UC Irvine
UC Irvine Electronic Theses and Dissertations

Title
Baryon Acoustic Oscillations in the Lyman-α Forest of High Redshift Quasars

Permalink
https://escholarship.org/uc/item/6cx4g0d1

Author
Margala, Daniel

Publication Date
2017

License
CC BY 4.0

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA,
IRVINE

Baryon Acoustic Oscillations in the Lyman-α Forest of High Redshift Quasars

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Daniel Margala

Dissertation Committee:
Professor David Kirkby, Chair
Professor Steve Barwick
Professor Manoj Kaplinghat

2017

© 2017 Daniel Margala

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

To Mom and Dad, thank you for your love, support, and everything else. Thank you also to Bob-o, Michael, Sara, and Jessica for, um, everything else :).

Thank you to Nick, Nadia, Natalie, and Jessica, who have been around throughout. To Sunil, Josh, Ricky, and PZ, I'm proud to call you friends, also, pizza bagels! Thank you Nick, not just for 3AM homework nights and 3PM happy hours, but for being my Britney. To Lisa, who once said there's a dinosaur in us all. To Dan, Joey, Matt, and Trevor, it's a cake in the park. And Tim too. Thank you to Alexa, who kept me clothed and fed. To Sasha, Collin, Colene, and the rest of the ESS crew, thank you for adopting me. And to Mackenzie, Alysha, and Pasha, for going to the beach. Thank you Alexandra, who could teach a penguin to fly. And to Yossof, need I say more? Thank you to Jonathan and Fausto, partners in my latest adventures. Also Chloe, of course, thank you.

To Mr. Bates, Mr. Pearsall, Prof. Regan, Prof. Saltzberg, Prof. Hamber, and every other teacher I've had, thank you for your dedication and for inspiring me. Thank you Prof. Ong for giving me an opportunity to get started with a key to a lab in a basement of the Physics and Astronomy building and the root passwords to a cluster of servers.

Finally, to David, I can't really express in words my gratitude for you and your endless supply of encouragement and support. I've learned so much while working with you over the years and I've had a grand experience. Thank you.

I would also like to thank the Department of Energy for GSR funding from grant DOE DE-SC0009920.

CURRICULUM VITAE

Daniel Margala

Data Scientist **2016–2017**
Silicon Valley Data Science LLC *Mountain View, California*

EDUCATION

Doctor of Philosophy in Physics **2017**
University of California, Irvine *Irvine, California*

Bachelor of Science in Physics **2009**
University of California, Los Angeles *Los Angeles, California*

RESEARCH

Graduate Research Assistant **2010–2017**
University of California, Irvine *Irvine, California*

TEACHING

Teaching Assistant **2009–2010**
University of California, Irvine *Irvine, California*

ABSTRACT OF THE DISSERTATION

Baryon Acoustic Oscillations in the Lyman-α Forest of High Redshift Quasars

By

Daniel Margala

Doctor of Philosophy in Physics

University of California, Irvine, 2017

Professor David Kirkby, Chair

One of the main goals of the third generation of the Sloan Digital Sky Survey (SDSS) Baryon Oscillation Spectroscopic Survey (BOSS) is to measure the baryon acoustic oscillation (BAO) scale in the Lyman-α forest from observations of high-redshift ($z > 2$) quasars. The transmitted flux fraction in the forest provides a measure of the neutral hydrogen density along the line of sight that can be used to infer the clustering of the underlying dark matter distribution. Measurements of the BAO scale constrain the expansion history of the universe and can be used to infer the characteristics of dark energy. During the period 2009-14, the BOSS observed $\sim 150,000$ ($z > 2$) quasars from a total sky area of 10,400 square degrees.

I present a model for spectrophotometric calibration errors these observations of quasars and describe the correction procedure we have developed and applied to this sample. Calibration errors are primarily due to atmospheric differential refraction and guiding offsets during each exposure. The corrections potentially reduce the systematics for any studies of BOSS quasars, including the measurement of baryon acoustic oscillations using the Lyman-α forest. The model suggests that, on average, the observed quasar flux in BOSS is overestimated by $\sim 19\%$ at 3600 Å and underestimated by $\sim 24\%$ at 10,000 Å. In addition, a new method for simultaneously fitting a universal quasar continuum along with observing frame systematics is discussed. I also describe the sub-sampling method used to estimate the correlation

function of the matter density tracer field and demonstrate this method using a new multi-threaded software package designed to efficiently process pixel-pairs. The technique used to fit cosmological models to the estimated correlation function is summarized as well.

The BOSS BAO observations using the SDSS DR12 sample yield measurements of $D_H(z)/r_d = 9.1 \pm 0.3$ and $D_M(z)/r_d = 37.8 \pm 2.1$ at $z = 2.34$. These measurements represent a 1-2σ tension with the prediction from a flat ΛCDM model with Planck best-fit parameters.

Chapter 1

Introduction

In this thesis, I describe my contributions to field of cosmology as a member of the BOSS collaboration. In this chapter, I give a brief overview of the historical context and the relevant parts of the modern cosmological model of our universe. In Chapter 2, I describe the SDSS telescope, focusing on the BOSS survey. In Chapter 3, I describe the procedure for processing the calibrated quasar spectra observations from the telescope and preparing the data set used in this work. In Chapter 4, I describe the method for estimating the correlation function of the Lyman-α forest flux density field. Next, in Chapter 5, I describe the fitting method used to measure the baryon acoustic oscillation (BAO) scale from the correlation function estimate described in the previous chapter. In Chapter 6, I provide a summary of BAO scale measurements and discuss their implications for cosmological models and dark energy. Finally, in Chapter 7, I give a brief summary and outlook.

1.1 History

The universe seems to be expanding. In 1929, Edwin Hubble first discovered this by measuring the redshift of several galaxies using Cepheid variable stars. The detection of the CMB a few decades later, along with precision follow-up study, has provided a wealth of information regarding the origins of our universe. In 1998, we discovered that the rate of expansion is actually increasing using observations of supernovae in distant galaxies. More recently, the 2dF and SDSS galaxy surveys pioneered a new technique for measuring the expansion rate of our universe using tracers of the matter density field. Using the expansion rate has proved to be a powerful and effective tool for better understanding the dynamics and constituents of our universe. The Big Bang cosmological model encapsulates our modern understanding of how the universe we observe today came to be. In short, it describes how the universe formed starting from an initial hot, dense state, expanding and cooling into the present. The Standard Model of particle physics is a theory concerning the electromagnetic, weak, and strong interactions, as well as classifying all the elementary particles known. he ΛCDM (Lambda cold dark matter) is a parametrization of the Big Bang cosmological model in which the universe contains a cosmological constant, denoted by Lambda (Greek Λ), associated with dark energy, and cold dark matter (abbreviated CDM). The model assumes that general relativity is the correct theory of gravity on cosmological scales. Measuring the expansion history allows us to measure the energy content and better understand the characteristics of the constituents of the universe.

1.2 Friedmann Equation

In Einstein's theory of general relativity, the curvature of space-time is related to the matter and energy content contained with it. The space-time metric tensor, $g_{\mu\nu}$, which characterizes

an invariant line element, $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$, is related to the matter and energy content by the field equation.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \quad . \tag{1.1}$$

The Einstein tensor, $G_{\mu\nu} = R_{\mu\nu} - \frac{R}{2} g_{\mu\nu}$, is a specific divergence-free combination of the Ricci tensor and the metric, where R is the curvature scalar $R = R^\mu_\mu$ and the Ricci tensor itself is related to the Riemann curvature tensor via $R_{\mu\nu} = R^\alpha_{\mu\alpha\nu}$. $T_{\mu\nu}$ is the energy-momentum tensor. G is Newton's gravitional constant and c is the speed of light.

The Friedmann–Lemaître–Robertson–Walker (FLRW) metric is an exact solution of Einstein's field equations of general relativity; it describes a homogeneous, isotropic expanding or contracting universe. The FLRW metric starts with the assumption of homogeneity and isotropy of space. It also assumes that the spatial component of the metric can be time-dependent. The metric can be written as:

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right] \quad . \tag{1.2}$$

where k is the curvature parameter. This equation governs the conversion between radial and transverse distances. The expansion or contraction of space is represented by the evolution of the scale factor $a(t) = (1 + z)^{-1}$. The Friedmann equation,

$$H^2 = \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{kc^2}{a^2} \quad , \tag{1.3}$$

describes how the evolution of the scale factor is related to the total energy density (radiation + matter + dark energy) ρ. The Friedmann equation can be re-written in terms of present values of individual density components,

$$\Omega_x = \frac{\rho_x}{\rho_c} = \frac{8\pi G}{3H^2} \rho_x \quad , \tag{1.4}$$

and the curvature parameter,

$$\Omega_k = 1 - \sum \Omega_x, \tag{1.5}$$

as:

$$\frac{H^2}{{H_0}^2} = \Omega_{0,R} a^{-4} + \Omega_{0,M} a^{-3} + \Omega_{0,k} a^{-2} + \Omega_{0,\Lambda} \tag{1.6}$$

Spatial curvature is nearly flat from observations ($k = 0$).

The energy components considered herein are pressureless (cold) dark matter, baryons, radiation, neutrinos, and dark energy. The densities of CDM and baryons scale as a^-3 and are typically represented together by Ω_{cb}. The energy density of neutrinos with non-zero mass scales like radiation at early times and like matter at late times, with

$$\frac{\rho_{\nu+r}(a)}{\rho_{\rm crit}} = \frac{8\pi^3 k_B^4 G}{45\hbar^3 v^5 H_0} \times \left[T_{\rm CMB}(a)^4 + T_\nu(a)^4 \sum_i I(m_i c^2 / k_B T_\nu(a)) \right] \tag{1.7}$$

where both CMB temperature $T_{\rm CMB}$ and neutrino temperature scale inversely with scale factor, and the neutrino temperature is given by $T_\nu = T_{\rm CMB} \left(\frac{4}{11}\right)^{1/3} g_c$, where $g_c = (3.046/3)^{1/4}$ accounts for small amount of heating of neutrinos due to electron-positron annihilation. The integral I is given by

$$I(r) = \frac{15}{\pi^4} \int_0^\infty \frac{\sqrt{x^2 + r^2}}{e^x + 1} x^2 dx \tag{1.8}$$

and must be evaluated numerically. When we refer to the $z = 0$ matter density parameter Ω_m, we include contributions of radiation (which is small compared to the uncertainties in Ω_m) and neutrinos (which are non-relativistic at $z = 0$), so

$$\Omega_m + \Omega_{\rm de} + \Omega_k \equiv 1 \quad . \tag{1.9}$$

1.2.1 Power Spectrum

The contribution to the variance of the density field per logarithmic interval k:

$$\Delta^2(k) \equiv (2\pi)^{-3} \times 4\pi k^3 P(k) = \frac{d\sigma}{d\ln k} \tag{1.10}$$

per linear k:

$$d\sigma^2 = (2\pi)^{-3} P(k) 4\pi k^2 dk \tag{1.11}$$

The total variance is given by:

$$\sigma^2 = (2\pi)^{-3} \int_0^\infty 4\pi k^2 P(k) dk \tag{1.12}$$

That describes the variance in a shell of $dV = 4\pi k^2 dk$, rewrite this for rectangular volume $dV = dk_1 dk_2 dk_3$

$$\sigma^2 = (2\pi)^{-3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dk_1 dk_2 dk_3 P(k) \tag{1.13}$$

For a finite volume on a grid with N points per axis and separation, d, between points, we can rewrite this as:

$$\sigma^2 = \frac{8}{(2\pi)^3} \left[\int_0^{k_{max}} \int_0^{k_{max}} \int_0^{k_{max}} dk_1 dk_2 dk_3 P(k) \right. \\ \left. - \int_0^{k_{min}} \int_0^{k_{min}} \int_0^{k_{min}} dk_1 dk_2 dk_3 P(k) \right] \tag{1.14}$$

where $k_{min} = 2\pi/(Nd)$ and $k_{max} = \pi/(Nd)$. The box in k-space spans $-k_{max}$ to k_{max}.

1.2.2 Correlation Function

The spatial correlation function $\xi(r)$ is related to the Fourier space power spectrum $P(k)$ as

$$\xi(r) = \frac{1}{2\pi} \int dk k^2 P(k) \frac{\sin kr}{kr} \tag{1.15}$$

The two-point auto-correlation function is a function that describes the probability that two overdensities are separated at a given distance.

1.2.3 Cosmological Models

We build the cosmological model starting from an isotropic linear power spectrum predication $\tilde{P}(k, z_0)$ at some reference redshift z_0, then embed this prediction in redshift space (we use tildes to denote linear-theory predictions without any redshift space distortions). In the general case of plane-parallel redshift-space distortion $r \to (r, \mu)$ we have:

$$\xi_{\rm cosmo}(r, \mu, z_0) = \sum_{\ell \, \rm even} L_\ell(\mu) \xi_{\ell, \rm cosmo}(r, z_0) \tag{1.16}$$

with

$$\xi_{\ell, \rm cosmo}(r, z_0) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) P_\ell(k, z_0) dk \tag{1.17}$$

where L_ℓ is the Legendre polynomial, j_ℓ is the spherical Bessel function, and $P_\ell(k, z_0)$ are the multipoles of the redshift-distorted power spectrum $P(k, \mu_k, z_0)$ with $\mu_k \equiv \hat{z} \cdot \hat{k}$:

$$P_\ell(k, z_0) = \frac{2\ell + 1}{2} \int_{-1}^{+1} P(k, \mu_k, z_0) L_\ell(\mu_k) d\mu_k \quad . \tag{1.18}$$

Specializing to linear theory and the distant observer approximation, the infinite series of $\xi_{\rm cosmo}(r,\mu,z_0)$ is truncated at $\ell=4$, with

$$\tilde{P}_\ell(k,z_0) = b^2(z_0) C_\ell(\beta(z_0)) \tilde{P}(k,z_0) \tag{1.19}$$

and

$$C_\ell(\beta) \equiv \frac{2\ell+1}{2} \int_{-1}^{+1} \left(1+\beta\mu_k^2\right)^2 L_\ell(\mu_k) d\mu_k = \begin{cases} 1+\frac{2}{3}\beta+\frac{1}{5}\beta^2 & \ell=0 \\ \frac{4}{3}\beta+\frac{4}{7}\beta^2 & \ell=2 \\ \frac{8}{35}\beta^2 & \ell=4 \end{cases} , \tag{1.20}$$

where $b(z)$ and $\beta(z)$ are the tracer bias and redshift-space distortion paramter at redshift z, respectively. We can therefore write

$$\xi_{\rm cosmo}(r,\mu,z_0) = b^2(z_0) \sum_{\ell=0,2,4} C_\ell(\beta(z_0)) L_\ell(\mu) \tilde{\xi}_{\ell,\rm cosmo}(r,z_0) \tag{1.21}$$

in terms of the undistorted linear-theory multipoles

$$\tilde{\xi}_{\ell,\rm cosmo}(r,z_0) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) \tilde{P}(k,z_0) dk \quad . \tag{1.22}$$

1.3 Baryon Acoustic Oscillations

1.3.1 Sound Horizon

In the early universe, photons are tightly coupled to the electron-proton fluid via Compton scattering. The comoving sound horizon, the distance a sound wave can travel during this

epoch, is defined as

$$r_s(z) = \int_0^{\eta(z)} \mathrm{d}\eta \, c_s(1+z) \tag{1.23}$$

with the sound speed, c_s, given by

$$c_s(z) = \frac{c}{\sqrt{3(1+R)}} \tag{1.24}$$

where $R \equiv 3\rho_b/4\rho_\gamma$ is the baryon to photon energy density ratio and c is the speed of light. The presence of baryons reduces the sound speed from the standard value for a relativistic fluid, $c_s = c/\sqrt{3}$.

1.3.2 Baryon Drag Depth

For a more detailed review, see Hu and Sugiyama [1996].

The baryon drag epoch, z_{drag}, is defined by the redshift at which the baryon drag depth equals one, $\tau_b(\eta(z_{drag})) = 1$. I suppose that the word *drag* is referring to the notion that the baryons are *dragged* along by the photons up until this point in time. As the universe cools, the photons eventually let go of the baryons and they sit there for the rest of eternity. The baryon drag depth is defined as

$$\tau_b(\eta) = \int_\eta^{\eta_0} \mathrm{d}\eta' \, \dot{\tau}_b \tag{1.25}$$

where η is the conformal time (η_0 is ...), $\dot{\tau}_b = \dot{\tau}/R$ and τ is the usual photon optical depth

$$\tau(\eta) = \int_\eta^{\eta_0} \mathrm{d}\eta' \, n_e \sigma_T a \, c \tag{1.26}$$

where n_e is the number density of free electrons, σ_T is the Thomson scattering cross-section,

and $a = 1/(1+z)$ is the scale factor.

1.3.3 Change of Variables

The following can be used to change the variable of integration from η to z:

$$\frac{\mathrm{d}\eta}{\mathrm{d}z} = -\frac{1}{H} \tag{1.27}$$

since η and H are defined as

$$\mathrm{d}\eta \equiv \mathrm{d}t/a,\ aH \equiv \frac{\mathrm{d}a}{\mathrm{d}t} \tag{1.28}$$

and the derivative of the scale factor a with respect to z is

$$\frac{\mathrm{d}a}{\mathrm{d}z} = -a^2 \tag{1.29}$$

Now we can rewrite Equation 1.25 (using Equation 1.26 and Equation 1.27) as

$$\tau_b(z) = \frac{4}{3}\int_0^z \frac{\mathrm{d}z'}{H(z)}\, n_e \sigma_T a \rho_\gamma / \rho_b\, c \tag{1.30}$$

1.3.4 Number Density of Free Electrons

This calculation follows the abundance calculations in Chapter 3 of Dodelson [2003].

Chemical Equilibrium

Assuming the reaction $e^- + p \leftrightarrow H + \gamma$ remains in equilibrium, x_e can be obtained by solving the chemical equilibrium equation

$$\frac{x_e^2}{1 - x_e} = \frac{1}{n_H} \left(\frac{m_e k_B T}{2\pi\hbar^2} \right)^{3/2} e^{-\epsilon_0 / k_B T} \tag{1.31}$$

where $\epsilon_0 = (m_e + m_p - m_H)c^2 = \frac{m_e c^2 \alpha^2}{2}$ is the ionization energy from the ground level, $T = T_0(1+z)$ is the radiation temperature, $x_e = n_e / n_H$ is the free electron fraction, and n_H is the Hydrogen number density, given by

$$n_H = (1 - Y_P)\frac{\rho_b}{m_p c^2} = (1 - Y_P)\frac{\Omega_b \rho_{crit,0}(1+z)^3}{m_p c^2} \tag{1.32}$$

where Y_P is the primordial Helium abundance, Ω_b is the baryon fraction today, and $\rho_{crit,0}$ is the critical energy density of the universe today.

Detailed Evolution

Detailed evolution of the free electron fraction is governed by the following differential equation

$$\frac{\mathrm{d}x_e}{\mathrm{d}z} = \frac{C}{(1+z)H(z)} \left\{ (1 - x_e)\beta_H e^{-E_{2\to1}/k_B T} - x_e^2 n_b \alpha_H \right\} \tag{1.33}$$

where $E_{2\to1}$ is the energy difference between the $n = 2$ and $n = 1$ levels. The case B recombination rate for Hydrogen, α_H, is given by

$$\alpha_H = F\frac{a t^b}{1 + c t^d} \times 10^{-19} m^3 s^{-1} \tag{1.34}$$

with $a = 4.309$, $b = -0.6166$, $c = 0.6703$, $d = 0.5300$ and $t = T/10^4 K$. The factor F is 1.14, and allows this simplified 3-level atom description to the multi-level calculation. The photoionization rate, β_H, is given by

$$\beta_H = \alpha_H \left(\frac{m_e k_B T}{2\pi\hbar^2} \right)^{3/2} e^{-E_{2s}/k_B T} \tag{1.35}$$

with E_{2s} being the ionization energy of 2s level. The factor C is given by

$$C = \frac{1 + K\Lambda_{2\gamma} n_H (1 - x_e)}{1 + K(\Lambda_{2\gamma} + \beta_H) n_H (1 - x_e)} \tag{1.36}$$

where $\Lambda_{2\gamma} = 8.22458 s^{-1}$ is the 2 photon decay rate and $K = \lambda_\alpha^3/(8\pi H(z))$ accounts for the cosmological redshifting of Lyα photons.

Matter Temperature

More accurate treatment .

The matter temperature should technically be used in the calculations above but the effect is small and is not necessary for the level of precision require here [Seager et al., 1999]. The differential equation for T_M to be solved simultaneously is given by

$$\frac{\mathrm{d}T_M}{\mathrm{d}z} = \left(\frac{8\sigma_T a_R T_R^4}{3H(z)(1+z)m_e c} \right) \left(\frac{n_e}{n_e + n_H + n_{He}} \right) (T_M - T_R) + \frac{2T_M}{(1+z)} \tag{1.37}$$

where $T_R = T_0(1+z)$ is the radiation temperature (T_0 is the temperature today), $a_R = 4\sigma_B/c$ is the radiation constant, and $n_{He} = Y_p \rho_b / 4m_p$ is the Helium number density.

1.3.5 Putting It All Together

Substituting $n_e = x_e n_H$ and $\rho_\gamma/\rho_b = a^{-1}\Omega_\gamma/\Omega_b$ into Equation 1.30 we have

$$\tau_b(z) = \frac{4\Omega_\gamma \, c \, \sigma_T}{3\Omega_b} \int_0^z \mathrm{d}z' \, \frac{x_e n_H}{H(z)} \tag{1.38}$$

Plugging in for the Hydrogen number density given by Equation 1.32, we have

$$\tau_b(z) = \frac{4(1-Y_P)\Omega_\gamma \, c \, \sigma_T \rho_{crit,0}}{3 m_p c^2} \int_0^z \mathrm{d}z' \, \frac{x_e (1+z)^3}{H(z)} \tag{1.39}$$

Solving this equation for $\tau_b(z) = 1$ yields the the drag epoch redshift z_{drag} which can be used along with Equation (1.23) to compute the BAO scale.

1.4 Summary

The field of cosmology involves physical models from the smallest to largest scales in the universe. A fascinating interaction between matter and radiation in the early universe provides a standard ruler for us to observe and measure the evolution of the the space-time metric. In the next chapter, I will provide an overview of the BOSS survey and the spectral observations of quasars which allow us to better understand and constrain the physical models of our universe touched upon in this chapter.

Chapter 2

The Baryon Oscillation Spectroscopic Survey

One of the main goals of the third generation of the Sloan Digital Sky Survey [SDSS-III; Eisenstein et al., 2011] Baryon Oscillation Spectroscopic Survey [BOSS; Dawson et al., 2013] is to measure the baryon acoustic oscillation (BAO) scale in the Lyman-α forest from observations of high-redshift ($z > 2$) quasars [Busca et al., 2013, Slosar et al., 2013, Kirkby et al., 2013, Font-Ribera et al., 2014, Delubac et al., 2015]. The transmitted flux fraction in the forest provides a measure of the neutral hydrogen density [Gunn and Peterson, 1965, Rauch, 1998, Meiksin, 2009] along the line of sight that can be used to infer the clustering of the underlying dark matter distribution [Cen et al., 1994, Bi et al., 1995, Zhang et al., 1995, Hernquist et al., 1996, Miralda-Escudé et al., 1996, Bi and Davidsen, 1997, Hui et al., 1997, Theuns et al., 1998]. Measurements of the BAO scale constrain the expansion history of the universe and can be used to infer the characteristics of dark energy [Weinberg et al., 2013].

During the period 2009-14, BOSS observed 294,512 quasars from a total sky area of 10,400 square degrees. As a result, the survey contains the largest sample of spectroscopic quasar

observations to date and enables an unprecedented view into the multiple areas of quasar science, e.g. clustering of quasars [White et al., 2012], quasar luminosity function [Palanque-Delabrouille et al., 2013, Ross et al., 2013, McGreer et al., 2013], and variability properties of broad absorption lines in quasar spectra caused by high-velocity outflows [Filiz Ak et al., 2012, 2013, Hall et al., 2013, Vikas et al., 2013, Filiz Ak et al., 2014].

