

1

Présentation des données

Dans le chapitre précédent, nous présentions l'instrument SDSS et sa prise de données, qui constitue le relevé eBOSS. Nous présentons ici, comment des photons acquis avec les CCD, les spectres sont déduits, puis comment du spectre de chaque quasar, nous construisons le champ d'absorption.

1 Réduction des données

L'observation de chaque plaque fournit des données en deux dimensions : la première correspond aux différentes longueurs d'onde, la seconde aux différentes fibres. La chaîne de réduction des données transforme ces informations en une liste de spectres. Premièrement, chaque spectre est extrait de l'image acquise par le CCD. Le spectre est étalonné en longueur d'onde à l'aide de lampes à arc. Le flux est calibré en utilisant les spectres des fibres dédiées aux étoiles standards. Puis, le fond du ciel est estimé dans chaque demi plaque grâce aux fibres dédiées, et soustrait à chaque spectre. La variance dans chaque pixel est ensuite estimée. Elle prend en compte le bruit de photon, le bruit de l'électronique, et aussi les éventuels défauts que présentent les CCD. Les pixels affectés par des rayons cosmiques sont rejetés. Enfin, pour chaque objet, toutes les expositions sont ajoutées pour former un seul et même spectre, avec un plus grand rapport signal sur bruit. Le nombre typique d'expositions par objet varie entre 4 et 6.

Les données que nous utilisons dans ce manuscript sont rendus publiques lors de la sixième publication de données SDSS (DR16¹). Elles sont décrites dans [1].

2 Le catalogue de quasar

Une fois les spectres extraits, il est important de les classer afin de construire des catalogues utilisables par les différentes analyses. Pour ce faire, chaque spectre est traité par la chaîne d'identification de SDSS, décrite dans [2]. A chaque spectre est ajusté plusieurs modèles d'étoiles, de galaxies et de quasars. Les modèles sont construits en utilisant les *PCA templates* : des spectres typiques qui sont combinés afin d'ajuster au mieux le spectre à classer. Une fois tous les modèles ajustés, ils sont triés selon leur χ^2 réduit. Le spectre est alors classifié selon le modèle possédant le plus faible χ^2 réduit. Tous les spectres mesurés par SDSS et classifiés comme quasar constituent l'échantillon *superset*. Il contient 1 440 627 objets.

Les quasars, à cause des absorptions intenses comme les DLA ou les BAL², peuvent être difficiles à classer. Lors de BOSS, tous les quasars ont été inspectés visuellement ([3]) afin de confirmer la classification, et d'estimer leur redshifts. Cet échantillon représente 297 301 objets. Cependant, à cause du grand nombre de quasars observés par eBOSS (voir figure 1.1), l'observation visuelle n'a pas pu être effectuée pour tous ces objets. Une seconde classification, fondée sur la précédente, est alors appliquée [4]. Elle est complétée par l'algorithme **QuasarNET** ([4]) afin de réduire le nombre d'inspection visuelle à 0,6 % des spectres, soit 8581 spectres. Une sous-échantillon du *superset* est alors construit. Il contient 750 426 objets, confirmés comme quasar par la chaîne de traitement précédente.

Pour chaque objet, le catalogue fournit plusieurs estimations de redshift. La chaîne de classification SDSS produit une première estimation. L'algorithme **QuasarNET** en fournit une seconde. Les spectres inspectés visuellement possèdent une autre estimation. Enfin, l'algorithme **redvsblue**³ produit plusieurs estimations de redshift, parmi lesquelles figure **Z_PCA** et **Z_LYAWG**.

Différents algorithmes sont alors appliqués au catalogue, afin d'identifier les DLA et les BAL présents. L'identification, qui prend place sur les spectres, utilise les différentes expositions additionnées

1. <https://www.sdss.org/dr16/>

2. Les BAL (Broad Absorption Line) sont des quasars qui présentent des absorptions intenses au voisinage de leurs raies d'émission. Cette absorption est interprétée comme étant due à un absorbeur dense situé juste devant le quasar.

