* 203, 21.

Aihara, H., Allende Prieto, C., An, D., Anderson, S. F., Aubourg, É., Balbinot, E., Beers, T. C., Berlind, A. A., Bickerton, S. J., Bizyaev, D., Blanton, M. R., Bochanski, J. J., Bolton, A. S., Bovy, J., Brandt, W. N., Brinkmann, J., Brown, P. J., Brownstein, J. R., Busca, N. G., Campbell, H., Carr, M. A., Chen, Y., Chiappini, C., Comparat, J., Connolly, N., Cortes, M., Croft, R. A. C., Cuesta, A. J., da Costa, L. N., Davenport, J. R. A., Dawson, K., Dhital, S., Ealet, A., Ebelke, G. L., Edmondson, E. M., Eisenstein, D. J., Escoffier, S., Esposito, M., Evans, M. L., Fan, X., Femenía Castellá, B., Font-Ribera, A., Frinchaboy, P. M., Ge, J., Gillespie, B. A., Gilmore, G., González Hernández, J. I., Gott, J. R., Gould, A., Grebel, E. K., Gunn, J. E., Hamilton, J.-C., Harding, P., Harris, D. W., Hawley, S. L., Hearty, F. R., Ho, S., Hogg, D. W., Holtzman, J. A., Honscheid, K., Inada, N., Ivans, I. I., Jiang, L., Johnson, J. A., Jordan, C., Jordan, W. P., Kazin, E. A., Kirkby, D., Klaene, M. A., Knapp, G. R., Kneib, J.-P., Kochanek, C. S., Koesterke, L., Kollmeier, J. A., Kron, R. G., Lampeitl, H., Lang, D., Le Goff, J.-M., Lee, Y. S., Lin, Y.-T., Long, D. C., Loomis, C. P., Lucatello, S., Lundgren, B., Lupton, R. H., Ma, Z., MacDonald, N., Mahadevan, S., Maia, M. A. G., Makler, M., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Maraston, C., Margala, D., Masters, K. L., McBride, C. K., McGehee, P. M., McGreer, I. D., Ménard, B., Miralda-Escudé, J., Morrison, H. L., Mullally, F., Muna, D., Munn, J. A., Murayama, H., Myers, A. D., Naugle, T., Neto, A. F., Nguyen, D. C., Nichol, R. C., O’Connell, R. W., Ogando, R. L. C., Olmstead, M. D., Oravetz, D. J., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Pandey, P., Pâris, I., Percival, W. J., Petitjean, P., Pfaffenberger, R., Pforr, J., Phleps, S., Pichon, C., Pieri, M. M., Prada, F., Price-Whelan, A. M., Raddick, M. J., Ramos, B. H. F., Reylé, C., Rich, J., Richards, G. T., Rix, H.-W., Robin, A. C., Rocha-Pinto, H. J., Rockosi, C. M., Roe, N. A., Rollinde, E., Ross, A. J., Ross, N. P., Rossetto, B. M., Sánchez, A. G., Sayres, C., Schlegel, D. J., Schlesinger, K. J., Schmidt, S. J., Schneider, D. P., Sheldon, E., Shu, Y., Simmerer, J., Simmons, A. E., Sivarani, T., Snedden, S. A., Sobeck, J. S., Steinmetz, M., Strauss, M. A., Szalay, A. S., Tanaka, M., Thakar, A. R., Thomas, D., Tinker, J. L., Tofflemire, B. M., Tojeiro, R., Tremonti, C. A., Vandenberg, J., Vargas Magaña, M., Verde, L., Vogt, N. P., Wake, D. A., Wang, J., Weaver, B. A., Weinberg, D. H., White, M., White, S. D. M., Yanny, B., Yasuda, N., Yèche, C. and Zehavi, I. (2011). The Eighth Data Release of the Sloan Digital Sky Survey : First Data from SDSS-III. *ApJS* 193, 29.

Anderson, L., Aubourg, E., Bailey, S., Bizyaev, D., Blanton, M., Bolton, A. S., Brinkmann, J., Brownstein, J. R., Burden, A., Cuesta, A. J., da Costa, L. A. N., Dawson, K. S., de Putter, R., Eisenstein, D. J., Gunn, J. E., Guo, H., Hamilton, J.-C., Harding, P., Ho, S., Honscheid, K., Kazin, E., Kirkby, D., Kneib, J.-P., Labatie, A., Loomis, C., Lupton, R. H., Malanushenko, E., Malanushenko, V., Mandelbaum, R., Manera, M., Maraston, C., McBride, C. K., Mehta, K. T., Mena, O., Montesano, F., Muna, D., Nichol, R. C., Nuza, S. E., Olmstead, M. D., Oravetz, D., Padmanabhan, N., Palanque-Delabrouille, N., Pan, K., Parejko, J., Pâris, I., Percival, W. J., Petitjean, P., Prada, F., Reid, B., Roe, N. A., Ross, A. J., Ross, N. P., Samushia, L., Sánchez, A. G., Schlegel, D. J., Schneider, D. P., Scóccola, C. G., Seo, H.-J., Sheldon, E. S., Simmons, A., Skibba, R. A., Strauss, M. A., Swanson, M. E. C., Thomas, D., Tinker, J. L., Tojeiro, R., Magaña, M. V., Verde, L., Wagner, C., Wake, D. A., Weaver, B. A., Weinberg, D. H., White, M., Xu, X., Yèche, C., Zehavi, I. and Zhao, G.-B. (2012). The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : baryon acoustic oscillations in the Data Release 9 spectroscopic galaxy sample. *MNRAS* 427, 3435–3467.

Antonucci, R. (1993). Unified models for active galactic nuclei and quasars. *ARA&A* 31, 473–521.