2.1 Survey design

The BOSS double spectrograph mounted at the Cassegrain focus of the 2.5-m SDSS telescope [Gunn et al., 2006] located at Apache Point Observatory[1] (APO) simultaneously records 1000 spectra over a 3° field of view. The light from each target is captured by an optical fiber plugged into an aluminum plate at the focal plane and transported to either one of two spectrographs for analysis [Smee et al., 2013]. The circular focal plane hole for each 2 arcsecond ($120\mu m$) diameter fiber is predrilled according to each target's sky position at a nominal observing time. Each BOSS plate includes 20 spectrophotometric standard star targets which are used for spectrophotometric calibration. The calibrations are derived by fitting stellar spectrum models to the recorded spectra for these targets. Spectrophotometric calibration errors in BOSS quasar spectra are primarily due to offsets in fiber hole positioning between quasar targets and spectrophotometric standard stars that are intentionally introduced to improve the signal to noise ratio of the Lyman-α forest region of high-redshift quasars. Additionally, there are 16 "guide" stars for each plate which are each observed with coherent fiber bundles and used for guiding the telescope during exposures; see Dawson et al. [2013] for more details.

[1] `http://www.apo.nmsu.edu/`

2.2 Improved calibration

One potential source of systematics for any study of BOSS quasars is spectrophotometric calibration errors. Calibration errors in BOSS are larger than for SDSS-I [York et al., 2000] due to a design tradeoff that improves throughput in the Lyman-α forest of quasar spectra. In this work, we describe the dominant source of these errors in BOSS spectra and our procedure for reducing them. The miscalibration of BOSS spectra, on average, accounts for a $\sim 19\%$ excess at 3600 Å and a $\sim 24\%$ decrement at 10,000 Å with a smooth transition between (see Figure 2.8).

2.3 Data Samples

We use data from the `v5_7_0` BOSS spectroscopic pipeline [Bolton et al., 2012] processing of the SDSS Data Release 12 [Alam et al., 2015][2]. There are two primary classes of targets used in this work: spectrophotometric standard stars (main sequence F stars used for calibration) and quasars; see Dawson et al. [2013] and Ross et al. [2012], respectively, for descriptions of the target selection for these samples.

The BOSS DR12 sample contains a total of 487,276 targets with focal plane offsets, distributed between 2,377 observations of 2,340 plates. Of those targets, the BOSS data processing pipeline [Bolton et al., 2012] identified 284,085 as quasars and 159,886 as stars. These stars are often referred to as "failed quasars", since they were targeted as quasars due to their photometric similarities [Ross et al., 2012]. Table 2.1 summarizes the different target samples.

In addition to the primary BOSS spectrophotometric standard stars, a small subset of observations contain a second sample of spectrophotometric standard stars which were selected

[2] `http://www.sdss.org/dr12/`

Target Sample	LAMBDA_EFF (Å)	OBJTYPE	CLASS	$N_{\rm DR12}$	$N_{\rm validation}$
Offset targets	4000	-	-	487,276	4,104
Quasars	4000	QSO	QSO	284,085	1,737
Failed quasars	4000	QSO	STAR	159,886	1,049
Spec. standards	5400	SPECTROPHOTO_STD	STAR	49,635	400
Offset standards	4000	-	STAR	1,770	486

Table 2.1: Target samples used in this work. For the DR12 samples, we include any potentially useful spectra, only requring that the fiber was plugged (i.e., that the UNPLUGGED bit of the ZWARNING bitmask is not set). For the validation samples, we additionally require that the spectra have no known problems (ZWARNING = 0) and that the spectra are from plates in the validation sample (PLATE keyword matches a plate in the validation set). The offset standards are identified via the QSO_STD bit of the ANCILLARY_TARGET2 bitmask.

and designed to have the same offsets in the focal plane as the BOSS quasar targets. This sample was collected as part of an ancillary program [Dawson et al., 2013, Alam et al., 2015]. These stars were selected and visually inspected to ensure a uniform distribution across the focal plane, in a manner similar to normal spectrophotometric standard stars targeted in BOSS. The algorithm for photometric selection applied to this sample was identical to the algorithm for primary spectrophotometric standard stars in BOSS.

Spectrophotometric standard stars offset in the focal plane are identified by bit 20 of the ANCILLARY_TARGET2 keyword in DR12 data model[3]. We will refer to these objects as "offset standards". There are 1,770 offset standard star targets confirmed as stars distributed between 161 observations. Of those observations, 79 contain at least 10 offset standards in total, and 20 contain at least 10 offset standards per spectrograph.

As part of this ancillary program, we modified the BOSS data processing pipeline to use the offset standards for spectrophotometric calibration, instead of the normal spectrophotometric standard stars. This sample provides a vital cross-check of the throughput correction model described below. In particular, we use the sample of 20 plates with at least 10 offset standards in each spectrograph for validation tests. In Table 2.1, the column labeled $N_{\rm validation}$ indicates the number of targets in the validation sample for each of the relevant target samples. We

[3] `http://www.sdss.org/dr12/algorithms/bitmasks/`

refer to the sample of self-calibrated offset standard stars as "offset standards".

The design wavelength λ_i for a target is encoded in the LAMBDA_EFF quantity of a `plateDesign` file and is propagated to a target's corresponding entry in `spPlate` and `spAll` files in the DR12 data model[4]. The design hour angle h_0 for each plate is encoded in the HA quantity of its `plateHoles` file, but is not propagated to `spPlate` or `spAll`. We calculate the design altitude a_0 using the design hour angle and the central sky position of a plate.

The distribution of point-spread function (PSF) full-width half-maxima (FWHM) for all observations in DR12 is shown in Figure 2.1. For each observation, we summarize the PSF FWHM using an unweighted mean calculated from the individual exposures of the observation. The PSF FWHM is available via the SEEING50 keyword in `spCFrame` files for each exposure and is estimated from guide star images. Similarly, we calculate the mean hour angle for an observation $h_{\rm obs}$ and the mean altitude $a_{\rm obs}$ of the individual exposures. We calculate the observing hour angle from the midpoint of the times stored in TAIBEG and TAIEND keywords in `spCFrame` files for each exposure. We display the distribution of $h_{\rm obs}$ for plates relative to their design observing hour angle h_0 in Figure 2.2.

The BOSS data processing pipeline combines individual 15-minute exposures, typically about four, into a single co-added spectrum. Ideally, the individual exposures should be corrected prior to co-addition, however, few BOSS analyses currently make use of individual exposures. Instead, we calculate a correction for co-added data using the mean PSF FWHM and observing hour angle of the exposures for an observation. There were 79 plates observed during the beginning of the survey which do not have PSF FWHM data recorded; for those observations, we use the sample median as an estimate for the PSF FWHM.

[4] `http://data.sdss3.org/datamodel/`

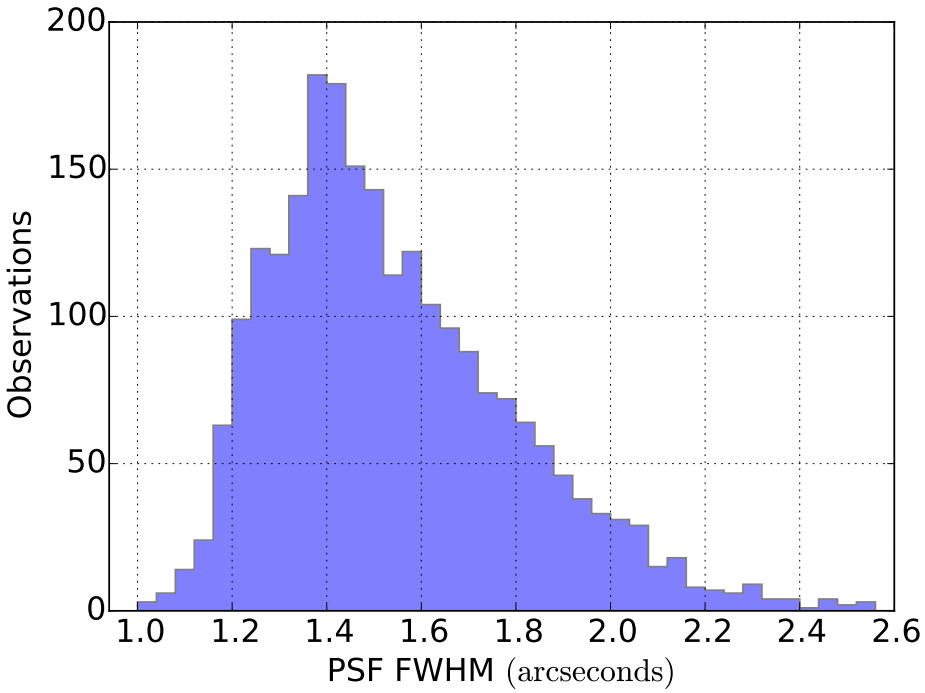

Figure 2.1: Distribution of focal plane PSF FWHM, as indicated by the SEEING50 keyword, for DR12 observations. A total of 79 observations are missing this information and omitted here. The mean (median) FWHM is 1.54 arcseconds (1.49 arcseconds).

2.4 Methods

2.4.1 Fiber illumination

The nominal transformation from sky coordinates (α, δ) to focal-plane coordinates $\mathbf{r} = (x, y)$ depends on:

- The sky coordinates (α_0, δ_0) of the focal plane origin $\mathbf{r} = 0$.
- The wavelength λ of incident light.
- The time of the observation, expressed as the local hour angle h for the right ascension α_0.

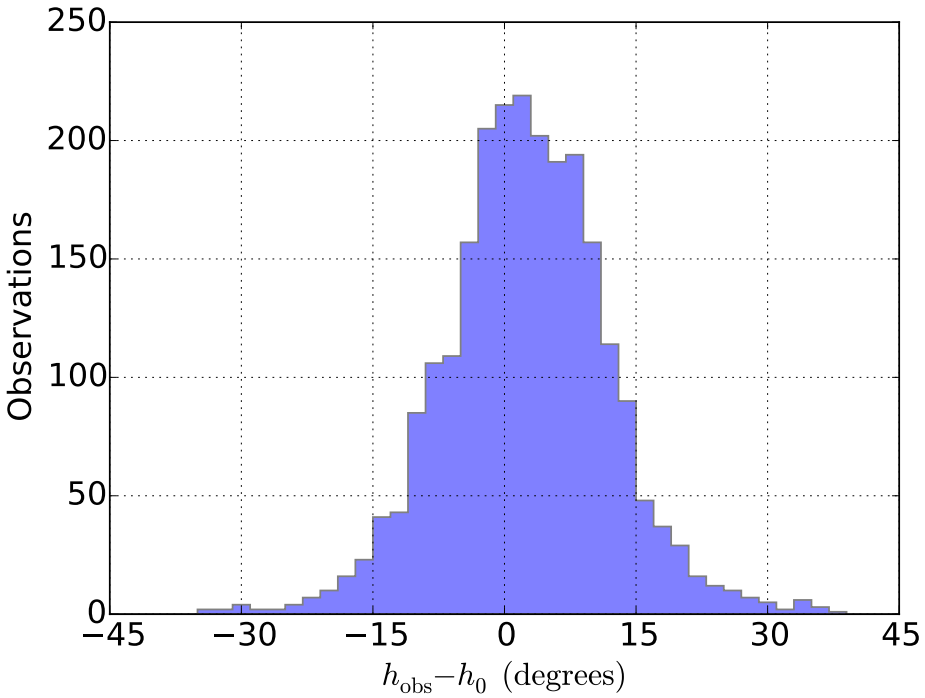

Figure 2.2: Distribution of $h_{\rm obs} - h_0$ for DR12 observations. The mean and RMS are 2.5° and 9.4°, respectively.

The last two are due to wavelength-dependent refraction through the atmosphere with a time-dependent angle of incidence equal to the telescope altitude. We assume constant nominal conditions for atmospheric temperature and pressure for all refraction calculations. The atmospheric differential refraction (ADR) angle between 4000 Å and 5400 Å is about 0.5 arcseconds (the significance of these wavelengths is explained below), which is comparable to the 2 arcsecond diameter of BOSS fibers and the typical point-spread function (PSF) full-width half-maximum (FWHM) during DR12 observations of 1.5 arcseconds. In Figure 2.3, we show differential refraction angles relative to both 4000 Å (blue) and 5400 Å (red) light at various observing altitudes. The magnitude of ADR decreases at higher observing altitudes.

Each BOSS plate is designed for a specific pointing (α_0, δ_0) at a nominal design hour angle h_0. The nominal orientation of the focal-plane coordinate system at hour angle h is fixed such that the $+\hat{y}$ direction is aligned with increasing δ and $+\hat{x}$ is pointing eastwards. Photons at wavelength λ observed at time h from a target i at sky coordinates (α_i, δ_i) then have a PSF

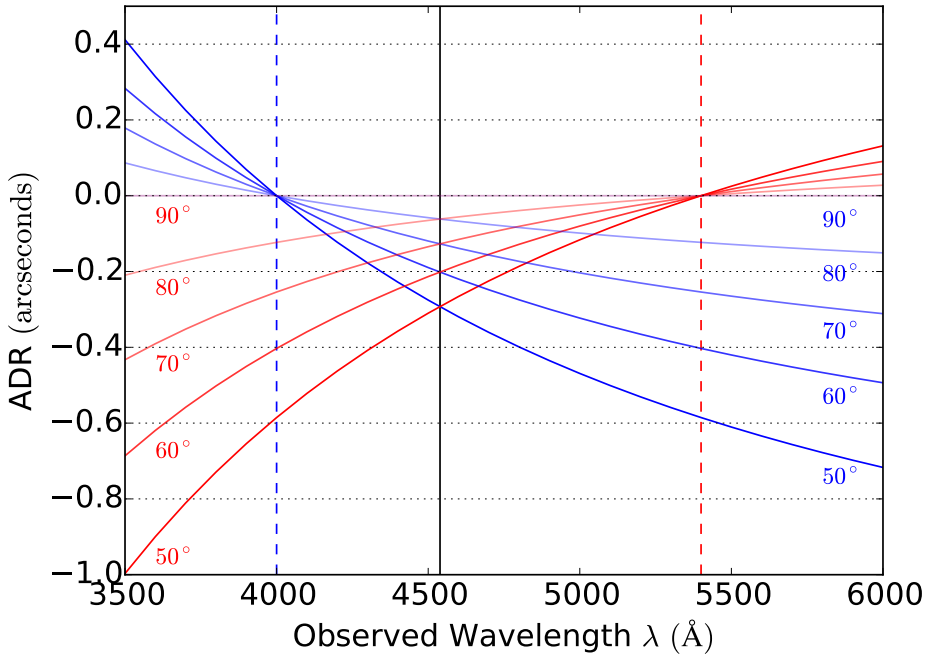

Figure 2.3: Differential refraction of light in the atmosphere. The blue (red) curves show the differential refraction of light relative to 4000 Å (5400 Å). Lines show observing altitudes of 50°, 60°, 70°, 80°, and 90°. At 90°, the differential refraction is zero in both cases. The blue and red vertical dashed lines indicate the two relevant design waveslengths, 4000 Å and 5400 Å, respectively. The black vertical solid line indicates the interesection wavelength which is essentially independent of observing altitude.

centered at the focal plane position

$$\mathbf{r}_i(\lambda, h) \equiv \mathbf{r}(\lambda, h; \alpha_i, \delta_i, \alpha_0, \delta_0) \ . \tag{2.1}$$

The fiber hole for each target i is positioned at $\mathbf{r}(\lambda_i, h_0)$ so that the the target's nominal design wavelength λ_i is centered on the fiber, and therefore has the maximum possible throughput, at time h_0. Any departures from $\lambda = \lambda_i$ or $h = h_0$ will therefore introduce a fiber centering offset

$$d_i(\lambda, \lambda_i, h) \equiv |\mathbf{r}_i(\lambda, h) - \mathbf{r}(\lambda_i, h_0)| \tag{2.2}$$

and a correspondingly reduced throughput that we calculate below.

In practice, the majority of fibers are positioned using the same central wavelength $\lambda_i = 5400$ Å. The exception are fibers assigned to high-redshift quasar targets, which use a bluer $\lambda_i = 4000$ Å in order to improve the throughput (and hence also signal-to-noise ratio) in the Lyman-α forest region ($\lambda_{\mathrm{Ly}\alpha}(1+z) = 4000$ Å at $z = 2.29$). Figure 2.4 shows an example of $\mathbf{r}_i^*(\lambda, h)$ trajectories (the asterisked quantity is explained in the following section) plotted for a single quasar target with $\lambda_i = 4000$ Å, covering a range of wavelengths λ and hour angles h. The origin of the calibration errors that we seek to fix in this work is this difference of design wavelengths. Focal-plane offsets were first introduced in the SDSS-III BOSS survey, so quasar spectra recorded during SDSS-I do not have these calibration errors.

The overall system throughput for target i includes a geometrical factor $A_i(\lambda, \lambda_i, h)$ that measures the fraction of incident light that enters its aperture. This fiber acceptance fraction depends on the PSF size and shape and its centroid offset $d_i(\lambda, h)$. In the following, we assume that all PSFs have circular Gaussian profiles described by a standard deviation σ that is independent of wavelength and constant during an exposure, so that

$$A_i(\lambda, \lambda_i, h) = A_{\mathrm{Gauss}}(\sigma, d_i(\lambda, \lambda_i, h)) \tag{2.3}$$

with

$$A_{\mathrm{Gauss}}(\sigma, d) \equiv \sigma^{-2} \int_0^{D/2} e^{-(d^2+r^2)/(2\sigma^2)} I_0(rd/\sigma^2)\, dr \ , \tag{2.4}$$

where $D = 2$ arcseconds is the BOSS fiber diameter and I_0 is a modified Bessel function of the first kind. We show later that our corrections are relatively insensitive to these assumptions since they involve fiber-acceptance ratios. Figure 2.1 shows the distribution of PSF sizes measured for DR12 observations and Figure 2.5 shows examples of acceptance fractions calculated under different assumptions.

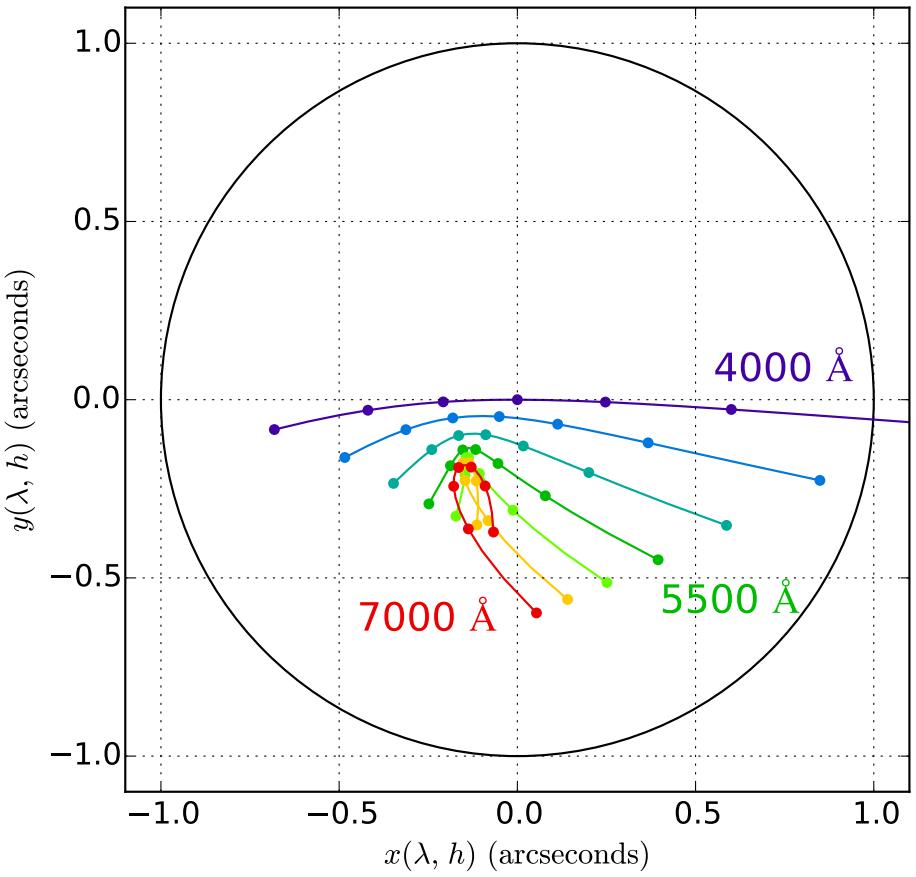

Figure 2.4: Target centroid position relative to a fixed fiber hole as a function of wavelength λ and observing hour angle h, as predicted for an ideal guiding model. The black circle represents a fiber hole (2 arcsecond diameter). The colored lines indicate the target's centroid position for specific wavelengths of incident light (4000 Å to 7000 Å in 500 Å intervals) as a function of observing hour angle h. The observing hour angle window shown here is $\pm$ 3 hours from the design hour angle h_0 of the plate. The spacing between points is 1 hour. The fiber hole is positioned such that $\lambda_i = 4000$ Å light is centered at the design hour angle h_0.

2.4.2 Telescope guiding

BOSS exposures are typically 15 minutes long, during which time an active telescope guiding loop uses 16 bright stars in the focal plane to maintain the pointing towards (α_0, δ_0). Without any additional adjustments, the variation of $d_i(\lambda, \lambda_i, h)$ during the exposure would noticeably reduce target throughputs relative to their optimum values at h_0 where $d_i(\lambda_i, \lambda_i, h_0) = 0$. Therefore the guide loop also makes small adjustments to the focal plane position $\delta\mathbf{r}$, radial scale s, and rotation θ during each exposure, which modify the transformation from sky to

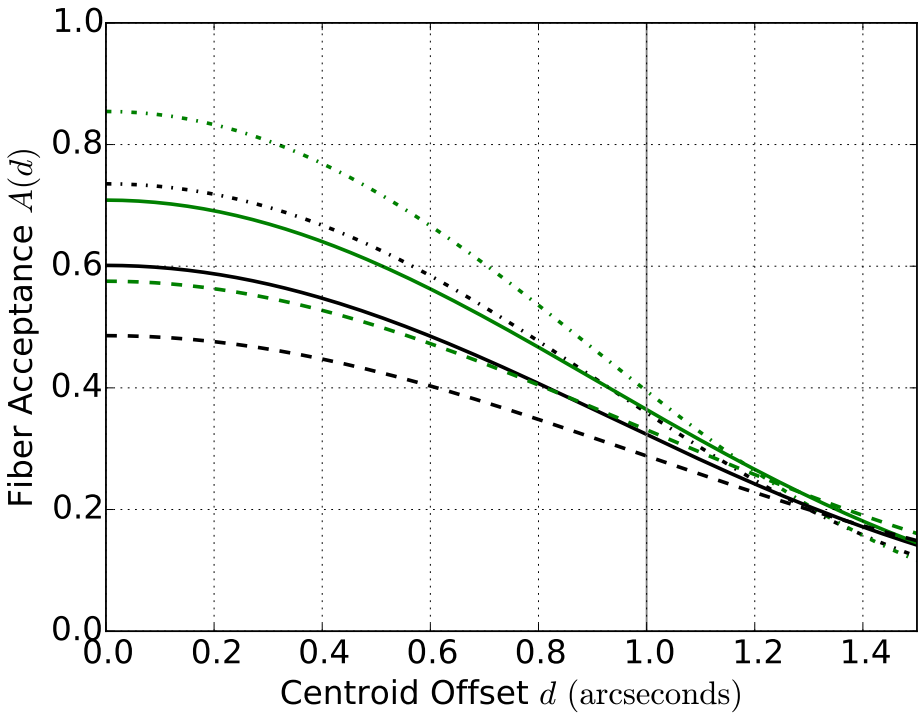

Figure 2.5: Acceptance fraction as a function of offset distance for typical PSF sizes. The dotted, solid, and dashed curves respectively correspond to acceptance fractions calculated using PSF FWHM values of 1.2, 1.5, and 1.8 arcseconds. The green and black curves respectively correspond to Gaussian and Kolmogorov PSF shapes. The solid vertical line indicates the BOSS fiber aperture radius.

focal-plane coordinates according to

$$\mathbf{r} \to G(s,\theta)\cdot\mathbf{r} + \delta\mathbf{r} \quad , \quad G(s,\theta) \equiv (1+s)\,R(\theta) \tag{2.5}$$

where R is a 2D rotation matrix acting on focal-plane coordinates. The corresponding centering offset for fiber i is then

$$d_i'(\lambda,\lambda_i,h) = |\,G(s(h),\theta(h))\cdot\mathbf{r}_i(\lambda,h) + \delta\mathbf{r}(h) - \mathbf{r}_i(\lambda_i,h_0)\,|\ , \tag{2.6}$$

where the prime denotes the effects of the guiding adjustments $\delta\mathbf{r}(h)$, $s(h)$, and $\theta(h)$.