3. <https://github.com/londumas/redvsblue>

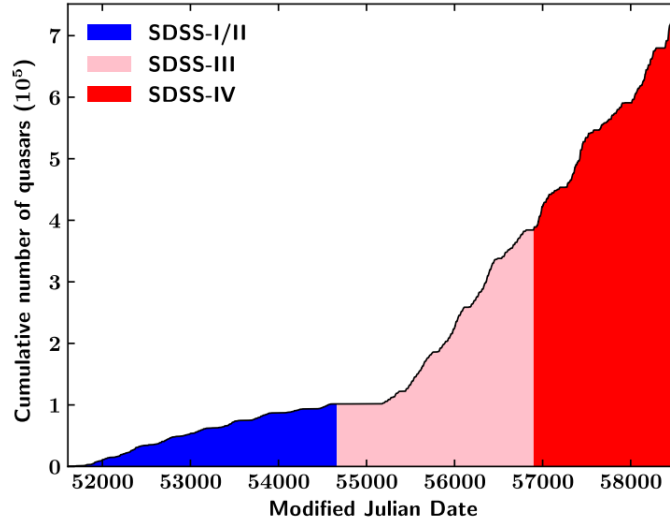


FIGURE 1.1 – Evolution du nombre de quasars observés par les différentes générations de SDSS en fonction du temps. Les générations SDSS I et II correspondent à $MJD < 54663$ (bleu), SDSS III à $54663 \leq MJD < 56898$ (rose) et SDSS IV à $56898 < MJD < 58543$ (rouge). Crédits : #prov Lyke in prep

pour chaque objet. Concernant les DLA, l’algorithme de détection est décrit dans [5]. Il est appliqué sur les quasars à un redshift $2 \leq Z_{PCA} \leq 6$, afin d’avoir suffisamment de pixel dans la zone $900 < \lambda_{RF} < 1346 \text{ \AA}$. Parmi les 270 315 spectres inspectés, 39 514 DLA ont été identifiés, distribués dans 35 686 spectres. Concernant les BAL, l’algorithme utilisé est très similaire à celui décrit dans [6]. Les BAL sont recherchés dans les spectres ayant un redshift entre 1,57 et 5,6. L’algorithme fournit la probabilité qu’un spectre possède un BAL. Le champ `BAL_PROB` du catalogue indique cette probabilité.

A la fin, le catalogue ainsi construit (DR16Q dans la suite de ce manuscrit) contient 750 426 quasars confirmés. Nous référons le lecteur à l’article [?] pour davantage d’informations.

L’analyse $\text{Ly}\alpha$ finale d’eBOSS ([?]) utilise le catalogue DR16Q. Le redshift des quasars est choisi comme étant `Z_LYAWG`. Les quasars sont sélectionnés avec un redshift $1.77 < z \leq 4$. L’échantillon correspondant représente alors 341 468 objets.

3 La sélection des forêts

L’analyse $\text{Ly}\alpha$ finale d’eBOSS, que nous utilisons dans ce manuscrit, utilise l’absorption $\text{Ly}\alpha$ présente dans les spectres de quasars. Les spectres produits par la chaîne de réduction des données SDSS sont rebinnés : 3 pixels du spectre original, d’une taille $\Delta \log_{10}(\lambda) \sim 10^{-4}$, sont combinés en 1 seul pixel d’analyse, d’une taille $\Delta \log_{10}(\lambda) \sim 3 \times 10^{-4}$. Ceci est fait afin de simplifier la mesure de l’absorption causée par le $\text{Ly}\alpha$. Dans la suite, l’utilisation de “pixel” réfère à ces pixels d’analyse.