Beutler, F., Blake, C., Colless, M., Jones, D. H., Staveley-Smith, L., Campbell, L., Parker, Q., Saunders, W. and Watson, F. (2011). The 6dF Galaxy Survey : baryon acoustic oscillations and the local Hubble constant. *MNRAS* 416, 3017–3032.

Bi, H. (1993). Lyman-alpha absorption spectrum of the primordial intergalactic medium. *ApJ* 405, 479–490.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press.

Blake, C., Kazin, E. A., Beutler, F., Davis, T. M., Parkinson, D., Brough, S., Colless, M., Contreras, C., Couch, W., Croom, S., Croton, D., Drinkwater, M. J., Forster, K., Gilbank, D., Gladders, M., Glazebrook, K., Jelliffe, B., Jurek, R. J., Li, I.-H., Madore, B., Martin, D. C., Pimblet, K., Poole, G. B., Pracy, M., Sharp, R., Wisnioski, E., Woods, D., Wyder, T. K. and Yee, H. K. C. (2011). The WiggleZ Dark Energy Survey : mapping the distance-redshift relation with baryon acoustic oscillations. *MNRAS* 418, 1707–1724.

Blanton, M. R., Lin, H., Lupton, R. H., Maley, F. M., Young, N., Zehavi, I. and Loveday, J. (2003). An Efficient Targeting Strategy for Multiobject Spectrograph Surveys : the Sloan Digital Sky Survey “Tiling” Algorithm. *AJ* 125, 2276–2286.

Busca, N. G., Delubac, T., Rich, J., Bailey, S., Font-Ribera, A., Kirkby, D., Le Goff, J.-M., Pieri, M. M., Slosar, A., Aubourg, É., Bautista, J. E., Bizyaev, D., Blomqvist, M., Bolton, A. S., Bovy, J., Brewington, H., Borde, A., Brinkmann, J., Carithers, B., Croft, R. A. C., Dawson, K. S., Ebelke, G., Eisenstein, D. J., Hamilton, J.-C., Ho, S., Hogg, D. W., Honscheid, K., Lee, K.-G., Lundgren, B., Malanushenko, E., Malanushenko, V., Margala, D., Maraston, C., Mehta, K., Miralda-Escudé, J., Myers, A. D., Nichol, R. C., Noterdaeme, P., Olmstead, M. D., Oravetz, D., Palanque-Delabrouille, N., Pan, K., Pâris, I., Percival, W. J., Petitjean, P., Roe, N. A., Rolin, E., Ross, N. P., Rossi, G., Schlegel, D. J., Schneider, D. P., Sheldon, A., Sheldon, E. S., Simmons, A., Snedden, S., Tinker, J. L., Viel, M., Weaver, B. A., Weinberg, D. H., White, M., Yèche, C. and York, D. G. (2013). Baryon acoustic oscillations in the Ly $\alpha$  forest of BOSS quasars. *A&A* 552, A96.

- Cen, R., Miralda-Escudé, J., Ostriker, J. P. and Rauch, M. (1994). Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest. *ApJ* 437, L9–L12.
- Chevallier, M. and Polarski, D. (2001). Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D* 10, 213–223.
- Conley, A., Guy, J., Sullivan, M., Regnault, N., Astier, P., Balland, C., Basa, S., Carlberg, R. G., Fouchez, D., Hardin, D., Hook, I. M., Howell, D. A., Pain, R., Palanque-Delabrouille, N., Perrett, K. M., Pritchet, C. J., Rich, J., Ruhlmann-Kleider, V., Balam, D., Baumont, S., Ellis, R. S., Fabbro, S., Fakhouri, H. K., Fourmanoit, N., González-Gaitán, S., Graham, M. L., Hudson, M. J., Hsiao, E., Kronborg, T., Lidman, C., Mourao, A. M., Neill, J. D., Perlmutter, S., Ripoche, P., Suzuki, N. and Walker, E. S. (2011). Supernova Constraints and Systematic Uncertainties from the First Three Years of the Supernova Legacy Survey. *ApJS* 192, 1.
- Croft, R. A. C., Weinberg, D. H., Katz, N. and Hernquist, L. (1997). Intergalactic Helium Absorption in Cold Dark Matter Models. *ApJ* 488, 532.
- Croom, S. M., Richards, G. T., Shanks, T., Boyle, B. J., Sharp, R. G., Bland-Hawthorn, J., Bridges, T., Brunner, R. J., Cannon, R., Carson, D., Chiu, K., Colless, M., Couch, W., de Propris, R., Drinkwater, M. J., Edge, A., Fine, S., Loveday, J., Miller, L., Myers, A. D., Nichol, R. C., Outram, P., Pimblet, K., Roseboom, I., Ross, N., Schneider, D. P., Smith, A., Stoughton, C., Strauss, M. A. and Wake, D. (2009). The 2dF-SDSS LRG and QSO Survey : the spectroscopic QSO catalogue. *MNRAS* 392, 19–44.
- Croom, S. M., Smith, R. J., Boyle, B. J., Shanks, T., Miller, L., Outram, P. J. and Loaring, N. S. (2004). The 2dF QSO Redshift Survey - XII. The spectroscopic catalogue and luminosity function. *MNRAS* 349, 1397–1418.
- de Vries, W. H., Becker, R. H., White, R. L. and Loomis, C. (2005). Structure Function Analysis of Long-Term Quasar Variability. *AJ* 129, 615–629.
- Efron, B. and Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *The American Statistician* 37, 36–48.
- Einstein, A. (1917). Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, Seite 142-152. , 142–152.
- Eisenstein, D. J. and Hu, W. (1998). Baryonic Features in the Matter Transfer Function. *ApJ* 496, 605.
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., Scoccimarro, R., Blanton, M. R., Nichol, R. C., Scranton, R., Seo, H.-J., Tegmark, M., Zheng, Z., Anderson, S. F., Annis, J., Bahcall, N., Brinkmann, J., Burles, S., Castander, F. J., Connolly, A., Csabai, I., Doi, M., Fukugita, M., Frieman, J. A., Glazebrook, K., Gunn, J. E., Hendry, J. S., Hennessy, G., Ivezić, Z., Kent, S., Knapp, G. R., Lin, H., Loh, Y.-S., Lupton, R. H., Margon, B., McKay, T. A., Meiksin, A., Munn, J. A., Pope, A., Richmond, M. W., Schlegel, D., Schneider, D. P., Shimasaku, K., Stoughton, C., Strauss, M. A., SubbaRao, M., Szalay, A. S., Szapudi, I., Tucker, D. L., Yanny, B. and York, D. G. (2005). Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ* 633, 560–574.
- Fan, X. (1999). Simulation of Stellar Objects in SDSS Color Space. *AJ* 117, 2528–2551.