It is not feasible to reconstruct the actual history of guiding adjustments $\delta\mathbf{r}(h)$, $s(h)$, and $\theta(h)$ during an exposure so we instead adopt an ideal guiding model, which assumes that

the plate center tracks (α_0, δ_0) optimally and that the guider makes continuous adjustments to minimize the chi-square figure of merit

$$\chi^2(s(h), \theta(h)) \equiv \sum_g d'_g(\lambda, \lambda_i, h)^2 \tag{2.7}$$

where the sum is over the 16 guide stars, indexed by g and given equal weight, and the resulting adjustments optimize throughput at $\lambda_g = 5400$ Å. We denote the trajectories resulting from ideal guiding as $\mathbf{r}_i^*(\lambda, h)$ and the corresponding centering offsets as $d_i^*(\lambda, \lambda_i, h)$. Figure 2.6 shows an example of trajectories $\mathbf{r}_i(\lambda_i, h)$ (left panel) and $\mathbf{r}_i^*(\lambda_i, h)$ (right panel) at each target's design wavelength λ_i for an entire plate. The middle panel shows trajectories when only the plate center (α_0, δ_0) is tracked and no additional adjustments are made to track the guide stars. We find that, under the ideal guiding assumption, the translation, scale, and rotation degrees of freedom completely remove the monopole and dipole contributions to centering offsets for targets with $\lambda_i = \lambda_g = 5400$ Å (shown in red), but lead to larger offsets with some residual monopole and dipole for the smaller set of quasar targets with $\lambda_i = 4000$ Å (shown in blue).

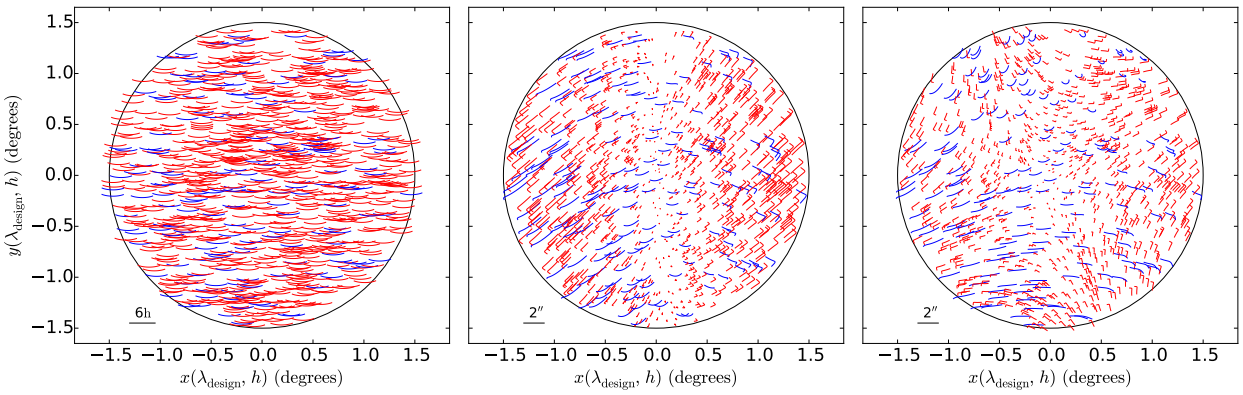

Figure 2.6: The ideal guiding model for an example plate (PLATE = 6114). The black circle indicates the plate's 3° field of view. The colored lines represent a target's λ_i light path spanning a range of ± 3 hours from the design hour angle of the plate. Red (blue) lines indicate $\lambda_i = 5400$ Å (4000 Å) targets. The fiber centroid offset distances have been exagerated relative to the plate dimensions with the scale indicated in the bottom left of each panel.

2.4.3 Spectrophotometric calibration

The mean rate of detected photons from source i at wavelength λ and hour angle h is related to the source's true spectral energy distribution (SED) $f_i(\lambda)$ by a fiber-dependent calibration factor C

$$\frac{dn_i}{dh}(\lambda, \lambda_i, h) = f_i(\lambda)\, C_i(\lambda, \lambda_i, h) \;, \tag{2.8}$$

which we split

$$C_i(\lambda, \lambda_i, h) = A_i^*(\lambda, \lambda_i, h) B_i(\lambda, h) \tag{2.9}$$

into the fiber-acceptance fraction $A_i^*(\lambda, \lambda_i, h)$ assuming ideal guiding (denoted by the asterisk), and a term B_i that includes all other contributions to the signal throughput but is independent of a target's design wavelength λ_i. The quantities that we actually measure from an exposure are the integrated photon counts

$$n_i(\lambda, \lambda_i, h_{\rm obs}) = \int_{h_{\rm obs}-\Delta h/2}^{h_{\rm obs}+\Delta h/2} \frac{dn_i}{dh}(\lambda, \lambda_i, h)\, dh \tag{2.10}$$

$$\simeq \frac{dn_i}{dh}(\lambda, \lambda_i, h_{\rm obs})\, \Delta h \;, \tag{2.11}$$

where $h_{\rm obs}$ is the exposure midpoint hour angle and $\Delta h = 3.75^\circ$ (15 minutes) is the exposure duration. The variation of the integrand is sufficiently small over Δh that we use the trapezoidal approximation to the integral in the following. Figure 2.2 shows the distribution of $h_{\rm obs} - h_0$ for all DR12 observations.

The 20 spectrophotometric calibration targets on each plate, indexed by c, are chosen so that their true SED $f_c(\lambda)$ can be directly estimated as $\tilde{f}_c(\lambda)$ using a stellar-model fit to their observed photon counts $n_c(\lambda, \lambda_c, h_{\rm obs})$. We use these SED estimates to provide estimates of

the calibration factors $\tilde{C}_c(\lambda, \lambda_c, h_{\rm obs})$ at 20 locations across each plate, and then interpolate these factors to the locations of every fiber, $\tilde{C}_i(\lambda, \lambda_i, h_{\rm obs})$. We then estimate the SED of all targets on a plate as

$$\tilde{f}_i(\lambda) = \tilde{C}_i(\lambda, \lambda_i, h_{\rm obs})^{-1}\, n_i(\lambda, \lambda_i, h_{\rm obs})/\Delta h \ . \tag{2.12}$$

A key point of this work is that *the method used by the BOSS data processing pipeline to interpolate the standard-star calibrations $\tilde{C}_c$ to other targets in the focal plane does not account for differences in each target's design wavelength λ_i.* Since the calibration stars have $\lambda_c = 5400$ Å, this only affects the quasar targets, which are designed with $\lambda_i = 4000$ Å. Under the ideal guiding assumption, the SED mis-calibration of targets with $\lambda_i \neq \lambda_c$ is the ratio of acceptance fractions at the actual (λ_i) and assumed (λ_c) design wavelengths

$$R_i^*(\lambda) \equiv \frac{A_i^*(\lambda, \lambda_c, h_{\rm obs})}{A_i^*(\lambda, \lambda_i, h_{\rm obs})} \ , \tag{2.13}$$

and our improved calibration consists of scaling the pipeline SED estimates by this factor

$$\tilde{f}_i(\lambda) \to R_i^*(\lambda)\, \tilde{f}_i(\lambda) \ . \tag{2.14}$$

To summarize, the corrections $R_i^*(\lambda)$ that we implement depend on:

- The actual fiber centering offsets $d_i^*(\lambda, \lambda_i, h_{\rm obs})$ and the assumed offsets $d_i^*(\lambda, \lambda_c, h_{\rm obs})$ at the exposure midpoint $h_{\rm obs}$ (both under ideal guiding conditions).
- The design hour angle h_0, which determines the locations of each fiber on the focal plane.
- The PSF size σ.

Figure 2.7 shows scatter plots of $d_i^*(\lambda, \lambda_i, h_{\rm obs})$ versus $d_i^*(\lambda, \lambda_c, h_{\rm obs})$ for all offset targets in

DR12 observations ($h_{\rm obs}$ and h_0), for six different wavelengths λ, with contours of $R_i^*(\lambda)$ superimposed for three different PSF sizes σ. Additionally, we show contours of $R_i^*(\lambda)$ for a nominal PSF size assuming a Kolmogorov atmospheric turbulence model [Tatarskii, 1961] which includes a wavelength dependent scaling of the nominal PSF FWHM $\propto \lambda^{-1/5}$ [Fried, 1966].

In Figure 2.8, we show the central 68% and 95% quantiles of $R_i^*(\lambda)$ for all 487,276 offset targets in DR12. In general, the corrections $R_i^*(\lambda)$ are smooth, monotonic functions that cross unity near 4500 Å(as expected, see Figure 2.3). Comparing with the subset of $R_i^*(\lambda)$ where $|h_{\rm obs} - h_0| < 1.25^\circ$ (5 minutes), approximately 10% of all observations, we see that the vertical spread at the crossover λ is caused by plates observed away from their design hour angle. On average, our model suggests that the observed quasar flux is overestimated by $\sim 19\%$ at and underestimated by $\sim 24\%$ near 3600 Å and 10,000 Å, respectively.

2.4.4 Correction Parametrization

In order to summarize the general properties of our correction model for all DR12 observations, we fit a parametrized model

$$1 + c_1 \log \lambda/\lambda_0 + c_2 \left(\log \lambda/\lambda_0\right)^2 \tag{2.15}$$

to each throughput correction prediction $R_i^*(\lambda)$. Figures 2.9-2.11 show the median best fit parameter per plate and its relationship to either $a_{\rm obs}$ or $(h_{\rm obs} - h_0)$. The points in each figure are colored by $(a_{\rm obs} - a_0)$ for the corresponding observation. The λ_0 parameter, which corresponds to the crossover point of the multiplicative correction, is strongly correlated with $(a_{\rm obs} - a_0)$, as seen in Figure 2.9. The value of c_1 sets the overall size of the correction and is strongly correlated with the altitude of the observation, as seen in Figure 2.10. Figure 2.11 shows a similar relationship betwen the magnitude c_2 and observing altitude. Observations

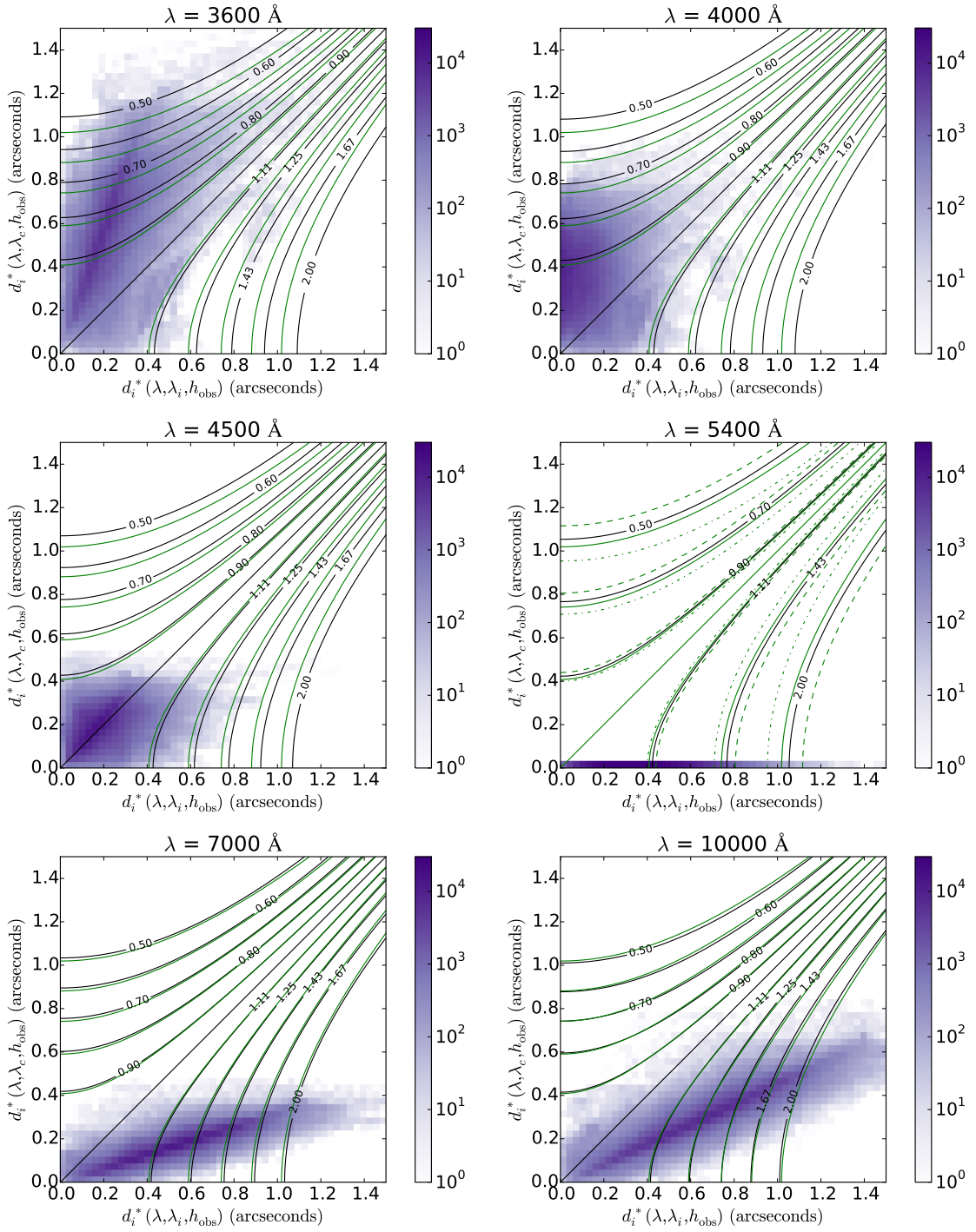

Figure 2.7: 2D histrogram of $d_i^*(\lambda, \lambda_i, h_{\rm obs})$ and $d_i^*(\lambda, \lambda_c, h_{\rm obs})$ focal plane position offsets relative to the fiber hole center for all offset targets. The color of each 2D bin in a panel corresponds to the number of entries as indicated by the panel's adjacent vertical color bar. The green and black contours represent levels of $R_i^*(\lambda)$ calculated assuming Gaussian and Kolmogorov PSF shapes, respectively. In the panel labeled $\lambda = 5400$ Å the additional dotted, solid, and dashed contours represent levels of $R_i^*(\lambda)$ for PSF FWHM of 1.2, 1.5, and 1.8 arcseconds, respectively.

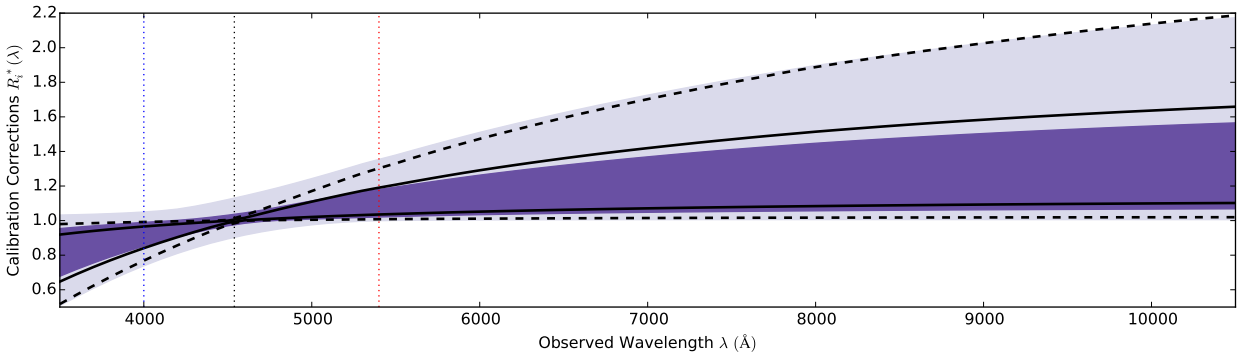

Figure 2.8: Throughput correction summary. The shaded regions correspond to the the central 68% and 95% quantiles of the predicted correction at each wavelength. The black solid and dashed lines similarily correspond to the central 68% and 95% quantiles of $R_i^*(\lambda)$ for observations where $|h_{\rm obs} - h_0| < 1.25^\circ$ (5 minutes), approximately 10% of all observations.

at higher altitude tend to have smaller c_1 and c_2 values, which corresponds to a smaller correction. For reference, we also display the relationship between $(h_{\rm obs} - h_0)$ and $(a_{\rm obs} - a_0)$ in Figure 2.12.

2.5 Validation

In order to validate our model, we use observations that have been calibrated with offset standards, as described previously. In Figure 2.13, we display the median flux ratio of DR12 failed quasar spectra, with (blue) corrections and without (red) corrections, compared to their corresponding spectra from the "validation" reduction. The corrections remove a wavelength-dependent systematic bias in the distribution of BOSS quasar spectra. The vertical spread in each distribution indicates that there are other factors affecting the calibration but that they are independent of the effect described in this work. Figure 2.14 displays a similar comparison between a set of quasar targets with spectroscopic observations in both BOSS and SDSS-I. In this case, the SDSS-I observation is used for reference since focal plane offsets were not introduced prior to BOSS. We note a significant drop off below 5000 Åbut do not necessarily expect agreement between SDSS-I and BOSS because of changes to the

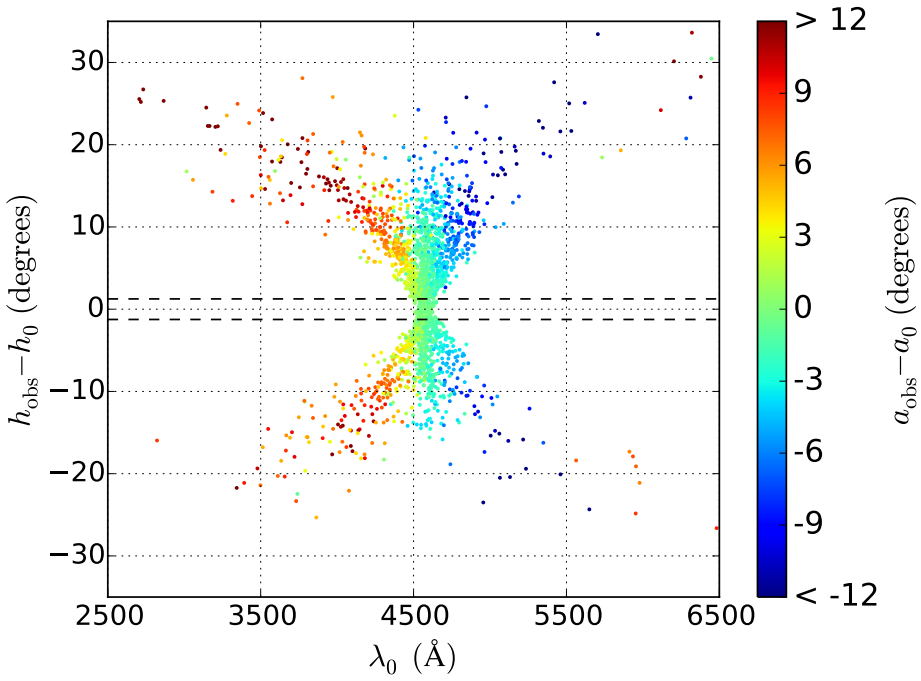

Figure 2.9: Relationship between median fit parameter λ_0 and hour angle difference ($h_{\rm obs} - h_0$) for all 2377 DR12 plates with offset targets. Points are colored by the corresponding altitude difference ($a_{\rm obs} - a_0$) for the observation.

spectrographs and the intrinsic variability of quasars on these timescales.

Following the assessment of spectrophotometric calibration performed in Dawson et al. [2013], we calculate synthetic AB magnitudes [Oke and Gunn, 1983] using SDSS filter-curve-weighted[5] integration of spectra for BOSS targets and compare them to their targets' corresponding SDSS imaging magnitudes [Gunn et al., 1998, Fukugita et al., 1996]. We apply AB magnitude corrections for SDSS imaging magnitudes ($m_{AB} - m_{SDSS}$) used by the BOSS data processing pipeline for the g-, r-, and i-band of 0.036, 0.015, and 0.013. The BOSS spectrophotometric calibration of standard stars yields $(g - r)$ colors that are on average 0.022 ± 0.003 magnitudes redder compared to their SDSS imaging magnitudes. Note that we calculate the sample dispersion using the normalized median absolute deviation. For the offset standards, the $(g-r)$ colors are 0.117 ± 0.004 magnitudes bluer before applying the correction and 0.006 ± 0.004 magnitudes redder after. Using spectra from the validation pipeline

[5] `http://www.sdss.org/instruments/camera/#Filters`

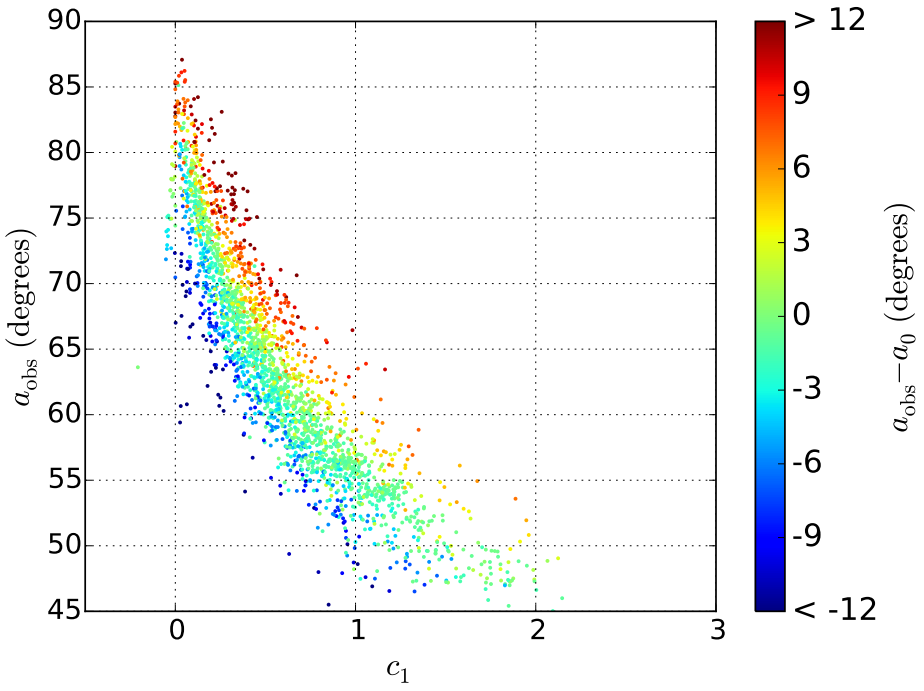

Figure 2.10: Relationship between median fit parameter c_1 and observing altitude $a_{\rm obs}$ for all 2377 DR12 plates with offset targets. Points are colored by the corresponding altitude difference $(a_{\rm obs} - a_0)$ for the observation.

yields $(g-r)$ colors that are 0.010 ± 0.003 magnitudes redder. Figure 2.15 displays the distribution of color residuals $\Delta(g-r)$ between SDSS imaging and BOSS synthetic magnitudes for the offset standards. The red and blue distributions correspond to synthetic magnitudes calculated from the BOSS DR12 spectra before and after applying the corrections, respectively. The black dashed line corresponds to synthetic magnitudes calculated using spectra from the validation pipeline that uses offset standard stars for spectrophotometric calibration.