L’absorption $\text{Ly}\alpha$ est mesurée dans deux régions distinctes du spectre. La première, dénommée région $\text{Ly}\alpha$, correspond aux longueurs d’onde $1040 \leq \lambda_{RF} \leq 1200 \text{ \AA}$, c’est à dire entre les raies d’émission $\text{Ly}\beta$ et $\text{Ly}\alpha$. La seconde, dénommée région $\text{Ly}\beta$, correspond aux longueurs d’onde $920 \leq \lambda_{RF} \leq 1020 \text{ \AA}$, c’est à dire entre la limite de la série de Lyman et la raie d’émission $\text{Ly}\beta$. Les pixels d’absorption $\text{Ly}\alpha$ dans la région $\text{Ly}\alpha$ sont dénommés pixels $\text{Ly}\alpha(\text{Ly}\alpha)$, et ceux dans la région $\text{Ly}\beta$ sont dénommés pixels $\text{Ly}\alpha(\text{Ly}\beta)$. De plus, l’analyse se limite aux pixels dont la longueur d’onde observée est comprise entre $3000 \leq \lambda_{obs} \leq 6000 \text{ \AA}$, à cause notamment des absorptions atmosphériques intenses dans l’UV, et des raies d’émissions du ciel dans le proche infrarouge. Ces limites sur λ_{RF} correspondent à un redshift

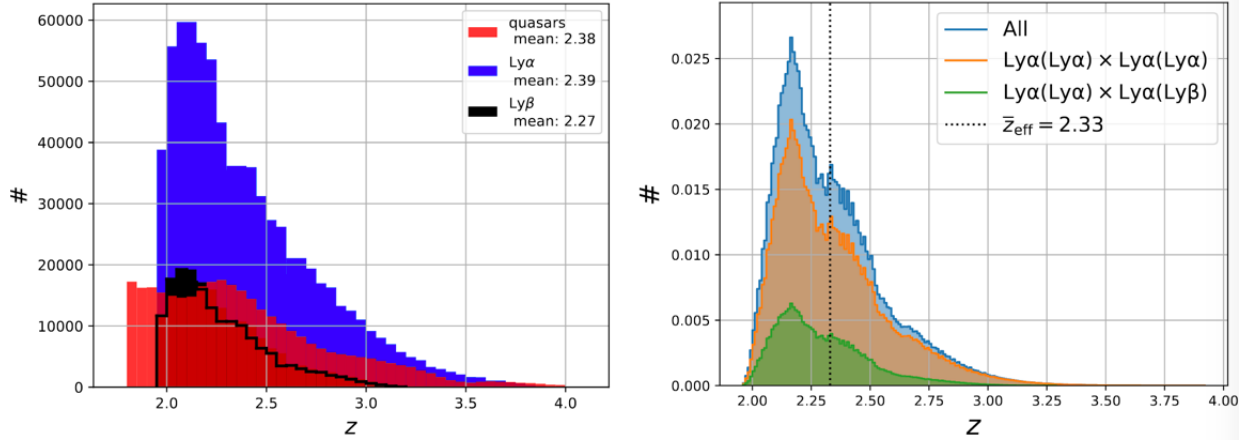


FIGURE 1.2 – Gauche : distribution en redshift des pixels $\text{Ly}\alpha(\text{Ly}\alpha)$ (bleu) et des pixels $\text{Ly}\alpha(\text{Ly}\beta)$ (noir) (nombres divisés par 50), ainsi que celle des quasars traceurs (rouge). Droite : $\#$ prov

minimal $z_{QSO} = 2$ pour les quasars $\text{Ly}\alpha$, et $z_{QSO} = 2.53$ pour les quasars $\text{Ly}\beta$. Parmi les 341 468 quasars utilisés comme traceur (dénommés quasars traceurs), 256 328 sont utilisés pour leurs absorptions $\text{Ly}\alpha$ dans la région $\text{Ly}\alpha$ (dénommés quasar $\text{Ly}\alpha$), et 103 080 sont utilisés pour leurs absorptions $\text{Ly}\beta$ dans la région $\text{Ly}\beta$ (dénommés quasar $\text{Ly}\beta$).