- Fixsen, D. J. (2009). The Temperature of the Cosmic Microwave Background. *ApJ* 707, 916–920.
- Font-Ribera, A., McDonald, P. and Miralda-Escudé, J. (2012). Generating mock data sets for large-scale Lyman- $\alpha$  forest correlation measurements. *JCAP* 1, 1.
- Frieman, J. A., Bassett, B., Becker, A., Choi, C., Cinabro, D., DeJongh, F., Depoy, D. L., Dilday, B., Doi, M., Garnavich, P. M., Hogan, C. J., Holtzman, J., Im, M., Jha, S., Kessler, R., Konishi, K., Lampeitl, H., Marriner, J., Marshall, J. L., McGinnis, D., Miknaitis, G., Nichol, R. C., Prieto, J. L., Riess, A. G., Richmond, M. W., Romani, R., Sako, M., Schneider, D. P., Smith, M., Takahashi, N., Tokita, K., van der Heyden, K., Yasuda, N., Zheng, C., Adelman-McCarthy, J., Annis, J., Assef, R. J., Barentine, J., Bender, R., Blandford, R. D., Boroski, W. N., Bremer, M., Brewington, H., Collins, C. A., Crotts, A., Dembicky, J., Eastman, J., Edge, A., Edmondson, E., Elson, E., Eyller, M. E., Filippenko, A. V., Foley, R. J., Frank, S., Goobar, A., Gueth, T., Gunn, J. E., Harvanek, M., Hopp, U., Ihara, Y., Ivezić, Ž., Kahn, S., Kaplan, J., Kent, S., Ketzeback, W., Kleinman, S. J., Kollatschny, W., Kron, R. G., Krzesiński, J., Lamenti, D., Leloudas, G., Lin, H., Long, D. C., Lucey, J., Lupton, R. H., Malanushenko, E., Malanushenko, V., McMillan, R. J., Mendez, J., Morgan, C. W., Morokuma, T., Nitta, A., Ostman, L., Pan, K., Rockosi, C. M., Romer, A. K., Ruiz-Lapuente, P., Saurage, G., Schlesinger, K., Snedden, S. A., Sollerman, J., Stoughton, C., Stritzinger, M., Subba Rao, M., Tucker, D., Vaisanen, P., Watson, L. C., Watters, S., Wheeler, J. C., Yanny, B. and York, D. (2008). The Sloan Digital Sky Survey-II Supernova Survey : Technical Summary. *AJ* 135, 338–347.
- Gunn, J. E., Carr, M., Rockosi, C., Sekiguchi, M., Berry, K., Elms, B., de Haas, E., Ivezić, Ž., Knapp, G., Lupton, R., Pauls, G., Simcoe, R., Hirsch, R., Sanford, D., Wang, S., York, D., Harris, F., Annis, J., Bartozek, L., Boroski, W., Bakken, J., Haldeman, M., Kent, S., Holm, S., Holmgren, D., Petravick, D., Prosapio, A., Rechenmacher, R., Doi, M., Fukugita, M., Shimasaku, K., Okada, N., Hull, C., Siegmund, W., Mannery, E., Blouke, M., Heidtman, D., Schneider, D., Lucinio, R. and Brinkman, J. (1998). The Sloan Digital Sky Survey Photometric Camera. *AJ* 116, 3040–3081.
- Gunn, J. E. and Peterson, B. A. (1965). On the Density of Neutral Hydrogen in Intergalactic Space. *ApJ* 142, 1633–1641.
- Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., Kent, S. M., Lupton, R. H., Rockosi, C. M., Evans, M. L., Waddell, P., Anderson, J. E., Annis, J., Barentine, J. C., Bartoszek, L. M., Bastian, S., Bracker, S. B., Brewington, H. J., Briegel, C. I., Brinkmann, J., Brown, Y. J., Carr, M. A., Czarapata, P. C., Drennan, C. C., Dombeck, T., Federwitz, G. R., Gillespie, B. A., Gonzales, C., Hansen, S. U., Harvanek, M., Hayes, J., Jordan, W., Kinney, E., Klaene, M., Kleinman, S. J., Kron, R. G., Kresinski, J., Lee, G., Limmongkol, S., Lindenmeyer, C. W., Long, D. C., Loomis, C. L., McGehee, P. M., Mantsch, P. M., Neilsen, Jr., E. H., Neswold, R. M., Newman, P. R., Nitta, A., Peoples, Jr., J., Pier, J. R., Prieto, P. S., Prosapio, A., Rivetta, C., Schneider, D. P., Snedden, S. and Wang, S.-i. (2006). The 2.5 m Telescope of the Sloan Digital Sky Survey. *AJ* 131, 2332–2359.
- Hamilton, A. J. S. (1992). Measuring Omega and the real correlation function from the redshift correlation function. *ApJ* 385, L5–L8.
- Hernquist, L., Katz, N., Weinberg, D. H. and Miralda-Escudé, J. (1996). The Lyman-Alpha Forest in the Cold Dark Matter Model. *ApJ* 457, L51.