A complete listing of magnitude and color residuals for the relevant data samples is given in Table 2.2. Our corrections generally remove a significant bias in the mean and reduce the dispersion of the magnitude and color residuals.

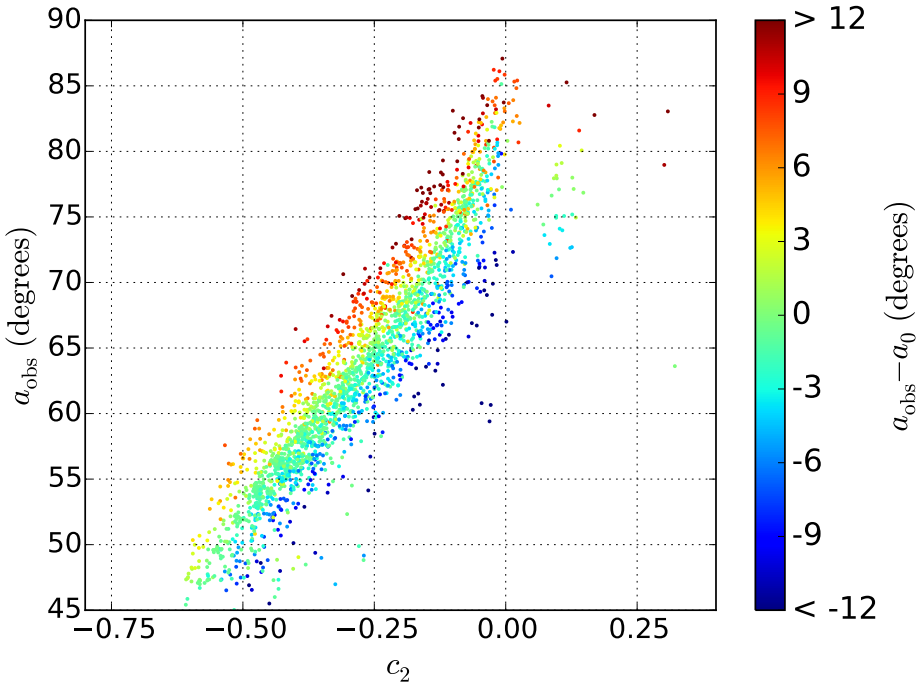

Figure 2.11: Relationship between median fit parameter c_2 and observing altitude $a_{\rm obs}$ for all 2377 DR12 plates with offset targets. Points are colored by the corresponding altitude difference $(a_{\rm obs} - a_0)$ for the observation.

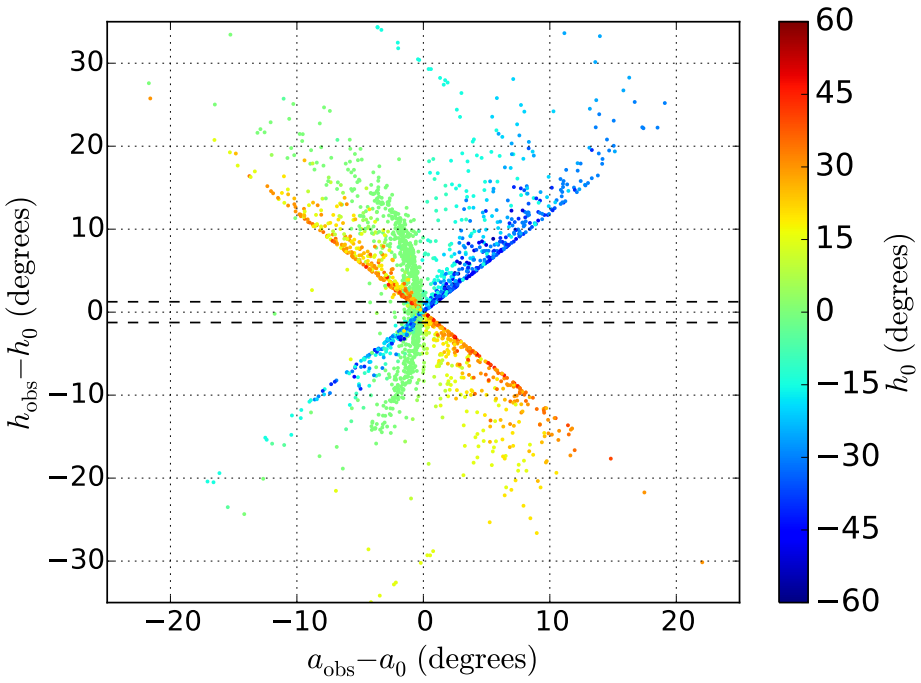

Figure 2.12: Relationship between the hour angle difference $(h_{\rm obs} - h_0)$ and altitude difference $(a_{\rm obs} - a_0)$ for all 2377 DR12 plates with offset targets. Points are colored by the corresponding design hour angle h_0 for the observation.

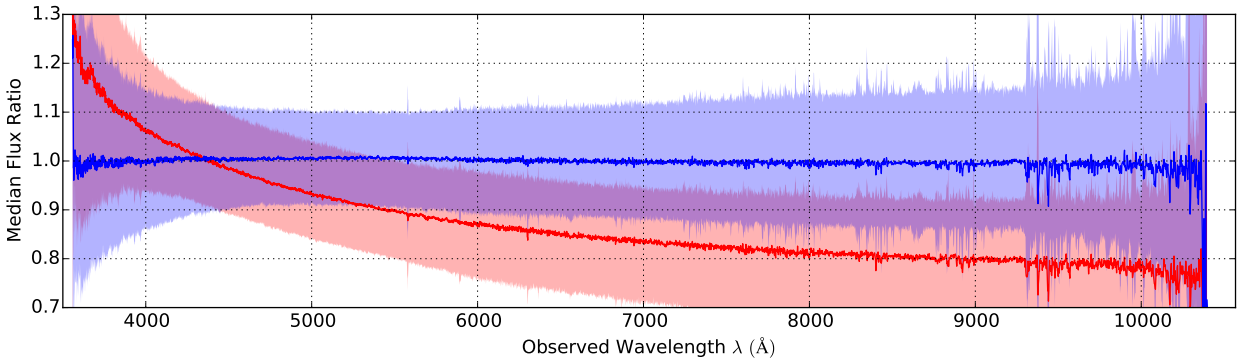

Figure 2.13: Median flux ratios of failed quasar spectra from the 20 validation plates. The blue (red) curve shows the median flux ratio between corrected (uncorrected) BOSS spectra and their corresponding "validation" calibrated spectra. The shaded region corresponds to the 68% level of the distribution in each wavelength bin.

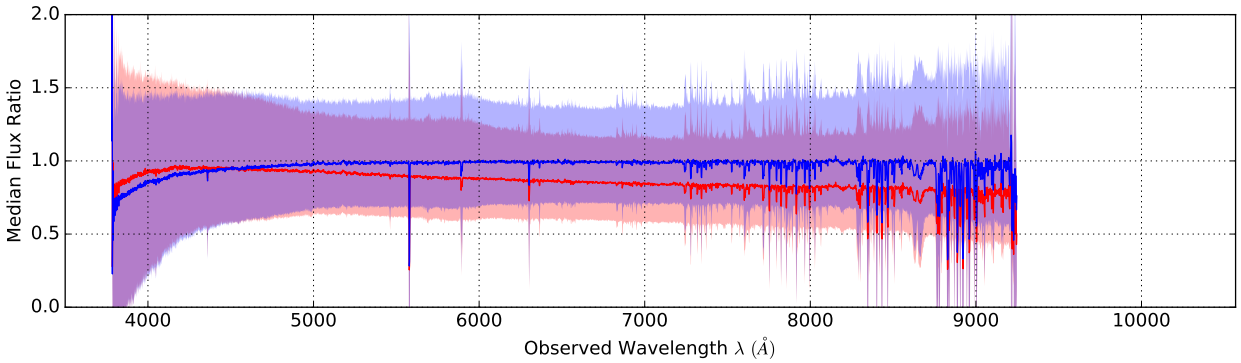

Figure 2.14: Median flux ratios of quasar spectra that were observed in both SDSS-I and BOSS. The blue (red) curve shows the median flux ratio between corrected (uncorrected) BOSS spectra and their corresponding SDSS-I spectra.

Data Sample	# Targets	$\Delta(g-r)$		$\Delta(r-i)$		Δg		Δr		Δi	
		Mean	Disp.	Mean	Disp.	Mean	Disp.	Mean	Disp.	Mean	Disp.
Spec. standards	400	0.022	0.057	0.014	0.032	0.032	0.057	0.011	0.052	−0.004	0.060
Offset standards	486	−0.117	0.098	−0.053	0.068	0.084	0.122	0.201	0.159	0.255	0.214
Corr. offset standards	486	0.006	0.085	0.004	0.062	0.047	0.124	0.041	0.163	0.037	0.198
Spec. offset standards	486	0.010	0.077	0.008	0.058	0.054	0.086	0.044	0.097	0.036	0.127
Failed quasars	1,049	−0.067	0.132	−0.065	0.127	0.071	0.158	0.138	0.187	0.203	0.245
Corr failed quasars	1,049	0.058	0.135	−0.006	0.116	0.029	0.157	−0.028	0.186	−0.023	0.238

Table 2.2: Band-pass filter magnitude and color differences between SDSS imaging and synthetic AB magnitudes calculated from BOSS spectra. The sample mean and dispersion are listed for various data samples. Dispersions are calculated as the normalized median absolute deviation.

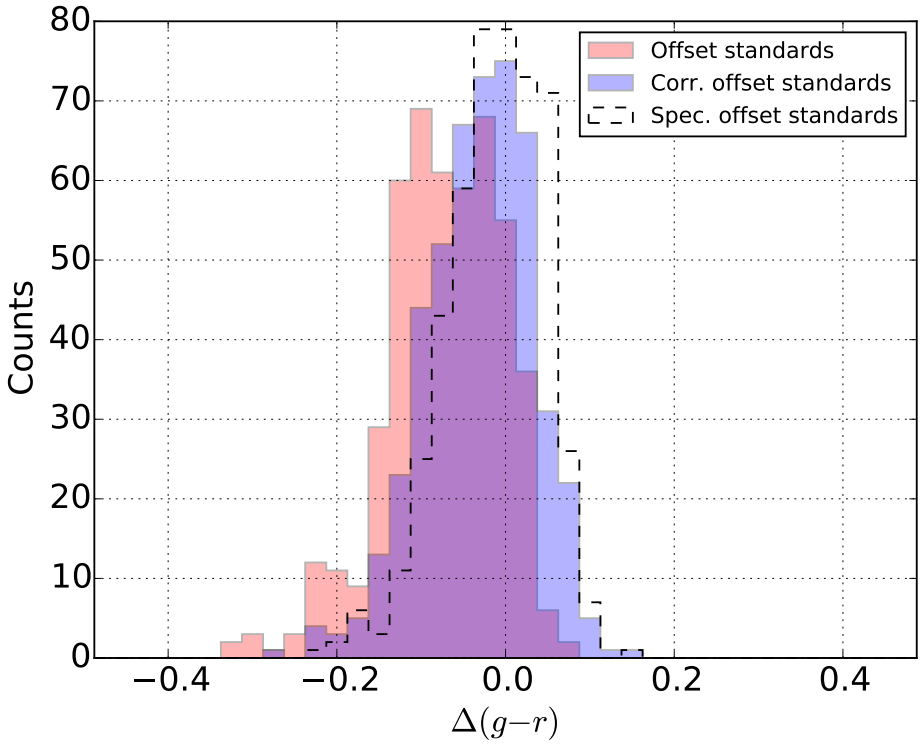

Figure 2.15: Residual $(g-r)$ color distribution between SDSS imaging and BOSS synthetic magnitudes for a sample of offset standard stars. The red and blue distributions correspond to synthetic magnitudes calculated from the BOSS DR12 spectra before and after the correction, respectively. The black dashed line corresponds to synthetic magnitudes calculated from the modified data processing pipeline that uses offset standard stars for spectrophotometric calibration.

2.6 Discussion

The corrections described here provide a significant improvement over the uncorrected quasar spectra in DR12 and are publicly available[6]. However, there are limitations to attempting to correct the quasar spectrophotometric calibration as a post-processing step after the data processing pipeline, since many simplifying assumptions are required. In order to achieve better results, it is necessary and desirable to use as much of the existing pipeline calibration algorithms as possible but with spectrophotometric standards observed with the same

[6] `http://darkmatter.ps.uci.edu/tpcorr/`.

focal-plane offsets as the quasar targets. Short of including spectrophotometric standards with focal plane offsets, the throughput corrections described here can be applied as a pre-processing step of the calibration process. Other improvements over the method described here include using a more realistic description for the PSF, using the actual guiding adjustments for determining centroid offsets, and applying corrections to individual exposures prior to co-addition. This is especially relevant for the extended Baryon Oscillation Spectroscopic Survey (eBOSS) [7] of SDSS-IV, which uses similar focal plane offsets for quasar targets. We are currently working with the eBOSS data processing team to implement the throughput corrections described for individual exposures into the data processing pipeline. The proposed Dark Energy Spectroscopic Instrument [DESI; Levi et al., 2013] includes an atmospheric dispersion compensator in its design so the effects of ADR should be much reduced for that survey.

2.7 Calibration Tests

The observed flux $f_i^{\rm obs}$ for a target i related to its true flux $f_i^{\rm true}$ by

$$f_i^{\rm obs} = f_i^{\rm true} D_i \tag{2.16}$$

where D_i represents the distortion of the true flux caused by the atmosphere, telescope, and/or spectrograph. If we could observe a "standard" target j under precisely the same conditions (same sky position, focal plane hole, fiber cable, camera pixels) as target i, we could fit for the distortion D_j using a model of the true flux $f_j^{\rm model}$ and observation $f_j^{\rm obs}$. Let's call the thing we learn from this exercise the calibration C_j where

$$C_j \equiv {D_j}^{-1} = f_j^{\rm model}/f_j^{\rm obs} \tag{2.17}$$

[7] `http://www.sdss.org/surveys/eboss/`

We could then estimate the true flux $f_{i,j}^{\rm est}$ using information from both observations as

$$f_{i,j}^{\rm est} = f_i^{\rm obs} C_j \left(\frac{D_j}{D_i} \right) \tag{2.18}$$

If i and j were observed under the exact same conditions we could assume $D_i = D_j$. Unfortunately, we can't observe i and j under the exact same conditions so $D_i \neq D_j$. The measured fluxes will have different losses due to camera pixels, fiber cables, focal plane positions, atmospheric transmission, etc. For example, if we assume that the $\lambda^{\rm design}$ for the fiber hole position is the only difference in the observations, the distortion ratio becomes $D_j/D_i = \Lambda_j/\Lambda_i$. More specifically, if the fiber for j is positioned for 5400 Å and the fiber i is positioned for 4000 Å) our true flux estimate for i becomes

$$f_{i,j}^{\rm est} = f_i^{\rm obs} C_j \left(\frac{\Lambda^{5400}}{\Lambda^{4000}} \right) \tag{2.19}$$

The pipeline does some fancy stuff to try to account for the ratio of distortions using multiple "standards" distributed across the field of view during the observations. At the end of the day we are left with something more like:

$$f_{i,\{j\}}^{\rm est} = f_i^{\rm obs} C_{i,\{j\}} \tag{2.20}$$

where $C_{i,\{j\}}$ is an estimate of the calibration for target i derived from a set of observed fluxes $\{j\}$.

2.8 How well do we measure $C_{i,\{j\}}$

If we estimate $f_i^{\rm true}$ using two sets of "standard" targets $\{j\}$ and $\{j'\}$ we can compare the calibrations using the ratio of flux estimates:

$$\frac{f_{i,\{j\}}^{\rm est}}{f_{i,\{j'\}}^{\rm est}} = \frac{C_{i,\{j\}}}{C_{i,\{j'\}}} \tag{2.21}$$

Then we can estimate the uncertainty in $C_{i,\{j\}}$ using the distribution of ratios from a sample of targets i. The uncertainty in the calibration is related to RMS of this distribution by a factor of $\sqrt{2}$. For two sets of standards, $\{j\} = 1$ and $\{j'\} = 2$, the RMS of the distribution of ratios is related to the RMS of each of the flux estimates:

$$\frac{\sigma^2_{f_1/f_2}}{|f_1/f_2|^2} = \frac{\sigma^2_{f_1}}{f_1^2} + \frac{\sigma^2_{f_2}}{f_2^2} \tag{2.22}$$

With the expectation that $f = f_1 = f_2$ and $\sigma_f = \sigma_{f_1} = \sigma_{f_2}$ we have:

$$\sigma^2_{f_1/f_2} = 2\sigma_f^2 \tag{2.23}$$

2.9 Half Plate Calibration Test

The calibration C_i for a target i is:

$$C_{i,\{j\}} = f_{i,\{j\}}^{\rm est}/f_i^{\rm obs} \tag{2.24}$$

where i belongs to a plate p and a spectrograph s and the set of "standards" $\{j\}$ is different for each spectrograph. The spectrograph is related to a target's fiber id by:

$$s = \begin{cases} 1 & : 1 \leq i \leq 500 \\ 2 & : 501 \leq i \leq 1000 \end{cases} \tag{2.25}$$

We can study the calibration due to the set of "standard" targets with the assumption that there are not any other spectrograph dependent sources of error in an observation. The mean correction for a spectrograph observation is obtained by averaging the corrections for all the targets in each spectrograph for an observations:

$$C_{p,s} = \langle C_{i\{j\}} \rangle_{i,\{j\} \in p,s} \tag{2.26}$$

The average mean correction for either spectrograph over all observations is:

$$C_s = \langle C_{p,s} \rangle_p \tag{2.27}$$

A deviation from the mean correction for a spectrograph observation is:

$$\Delta C_{p,s} = C_{p,s} - C_s \tag{2.28}$$

The RMS of the mean correction is:

$$\sigma_s = \sqrt{\langle C_{p,s}^2 \rangle_p - C_s^2} \equiv \sqrt{\langle \Delta C_{p,s}^2 \rangle_p} \tag{2.29}$$

For an observation we can compare the relative calibrations in each spectrograph:

$$\Delta C_p = \Delta C_{p,2} - \Delta C_{p,1} \tag{2.30}$$

The RMS of this distribution tells us about relative difference in calibration between spectrographs.

Chapter 3

The Lyman-α Forest Sample

In this chapter, I describe the procedure for estimating the matter density field tracer $\delta_F(\vec{r})$ using quasar observations from the BOSS survey discussed in the previous chapter. I start by summarizing the sample of quasar observations in Section 3.1. In Section 3.2, I describe the pixel selection criteria and processing steps to set up efficient computation in the following sections. Then, in Section 3.3, I detail the continuum fitting procedure that is used to estimate the unabsorbed flux in the Lyman-α forest region of the quasar spectrum observations. Finally, the transformation from flux measurements to the matter density field tracer $\delta_F(\vec{r})$ is explained in Section 3.4.

3.1 DR12Q Sample

This work uses BOSS quasar observations from the SDSS Data Release 12 (DR12) sample [Dawson et al., 2013]. The quasar target selection of BOSS is described in Ross et al. [2012]. The DR12 Quasar Catalog (DR12Q) is used to select observations for this analysis. The DR12Q is a manually curated catalog of BOSS quasar observations that have been visually

identified and confirmed as described in Pâris et al. [2016]. In total, 155,446 high redshift ($2.1 < z < 3.5$) quasars with optical spectra are used in this analysis. The distribution of observations on the sky is shown in Figure 3.1 and spans the entire BOSS survey footprint, which is approximately 10,400 square degrees or one quarter of the entire sky.

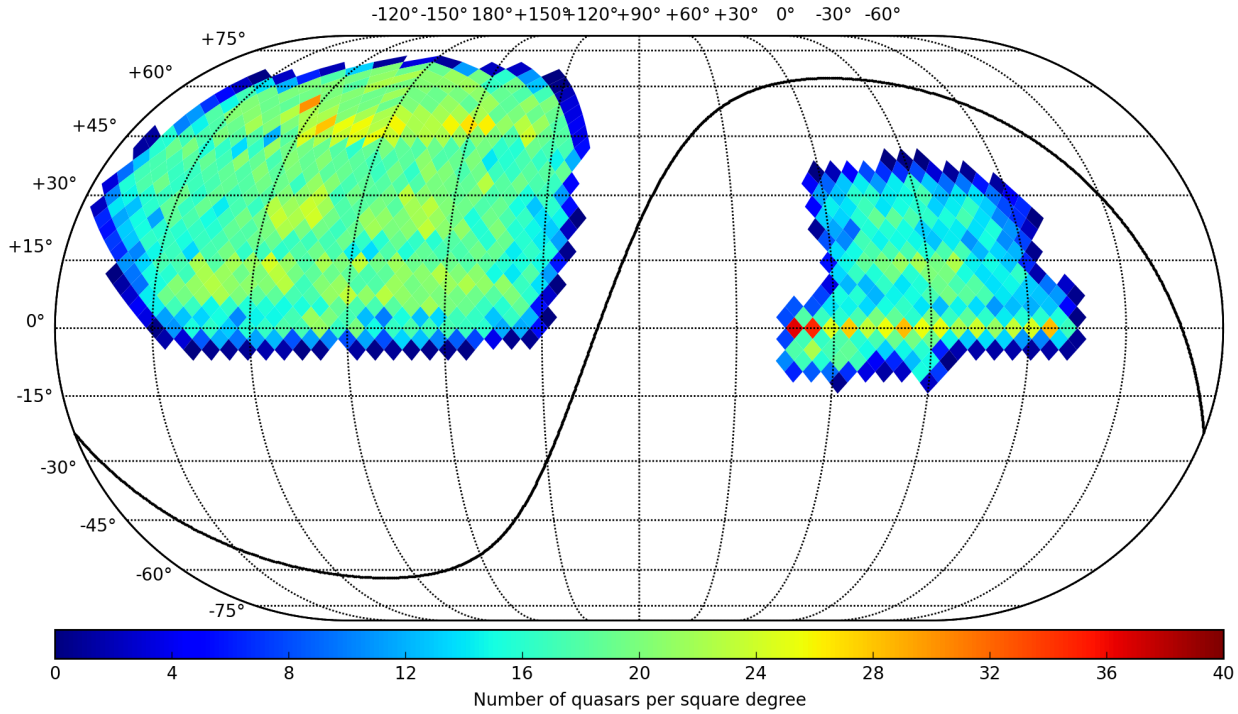

Figure 3.1: Distribution of DR12 high redshift ($2.1 < z < 3.5$) quasar observations on the sky (Eckert IV projection). The data are binned in equal area pixels corresponding to approximately 13.42 square degrees using HEALPix [Górski et al., 2005], with grid resolution parameter $N_{\rm side} = 16$. The color bar indicates the corresponding number of quasars per square degree. The thick black line indicates the position of the galactic plane.

