

Les forêts Lyman- α du relevé eBOSS : comprendre les fonctions de corrélation et les systématiques

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 576, Particules, Hadrons, Énergie, Noyau :
Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat : astroparticules et cosmologie
Unité de recherche : Université Paris-Saclay, CEA, Département de Physique
des Particules, 91191, Gif-sur-Yvette, France.
Référent : faculté des sciences d'Orsay

**Thèse présentée et soutenue à l'Orme des merisiers,
le 24 septembre 2020, par**

Thomas ETOURNEAU

Thèse de doctorat

NNT : 2020UPASP029

Composition du jury :

Réza Ansari Professeur, Université Paris-Saclay	Président
Delphine Hardin Professeure, Sorbonne Université	Rapporteuse
Matthew Pieri Professeur, Université Aix-Marseille	Rapporteur
Andreu Font-Ribera Docteur, Institut de Física d'Altes Energies	Examineur

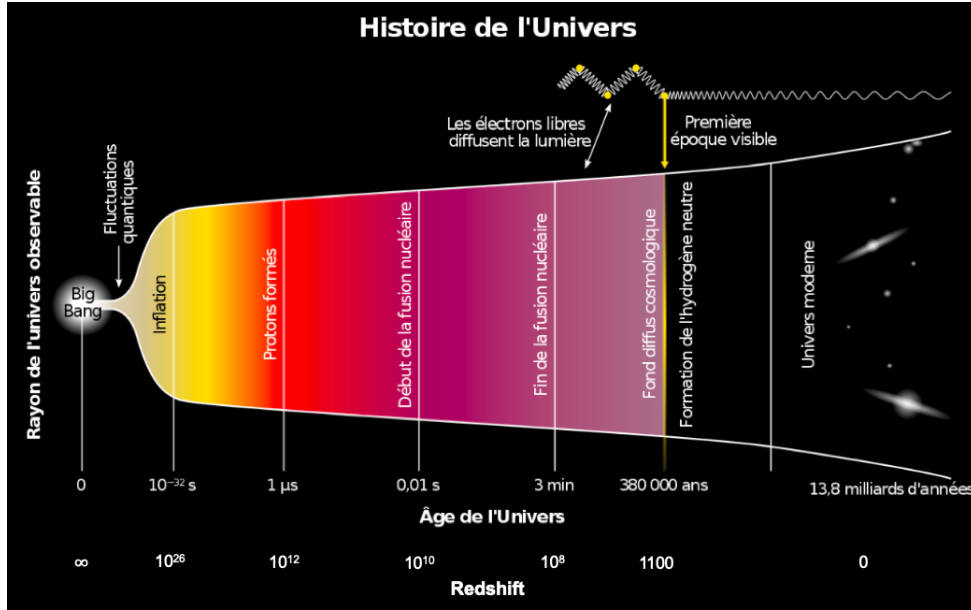


FIGURE 1.2 – Illustration de l’histoire de l’univers depuis ses origines jusqu’à aujourd’hui. Les principales étapes sont représentées : la formation des premiers protons et neutrons, puis des premiers noyaux, et enfin l’émission du CMB.

du milieu intergalactique

la phase de réionisation, la quasi-totalité des atomes d’hydrogène sont ionisés. Seul 1 atome sur 10 000 à 1 atome sur 1000 demeure neutre ~~au sein du milieu intergalactique.~~

En 1965, 17 ans après sa prédiction, le CMB est détecté par Penzias et Wilson, établissant ainsi le consensus sur les modèles de big bang. A partir de ce moment là, un certain nombre d’observations ont été menées par les cosmologistes afin de contraindre et distinguer les différents modèles de big bang. Nous retraçons dans la section suivante comment le modèle Λ CDM s’est établi en tant que modèle standard de la cosmologie, puis nous le décrivons en détail.

1.2 Le modèle Λ CDM

Le modèle Λ CDM est aujourd’hui le modèle cosmologique qui fait consensus dans la communauté scientifique. Il est souvent désigné comme le modèle standard de la cosmologie. C’est un modèle de big bang, décrivant un univers composé principalement d’énergie noire, aussi appelée constante cosmologique (Λ), et de matière noire froide (CDM : Cold Dark Matter). La figure 1.3 présente la répartition de ses différentes composantes.