- Hui, L. and Gnedin, N. Y. (1997). Equation of state of the photoionized intergalactic medium. *MNRAS* 292, 27.
- Kaiser, N. (1987). Clustering in real space and in redshift space. *MNRAS* 227, 1–21.
- Kirkby, D., Margala, D., Slosar, A., Bailey, S., Busca, N. G., Delubac, T., Rich, J., Bautista, J. E., Blomqvist, M., Brownstein, J. R., Carithers, B., Croft, R. A. C., Dawson, K. S., Font-Ribera, A., Miralda-Escudé, J., Myers, A. D., Nichol, R. C., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Rossi, G., Schlegel, D. J., Schneider, D. P., Viel, M., Weinberg, D. H. and Yèche, C. (2013). Fitting methods for baryon acoustic oscillations in the Lyman- $\alpha$  forest fluctuations in BOSS data release 9. *JCAP* 3, 24.
- Le Goff, J. M., Magneville, C., Rollinde, E., Peirani, S., Petitjean, P., Pichon, C., Rich, J., Yèche, C., Aubourg, E., Busca, N., Charlassier, R., Delubac, T., Hamilton, J. C., Palanque Delabrouille, N., Pâris, I. and Vargas Magaña, M. (2011). Simulations of BAO reconstruction with a quasar Ly- $\alpha$  survey. *A&A* 534, A135.
- Lee, K.-G., Suzuki, N. and Spergel, D. N. (2012). Mean-flux-regulated Principal Component Analysis Continuum Fitting of Sloan Digital Sky Survey Ly $\alpha$  Forest Spectra. *AJ* 143, 51.
- Lewis, A., Challinor, A. and Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ* 538, 473–476.
- Linder, E. V. (2003). Mapping the Dark Energy Equation of State. *ArXiv Astrophysics e-prints* .
- Lineweaver, C. H., Tenorio, L., Smoot, G. F., Keegstra, P., Banday, A. J. and Lubin, P. (1996). The Dipole Observed in the COBE DMR 4 Year Data. *ApJ* 470, 38.
- LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S. and et al. (2009). LSST Science Book, Version 2.0. *ArXiv e-prints* .
- MacLeod, C., Ivezić, Ž., de Vries, W., Sesar, B. and Becker, A. (2008). Time Variability of Quasars : the Structure Function Variance. In American Institute of Physics Conference Series, (Bailer-Jones, C. A. L., ed.), vol. 1082, of American Institute of Physics Conference Series pp. 282–286,.
- McDonald, P. (2003). Toward a Measurement of the Cosmological Geometry at  $z \sim 2$  : Predicting Ly $\alpha$  Forest Correlation in Three Dimensions and the Potential of Future Data Sets. *ApJ* 585, 34–51.
- McDonald, P. and Eisenstein, D. J. (2007). Dark energy and curvature from a future baryonic acoustic oscillation survey using the Lyman- $\alpha$  forest. *Phys. Rev. D* 76, 063009.
- McDonald, P., Seljak, U., Burles, S., Schlegel, D. J., Weinberg, D. H., Cen, R., Shih, D., Schaye, J., Schneider, D. P., Bahcall, N. A., Briggs, J. W., Brinkmann, J., Brunner, R. J., Fukugita, M., Gunn, J. E., Ivezić, Ž., Kent, S., Lupton, R. H. and Vanden Berk, D. E. (2006). The Ly $\alpha$  Forest Power Spectrum from the Sloan Digital Sky Survey. *ApJS* 163, 80–109.
- McQuinn, M. and White, M. (2011). On estimating Ly $\alpha$  forest correlations between multiple sightlines. *MNRAS* 415, 2257–2269.