The DR12Q redshift estimate (indicated by `Z_VI` in the catalog) is used to select quasars with redshift $2.1 < z < 3.5$. The redshifts in the DR12Q catalog have been fit manually starting from the automated estimate of the BOSS data reduction pipeline. Manual fitting is sometimes necessary in cases where an emission line is misidentified, a dubious feature is considered an emission line, or when emission lines are correctly identified but not properly centered. The lower limit of the redshift cut is imposed by the lower limit of the wavelength range of the BOSS spectrograph 3,600 Å $< \lambda_{\rm obs} <$ 10,500 Å. The upper limit is applied

to the sample because the number density of quasars becomes low enough such that there are not enough neighboring lines of sight to effectively probe the BAO scale. The redshift distribution of all quasars in the DR12Q sample is shown in Figure 3.2, the vertical dashed lines indicate the redshift limits described previously. In Figure 3.3, I show a few example high signal-to-noise quasar observations with redshifts spanning the limits of the sample. The most prominent emission line in all cases is Lyman-α. The Lyman-α forest is notably visible on the blue side of the Lyman-α emission peak.

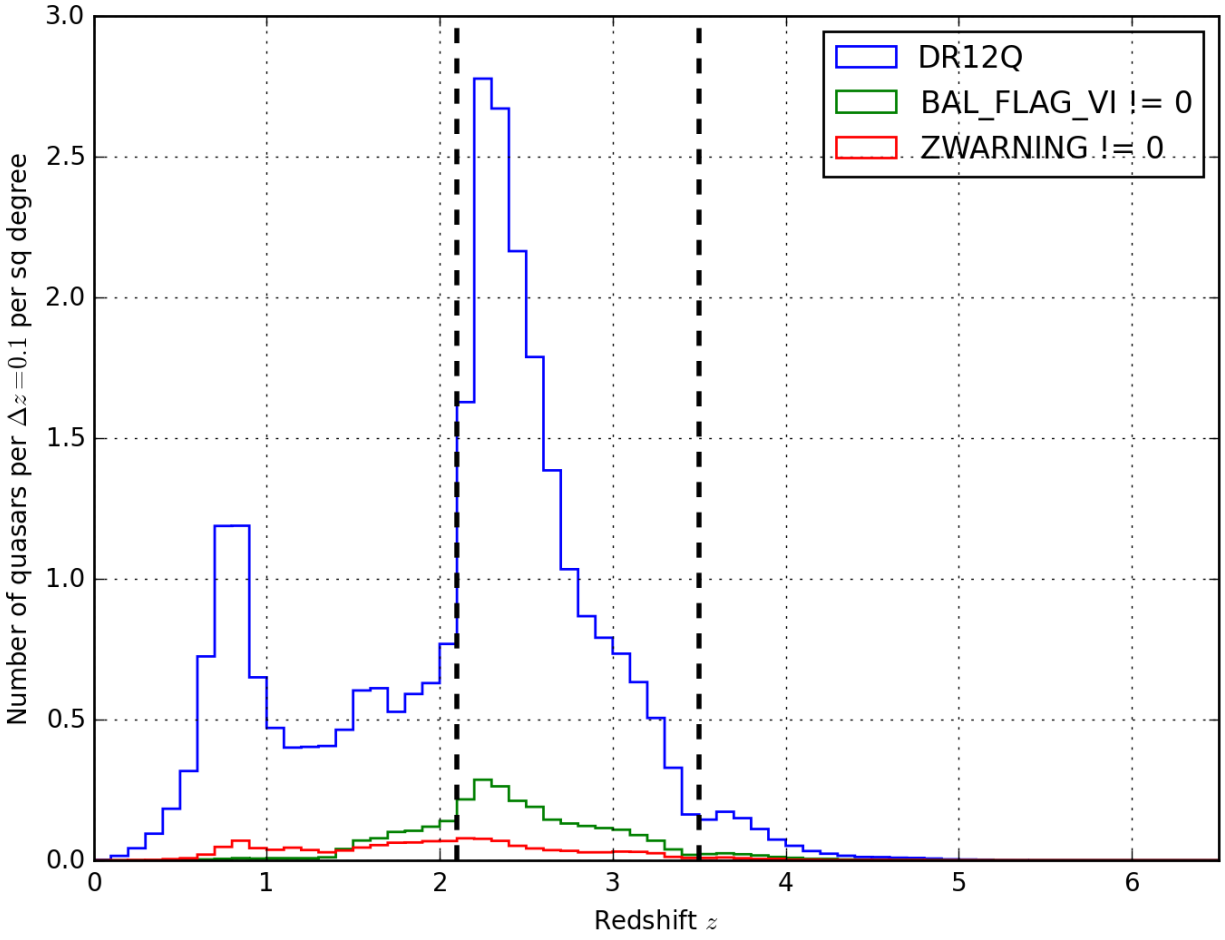

Figure 3.2: Redshift distribution of the DR12 quasar sample. The blue line shows the redshift distribution for all visually confirmed quasars in DR12. The green and red lines show the redshift distribution of visually confirmed quasars in DR12 that have been flagged with `ZWARNING` and `BAL_FLAG_VI`, respectively. The vertical dashed lines indicate the boundaries of the high redshift quasar sample ($2.1 < z < 3.5$) used in this work.

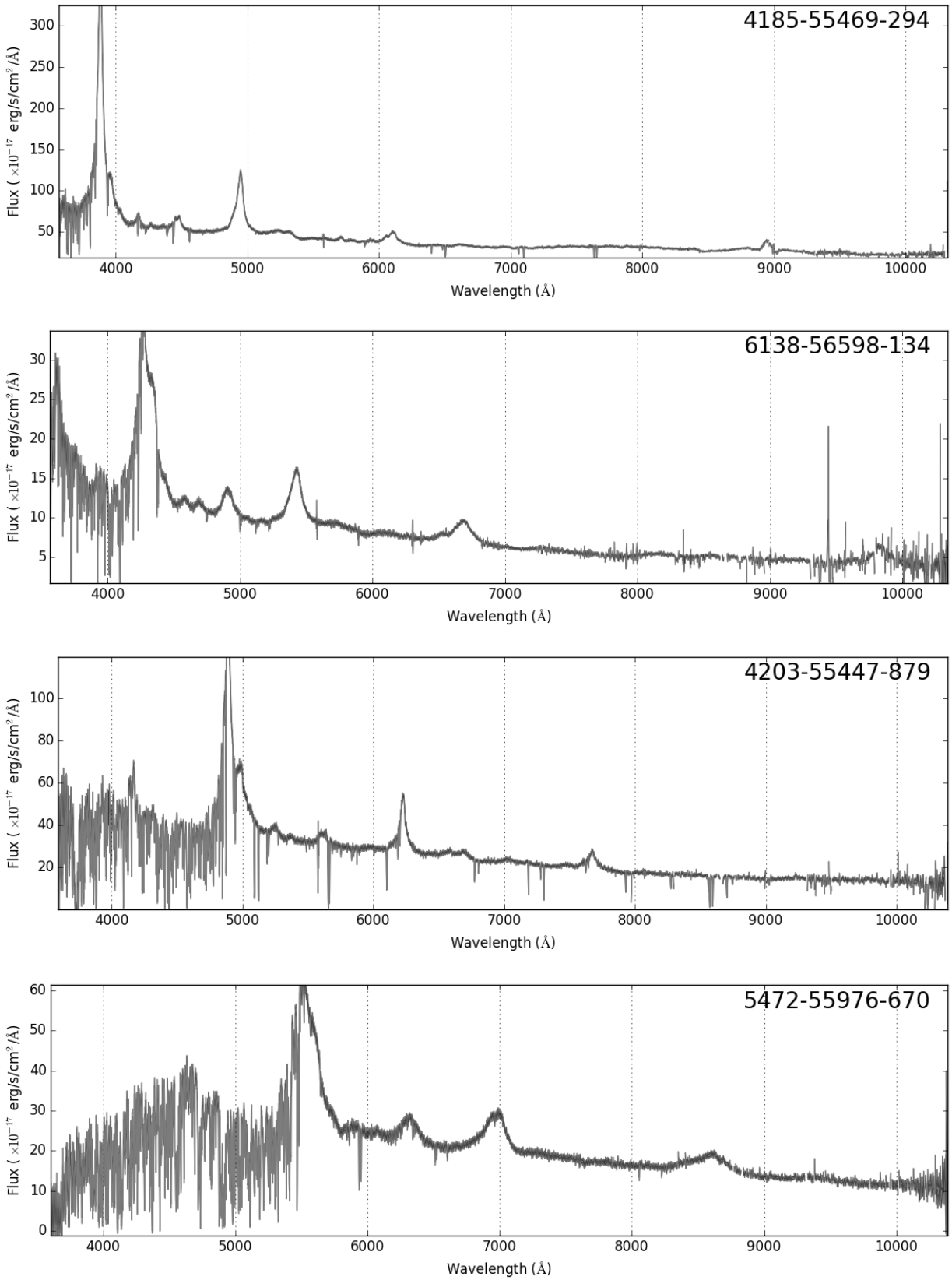

Figure 3.3: Example high signal to noise quasar spectra. The corresponding redshifts are approximately 2.2, 2.5, 3, and 3.5 from top to bottom.

In addition to the quasar redshift, a number of spectral features can potentially be used to filter a sample for analysis. The most prominent and commonly used features to characterize potentially biased spectra are Broad Absorption Lines (BAL) and Damped Lyman-α systems (DLA). Observations that have been identified in the DR12Q catalog (via the `BAL_FLAG_VI` flag) to contain BAL features are excluded. Observations with DLA are also indicated in the DR12Q but usually one would mask or attempt to correct the region the spectrum that contains the DLA. Other spectral features that potentially could systematically affect a BAO measurement are metal absorption lines and sky lines. Attempts to mask or filter pixels that are affected by these sources of contamination has been carried out by the BOSS Lyman-α working group but have not seen any significant evidence that they impact the final measurement of the BAO peak position. In the few cases where an object has been observed on multiple occasions, the observation marked `SPEC_PRIMARY` is used. Finally, we require that spectra have `ZWARNING` equal to zero from the BOSS data reduction pipeline which indicates that a spectrum has no known problems.

3.2 Pixel Selection

The goal of this phase of the analysis is to process and organize the data so that the following operations can be performed efficiently in parallel whenever possible. First, each spectrum is normalized by the integrated inverse variance weighted flux between 1,275 Å $< \lambda_{\rm rest} <$ 1,285 Å, where $\lambda_{\rm rest} = \lambda_{\rm obs}/(1 + z_{\rm quasar})$. Then, spectral observations are trimmed to only contain pixels in the Lyman-α forest, defined as 1,040 Å $< \lambda_{\rm rest} <$ 1,200 Å (conservatively between the Lyman-α and Lyman-β hydrogen transition wavelengths). Additionally pixels with wavelength less than 3,600 Å are excluded due to the significant drop off of signal-to-noise in the BOSS spectrographs. Pixels that are masked by the BOSS data reduction pipeline are ignored, as well as those with inverse variance equal to zero. Figure 3.4 shows the

observed flux in the forest region after normalizing each of the spectra. Each row corresponds to a unique quasar spectrum, using one of every hundredth quasar in redshift sorted order.

After pixel selection, the spectra are down-sampled by a factor of three, using an inverse variance weighted average. This is done to reduce the overall number of pixels in the correlation function estimate which naively has a computational complexity of $\mathcal{O}(N^2)$ where N is the total number of pixels in the sample and N^2 is effectively the number of pixel pairs, approximately $(150,000 \times 400)^2 = 3.6 \times 10^{15}$. Down-sampling by a factor of three reduces the computational cost by nearly an order of magnitude, with negligible impact on the measurement.

Finally, the spectra are transformed to their corresponding rest frames, effectively aligning the sample for efficient use matrix operations in the following steps. Figure 3.5 shows the sample after aligning the data in the quasar rest-frame.

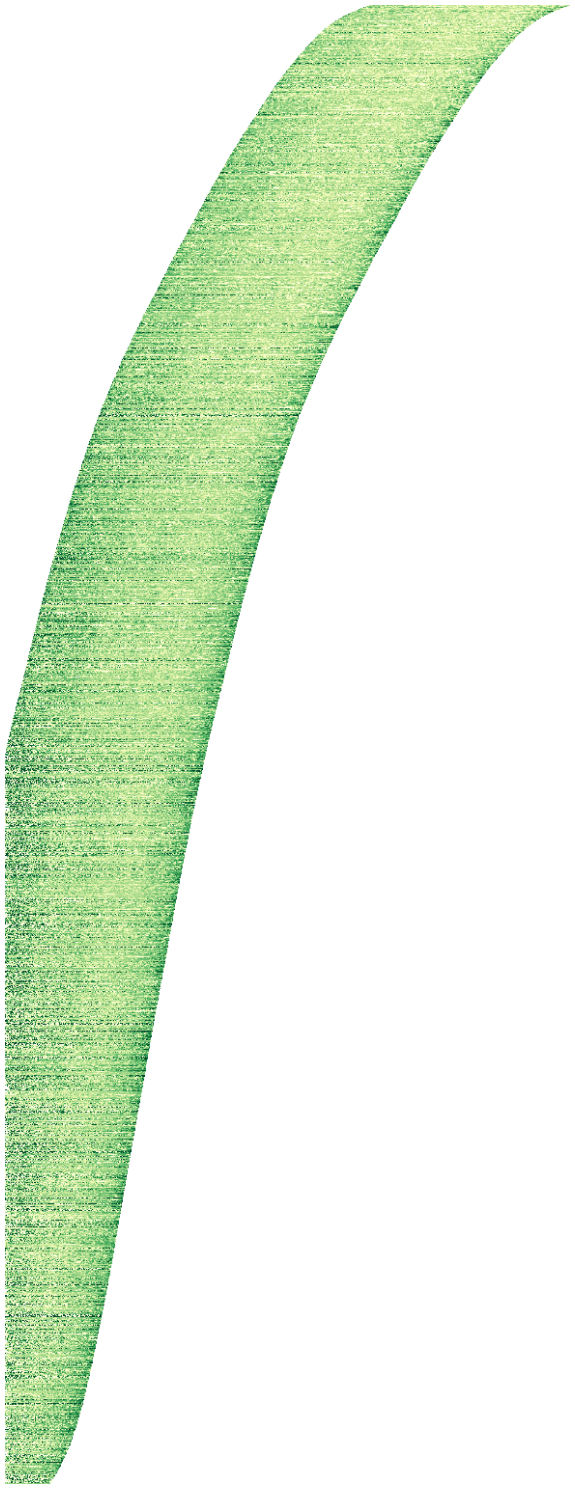

Figure 3.4: One percent of the DR12 Lyman-α forest sample aligned in the observing frame. The observed flux in the forest region has been normalized by the inverse variance weighted average flux observed between $1,275$ Å $< \lambda_{\rm rest} < 1,285$ Å of the corresponding quasar. Each row corresponds to a unique quasar spectrum, using one of every hundredth quasar in redshift sorted order.

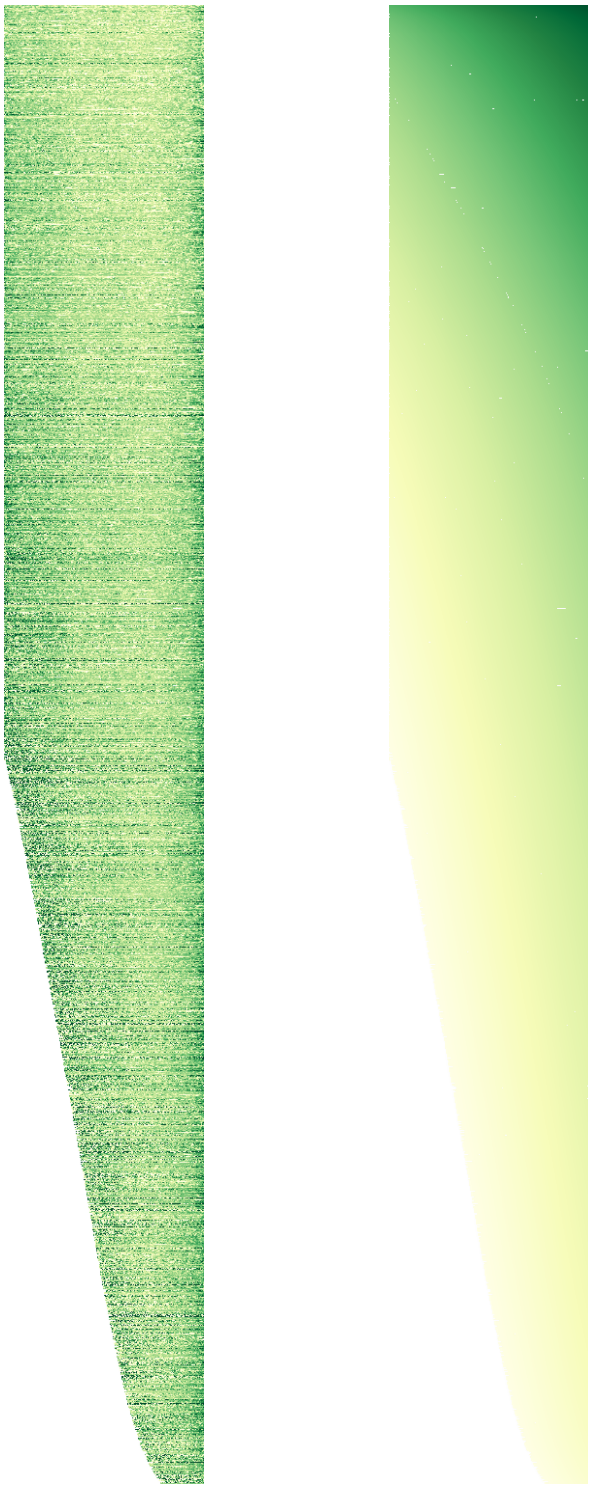

Figure 3.5: One percent of the DR12 forest sample aligned in the rest frame. The plot on the left shows the normalized flux values of pixels aligned in the quasar rest frame. The redshift of the corresponding absorbers is shown on the right where the color scale from lightest to darkest corresponds to the interval $1.9 < z < 3.5$.

3.3 Continuum Fitting

The approach described here has been developed to emphasize computational efficiency. The aim is to establish a baseline analysis that will serve as a starting point so that future analyses can test and assess a multitude of potential systematic effects and analysis design decisions.

The basic idea is to turn the continuum fitting problem into one that can be solved using the method of linear least squares. Note that, for given quasar observations, the observed flux at wavelength λ_o from the quasar q can be modeled as

$$f(q, \lambda_o) = a(q) \left(\lambda_r / \lambda_\star\right)^{b(q)} C(\lambda_r) \bar{F}(\lambda_o)(1 + \delta_F(q, \lambda_o))T(\lambda_o) \tag{3.1}$$

where $\lambda_r = \lambda_o/(1 + z_q)$ is the rest-frame wavelength of the quasar. The term $C(\lambda_r)$ represents a universal mean unabsorbed quasar spectrum. The function $T(\lambda_o)$ represents an observed frame transmission function. The parameters $a(q)$ and $b(q)$ allow for variation in the normalization and slope of individual quasars. The mean quasar transmission in the forest is given by $\bar{F}(\lambda_o)$. The observed wavelength is interchangeable with the redshift of the intervening Lyman-α absorbing neutral hydrogen via $\lambda_o = (1 + z)\lambda_\alpha$. The fluctuations in transmission that we seek to measure from the data are represented by δ_F. In order to simultaneously fit for individual quasar shape parameters and the universal continuum and transmission model, we treat the fluctuations as noise and linearize the model by taking the log. Equation (3.1) becomes

$$\log f(q, \lambda_o) = \log a(q) + b(q) \log \lambda_r/\lambda_\star + \log \bar{F}(\lambda_o) + \log C(\lambda_r) + \log T(\lambda_o) \; . \tag{3.2}$$

To keep the baseline analysis as simple as possible, we model the mean absorption in the Lyman-α forest using $\bar{F}(\lambda_o) = \exp(-\tau(\lambda_o))$ and assume a power law model for the redshift

evolution of $\tau(\lambda_o) = \alpha(1+z)^\beta$ following FG08, which leads to

$$\log f(q, \lambda_o) = \log a(q) + b(q) \log \lambda_r/\lambda_\star + \log C(\lambda_r) - \alpha(1+z)^\beta - \log(1+z_q) + \log T(\lambda_o) \ . \quad (3.3)$$

Figure 3.6 shows the continuum $C(\lambda_r)$ in the lyman alpha forest region ($1040 < \lambda_r < 1200$) DR12 high redshift quasars. Figure 3.7 shows the distributed of fitted shape parameters a and b for the sample.

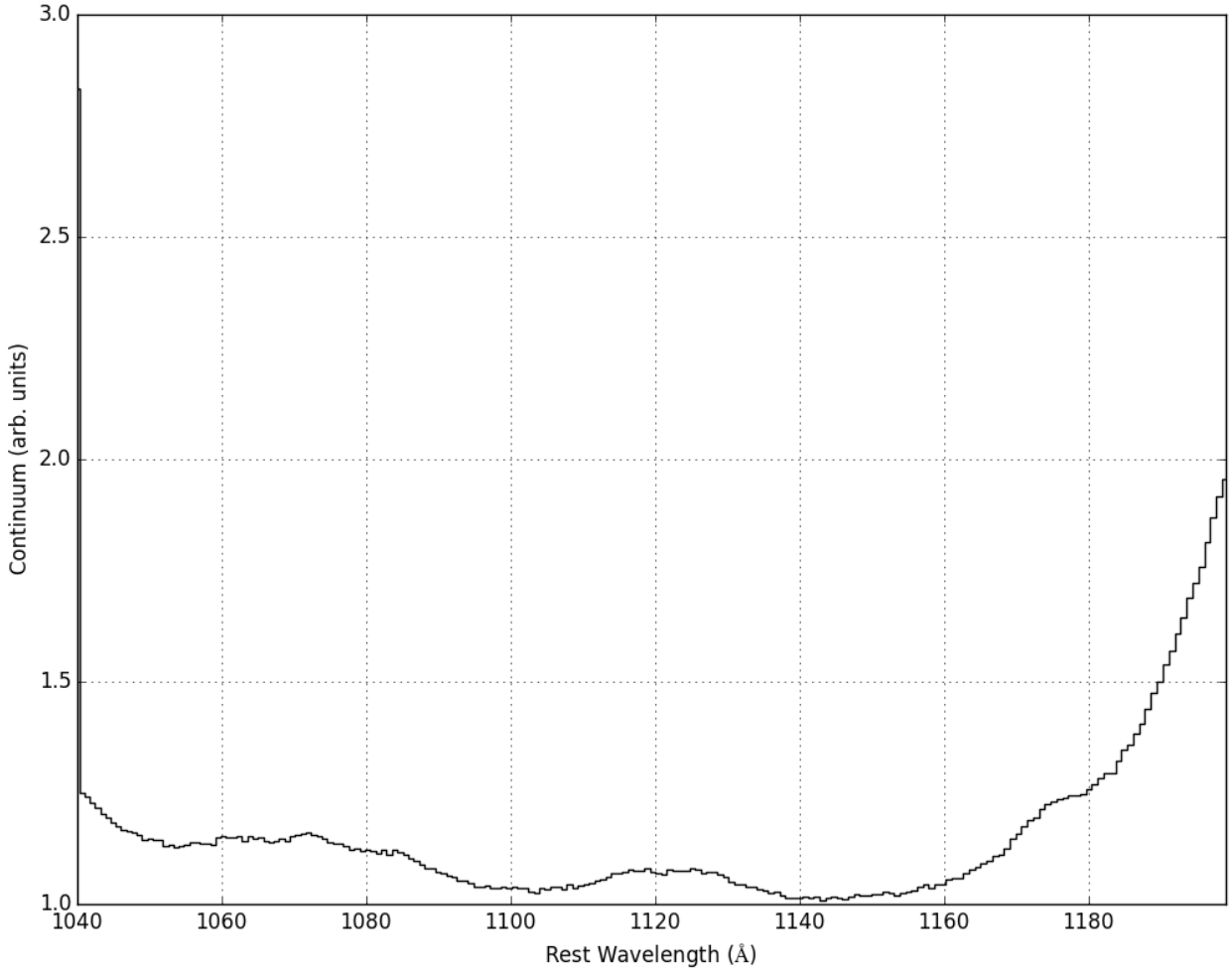

Figure 3.6: Fitted continuum $C(\lambda_r)$ in the lyman alpha forest region (1040 Å $< \lambda_r <$ 1200 Å) DR12 high redshift quasars.