1.2.1 Mise en place du modèle

Le modèle s’est établi suite à un certain nombre d’observations. D’abord, la détection du CMB en 1965, qui confirme les modèles de big bang. Puis l’introduction de la matière noire dans les modèles au cours des années 70 et 80, notamment grâce aux travaux de Vera Rubin sur le problème de la masse manquante dans les galaxies. Déjà en 1933, Fritz Zwicky remarquait que la masse visible dans les amas n’était pas suffisante pour expliquer leur cohésion, et supposa donc l’existence d’une matière invisible.

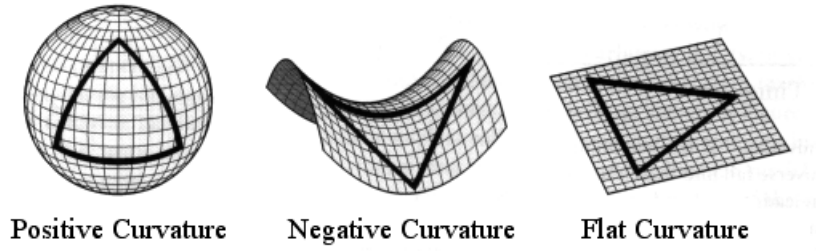


FIGURE 1.4 – Représentation de la courbure de l'univers : positive à gauche, négative au centre et nulle à droite.

Dans le cadre du modèle Λ CDM, la métrique utilisée est la métrique FLRW (pour Friedmann Lemaître Robertson Walker), elle s'exprime comme :

$$ds^2 = -dt^2 + R(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\Omega \right], \quad (1.2)$$

où $d\Omega = d\theta \sin(\theta) d\phi$, $R(t)$ rend compte de l'expansion de l'univers à l'instant t , et k vaut soit 1, 0 ou -1 selon que l'univers possède une courbure positive, nulle ou négative (figure 1.4). A l'aide d'un changement de coordonnées, il est possible de se ramener à la formule suivante :

$$ds^2 = -dt^2 + a(t) \left[d\chi^2 + S_k^2(\chi) d\Omega \right], \quad (1.3)$$

où $a(t) = R(t)/R(t_0)$, t_0 est le temps présent et S_k est défini comme

$$S_k(\chi) = R(t_0) \begin{cases} \sin(\chi/R(t_0)) & \text{si } k = 1 \\ \chi/R(t_0) & \text{si } k = 0 \\ \sinh(\chi/R(t_0)) & \text{si } k = -1 \end{cases}. \quad (1.4)$$

Cette formulation permet de mettre en évidence le rapport $a(t)$, appelé facteur d'échelle. Par définition il vaut 1 aujourd'hui. Afin de rendre compte de l'expansion, $a(t) < 1$ pour $t < t_0$ (passé) et $a(t) > 1$ pour $t > t_0$ (futur).

Le redshift

Le décalage vers le rouge, ou *redshift* en anglais, est la mesure du décalage du spectre des objets vers les grandes longueurs d'onde. Le redshift z est défini comme

$$1 + z = \frac{\lambda_o}{\lambda_e}, \quad (1.5)$$

où λ_e est la longueur d'onde émise et λ_o la longueur d'onde observée. **Le redshift trouve son origine dans deux effets. Le premier : le redshift est causé par la vitesse particulière de l'objet. Dans ce cas, nous parlons d'effet Doppler¹. Le second : dans le cadre des modèles de big**

1. L'effet Doppler est l'augmentation ou la diminution de la longueur d'onde d'une onde lorsque l'émetteur de cette dernière s'approche ou s'éloigne de l'observateur. L'exemple le plus connu est celui de l'ambulance : le son entendu est

bang, l'expansion de l'univers produit un décalage vers le rouge du spectre des objets distants. Dans ce cas, nous parlons de *redshift cosmologique*. Pour les objets lointains que nous considérons dans ce manuscrit, le redshift induit par les vitesses particulières est **faible** **négligeable** devant le redshift causé par l'expansion de l'univers. Nous confondons donc dans la suite redshift et redshift cosmologique. **Mais c'est une approx ... RSD**