- Miralda-Escude, J. and Rees, M. J. (1993). Tests for the minihalo model of the Lyman-alpha forest. *MNRAS* 260, 617–624.
- Noterdaeme, P., Petitjean, P., Ledoux, C. and Srianand, R. (2009). DLAs in SDSS-DR7 (Noterdaeme+, 2009). *VizieR Online Data Catalog* 350, 51087.
- Palanque-Delabrouille, N., Yeche, C., Myers, A. D., Petitjean, P., Ross, N. P., Sheldon, E., Aubourg, E., Delubac, T., Le Goff, J.-M., Pâris, I., Rich, J., Dawson, K. S., Schneider, D. P. and Weaver, B. A. (2011). Variability selected high-redshift quasars on SDSS Stripe 82. *A&A* 530, A122.
- Pâris, I., Petitjean, P., Aubourg, É., Bailey, S., Ross, N. P., Myers, A. D., Strauss, M. A., Anderson, S. F., Arnau, E., Bautista, J., Bizyaev, D., Bolton, A. S., Bovy, J., Brandt, W. N., Brewington, H., Browstein, J. R., Busca, N., Capellupo, D., Carithers, W., Croft, R. A. C., Dawson, K., Delubac, T., Ebelke, G., Eisenstein, D. J., Engelke, P., Fan, X., Filiz Ak, N., Finley, H., Font-Ribera, A., Ge, J., Gibson, R. R., Hall, P. B., Hamann, F., Hennawi, J. F., Ho, S., Hogg, D. W., Ivezić, Ž., Jiang, L., Kimball, A. E., Kirkby, D., Kirkpatrick, J. A., Lee, K.-G., Le Goff, J.-M., Lundgren, B., MacLeod, C. L., Malanushenko, E., Malanushenko, V., Maraston, C., McGreer, I. D., McMahon, R. G., Miralda-Escudé, J., Muna, D., Noterdaeme, P., Oravetz, D., Palanque-Delabrouille, N., Pan, K., Perez-Fournon, I., Pieri, M. M., Richards, G. T., Rollinde, E., Sheldon, E. S., Schlegel, D. J., Schneider, D. P., Slosar, A., Shelden, A., Shen, Y., Simmons, A., Snedden, S., Suzuki, N., Tinker, J., Viel, M., Weaver, B. A., Weinberg, D. H., White, M., Wood-Vasey, W. M. and Yèche, C. (2012). The Sloan Digital Sky Survey quasar catalog : ninth data release. *A&A* 548, A66.
- Percival, W. J., Reid, B. A., Eisenstein, D. J., Bahcall, N. A., Budavari, T., Frieman, J. A., Fukugita, M., Gunn, J. E., Ivezić, Ž., Knapp, G. R., Kron, R. G., Loveday, J., Lupton, R. H., McKay, T. A., Meiksin, A., Nichol, R. C., Pope, A. C., Schlegel, D. J., Schneider, D. P., Spergel, D. N., Stoughton, C., Strauss, M. A., Szalay, A. S., Tegmark, M., Vogeley, M. S., Weinberg, D. H., York, D. G. and Zehavi, I. (2010). Baryon acoustic oscillations in the Sloan Digital Sky Survey Data Release 7 galaxy sample. *MNRAS* 401, 2148–2168.
- Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., Ellis, R. S., Irwin, M., McMahon, R. G., Ruiz-Lapuente, P., Walton, N., Schaefer, B., Boyle, B. J., Filippenko, A. V., Matheson, T., Fruchter, A. S., Panagia, N., Newberg, H. J. M., Couch, W. J. and Supernova Cosmology Project (1999). Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *ApJ* 517, 565–586.
- Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Bacigalupi, C., Balbi, A., Banday, A. J., Barreiro, R. B., Battaner, E., Benabed, K., Benoit-Levy, A., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bikmaev, I., Bobin, J., Bock, J. J., Bonaldi, A., Bond, J. R., Borrill, J., Bouchet, F. R., Burigana, C., Butler, R. C., Cabella, P., Cardoso, J.-F., Catalano, A., Chamballu, A., Chiang, L.-Y., Chon, G., Christensen, P. R., Clements, D. L., Colombi, S., Colombo, L. P. L., Crill, B. P., Cuttaia, F., Da Silva, A., Dahle, H., Davies, R. D., Davis, R. J., de Bernardis, P., de Gasperis, G., de Zotti, G., Delabrouille, J., Democles, J., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Dore, O., Doerl, U., Douspis, M., Dupac, X., Ensslin, T. A., Finelli, F., Flores-Cacho, I., Forni, O., Frailis, M., Frommert, M., Galeotta, S., Ganga, K., Genova-Santos, R. T., Giard, M., Giardino, G., Gonzalez-Nuevo, J., Gregorio, A., Gruppuso, A., Hansen, F. K., Harrison, D., Hernandez-Monteagudo, C., Herranz, D., Hildebrandt,

S. R., Hivon, E., Holmes, W. A., Hovest, W., Huppenberger, K. M., Hurier, G., Jaffe, T. R., Jaffe, A. H., Jasche, J., Jones, W. C., Juvela, M., Keihanen, E., Keskitalo, R., Khamitov, I., Kisner, T. S., Knoche, J., Kunz, M., Kurki-Suonio, H., Lagache, G., Lahteenmaki, A., Lamarre, J.-M., Lasenby, A., Lawrence, C. R., Le Jeune, M., Leonardi, R., Lilje, P. B., Linden-Vornle, M., Lopez-Caniego, M., Macias-Perez, J. F., Maino, D., Mak, D. S. Y., Mandolesi, N., Maris, M., Marleau, F., Martinez-Gonzalez, E., Masi, S., Matarrese, S., Mazzotta, P., Melchiorri, A., Melin, J.-B., Mendes, L., Mennella, A., Migliaccio, M., Mitra, S., Miville-Deschenes, M.-A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Norgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Osborne, S., Pagano, L., Paoletti, D., Perdureau, O., Perrotta, F., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Popa, L., Poutanen, T., Pratt, G. W., Prunet, S., Puget, J.-L., Puisieux, S., Rachen, J. P., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Ricciardi, S., Roman, M., Rubino-Martin, J. A., Rusholme, B., Sandri, M., Savini, G., Scott, D., Spencer, L., Sunyaev, R., Sutton, D., Suur-Uski, A.-S., Sygnet, J.-F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Tucci, M., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Vittorio, N., Wade, L. A., Welikala, N., Yvon, D., Zacchei, A., Zibin, J. P. and Zonca, A. (2013a). Planck intermediate results. XIII. Constraints on peculiar velocities. ArXiv e-prints .