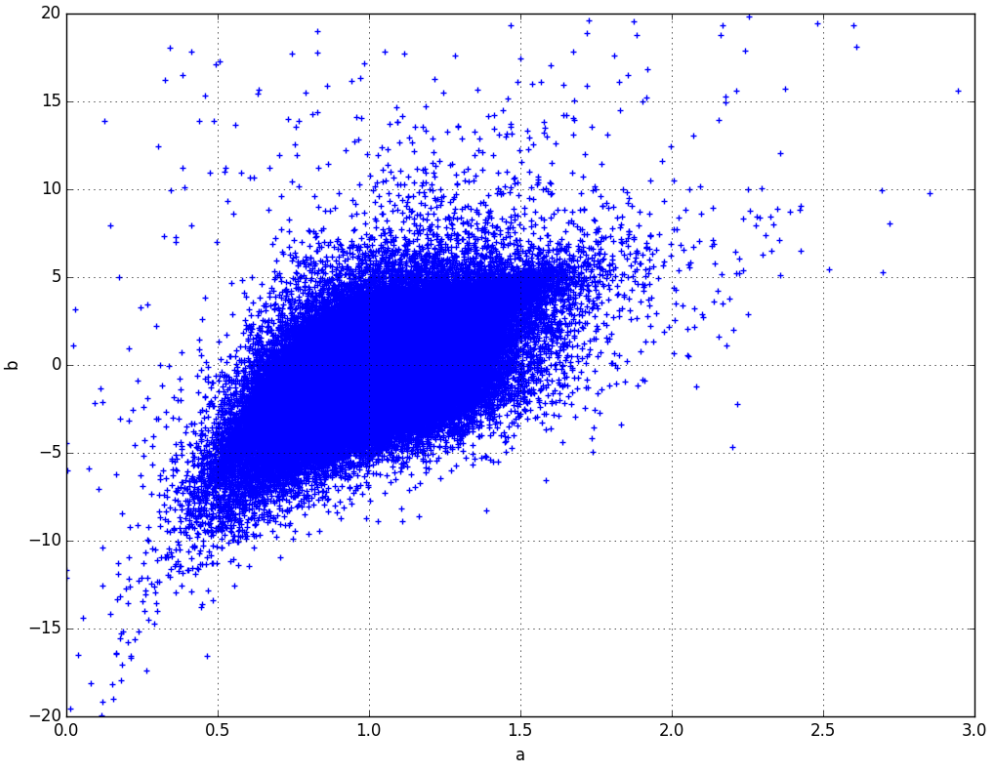

Figure 3.7: Distribution of fitted shape parameters a and b.

3.3.1 Method of Least Squares

For N observations, y, with covariance C, and M model parameters, β, the least squares solution to the overdetermined ($N > M$) set of linear equations,

$$X\beta = y \ , \tag{3.4}$$

is given by

$$\hat{\beta} = [X^T C^{-1} X]^{-1} [X^T C^{-1} y] \ , \tag{3.5}$$

where the matrix X encodes the structure of the linear model. The matrix X is N by M, where N is the total number of quasar pixels and M is the total number of model parameters. This matrix is sparse; there are fewer than $5N$ nonzero entries.

The estimate, $\hat{\beta}$, minimizes the χ^2 objective function,

$$\chi^2 = [y - X\beta]^T C^{-1} [y - X\beta] \ . \tag{3.6}$$

The model matrix X is computed for the pixel sample described in the previous section for corresponding y pixel fluxes. The `linear_model` solver from the `scikit-learn` python package is used to estimate $\hat{\beta}$.

Assuming that the covariance matrix is diagonal (uncertainties are uncorrelated), a weighted least squares fit can be performed by substituting $X = \sqrt{C^{-1}}\ \tilde{X}$ and $y = \sqrt{C^{-1}}\ \tilde{y}$, where the tilde represents the unweighted quantity. However, in this analysis, we assume equal weights since further study must be done to properly understand how the actual pixel variances transform after linearization.

3.4 The Matter Density Tracer Field δ_F

Rearranging Equation (3.1) yields our estimate for the field δ_F along a line of sight

$$\delta_F(q, \lambda_o) = f(q, \lambda_o) / \left(a(q) \left(\lambda_r / \lambda_\star\right)^{b(q)} \bar{F}(\lambda_o) C(\lambda_r) \right) - 1 \tag{3.7}$$

An example delta field along a single line of sight is shown in Figure 3.8. The pixel flux variances are similarly transformed to be variances of our δ_F values. The observed variance in the distribution of δ_F due to cosmic variance is added in quadrature. The mean δ_F as a function of observed wavelength for the entire data sample is shown in Figure 3.9. Ideally, the mean δ_F should be zero but we clearly see structure at wavelengths associated with absorption and emission in the observing frame. For example, the calcium H and K absorption lines are apparent in Figure 3.9. Although there is residual structure, Slosar et al. [2013] and Busca et al. [2013] showed that there is not a significant impact on the position of the BAO peak position.

The observed matter density tracer field along the quasar line of sight $\delta_F(q, \lambda_o)$ can be written as $\delta_F(\vec{r})$ where

$$\vec{r} \equiv (r, \alpha, \delta) \equiv (d_C(z), \alpha, \delta) \tag{3.8}$$

with the co-moving line of sight distance

$$d_C(z) = d_H \int_0^z \frac{dz'}{E(z')} \tag{3.9}$$

and the Hubble distance $d_H = c/H_0$ and

$$E(z) = \sqrt{\Omega_m (1+z)^3 + \Omega_k (1+z)^2 + \Omega_\Lambda} \tag{3.10}$$

where the redshift z is determined by

$$z = \lambda_o/\lambda_{Ly\alpha} - 1 \quad . \tag{3.11}$$

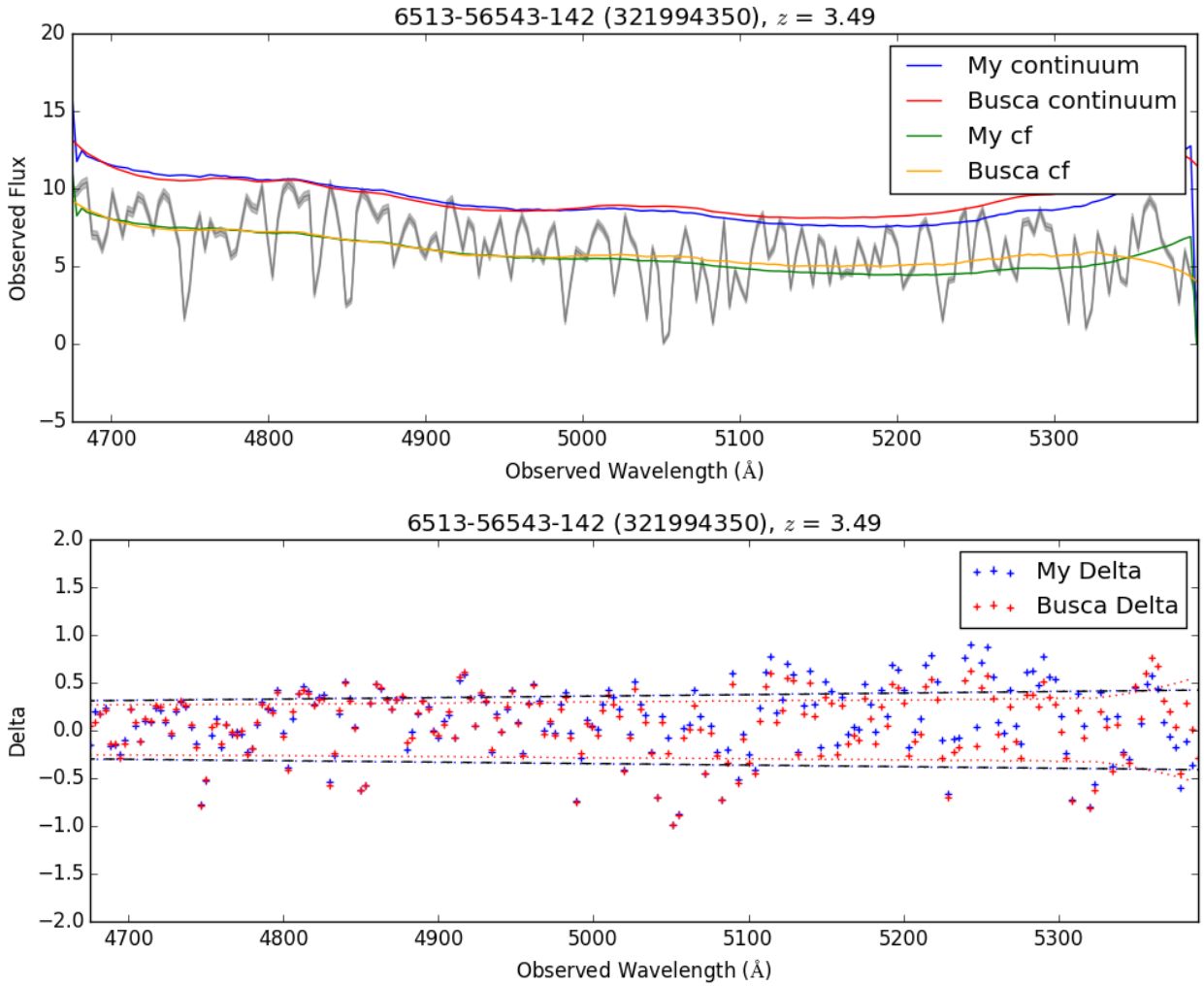

Figure 3.8: Example fit for a single quasar. The top panel shows the observed flux in the Lyman alpha forest for this quasar. The fitted unabsorbed continuum is shown in blue. The bottom panel shows the corresponding delta field.

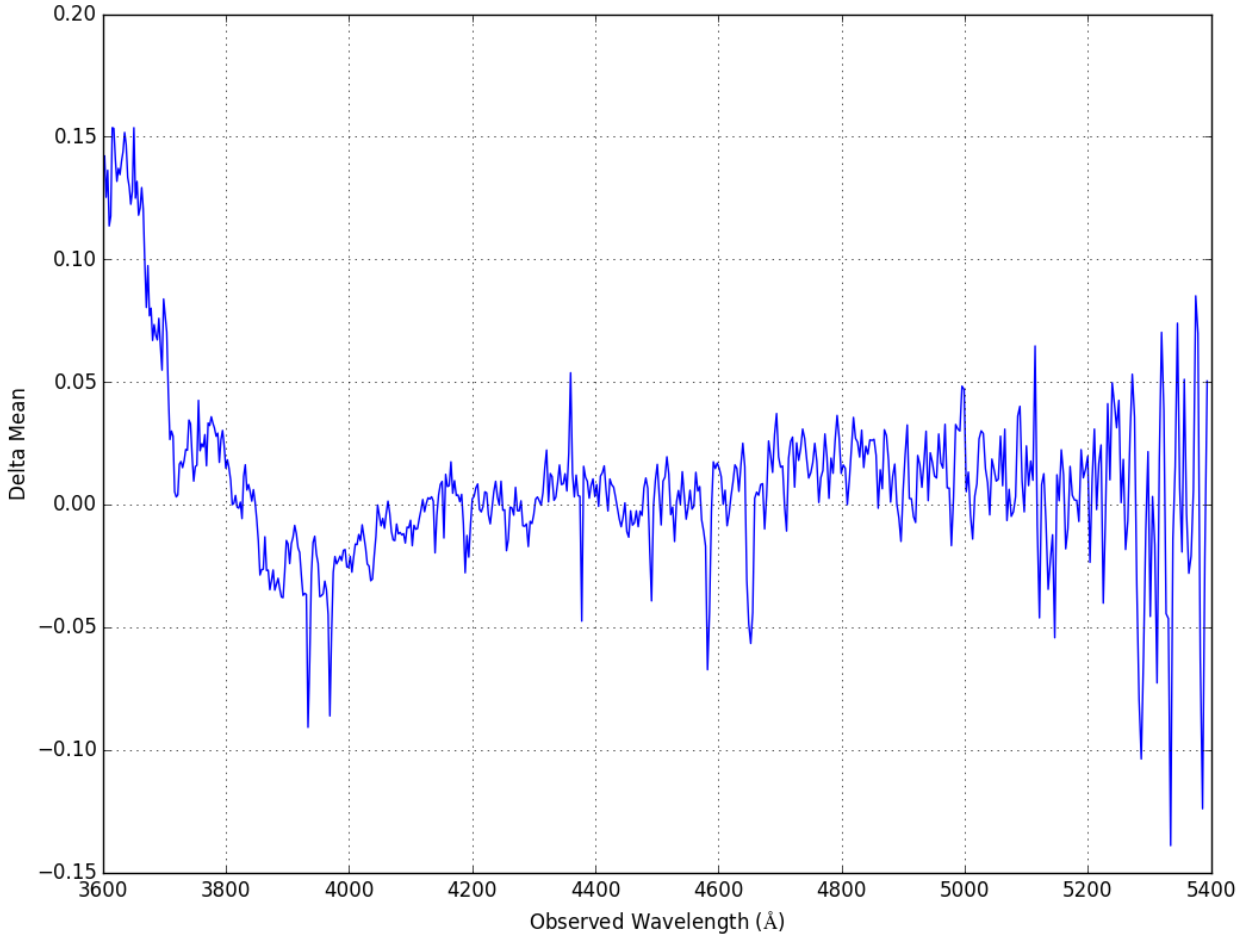

Figure 3.9: Mean δ_F per observed wavelength. The calcium H and K absorption lines are clearly visible.

Chapter 4

The correlation function and its covariance

In the previous chapter, I described the procedure for taking observed spectra of quasars and transforming them into measurements of $\delta_F(\alpha, \delta, z)$. The objective of this chapter is to compute the correlation of δ_F and its covariance.

4.1 Correlation function estimate

After estimating the transmitted flux fraction field, δ_F, we estimate its auto-correlation. A correlation function is a statistical correlation between random variables at two different points in space or time, usually as a function of the spatial or temporal distance between the points. We can use correlation functions to characterize how far and on what scales the distribution of points differs from a random distribution [Ivezić et al., 2014]. The probability of finding a point in a volume element, dV, is directly proportional to the density of points, ρ. The probability of finding a pair of points in two volume elements, dV_1 and dV_2, separated

Quantity	Count	% Reduction
Lines of sight	152,768	-
LOS pairs	23,338,061,824	-
Distinct LOS pairs	11,668,954,528	50.0%
Distinct LOS pairs considered	69,332,985	99.4%
Distinct LOS pairs used	34,356,219	50.4%
Pixels	26,147,581	-
Pixel pairs	683,695,992,151,561	-
Distinct pixel pairs	341,847,983,001,990	50.0%
Distinct pixel pairs considered	1,005,960,788,725	99.7%
Distinct pixel pairs used	260,122,842,193	74.1%

Table 4.1: Pixel pair counts

by a distance, r, is given by

$$dP_{12} = \rho^2 dV_1 dV_2 \left(1 + \xi(r)\right) \tag{4.1}$$

where $\xi(r)$ is known as the *two-point correlation function*. The two-point correlation function relates directly to the power spectrum, $P(k)$, through the Fourier transform,

$$\xi(r) = \frac{1}{2\pi} \int dk k^2 P(k) \frac{\sin(kr)}{kr} \tag{4.2}$$

where the scale of the function, λ, is related to the wave number k by $k = 2\pi/\lambda$. The correlation function can be used to describe the density of fluctuations of sources by

$$\xi(r) = \langle \delta(x)\delta(x+r) \rangle \quad , \tag{4.3}$$

where $\delta(x) = (\rho - \bar{\rho})/\rho$ is the density contrast, relative to the mean value of $\bar{\rho}$, at position x. The naive computation scaling of the n-point correlation function, where we evaluate all permutations of points is $\mathcal{O}(N^n)$, with N the size of the data and n the order of the correlation function. The computational expense of this operation can become prohibitive for large samples of points.

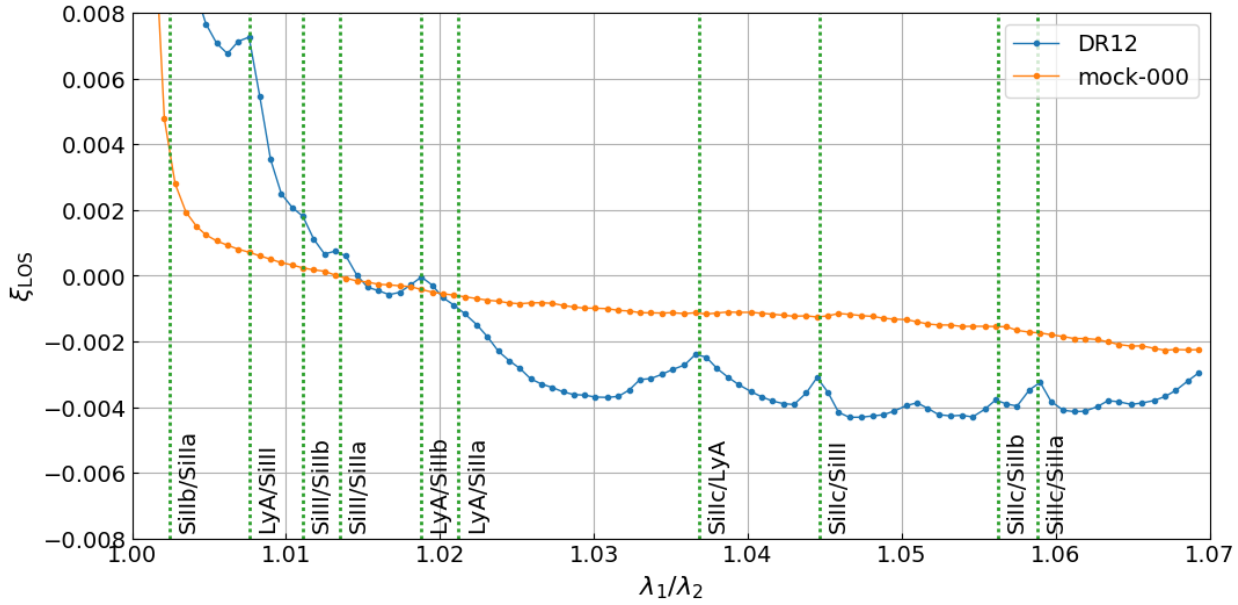

Figure 4.1: Line of sight correlation function of $\delta_F(\lambda_r)$. The correlation function is binned as a function of wavelength ratio.

4.2 Binning coordinates

We seek to measure:

$$\xi(r, \mu, z) \equiv \langle \delta(\mathbf{s}_1)\delta(\mathbf{s}_2)\rangle - \langle \delta(\mathbf{s}_1)\rangle\langle \delta(\mathbf{s}_2)\rangle \tag{4.4}$$

where

$$|\mathbf{s}_1 - \mathbf{s}_2| = r \quad , \quad |\mathbf{s}_1| - |\mathbf{s}_2| = \mu r \quad , \quad \frac{1}{2}|\mathbf{s}_1 + \mathbf{s}_2| = c\int_0^z \frac{dz'}{H_{\rm fid}(z')} \tag{4.5}$$

We estimate the correlation function in bins of

$$\xi_{ijk} = \xi(r_{ijk}, \mu_{ijk}, z_k) \tag{4.6}$$

where

$$r_{ijk} = \sqrt{r_{\parallel}^2(\Delta v_i, z_k) + r_{\perp}^2(\Delta\theta_j, z_k)} \tag{4.7}$$

and

$$\mu_{ijk} = r_{\parallel}(\Delta v_i, z_k)/r_{ijk} \tag{4.8}$$

A pair of pixels (i, j) with line of sight separation angle $\Delta\theta_{ij}$, line of sight separation $\Delta v_{ij} = c \log \lambda_j/\lambda_i$ and redshift $z_{ij} = \frac{\sqrt{\lambda_i \lambda_j}}{\lambda_\alpha} - 1$. Where $\lambda_\alpha \approx 1216$ Å is the wavelength of light associated with the energy of the $n = 2$ to $n = 1$ transition of neutral Hydrogen.

There are three sets of binning coordinates: polar (r, μ, z), Cartesian $(r_{\parallel}, r_{\perp}, z)$, and observing $(\Delta \log \lambda, \Delta\theta, \lambda)$. A fiducial cosmology must be chosen to translate observing coordinates into polar or Cartesian. We assume a homogeneous ΛCDM universe with $\Omega_\Lambda = 0.73$ and $\Omega_M = 1 - \Omega_\Lambda$. We compute the co-moving distance for all observed pixels in the Lyman alpha forest assuming this cosmology.

4.3 Sky bins

In order to facilitate computation of the correlation function over a finite volume, we assign observations into small sky bins. A natural binning scheme is to use the observations plates for this, however, these are of fixed size and are overlapping. We additionally employ HEALPix, a technique commonly used in the CMB community. Now instead of searching for neighboring sightlines over the entire survey, we only have to look in neighboring sky bins. This significantly reduces the computational complexity of the correlation function estimate. We calculate the minimum redshift of our survey volume as $z_{\min} = 10\,(\log \lambda_{\min} - \log \lambda_{\mathrm{Ly}\alpha}) - 1$ and calculate the transverse co-moving distance, $D_{M,\max}$ at this redshift for the appropriate

binning scheme. For observing coordinates, this is specified. For Cartesian this corresponds to $\theta_{\max} = r_{\perp,\max}/D_{M,\max}$. And for polar coordinates, $\theta_{\max} = r_{\max}\sqrt{1-\mu^2_{\min}}/D_{M,\max}$. For N objects, there are N^2 possible pairs. If we exclude cases where $i == j$ then we have $N(N-1)$ and not repeated pairs $i,j == j,i$ then we have $N(N-1)/2$ distinct pairs. For example, using just the observations from the southern galactic cap, we have 39,118 lines of sight. There are 6,759,094 total analysis pixels yielding an average of 172.787 per LOS. There are 765,089,403 distinct LOS Pairs and 22,842,672,470,871 distinct pixel pairs. Whole sky, 157,843 sight lines. 12,457,127,403 distinct LOS pairs, 74,395,900 considered (0.6%), 36,861,872 used (50%). Distinct pixel pairs 370,632,343,404,576, considered 1,095,587,702,034 (0.3%), used 281,469,767,392 (25%). Runtime is 183m10.299s on laptop.

We can exploit the spatial structure in the observations to trim this down.