Les objets distants s'éloignent de nous du fait de l'expansion. Similairement à l'effet Doppler, le spectre observé de ces objets est décalé vers les grandes longueurs d'onde. Mais contrairement à l'effet Doppler, le redshift n'est pas directement dû à la vitesse de récession de l'objet : à cause de l'expansion, les photons, lors de leur propagation, voient leur longueur d'onde augmenter. On peut montrer que

$$\frac{\lambda_o}{\lambda_e} = \frac{a(t_e)}{a(t_o)}, \quad (1.6)$$

où t_e et t_o sont les temps d'émission et d'observation du photon, λ_e et λ_o sa longueur d'onde lors de l'émission et de l'observation. Le redshift est donc relié au facteur d'échelle via la relation

$$1 + z = \frac{1}{a(t)}. \quad (1.7)$$

Le redshift (cosmologique) est donc directement dû à l'expansion de l'univers. Il peut servir de mesure de temps (et aussi de distance, voir section 1.2.2 paragraphe *Les distances*) : le spectre d'un objet avec un redshift $z = 2$ est décalé vers le rouge d'un facteur 3. Il en découle que sa lumière observée aujourd'hui a été émise lorsque l'univers avait une taille 3 fois plus petite qu'aujourd'hui, soit il y a environ 12 milliards d'années.

Les équations d'Einstein

Lorsque Einstein publie sa théorie en 1915, la façon d'introduire les équations d'Einstein, le cœur de la théorie, est différente de la façon de les introduire aujourd'hui. Nous nous proposons ici de suivre l'approche de la physique moderne, qui formule toutes les théories en termes d'un seul et même principe : le *principe de moindre action*. Ce principe stipule que l'action mis en œuvre lors de l'évolution d'un système entre deux instants est toujours extrémale¹. L'action est une quantité caractérisant globalement un système, elle est définie comme

$$\mathcal{S} = \int_{t_0}^{t_1} L dt, \quad (1.8)$$

où L est le lagrangien du système. En mécanique newtonienne, ce dernier est défini comme la différence de l'énergie cinétique et de l'énergie potentiel. En relativité générale, tout comme dans les théories de champs², le terme du lagrangien est représenté plutôt par une densité de lagrangien. Cette densité de lagrangien est alors intégrée sur l'espace-temps afin d'obtenir l'action. Dans le cas de la relativité

plus aigu lorsque l'ambulance s'approche, puis plus grave lorsqu'elle s'éloigne. Le même processus est à l'œuvre avec la lumière.

1. Elle est minimale dans la grande majorité des cas.
2. En particulier la théorie quantique des champs.

de corrélation obtenue à l'aide de ces traceurs n'est pas la fonction de corrélation de la matière. Dans cette section, nous décrivons les différents traceurs utilisés dans l'analyse présentée dans ce manuscrit, et comment, à l'aide de leur fonction de corrélation, déduire la fonction de corrélation de la matière.

1.5.1 Traceurs et biais

Le moyen le plus évident auquel nous pouvons penser pour sonder la matière est l'observation de la matière baryonique via la lumière qu'elle émet. Les étoiles en sont un bon exemple. La lumière qu'elles émettent nous permet de les localiser dans notre galaxies, à la différence des planètes qui sont quasiment invisibles. C'est notamment grâce à la lumière émise par les milliards d'étoiles présentes dans les galaxies que nous pouvons observer ces dernières. Mais ceci présente un défaut : la distribution que tracent les étoiles ou les galaxies n'est pas la distribution de matière sous-jacente.

Prenons un exemple. Supposons que nous voulons reconstruire la distribution de la matière dans une région de l'espace. Pour ce faire, nous observons toutes les galaxies, que nous supposons identiques, dans cette région. Les galaxies, comme tous les objets effondrés, se forment dans les endroits les plus denses de l'univers. **En dessous d'un certain seuil de densité, les galaxies ne peuvent se former**¹. Ainsi, sonder la distribution de matière via l'observation des galaxies nous fait manquer toute la distribution de matière pour laquelle ρ est inférieur à ce seuil. La distribution reconstruite sera donc plus structurée que la distribution sous-jacente. ~~les surdensités seront mieux échantillonnées que les sous-densités.~~ Ce phénomène est décrit par ce qu'on nomme le *biais*. Le biais $b_i(z)$ du traceur i au redshift z est défini par

$$\delta_i(\vec{r}, z) = b_i(z) \delta_{\text{matière}}(\vec{r}, z), \quad (1.57)$$