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Bacigalupi, C., Balbi, A., Banday, A. J., Barreiro, R. B., Battaner, E., Benabed, K., Benoit-Levy, A., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bikmaev, I., Bobin, J., Bock, J. J., Bonaldi, A., Bond, J. R., Borrill, J., Bouchet, F. R., Burigana, C., Butler, R. C., Cabella, P., Cardoso, J.-F., Catalano, A., Chamballu, A., Chiang, L.-Y., Chon, G., Christensen, P. R., Clements, D. L., Colombi, S., Colombo, L. P. L., Crill, B. P., Cuttaia, F., Da Silva, A., Dahle, H., Davies, R. D., Davis, R. J., de Bernardis, P., de Gasperis, G., de Zotti, G., Delabrouille, J., Democles, J., Diego, J. M., Dolag, K., Dole, H., Donzelli, S., Dore, O., Doerl, U., Douspis, M., Dupac, X., Ensslin, T. A., Finelli, F., Flores-Cacho, I., Forni, O., Frailis, M., Frommert, M., Galeotta, S., Ganga, K., Genova-Santos, R. T., Giard, M., Giardino, G., Gonzalez-Nuevo, J., Gregorio, A., Gruppuso, A., Hansen, F. K., Harrison, D., Hernandez-Monteagudo, C., Herranz, D., Hildebrandt, S. R., Hivon, E., Holmes, W. A., Hovest, W., Huppenberger, K. M., Hurier, G., Jaffe, T. R., Jaffe, A. H., Jasche, J., Jones, W. C., Juvela, M., Keihanen, E., Keskitalo, R., Khamitov, I., Kisner, T. S., Knoche, J., Kunz, M., Kurki-Suonio, H., Lagache, G., Lahteenmaki, A., Lamarre, J.-M., Lasenby, A., Lawrence, C. R., Le Jeune, M., Leonardi, R., Lilje, P. B., Linden-Vornle, M., Lopez-Caniego, M., Macias-Perez, J. F., Maino, D., Mak, D. S. Y., Mandolesi, N., Maris, M., Marleau, F., Martinez-Gonzalez, E., Masi, S., Matarrese, S., Mazzotta, P., Melchiorri, A., Melin, J.-B., Mendes, L., Mennella, A., Migliaccio, M., Mitra, S., Miville-Deschenes, M.-A., Moneti, A., Montier, L., Morgante, G., Mortlock, D., Moss, A., Munshi, D., Murphy, J. A., Naselsky, P., Nati, F., Natoli, P., Netterfield, C. B., Norgaard-Nielsen, H. U., Noviello, F., Novikov, D., Novikov, I., Osborne, S., Pagano, L., Paoletti, D., Perdureau, O., Perrotta, F., Piacentini, F., Piat, M., Pierpaoli, E., Pietrobon, D., Plaszczynski, S., Pointecouteau, E., Polenta, G., Popa, L., Poutanen, T., Pratt, G. W., Prunet, S., Puget, J.-L., Puisieux, S., Rachen, J. P., Rebolo, R., Reinecke, M., Remazeilles, M., Renault, C., Ricciardi, S., Roman, M., Rubino-Martin, J. A., Rusholme, B., Sandri, M., Savini, G., Scott, D., Spencer, L., Sunyaev, R., Sutton, D., Suur-Uski, A.-S., Sygnet, J.-F., Tauber, J. A., Terenzi, L., Toffolatti, L., Tomasi, M., Tristram, M., Tucci, M., Valenziano, L., Valiviita, J., Van Tent, B., Vielva, P., Villa, F., Vittorio, N., Wade, L. A., Welikala, N., Yvon, D., Zacchei, A., Zibin, J. P. and Zonca, A. (2013b). Planck 2013 results. XVI. Cosmological parameters. ArXiv e-prints .

- Rau, A., Kulkarni, S. R., Law, N. M., Bloom, J. S., Ciardi, D., Djorgovski, G. S., Fox, D. B., Gal-Yam, A., Grillmair, C. C., Kasliwal, M. M., Nugent, P. E., Ofek, E. O., Quimby, R. M., Reach, W. T., Shara, M., Bildsten, L., Cenko, S. B., Drake, A. J., Filippenko, A. V., Helfand, D. J., Helou, G., Howell, D. A., Poznanski, D. and Sullivan, M. (2009). Exploring the Optical Transient Sky with the Palomar Transient Factory. *PASP* *121*, 1334–1351.
- Reid, B. A., Samushia, L., White, M., Percival, W. J., Manera, M., Padmanabhan, N., Ross, A. J., Sánchez, A. G., Bailey, S., Bizyaev, D., Bolton, A. S., Brewington, H., Brinkmann, J., Brownstein, J. R., Cuesta, A. J., Eisenstein, D. J., Gunn, J. E., Honscheid, K., Malanushenko, E., Malanushenko, V., Maraston, C., McBride, C. K., Muna, D., Nichol, R. C., Oravetz, D., Pan, K., de Putter, R., Roe, N. A., Ross, N. P., Schlegel, D. J., Schneider, D. P., Seo, H.-J., Sheldon, A., Sheldon, E. S., Simmons, A., Skibba, R. A., Snedden, S., Swanson, M. E. C., Thomas, D., Tinker, J., Tojeiro, R., Verde, L., Wake, D. A., Weaver, B. A., Weinberg, D. H., Zehavi, I. and Zhao, G.-B. (2012). The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey : measurements of the growth of structure and expansion rate at  $z = 0.57$  from anisotropic clustering. *MNRAS* *426*, 2719–2737.
- Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B. and Tonry, J. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ* *116*, 1009–1038.
- Riess, A. G., Macri, L., Casertano, S., Lampeitl, H., Ferguson, H. C., Filippenko, A. V., Jha, S. W., Li, W. and Chornock, R. (2011). A 3% Solution : Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3. *ApJ* *730*, 119.
- Ross, N. P., Myers, A. D., Sheldon, E. S., Yèche, C., Strauss, M. A., Bovy, J., Kirkpatrick, J. A., Richards, G. T., Aubourg, É., Blanton, M. R., Brandt, W. N., Carithers, W. C., Croft, R. A. C., da Silva, R., Dawson, K., Eisenstein, D. J., Hennawi, J. F., Ho, S., Hogg, D. W., Lee, K.-G., Lundgren, B., McMahon, R. G., Miralda-Escudé, J., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Pieri, M. M., Rich, J., Roe, N. A., Schiminovich, D., Schlegel, D. J., Schneider, D. P., Slosar, A., Suzuki, N., Tinker, J. L., Weinberg, D. H., Weyant, A., White, M. and Wood-Vasey, W. M. (2012). The SDSS-III Baryon Oscillation Spectroscopic Survey : Quasar Target Selection for Data Release Nine. *ApJS* *199*, 3.
- Sachs, R. K. and Wolfe, A. M. (1967). Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *ApJ* *147*, 73.
- Schmidt, K. B., Marshall, P. J., Rix, H.-W., Jester, S., Hennawi, J. F. and Dobler, G. (2010). Selecting Quasars by Their Intrinsic Variability. *ApJ* *714*, 1194–1208.
- Schmidt, M. (1963). 3C 273 : A Star-Like Object with Large Red-Shift. *Nature* *197*, 1040.
- Schneider, D. P., Richards, G. T., Hall, P. B., Strauss, M. A., Anderson, S. F., Boroson, T. A., Ross, N. P., Shen, Y., Brandt, W. N., Fan, X., Inada, N., Jester, S., Knapp, G. R., Krawczyk, C. M., Thakar, A. R., Vanden Berk, D. E., Voges, W., Yanny, B., York, D. G., Bahcall, N. A., Bizyaev, D., Blanton, M. R., Brewington, H., Brinkmann, J., Eisenstein, D., Frieman, J. A., Fukugita, M., Gray, J., Gunn, J. E., Hibon, P., Ivezić, Ž., Kent, S. M., Kron, R. G., Lee, M. G., Lupton, R. H., Malanushenko, E., Malanushenko, V., Oravetz, D., Pan, K., Pier, J. R., Price, III, T. N., Saxe, D. H., Schlegel, D. J., Simmons, A., Snedden, S. A., SubbaRao, M. U., Szalay, A. S.