D'autres sélections sont aussi appliquées. Les quasars pour lesquels la probabilité d'avoir un BAL est supérieure à 0.9 sont écartés. Les mauvaises observations durant BOSS ou eBOSS sont mises de côté. De plus, chaque région nécessite au moins 50 pixels. Cette sélection décale le redshift minimal à $z_{QSO} = 2.10$ pour les quasars $\text{Ly}\alpha$, et à $z = 2.65$ pour les quasars $\text{Ly}\beta$. Enfin, l'ajustement du continu (voir section suivante) échoue pour certains spectres. Ces spectres sont aussi écartés. Après toutes ces sélections, l'échantillon final contient 210 005 quasars $\text{Ly}\alpha$ et 69 656 quasars $\text{Ly}\beta$. La distribution en redshift des pixels et des quasars traceurs est présentée dans le graphique de gauche de la figure 1.2.

Bibliographie

- [1] Romina Ahumada, Carlos Allende Prieto, Andres Almeida, Friedrich Anders, Scott F. Anderson, Brett H. Andrews, Borja Anguiano, Riccardo Arcodia, Eric Armengaud, Marie Aubert, Santiago Avila, Vladimir Avila-Reese, Carles Badenes, Christophe Balland, Kat Barger, Jorge K. Barrera-Ballesteros, Sarbani Basu, Julian Bautista, Rachael L. Beaton, Timothy C. Beers, B. Izamar T. Benavides, Chad F. Bender, Mariangela Bernardi, Matthew Bershad, Florian Beutler, Christian Moni Bidin, Jonathan Bird, Dmitry Bizyaev, Guillermo A. Blanc, Michael R. Blanton, Mederic Boquien, Jura Borissova, Jo Bovy, W. N. Brandt, Jonathan Brinkmann, Joel R. Brownstein, Kevin Bundy, Martin Bureau, Adam Burgasser, Etienne Burtin, Mariana Cano-Diaz, Raffaella Capasso, Michele Cappellari, Ricardo Carrera, Solene Chabanier, William Chaplin, Michael Chapman, Brian Cherinka, Cristina Chiappini, Peter Doohyun Choi, S. Drew Chojnowski, Haeun Chung, Nicolas Clerc, Damien Coffey, Julia M. Comerford, Johan Comparat, Luiz da Costa, Marie-Claude Cousinou, Kevin Covey, Jeffrey D. Crane, Katia Cunha, Gabriele da Silva Ilha, Yu Sophia Dai, Sanna B. Damsted, Jeremy Darling, Danny Horta Darrington, James W. Davidson, Roger Davies, Kyle Dawson, Nikhil De, Axel de la Macorra, Nathan De Lee, Anna Barbara de Andrade Queiroz, Alice Deconto Machado, Sylvain de la Torre, Flavia Dell’Agli, Helion du Mas des Bourboux, Aleksandar M. Diamond-Stanic, Sean Dillon, John Donor, Niv Drory, Chris Duckworth, Tom Dwelly, Garrett Ebelke, Sarah Eftekharzadeh, Arthur Davis Eigenbrot, Yvonne P. Elsworth, Mike Eracleous, Ghazaleh Erfanianfar, Stephanie Escoffier, Xiaohui Fan, Emily Farr, Jose G. Fernandez-Trincado, Diane Feuillet, Alexis Finoguenov, Patricia Fofie, Amelia Fraser-McKelvie, Peter M. Frinchaboy, Sebastien Fromenteau, Hai Fu, Lluís Galbany, Rafael A. Garcia, D. A. Garcia-Hernandez, Luis Alberto Garma Oehmichen, Junqiang Ge, Marcio Antonio Geimba Maia, Doug Geisler, Joseph Gelfand, Julian Goddy, Jean-Marc Le Goff, Violeta Gonzalez-Perez, Kathleen Grabowski, Paul Green, Catherine J. Grier, Hong Guo, Julien Guy, Paul Harding, Sten Hasselquist, Adam James Hawken, Christian R. Hayes, Fred Hearty, S. Hekker, David W. Hogg, Jon Holtzman, Jiamin Hou, Bau-Ching Hsieh, Daniel Huber, Jason A. S. Hunt, J. Ider Chitham, Julie Imig, Mariana Jaber, Camilo Eduardo Jimenez Angel, Jennifer A. Johnson, Amy M. Jones, Henrik Jonsson, Eric Jullo, Yerim Kim, Karen Kinemuchi, Charles C. Kirkpatrick, George W. Kite, Mark Klaene, Jean-Paul Kneib, Juna A. Kollmeier, Hui Kong, Marina Kounkel, Dhanesh Krishnarao, Ivan Lacerna, Ting-Wen Lan, Richard R. Lane, David R. Law, Henry W. Leung, Hannah Lewis, Cheng Li, Jianhui Lian, Lihwai Lin, Dan Long, Penelope Longa-Pena, Britt Lundgren, Brad W. Lyke, J. Ted Mackereth, Chelsea L. MacLeod, Steven R. Majewski, Arturo Manchado, Claudia Maraston, Paul Martini, Thomas Masseron, Karen L. Masters, Savita Mathur, Richard M. McDermid, Andrea Merloni, Michael Merrifield, Szabolcs Meszaros, Andrea Miglio, Dante Minniti, Rebecca Minsley, Takamitsu Miyaji, Faizan Gohar Mohammad, Benoit Mosser, Eva-Maria Mueller, Demitri Muna, Andrea Munoz-Gutierrez, Adam D. Myers, Seshadri Nadathur, Preethi Nair, Janaina Correa do Nascimento, Rebecca Jean Nevin, Jeffrey A. Newman, David L. Nidever, Christian Nitschelm, Pasquier Noterdaeme, Julia E. O’Connell, Matthew D Olmstead, Daniel Oravetz, Audrey Oravetz, Yoisson Osorio, Zachary J. Pace, Nelson Padilla, Nathalie Palanque-Delabrouille, Pedro A. Palicio, Hsi-An Pan, Kaike Pan, James Parker, Romain Paviot, Sebastien