où δ_i est le contraste de densité du traceur i , et $\delta_{\text{matière}}$ le contraste de densité de la distribution de matière sous-jacente. Cette expression est valable uniquement aux grandes échelles, pour $r \gtrsim 15 h^{-1} \text{ Mpc}$. Le biais relie donc les fluctuations de densité du traceur i , aux fluctuations de densité de la matière. Pour les objets compacts, telles les galaxies, le biais est supérieur à 1. Plus le traceur se forme dans des régions denses, et plus le biais est grand. Grâce à la relation précédente, nous pouvons relier la fonction de corrélation du traceur i à celle de la matière par

$$\xi_i(r, z) = b_i^2(z) \xi_{\text{matière}}(r, z). \quad (1.58)$$

Ainsi, par l'intermédiaire du traceur i , nous pouvons mesurer la fonction de corrélation de la matière, à un facteur multiplicatif près. De plus, en comparant la fonction de corrélation du traceur i à celle de la matière prédite par les modèles, nous pouvons mesurer le biais b_i de ce traceur. Comme indiqué dans l'équation précédente, la fonction de corrélation associée au traceur i est amplifiée par un facteur b_i^2 . **Il est donc avantageux de choisir un traceur avec un biais important, afin d'obtenir une fonction de corrélation avec une amplitude importante, et ainsi un rapport signal sur bruit plus grand. La première détection du pic BAO dans la distribution de matière**

1. Pour être exact, la probabilité qu'une galaxie se forme dans une région où la densité moyenne est inférieure à ce seuil est extrêmement faible.

a ainsi été faite avec l'observation des LRG (des galaxies rouges) dont la mesure du biais donne $b_{LRG} \sim 2$ (D. J. Eisenstein et al. 2005).

1.5.2 Distorsions dans l'espace des redshifts idéalement, mais en pratique non

Afin de construire la fonction de corrélation du traceur i , il est **nécessaire** de connaître le redshift cosmologique de chaque traceur. Dans le cas des traceurs booléens, comme par exemple les galaxies, le redshift est obtenu en mesurant le spectre de l'objet, puis en comparant les longueurs d'onde des raies d'émission présentes dans le spectre aux longueurs d'onde mesurées en laboratoire. Le redshift mesuré est donc le suivant :

$$z_{\text{mesure}} = z_{\text{cosmo}} + \Delta z_v + \delta z_{\text{sys}} + \delta z_{\text{stat}} , \quad (1.59)$$

où z_{mesure} est le redshift mesuré, δz_{sys} et δz_{stat} sont les erreurs statistiques et systématiques sur la mesure du redshift, z_{cosmo} est le redshift cosmologique, et Δz_v le redshift induit par la vitesse particulière v_{\parallel} du traceur le long de la ligne de visée (effet Doppler) :

$$\Delta z_v = \frac{v_{\parallel}}{c} . \quad \text{pas de facteur } a \text{ ou } H ? \quad (1.60)$$

Comme nous l'avons vu dans la section 1.2.2, l'effet Doppler induit par les vitesses particulières des traceurs s'ajoute au redshift cosmologique. Ces deux effets sont indiscernables, et nous n'avons donc pas accès au redshift cosmologique des traceurs observés. Les vitesses particulières des traceurs étant provoquées par la gravité, celles-ci sont corrélées avec le champ de matière sous-jacent : les traceurs ont tendance à se déplacer vers les surdensités par effet gravitationnel. La figure 1.14 illustre la situation : au centre se trouve une surdensité, représentée par un amas de galaxies. Quatre galaxies se trouvent autour de cette surdensité, leur vitesse particulière est représentée par une flèche noire, qui est dirigée vers le centre à cause du puits de potentiel créé par la surdensité. La galaxie se trouvant derrière cette surdensité se déplace donc vers l'observateur, son redshift mesuré est ainsi plus petit que son redshift cosmologique. Similairement, la galaxie se trouvant devant la surdensité est reconstruite avec un redshift plus grand. Les objets se déplaçant perpendiculairement à la ligne de visée ne sont pas affectés. La ligne en pointillés bleus indique la distribution reconstruite dans l'espace des redshifts : cette distribution n'est plus circulaire, elle est aplatie selon la direction de la ligne de visée. Cet effet est appelé *RSD* (Redshift Space Distortions), ou *distorsions dans l'espace des redshifts*.