- and Weinberg, D. H. (2010). The Sloan Digital Sky Survey Quasar Catalog. V. Seventh Data Release. *AJ* 139, 2360.
- Scrimgeour, M. I., Davis, T., Blake, C., James, J. B., Poole, G. B., Staveley-Smith, L., Brough, S., Colless, M., Contreras, C., Couch, W., Croom, S., Croton, D., Drinkwater, M. J., Forster, K., Gilbank, D., Gladders, M., Glazebrook, K., Jelliffe, B., Jurek, R. J., Li, I.-h., Madore, B., Martin, D. C., Pimblet, K., Pracy, M., Sharp, R., Wisnioski, E., Woods, D., Wyder, T. K. and Yee, H. K. C. (2012). The WiggleZ Dark Energy Survey : the transition to large-scale cosmic homogeneity. *MNRAS* 425, 116–134.
- Seljak, U. and Zaldarriaga, M. (1996). A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *ApJ* 469, 437.
- Sesar, B., Ivezić, Ž., Lupton, R. H., Jurić, M., Gunn, J. E., Knapp, G. R., DeLee, N., Smith, J. A., Miknaitis, G., Lin, H., Tucker, D., Doi, M., Tanaka, M., Fukugita, M., Holtzman, J., Kent, S., Yanny, B., Schlegel, D., Finkbeiner, D., Padmanabhan, N., Rockosi, C. M., Bond, N., Lee, B., Stoughton, C., Jester, S., Harris, H., Harding, P., Brinkmann, J., Schneider, D. P., York, D., Richmond, M. W. and Vanden Berk, D. (2007). Exploring the Variable Sky with the Sloan Digital Sky Survey. *AJ* 134, 2236–2251.
- Slosar, A., Font-Ribera, A., Pieri, M. M., Rich, J., Le Goff, J.-M., Aubourg, É., Brinkmann, J., Busca, N., Carithers, B., Charlassier, R., Cortès, M., Croft, R., Dawson, K. S., Eisenstein, D., Hamilton, J.-C., Ho, S., Lee, K.-G., Lupton, R., McDonald, P., Medolin, B., Muna, D., Miralda-Escudé, J., Myers, A. D., Nichol, R. C., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Piškur, Y., Rollinde, E., Ross, N. P., Schlegel, D. J., Schneider, D. P., Sheldon, E., Weaver, B. A., Weinberg, D. H., Yeche, C. and York, D. G. (2011). The Lyman- $\alpha$  forest in three dimensions : measurements of large scale flux correlations from BOSS 1st-year data. *JCAP* 9, 1.
- Slosar, A., Iršič, V., Kirkby, D., Bailey, S., Busca, N. G., Delubac, T., Rich, J., Aubourg, É., Bautista, J. E., Bhardwaj, V., Blomqvist, M., Bolton, A. S., Bovy, J., Brownstein, J., Carithers, B., Croft, R. A. C., Dawson, K. S., Font-Ribera, A., Le Goff, J.-M., Ho, S., Honscheid, K., Lee, K.-G., Margala, D., McDonald, P., Medolin, B., Miralda-Escudé, J., Myers, A. D., Nichol, R. C., Noterdaeme, P., Palanque-Delabrouille, N., Pâris, I., Petitjean, P., Pieri, M. M., Piškur, Y., Roe, N. A., Ross, N. P., Rossi, G., Schlegel, D. J., Schneider, D. P., Suzuki, N., Sheldon, E. S., Seljak, U., Viel, M., Weinberg, D. H. and Yèche, C. (2013). Measurement of baryon acoustic oscillations in the Lyman- $\alpha$  forest fluctuations in BOSS data release 9. *JCAP* 4, 26.
- Smee, S., Gunn, J. E., Uomoto, A., Roe, N., Schlegel, D., Rockosi, C. M., Carr, M. A., Leger, F., Dawson, K. S., Olmstead, M. D., Brinkmann, J., Owen, R., Barkhouser, R. H., Honscheid, K., Harding, P., Long, D., Lupton, R. H., Loomis, C., Anderson, L., Annis, J., Bernardi, M., Bhardwaj, V., Bizyaev, D., Bolton, A. S., Brewington, H., Briggs, J. W., Burles, S., Burns, J. G., Castander, F., Connolly, A., Davenport, J. R., Ebelke, G., Epps, H., Feldman, P. D., Friedman, S., Frieman, J., Heckman, T., Hull, C. L., Knapp, G. R., Lawrence, D. M., Loveday, J., Mannery, E. J., Malanushenko, E., Malanushenko, V., Merrelli, A., Muna, D., Newman, P., Nichol, R. C., Oravetz, D., Pan, K., Pope, A. C., Ricketts, P. G., Shelden, A., Sandford, D., Siegmund, W., Simmons, A., Smith, D., Snedden, S., Schneider, D. P., Strauss, M., SubbaRao, M., Tremonti, C., Waddell, P. and York, D. G. (2012). The Multi-Object, Fiber-Fed Spectrographs for SDSS and the Baryon Oscillation Spectroscopic Survey. *ArXiv e-prints* .
- Suzuki, N. (2006). Quasar Spectrum Classification with Principal Component Analysis (PCA) : Emission Lines in the Ly $\alpha$  Forest. *ApJS* 163, 110–121.