- Peirani, Karla Pena Ramrez, Samantha Penny, Will J. Percival, Ismael Perez-Fournon, Ignasi Perez-Rafols, Patrick Petitjean, Matthew M. Pieri, Marc Pinsonneault, Vijith Jacob Poovelil, Joshua Tyler Povich, Abhishek Prakash, Adrian M. Price-Whelan, M. Jordan Raddick, Anand Raichoor, Amy Ray, Sandro Barboza Rembold, Mehdi Rezaie, Rogemar A. Riffel, Rogerio Riffel, Hans-Walter Rix, Annie C. Robin, A. Roman-Lopes, Carlos Roman-Zuniga, Benjamin Rose, Ashley J. Ross, Graziano Rossi, Kate Rowlands, Kate H. R. Rubin, Mara Salvato, Ariel G. Sanchez, Laura Sanchez-Menguiano, Jose R. Sanchez-Gallego, Conor Sayres, Adam Schaefer, Ricardo P. Schiavon, Jaderson S. Schimoia, Edward Schlafly, David Schlegel, Donald P. Schneider, Mathias Schultheis, Axel Schwope, Hee-Jong Seo, Aldo Serenelli, Arman Shafeloo, Shoaib Jamal Shamsi, Zhengyi Shao, Shiyin Shen, Matthew Shetrone, Raphael Shirley, Victor Silva Aguirre, Joshua D. Simon, M. F. Skrutskie, Anze Slosar, Rebecca Smethurst, Jennifer Sobeck, Bernardo Cervantes Sodi, Diogo Souto, David V. Stark, Keivan G. Stassun, Matthias Steinmetz, Dennis Stello, Julianna Stermer, Thaisa Storchi-Bergmann, Alina Streblyanska, Guy S. Stringfellow, Amelia Stutz, Genaro Suarez, Jing Sun, Manuchehr Taghizadeh-Popp, Michael S. Talbot, Jamie Tayar, Aniruddha R. Thakar, Riley Theriault, Daniel Thomas, Zak C. Thomas, Jeremy Tinker, Rita Tojeiro, Hector Hernandez Toledo, Christy A. Tremonti, Nicholas W. Troup, Sarah Tuttle, Eduardo Unda-Sanzana, Marica Valentini, Jaime Vargas-Gonzalez, Mariana Vargas-Magana, Jose Antonio Vazquez-Mata, M. Vivek, David Wake, Yuting Wang, Benjamin Alan Weaver, Anne-Marie Weijmans, Vivienne Wild, John C. Wilson, Robert F. Wilson, Nathan Wolthuis, W. M. Wood-Vasey, Renbin Yan, Meng Yang, Christophe Yèche, Olga Zamora, Pauline Zarrouk, Gail Zasowski, Kai Zhang, Cheng Zhao, Gongbo Zhao, Zheng Zheng, Zheng Zheng, Guangtun Zhu, and Hu Zou. The Sixteenth Data Release of the Sloan Digital Sky Surveys : First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra. dec 2019. URL <http://arxiv.org/abs/1912.02905>.