Ces distorsions résultent d'un effet gravitationnel : les traceurs acquièrent leur vitesse en tombant dans les puits de potentiel créés par les surdensités. L'effet peut donc être modélisé. La formule de Kaiser (KAISER 1987) relie le contraste de densité $\delta^s(\vec{k})$ dans l'espace des redshifts au contraste de densité $\delta(\vec{k})$ dans l'espace réel

$$\delta^s(\vec{k}, z) = (1 + f(z)\mu_k^2)\delta(\vec{k}, z) , \quad (1.61)$$

où $f = \frac{d \ln G}{d \ln a}$ est le taux de croissance des structures, et $\mu_k = \frac{\vec{k} \cdot \vec{u}}{k}$, où \vec{u} est la direction de la ligne de visée. Le vecteur \vec{k} peut être décomposé comme $\vec{k} = k_{\parallel} \vec{u} + \vec{k}_{\perp}$, où \vec{k}_{\perp} est perpendiculaire à \vec{u} . La

A

Fluctuating Gunn Peterson Approximation

rappeler que l'on néglige le thermal broadening

L'approximation FGPA (Fluctuating Gunn Peterson Approximation), présentée dans la section 4.2.1, relie le contraste de densité d'absorption $\text{Ly}\alpha$ δ_F à la profondeur optique τ :

$$\tau = a \exp(b\delta_F), \quad (\text{A.1})$$

où a et b sont deux paramètres.

Afin d'obtenir cette relation, nous débutons avec l'équation qui donne l'équilibre de photo-ionisation de l'hydrogène neutre :

$$\gamma + \text{H}_I \leftrightarrow e^- + p^+, \quad (\text{A.2})$$

qui se traduit comme

$$n_\gamma n_{\text{H}_I} \langle \sigma_{\text{ioni}} c \rangle = n_e n_p \langle \sigma_{\text{rec}} v \rangle_T, \quad (\text{A.3})$$

où n_γ , n_{H_I} , n_e et n_p donnent respectivement les densités de photons, d'hydrogènes neutres, d'électrons et de protons. σ_{ioni} et σ_{rec} donnent respectivement les sections efficaces de l'ionisation d'un atome d'hydrogène neutre et de la recombinaison d'un électron et d'un proton. En sachant que les densités d'électron et de protons sont égales à la densité de baryons, nous pouvons exprimer la densité d'hydrogène neutre comme

$$n_{\text{H}_I} = n_b^2 \frac{\langle \sigma_{\text{rec}} v \rangle_T}{n_\gamma \langle \sigma_{\text{ioni}} c \rangle}. \quad (\text{A.4})$$

De plus, le terme $\langle \sigma_{\text{rec}} v \rangle_T$ est proportionnel à $T^{-0.7}$, nous avons donc la relation de proportionnalité

$$n_{\text{H}_I} \propto \frac{(1+z)^6 \Omega_b^2 (1+\delta)^2 T^{-0.7}}{n_\gamma \langle \sigma_{\text{ioni}} c \rangle}, \quad (\text{A.5})$$

où nous avons explicité la densité de baryon $n_b^2 = (1+z)^6 \Omega_b^2 (1+\delta)^2$ en fonction du contraste de densité δ . Nous pouvons relier de la même manière la température T au contraste de densité δ comme

$$T = \bar{T}(z)(1+\delta)^{\gamma(z)-1}, \quad (\text{A.6})$$

ref ? Hui and Gnedin ? ask Jim

où $\gamma(z \sim 3) \sim 1,6$. Nous pouvons ainsi relier la profondeur optique τ au contraste de densité comme

$$\tau(z) \propto \Omega_b^2 \frac{(1+z)^6 \overline{T}(z)^{-0,7}}{H(z) J_\gamma} \frac{(1+\delta)^\beta}{1+\eta}, \quad (\text{A.7})$$

où η est le gradient de vitesse $\eta = v'_p(z)/H(z)$, $H(z)J_\gamma$ correspond au flux ionisant de photons, et $\beta = 2 - 0,7(\gamma(z) - 1) \sim 1,6$ à $z \sim 3$. De l'équation précédente, nous déduisons que, pour des faibles gradients de vitesse et pour $z \sim 3$

$$\tau(z) \propto \exp(1,6 \times \delta)(1 - \eta). \quad (\text{A.8})$$

Cette équation correspond à l'approximation FGPA (Fluctuating Gunn Peterson Approximation).