4.3.1 HEALPix

We do this by splitting the sky into equal area quadrilaterals using HEALPix. Then, for a given quasar, we can restrict our comparison to the nearby HEALPix bins. Check out Figure 4.2 for a killer example. We set the distance of HEALPix bins to include equal to our maximum scale ($r = 200$ Mpc/h) at the minimum redshift of our δ-field ($z \approx 1.96$) where the transverse comoving scale is ~ 3681.2 Mpc/h. This scale is $0.0543301 radians$. Using this technique, we cut our comparisons done by $\sim 98\%$. There is still a small amount of related to the HEALPix bin size, in this case, only $\sim 50\%$ of the LOSes are actually within 200 Mpc/h. This similarly cuts down the number of pixels that we have to compare. There is another overhead in pixel comparison because our quasars are at various redshifts. At the end of the day, $\sim 0.28\%$ of distinct pixel pairs are used to estimate $\xi(r)$. For a pair of pixels, we check that they are within the binning grid. The first check is on the separation along the line of sight. The check is on the

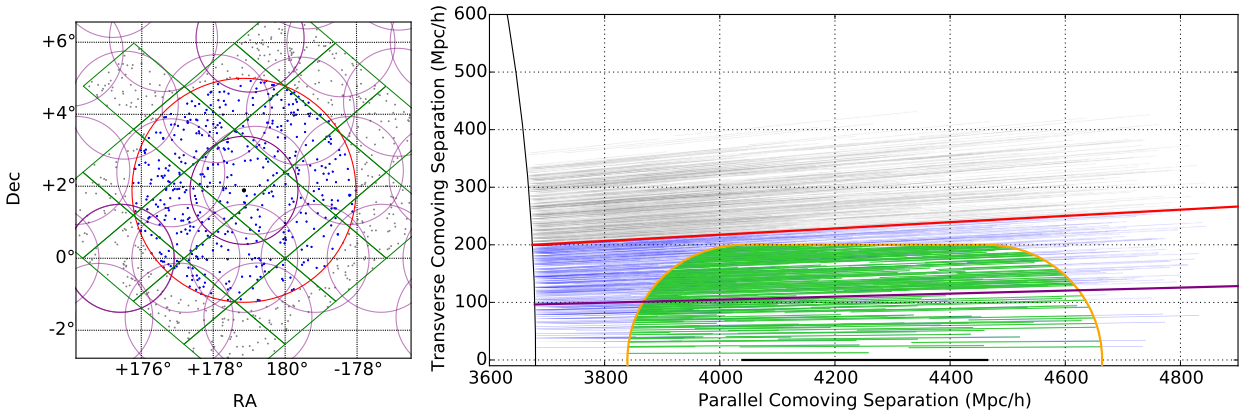

Figure 4.2: Example sight line distribution in small patch of sky. The left panel shows the 2D distribution of sight lines in a region of sky around a particular line of sight (black dot). The green quadrilaterals correspond to HEALPix boundaries. The purple circles correspond to the plate outline of BOSS observations. The blue (gray) points indicate neighbors in nearby sky pixels inside (outside) the maximum angular scale, which is indicated by the red circle. The right panel shows the sight line distribution in the parallel and transverse directions relative to a single line of sight (black line). The segments of sightlines within 200 Mpc/h of the example sight line are colored green.

4.4 Covariance matrix

In order to fit for the BAO peak position, we must also estimate the covariance matrix of the binned correlation function. We can estimate the covariance matrix $C(A, B)$ via sub-sampling by re-using the correlation function estimates in separate sky regions described above. Another method used by the BOSS Lyman-*alpha* working group is a Rick expansion of the four-point function of the δ_F field.

4.4.1 Sub-sampling

We estimate the covariance of the correlation function

$$\langle \xi_{ijk} \xi_{i'j'k'} \rangle - \langle \xi_{ijk} \rangle \langle \xi_{i'j'k'} \rangle \tag{4.9}$$

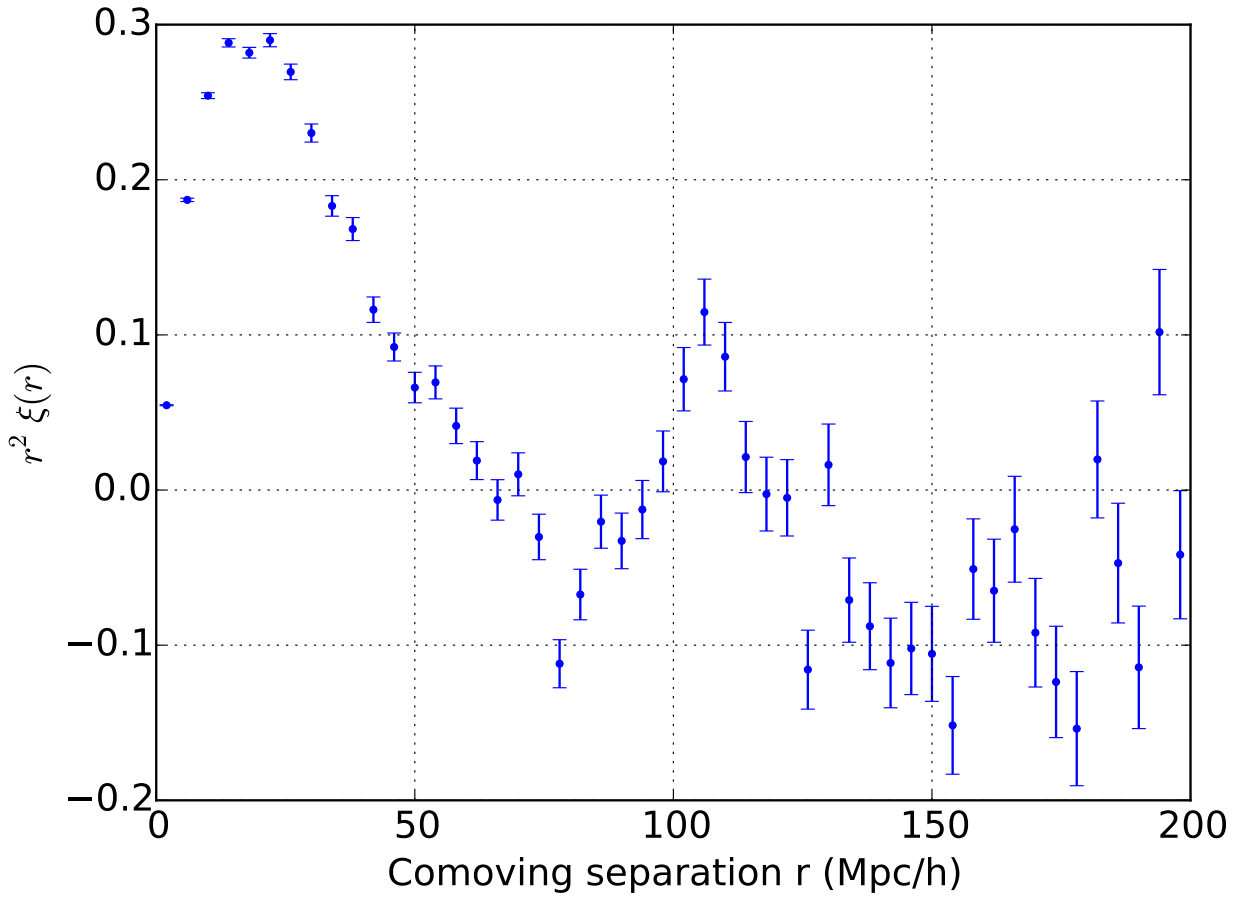

Figure 4.3: 1D correlation function.

via sub-sampling by rewriting this as

$$C_{AB} = \frac{1}{W_A W_B} \sum_{ij \in A} \sum_{kl \in B} w_i w_j w_k w_l \left(\langle \hat{\delta}_i \hat{\delta}_j \hat{\delta}_k \hat{\delta}_l \rangle - \langle \hat{\delta}_i \hat{\delta}_j \rangle \langle \hat{\delta}_k \hat{\delta}_l \rangle \right) \tag{4.10}$$

$$W_A = \sum_{ij \in A} w_i w_j, \quad \textit{and}, \quad W_B = \sum_{kl \in B} w_k w_j \tag{4.11}$$

Estimate C_{AB} by dividing footprint into subsamples ξ^s_A, ξ^s_B

$$C_{AB} = \frac{1}{W_A W_B} \sum_s W_A W_B \left(\hat{\xi}^s_A \hat{\xi}^s_B \hat{\xi}_A - \hat{\xi}_B \right) \tag{4.12}$$

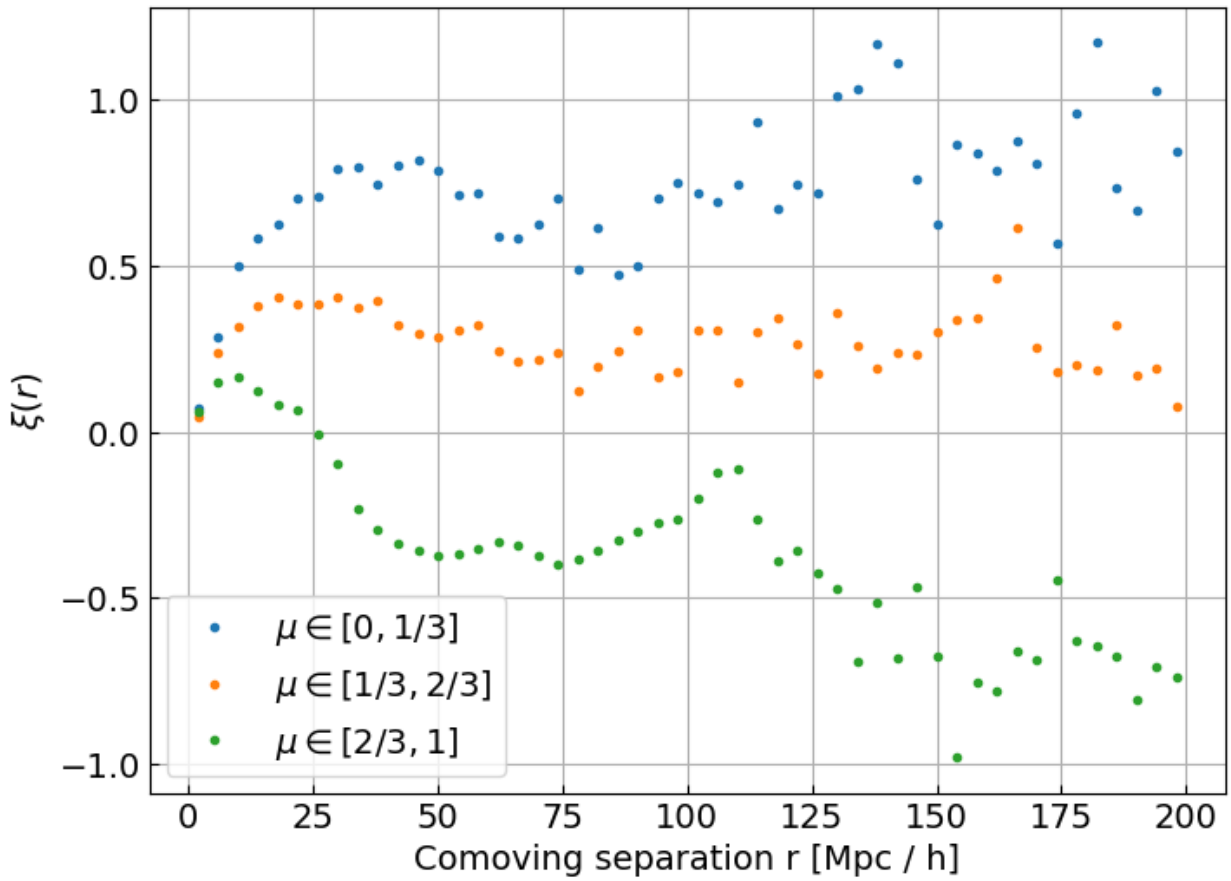

Figure 4.4: Correlation function estimated in (r, μ) bins.

where W_A^s are the sums of weights in subsample s

neglect inter-forest sub-samples:

$$C_{AB} = \frac{1}{W_A W_B} \sum_{ij \in A} \sum_{kl \in B} w_i w_j w_k w_l \xi_{1d}(\lambda_i/\lambda_k) \xi_{1d}(\lambda_j/\lambda_l) \tag{4.13}$$

4.4.2 Smoothing

Smooth the covariance matrix with the assumption that the correlation coefficient,

$$c(A, B) = \frac{C(A, B)}{\sqrt{C(A, A) C(B, B)}} \quad , \tag{4.14}$$

depends only on the difference of separations of bins A and B, on $\Delta r_\parallel = r_\parallel - r'_\parallel$ and $\Delta r_\perp = r_\perp - r'_\perp$. Therefore, correlation coefficients with the same $\Delta r_\parallel$ and $\Delta r_\perp$ are averaged and assigned to each element of $c(A, B)$.

Elements close to zero are clipped to zero to avoid numerical representation issues.

The final smoothed covariance matrix is given by:

$$\tilde{C}(A, B) = \bar{c}(\Delta r_\parallel, \Delta r_\perp)\sqrt{C(A, A)C(B, B)} \quad . \tag{4.15}$$

These assumptions as justified and validated in xyz which show estimates of the covariance via this procedure accurate at the few percent level.

4.5 Summary

In this section I have described how to estimate the transmitted flux density field's correlation function the corresponding covariance. using three binning techniques and two sub sampling methods. In the next chapter, I will describe how the BAO peak position is measured from the estimated correlation function and covariance.

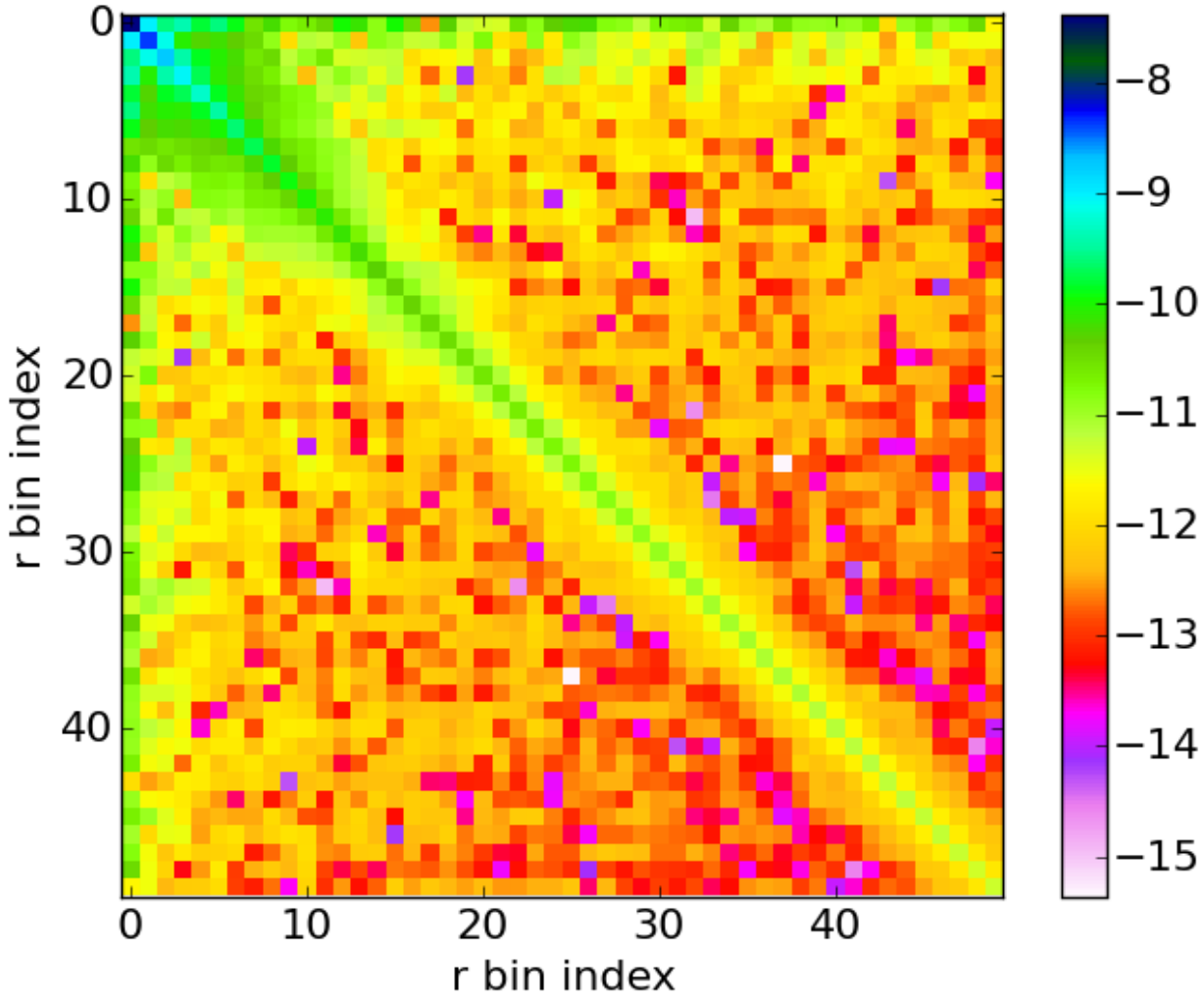

Figure 4.5: Covariance matrix of the 1D correlation function.

Chapter 5

Cosmological Fits

In this chapter I describe the fitting method procedure used to measure the BAO scale from the correlation function estimate described in the previous chapter. Much of this is described in more detail in Kirkby et al. [2013]. We have a correlation function estimate $\xi(r_\parallel, r_\perp)$, now fit a cosmological model to determine $\alpha_\parallel$ and $\alpha_\perp$. Summarize method for fitting correlation function.

5.1 Models and Parameters

We model a measurement of the correlation function

$$\xi(r, \mu, z) \equiv \langle \delta(\mathbf{s}_1)\delta(\mathbf{s}_2)\rangle - \langle \delta(\mathbf{s}_1)\rangle\langle \delta(\mathbf{s}_2)\rangle \tag{5.1}$$

where the ensemble averages are taken over realizations of a (possibly biased) tracer $\delta(\mathbf{s})$ of the large-scale distribution of matter in redshift space $\mathbf{s}$, with (r, μ, z) defined via[1]

$$|\mathbf{s}_2 - \mathbf{s}_1| = r \quad , \quad |\mathbf{s}_2| - |\mathbf{s}_1| = \mu \cdot r \quad , \quad \frac{1}{2}|\mathbf{s}_2 - \mathbf{s}_1| = c \int_0^z \frac{dz'}{H_{\rm fid}(z')} \tag{5.2}$$

for some fiducial cosmology with Hubble function $H_{\rm fid}(z)$. The model combines a cosmological prediction $\xi_{\rm cosmo}$ with a parametrization of possible multiplicative and additive broadband distortions introduced by the analysis method.

5.1.1 Physical Coordinates

The physical coordinates for a pair of pixels (i, j) measured in the absorption spectra of two quasars are the separation angle $\Delta\theta_{ij}$ between the quasar lines of sight ($\Delta\theta_{ij} = 0$ if the pixels are taken from the same quasar's spectrum) and the observed absorption wavelengths λ_i and λ_j. We convert the observed wavelengths to a relative velocity for the absorption systems[2]

$$\Delta v_{ij} = c \log(\lambda_j / \lambda_i) \tag{5.3}$$

and an average absorption redshift

$$z_{ij} = \frac{\sqrt{\lambda_i \lambda_j}}{\lambda_\alpha} - 1 \tag{5.4}$$

where $\lambda_\alpha \simeq 1216$ Å is the rest wavelength of the Lyman$-\alpha$ transition which determines pixel redshifts $z_i = \lambda_i / \lambda_\alpha - 1$. We calculate the corresponding co-moving separations along $(r_\parallel)$

[1] We use the notation $\mu \equiv \hat{\mathbf{z}} \cdot \hat{\mathbf{r}}$ and $\mu_k \equiv \hat{\mathbf{z}} \cdot \hat{\mathbf{k}}$

[2] This definition is not identical to the Doppler velocity that an observer at one absorber would measure for the other absorber, but does agree to second order in the wavelength ratio

and perpendicular to ($r_\perp$) the line of sight as

$$r_\parallel = D_{C,\mathrm{fid}}(z_j) - D_{C,\mathrm{fid}}(z_i) = \frac{1+z_{ij}}{H_{\mathrm{fid}}(z_{ij})} \cdot \left[1 + \mathcal{O}(\Delta v_{ij}/c)^2\right] \tag{5.5}$$

$$r_\perp = D_{A,\mathrm{fid}}(z_{ij}) \cdot \Delta\theta_{ij} = c \int_0^{z_{ij}} \frac{dz'}{H_{\mathrm{fid}}(z')} \cdot \Delta\theta_{ij} \tag{5.6}$$

where $D_{C,\mathrm{fid}}(z)$ and $D_{A,\mathrm{fid}}(z)$ are the co-moving line of sight and angular distance functions, respectively, for the assumed fiducial cosmology $H_{\mathrm{fid}}(z)$.

5.1.2 Cosmological Models

We build the cosmological model starting from an isotropic linear power spectrum predication $\tilde{P}(k, z_0)$ at some reference redshift z_0, then embed this prediction in redshift space (we use tildes to denote linear-theory predictions without any redshift space distortions). In the general case of plane-parallel redshift-space distortion $r \to (r, \mu)$ we have:

$$\xi_{\mathrm{cosmo}}(r, \mu, z_0) = \sum_{\ell\,\mathrm{even}} L_\ell(\mu)\xi_{\ell,\mathrm{cosmo}}(r, z_0) \tag{5.7}$$

with

$$\xi_{\ell,\mathrm{cosmo}}(r, z_0) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) P_\ell(k, z_0) dk \tag{5.8}$$

where L_ℓ is the Legendre polynomial, j_ℓ is the spherical Bessel function, and $P_\ell(k, z_0)$ are the multipoles of the redshift-distorted power spectrum $P(k, \mu_k, z_0)$ with $\mu_k \equiv \hat{\mathbf{z}} \cdot \hat{\mathbf{k}}$:

$$P_\ell(k, z_0) = \frac{2\ell+1}{2} \int_{-1}^{+1} P(k, \mu_k, z_0) L_\ell(\mu_k) d\mu_k \quad . \tag{5.9}$$

Specializing to linear theory and the distant observer approximation, the infinite series of $\xi_{\rm cosmo}(r,\mu,z_0)$ is truncated at $\ell=4$, with

$$\tilde{P}_\ell(k,z_0) = b^2(z_0) C_\ell(\beta(z_0)) \tilde{P}(k,z_0) \tag{5.10}$$

and

$$C_\ell(\beta) \equiv \frac{2\ell+1}{2}\int_{-1}^{+1} \left(1+\beta\mu_k^2\right)^2 L_\ell(\mu_k) d\mu_k = \begin{cases} 1+\frac{2}{3}\beta+\frac{1}{5}\beta^2 & \ell=0 \\ \frac{4}{3}\beta+\frac{4}{7}\beta^2 & \ell=2 \\ \frac{8}{35}\beta^2 & \ell=4 \end{cases} , \tag{5.11}$$

where $b(z)$ and $\beta(z)$ are the tracer bias and redshift-space distortion paramter at redshift z, respectively. We can therefore write

$$\xi_{\rm cosmo}(r,\mu,z_0) = b^2(z_0) \sum_{\ell=0,2,4} C_\ell(\beta(z_0)) L_\ell(\mu) \tilde{\xi}_{\ell,\rm cosmo}(r,z_0) \tag{5.12}$$

in terms of the undistorted linear-theory multipoles

$$\tilde{\xi}_{\ell,\rm cosmo}(r,z_0) = \frac{i^\ell}{2\pi^2}\int_0^\infty k^2 j_\ell(kr)\tilde{P}(k,z_0) dk \quad . \tag{5.13}$$

Non-linear Effects

The expected effects of non-linear structure growth on the BAO feature can be modeled with an anisotropic Gaussian roll-off of the linear power spectrum [Eisenstein et al., 2007]:

$$\tilde{P}_{\rm NL}(k,\mu_k,z_0) = \exp(-k^2\Sigma^2(\mu_k)/2 \cdot \tilde{P}(k,z_0)) \tag{5.14}$$

where

$$\Sigma^2(\mu_k) = \mu_k^2 \Sigma_\parallel^2 + (1+\mu_k^2)\Sigma_\perp^2 \quad . \tag{5.15}$$

Redshift Evolution

In general, we model the redshift evolution of a parameter $p(z)$ for z near z_0 in terms of two parameters p_0 and γ_p via

$$p(z) = p_0 \left(\frac{1+z}{1+z_0} \right)^{\gamma_p} \quad . \tag{5.16}$$

We apply this evolution to the parameters $b^2(z)$ and $\beta(z)$, introduced above, and to the BAO scale parameters $\alpha_{\rm iso}(z)$, $\alpha_\parallel(z)$, and $\alpha_\perp(z)$ introduced below. Given a covariance matrix for the parameters p_0 and γ_p,

$$C_p = \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_\gamma \\ \rho\sigma_0\sigma_\gamma & \sigma_\gamma^2 \end{pmatrix} \tag{5.17}$$

where

$$J = \left(\frac{\delta p(z)}{\delta p_0}, \frac{\delta p(z)}{\delta \gamma_p} \right) \tag{5.18}$$

is the Jacobian. The error on $p(z)$ is smallest at

$$\log\left(\frac{1+z}{1+z_0} \right) = -b - \rho a + \sqrt{b^2 - (1-\rho^2)a^2} \tag{5.19}$$

with

$$a \equiv \frac{\sigma_0}{\rho_0 \sigma_\gamma} \quad , \quad b \equiv \frac{1}{2\gamma_p} \quad . \tag{5.20}$$

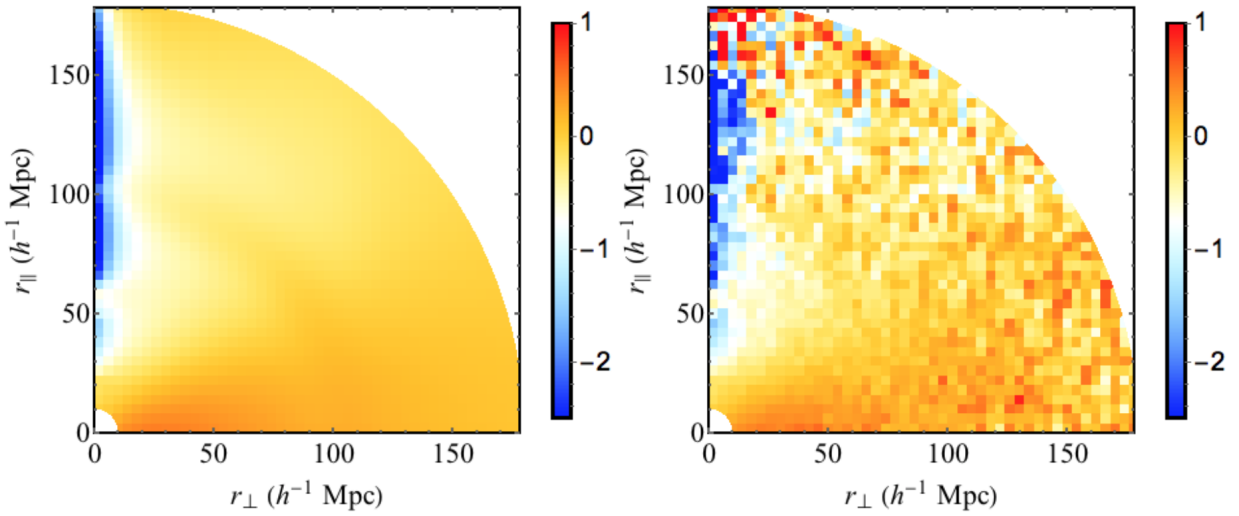

Figure 5.1: Two-dimensional representation of $r^2\xi(r_\perp, r_\parallel)$ in units of $(h^{-1}\text{Mpc})^2$. The right panel shows the measurement and the left panel shows the best-fit model. The BAO feature is at $r \sim 100h^{-}1\text{Mpc}$. The effects of metal-Lyα correlations are seen in the lowest $r_\perp$ bin, in particular the peak at $50 < r_\parallel < 70h^{-}1\text{Mpc}$ due to SiIIa and SiIIb.