- [2] Adam S. Bolton, David J. Schlegel, Eric Aubourg, Stephen Bailey, Vaishali Bhardwaj, Joel R. Brownstein, Scott Burles, Yan-Mei Chen, Kyle Dawson, Daniel J. Eisenstein, James E. Gunn, G. R. Knapp, Craig P. Loomis, Robert H. Lupton, Claudia Maraston, Demitri Muna, Adam D. Myers, Matthew D. Olmstead, Nikhil Padmanabhan, Isabelle Paris, Will J. Percival, Patrick Petitjean, Constance M. Rockosi, Nicholas P. Ross, Donald P. Schneider, Yiping Shu, Michael A. Strauss, Daniel Thomas, Christy A. Tremonti, David A. Wake, Benjamin A. Weaver, and W. Michael Wood-Vasey. Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey. jul 2012. doi : 10.1088/0004-6256/144/5/144. URL <http://arxiv.org/abs/1207.7326><http://dx.doi.org/10.1088/0004-6256/144/5/144>.
- [3] Isabelle Pâris, Patrick Petitjean, Nicholas P. Ross, Adam D. Myers, Éric Aubourg, Alina Streblyanska, Stephen Bailey, Éric Armengaud, Nathalie Palanque-Delabrouille, Christophe Yèche, Fred Hamann, Michael A. Strauss, Franco D. Albareti, Jo Bovy, Dmitry Bizyaev, W. Niel Brandt, Marcella Brusa, Johannes Buchner, Johan Comparat, Rupert A. C. Croft, Tom Dwelly, Xiaohui Fan, Andreu Font-Ribera, Jian Ge, Antonis Georgakakis, Patrick B. Hall, Linhua Jian, Karen Kinemuchi, Elena Malanushenko, Viktor Malanushenko, Richard G. McMahon, Marie-Luise Menzel, Andrea Merloni, Kirpal Nandra, Pasquier Noterdaeme, Daniel Oravetz, Kaike Pan, Matthew M. Pieri, Francisco Prada, Mara Salvato, David J. Schlegel, Donald P. Schneider, Audrey Simmons, Matteo Viel, David H. Weinberg, and Liu Zhu. The Sloan Digital Sky Survey Quasar Catalog : twelfth data release. aug 2016. doi : 10.1051/0004-6361/201527999. URL <https://arxiv.org/abs/1608.06483>.
- [4] Nicolas Busca and Christophe Balland. QuasarNET : Human-level spectral classification and redshifting with Deep Neural Networks. aug 2018. URL <http://arxiv.org/abs/1808.09955>.
- [5] David Parks, J. Xavier Prochaska, Shawfeng Dong, and Zheng Cai. Deep Learning of Quasar

Spectra to Discover and Characterize Damped Lya Systems. sep 2017. doi : 10.1093/mnras/sty196. URL <http://arxiv.org/abs/1709.04962><http://dx.doi.org/10.1093/mnras/sty196>.

- [6] Zhiyuan Guo and Paul Martini. Classification of Broad Absorption Line Quasars with a Convolutional Neural Network. jan 2019. doi : 10.3847/1538-4357/ab2590. URL <http://arxiv.org/abs/1901.04506><http://dx.doi.org/10.3847/1538-4357/ab2590>.