- Suzuki, N., Tytler, D., Kirkman, D., O'Meara, J. M. and Lubin, D. (2005). Predicting QSO Continua in the Ly $\alpha$  Forest. *ApJ* 618, 592–600.
- Urry, C. M. and Padovani, P. (1995). Unified Schemes for Radio-Loud Active Galactic Nuclei. *PASP* 107, 803.
- Vanden Berk, D. E., Wilhite, B. C., Kron, R. G., Anderson, S. F., Brunner, R. J., Hall, P. B., Ivezić, Ž., Richards, G. T., Schneider, D. P., York, D. G., Brinkmann, J. V., Lamb, D. Q., Nichol, R. C. and Schlegel, D. J. (2004). The Ensemble Photometric Variability of  $\sim$ 25,000 Quasars in the Sloan Digital Sky Survey. *ApJ* 601, 692–714.
- Weymann, R. J., Carswell, R. F. and Smith, M. G. (1981). Absorption lines in the spectra of quasistellar objects. *ARA&A* 19, 41–76.
- White, M., Myers, A. D., Ross, N. P., Schlegel, D. J., Hennawi, J. F., Shen, Y., McGreer, I., Strauss, M. A., Bolton, A. S., Bovy, J., Fan, X., Miralda-Escude, J., Palanque-Delabrouille, N., Paris, I., Petitjean, P., Schneider, D. P., Viel, M., Weinberg, D. H., Yeche, C., Zehavi, I., Pan, K., Snedden, S., Bizyaev, D., Brewington, H., Brinkmann, J., Malanushenko, V., Malanushenko, E., Oravetz, D., Simmons, A., Sheldon, A. and Weaver, B. A. (2012). The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey. *MNRAS* 424, 933–950.
- Xu, X., Cuesta, A. J., Padmanabhan, N., Eisenstein, D. J. and McBride, C. K. (2013). Measuring  $D_A$  and  $H$  at  $z=0.35$  from the SDSS DR7 LRGs using baryon acoustic oscillations. *MNRAS* 431, 2834–2860.
- Yèche, C., Petitjean, P., Rich, J., Aubourg, E., Busca, N., Hamilton, J.-C., Le Goff, J.-M., Paris, I., Peirani, S., Pichon, C., Rollinde, E. and Vargas-Magaña, M. (2010). Artificial neural networks for quasar selection and photometric redshift determination. *A&A* 523, A14.
- Zhang, Y., Anninos, P. and Norman, M. L. (1995). A Multispecies Model for Hydrogen and Helium Absorbers in Lyman-Alpha Forest Clouds. *ApJ* 453, L57.





## Titre

Détections des oscillations acoustiques de baryons grâce aux forêts Lyman- $\alpha$  des spectres de quasars de l'expérience BOSS.

## Résumé

Les oscillations acoustiques de baryons (BAO) constituent une règle standard permettant de contraindre les différents modèles cosmologiques. Cette thèse rend compte de la première mesure des BAO dans la fonction de corrélation de la fraction de flux transmise des forêts Lyman- $\alpha$  des quasars à grands décalages spectraux. Cette détection utilise 89322 spectres de quasars mesurés par le Baryon Oscillation Spectroscopic Survey (BOSS) de la troisième génération du Sloan Digital Sky Survey (SDSS-III). Les quasars considérés possèdent des décalages spectraux compris dans l'intervalle  $2,1 < z < 3,5$ . Un pic dans la fonction de corrélation est détecté à  $1,043^{+0,021}_{-0,020}$  fois la position attendu du pic BAO pour le modèle de concordance  $\Lambda$ CDM. Cette mesure permet de contraindre la distance angulaire  $D_A$  ainsi que le paramètre de Hubble  $H$  à un décalage spectral moyen  $z = 2,38$ .

Par ailleurs cette thèse présente une nouvelle méthode de sélection des quasars par variabilité. Cette méthode est appliquée à la région du *Stripe* 82 où un grand nombre de données photométriques multi-époque est disponible. Sur cette région, elle permet d'atteindre une densité d'environ  $30 \text{ deg}^{-2}$  quasars contre  $18 \text{ deg}^{-2}$  pour les sélections usuelles par couleur.

---

## Title

Detection of baryon acoustic oscillations in the Lyman- $\alpha$  forests of BOSS quasar spectra.

## Abstract

Baryon acoustic oscillations (BAO) form a standard ruler that can be used to constrain different cosmological models. This thesis reports the first measurement of the BAO feature in the correlation function of the transmitted flux fraction in the Lyman- $\alpha$  forests of high redshift quasars. This detection uses 89322 quasar spectra measured by the Baryon Oscillation Spectroscopic Survey (BOSS) of the third generation of the Sloan Digital Sky Survey (SDSS-III). Redshift of used quasars belong to the range  $2.1 < z < 3.5$ . A peak in the correlation function is seen at  $1.043^{+0.021}_{-0.020}$  times the expected BAO peak position for a concordance  $\Lambda$ CDM model.

In addition this thesis presents a new method of quasar selection through their variability. This method is applied to the *Stripe* 82 region where an important number of multi-epoch photometric data is available. On this region it achieves a quasar density of  $30 \text{ deg}^{-2}$  to be compared with the  $18 \text{ deg}^{-2}$  of usual color selections.