si on ne s'intéresse qu'à δ : $\tau \propto (1 + \delta)^\beta$
approx lognormale [Cole and Jones] $\tau \propto \exp(\beta \delta)$
ce qui est généralement appelé FGPA

Si on prend η en compte $\tau \propto 1 + \beta \delta - \eta$
puis $\tau \propto \exp(\beta \delta - \eta)$
on rajoute juste un paramètre devant η et on laisse β libre

Titre : Les forêts Lyman- α du relevé eBOSS : comprendre les fonctions de corrélation et les systématiques

Mots clés : cosmologie, relevé à grand redshift, énergie noire, forêts Lyman- α , simulations, champs aléatoires gaussiens

Résumé : Cette thèse s'inscrit dans le cadre des projets eBOSS et DESI. Ces projets utilisent, entre autres, l'absorption Lyman- α ($\text{Ly}\alpha$) afin de sonder la répartition de matière dans l'univers et ainsi mesurer l'échelle des oscillations acoustiques de baryon (BAO). La mesure du rapport de la taille de l'échelle BAO et de la taille de l'horizon acoustique lors du découplage des baryons et des photons permet de contraire l'expansion de l'univers, et donc le modèle Λ CDM, modèle standard de la cosmologie.

les paramètres de l'équation d'état de l'énergie noire

Cette thèse présente le développement de simulations utilisées afin de vérifier les analyses BAO des groupes $\text{Ly}\alpha$ de eBOSS et DESI. Ces simulations ont recours aux champs aléatoires gaussiens (GRF). Les GRF permettent de générer un champ de densité δ . A partir de ce champ de densité, les positions des quasars (QSO) sont tirées, puis à partir de chaque quasar, les lignes de visées sont constituées. Le champ de densité δ est interpolé le long de ces lignes de visées. Enfin, à l'aide de l'approximation FGPA (Fluctuating Gunn Peterson Approximation), la densité interpolée est transformée en profondeur optique τ , puis en absorption $\text{Ly}\alpha$. Grâce à un programme développé par la communauté de DESI, un continuum est ajouté aux forêts $\text{Ly}\alpha$ afin de créer des spectres de quasars synthétiques. Les simulations présentées dans ce manuscrit fournissent donc un relevé de quasars dont les forêts $\text{Ly}\alpha$ présentes dans les spectres possèdent les bonnes

fonctions d'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$, de corrélation croisée $\text{Ly}\alpha \times \text{QSO}$, ainsi que d'auto-corrélation $\text{QSO} \times \text{QSO}$ et $\text{HCD} \times \text{HCD}$ (High Column Density systems).

L'étude de ces simulations permet de montrer que l'analyse BAO menée sur l'ensemble des données $\text{Ly}\alpha$ du relevé eBOSS produit une mesure non biaisée des paramètres BAO α_{\parallel} et α_{\perp} . Par ailleurs, une étude approfondie du modèle utilisé pour ajuster les fonctions de corrélation montre que la forme de la fonction d'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$, c'est à dire les mesures du biais $b_{\text{Ly}\alpha}$ et du paramètre RSD (Redshift Space Distorsions) $\beta_{\text{Ly}\alpha}$, est comprise à environ 20 % près. Les systématiques qui affectent les mesures des paramètres $\text{Ly}\alpha$ ($b_{\text{Ly}\alpha}$ et $\beta_{\text{Ly}\alpha}$) sont issues de deux effets. Le premier effet provient de la matrice de distorsion qui ne capture pas l'intégralité des distorsions produites par l'ajustement du continuum des quasars. Le second effet est lié à la modélisation des HCD. La modélisation de ces absorbeurs denses n'est pas parfaite et affecte la mesure des paramètres $\text{Ly}\alpha$, en particulier le paramètre RSD $\beta_{\text{Ly}\alpha}$.

L'analyse de ces simulations permet donc de valider un bon contrôle des systématiques pour les analyses BAO avec le $\text{Ly}\alpha$. Cependant, une meilleure compréhension des mesures des paramètres $\text{Ly}\alpha$ est nécessaire afin d'envisager une analyse RSD à l'aide de la combinaison de l'auto-corrélation $\text{Ly}\alpha \times \text{Ly}\alpha$ et de la corrélation croisée $\text{Ly}\alpha \times \text{QSO}$.