5.1.3 DR12

The DR12 two-dimensional representation of $r^2\xi(r_\perp, r_\parallel)$ is shown in Figure 5.1. The right panel shows the measurement and the left panel shows the best-fit model. The BAO feature is at $r \sim 100h^{-}1\text{Mpc}$. The effects of metal-Lyα correlations are seen in the lowest $r_\perp$ bin, in particular the peak at $50 < r_\parallel < 70h^{-}1\text{Mpc}$ due to SiIIa and SiIIb.

5.2 Results

The two-dimensional $\Delta\chi^2$ surface for $\alpha_\parallel$ and $\alpha_\perp$ for DR12 and a mock catalog are presented in Figure 5.3 and Figure 5.4, respectively. Note the significant extended plateau in the $\alpha_\perp$ direction in Figure 5.3. This plateau resembles those seen using less sophisticated continuum fitting procedures explored by the BOSS LyaF working group [Delubac et al., 2015]. Figure 5.4, the surface seems to be somewhat more sharply peaked near the fiducial values

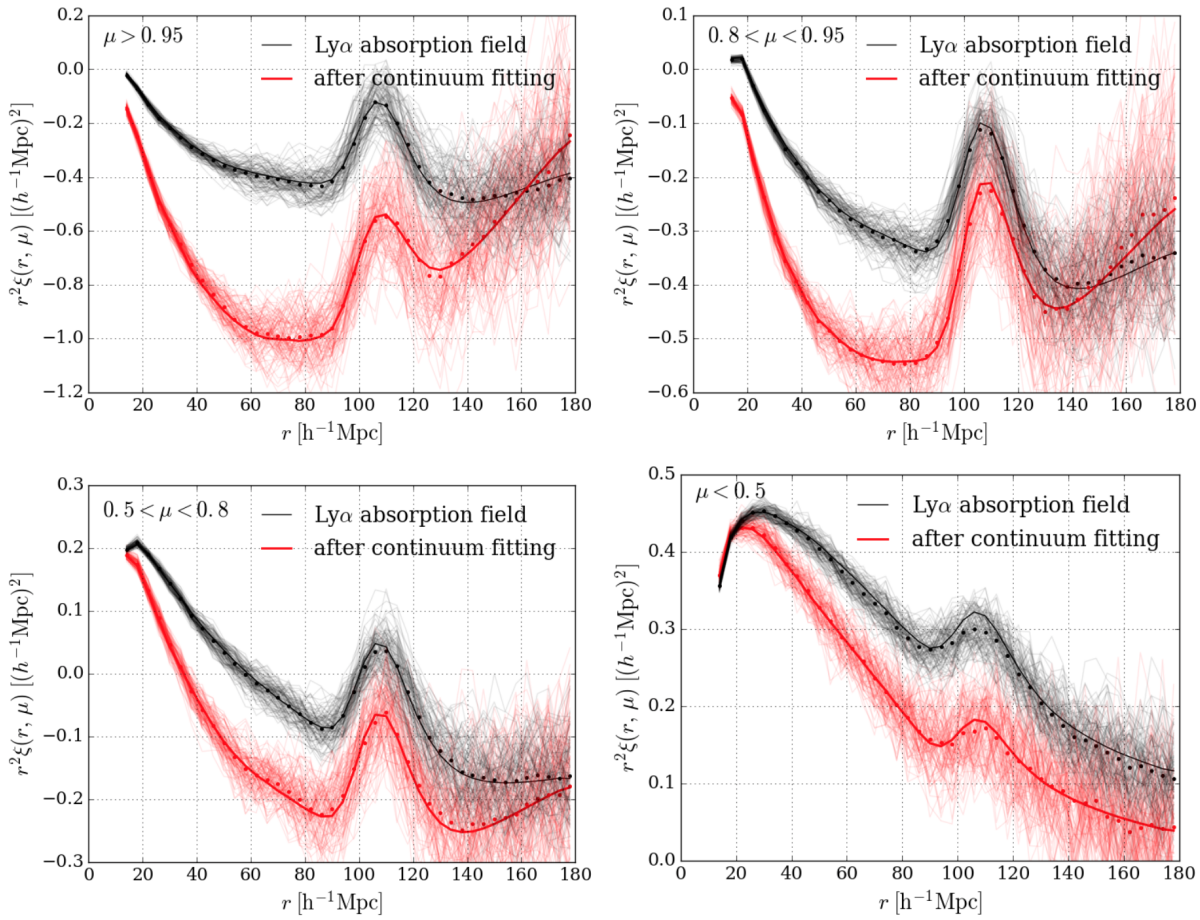

Figure 5.2: Correlation function for the metal-free mocks in four ranges of μ. The black curves correspond to mocks with Lyα absorption but without the addition of a quasar continuum. The red points and curves correspond to mocks with the addition of a continuum. The points correspond to stacks of 100 mocks and the light curves to individual mocks. The heavy curves correspond to the input model.

for $\alpha_\perp$ and $\alpha_\parallel$, although the faint blue region indicates a similar extended plateau as in Figure 5.3.

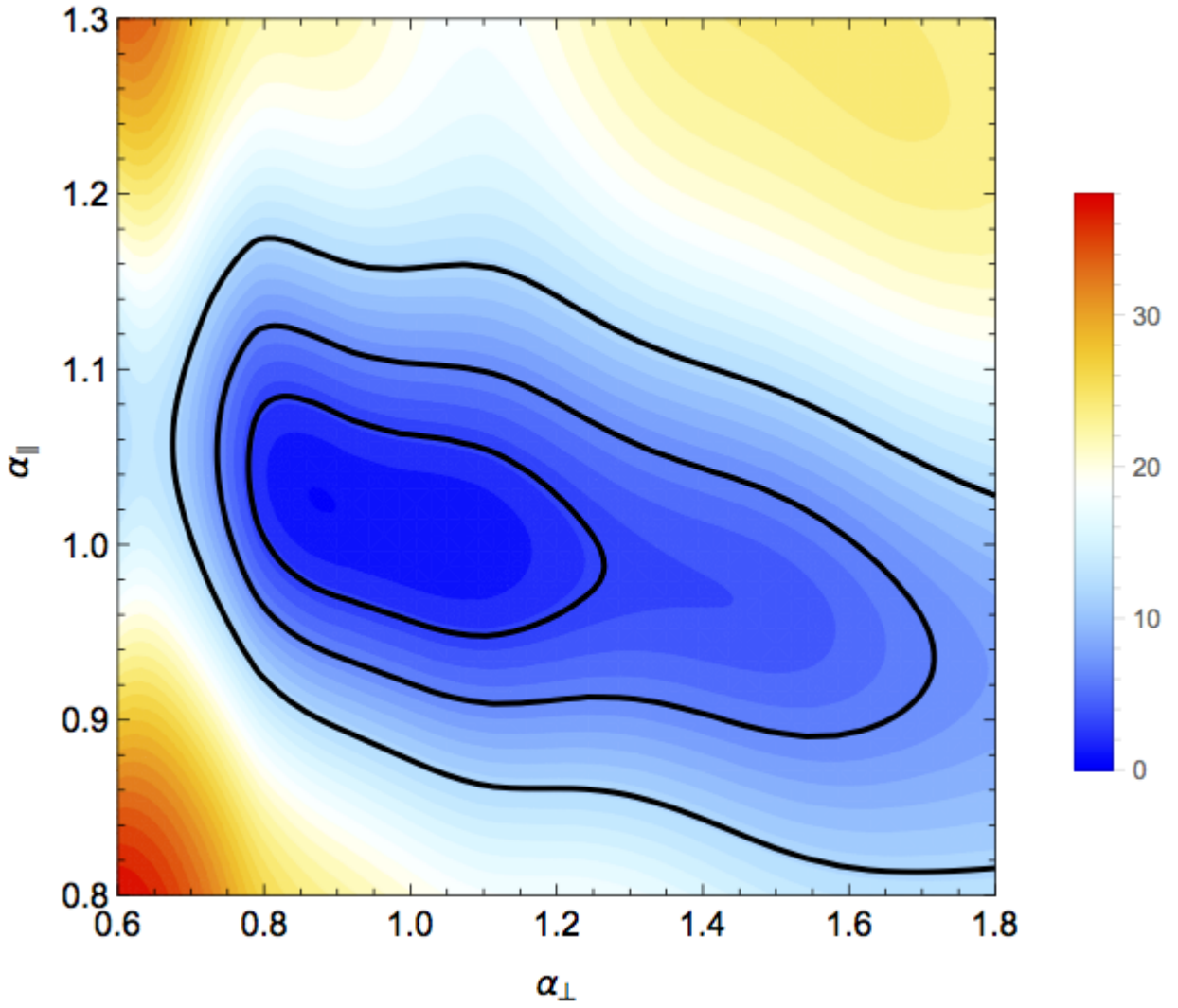

Figure 5.3: 2D $\Delta\chi^2$ surface of $\alpha_\perp$ and $\alpha_\parallel$ for DR12. The $\Delta\chi^2$ surface values are indicated by the colorbar on the right. Contours corresponding to 1-, 2-, and 3-σ are shown in black. Note the significant extended plateau in the $\alpha_\perp$ direction. This plateau resembles those seen using less sophisticated continuum fitting procedures explored by the BOSS LyaF working group [Delubac et al., 2015]

.

A comparison of the fitted $\alpha_\perp$ and $\alpha_\parallel$ on a set of ten mock catalogs is shown in Figure 5.5. The figure compares the results of the BOSS analysis on noiseless mocks, the BOSS analysis on full mocks (noise + metals), and the analysis described in this work on noisey mocks without metals. The spread is comparable between the various analyses, however a sample

of 10 points is too small to draw any firm conclusions.

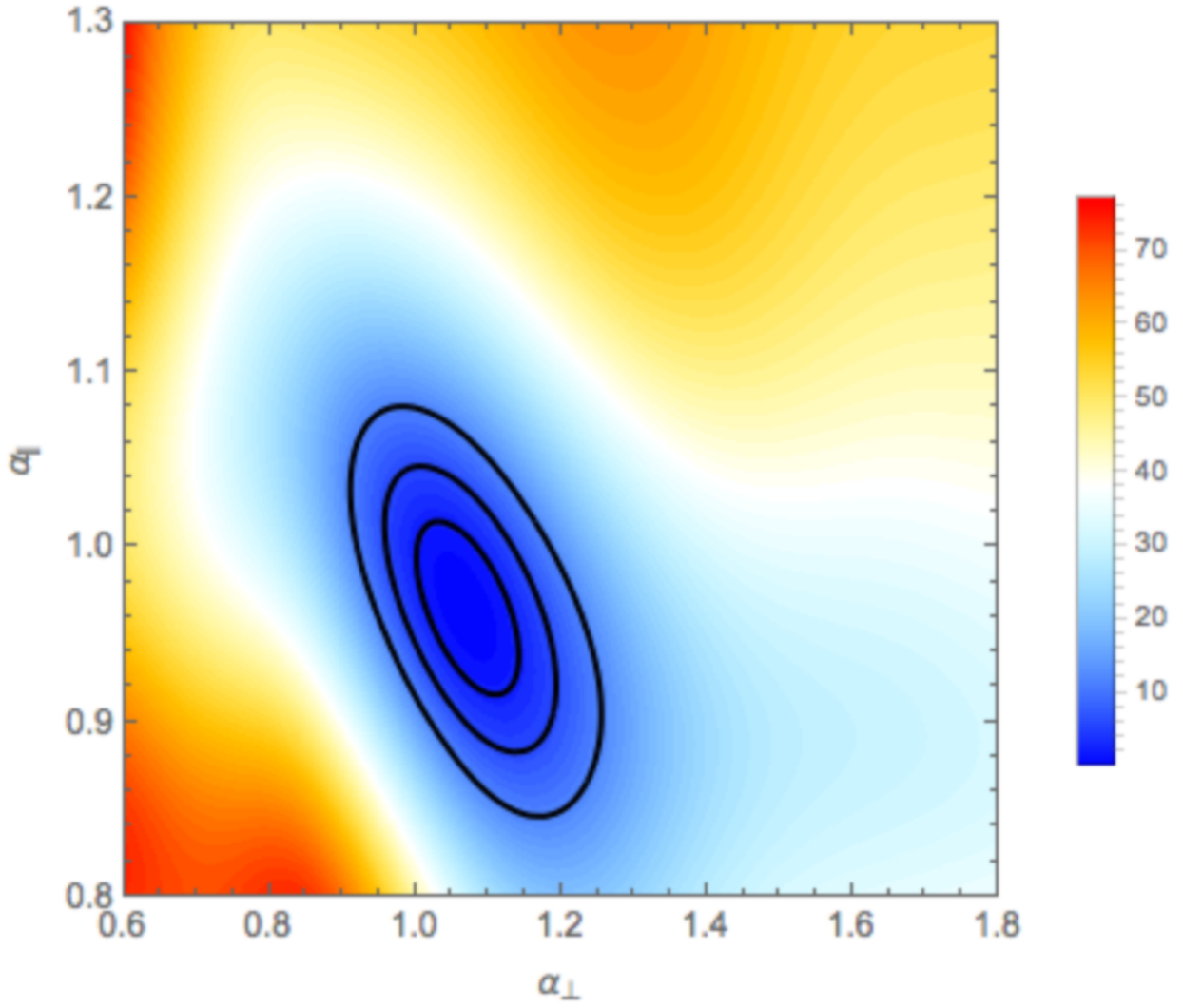

Figure 5.4: 2D $\Delta\chi^2$ surface of $\alpha_\perp$ and $\alpha_\parallel$ for mock-000. The $\Delta\chi^2$ surface values are indicated by the colorbar on the right. Contours corresponding to 1-, 2-, and 3-σ are shown in black. While, the surface seems to be somewhat more sharply peaked near the fiducial values for $\alpha_\perp$ and $\alpha_\parallel$, the faint blue region indicates a similar extended plateau as in Figure 5.3.

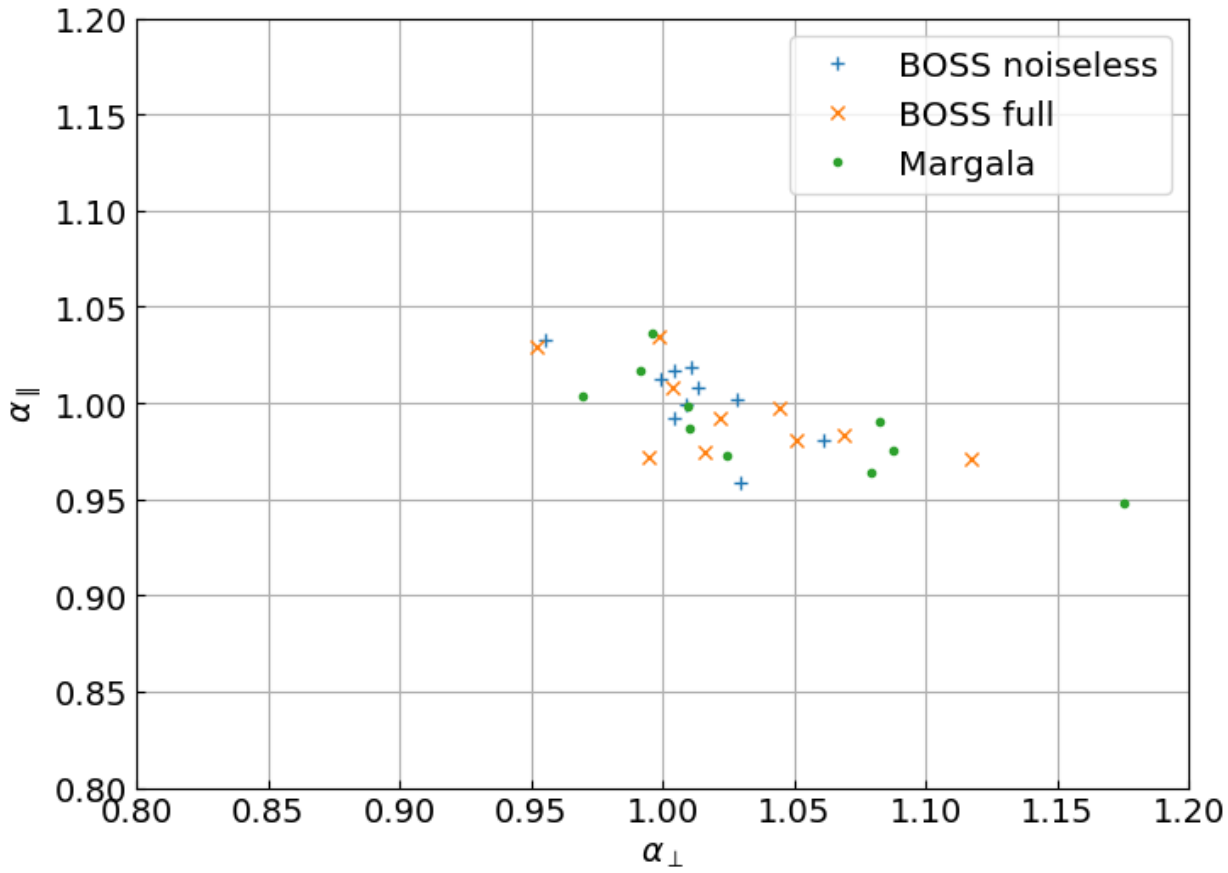

Figure 5.5: Measured $\alpha_\perp$ and $\alpha_\parallel$ for 10 mock catalogs. The blue pluses, orange cross, and green circles indicate the fitted $\alpha_\perp$ and $\alpha_\parallel$ on 10 mock catalogs using the BOSS analysis on noiseless mocks, the BOSS analysis on full mocks (noise + metals), and the analysis described in this work on noisey mocks without metals. The spread is comparable between the various analyses, however a sample of 10 points is too small to draw any firm conclusions.

Chapter 6

Cosmological implications of baryon acoustic oscillation measurements

In this chapter I discuss the cosmological implications of baryon acoustic oscillation measurements. First I relate the measurements of $\alpha_{\parallel}$ and $\alpha_{\perp}$ back to the standard ruler r_d. Next, I summarize BAO scale measurements form the BOSS galaxy and LyaF surveys, focusing on the combination of the LyaF auto-correlation and LyaF-QSO cross-correlation measurements. Finally, I discuss constraints on cosmological models using Planck MCMC chains.

6.1 From alpha to distance measures

As discussed in Chapter 1, the baryon acoustic oscillation (BAO) scale is computed from first principles, using physical parameters that are well constrained by cosmic microwave background (CMB) data. BAO measurements in the line-of-sight dimension allow direct determination of the expansion rate $H(z)$. The comoving angular diameter distance $D_M \propto$

$\int_0^z cH^{-1}(z)dz$ is also constrained from transverse clustering.

The BAO scale is set by the radius of the sound horizon at the drag epoch z_d when photons and baryons decouple,

$$r_d = \int_{z_d}^{\infty} \frac{c_s(z)}{H(z)} dz \tag{6.1}$$

For reasonable variations relative to a fiducial model, the *ratio* of BAO scales is given accurately by the ratio of r_d values from the equation above. Thus, a measurement of $\alpha_\perp$ from clustering at redshift z constrains the ratio of the comoving angular diameter distance to the sound horizon:

$$D_M(z)/r_d = \alpha_\perp D_{M,\mathrm{fid}}(z)/r_{d,\mathrm{fid}} \quad . \tag{6.2}$$

A measurement of $\alpha_\parallel$ constrains the Hubble parameter $H(z)$ and the analogous quantity:

$$D_H(z) = c/H(z) \quad , \tag{6.3}$$

with

$$D_H(z)/r_d = \alpha_\parallel D_{H,\mathrm{fid}}(z)/r_{d,\mathrm{fid}} \quad . \tag{6.4}$$

An isotropic BAO analysis measures some effective combination of these two distances. If redshift-space distortions are weak, then the constrained quantity is the volume average distance

$$D_V(z) = \left[z D_H(z) D_M^2(z) \right]^{1/3} \quad , \tag{6.5}$$

Name	Redshift	D_V/r_d	D_M/r_d	D_H/r_d
6dFGS	0.106	3.047 ± 0.137	-	-
MGS	0.15	4.480 ± 0.168	-	-
BOSS LOWZ Sample (DR12)	0.32	8.611 ± 0.095	8.774 ± 0.142	25.926 ± 0.761
BOSS CMASS Sample (DR12)	0.57	13.784 ± 0.142	14.774 ± 0.237	21.042 ± 0.523
Margala auto-correlation (DR12)	2.33	-	38.673 ± 6.915	8.870 ± 0.392
LyaF auto-correlation (DR12)	2.33	-	37.770 ± 2.130	9.070 ± 0.310
LyaF-QSO cross-correlation (DR11)	2.36	-	36.288 ± 1.344	9.000 ± 0.300
Combined LyaF (DR11)	2.34	-	36.489 ± 1.152	9.145 ± 0.204

Table 6.1: Summary of BAO constraints. These values are taken from Jones et al. [2009] (6dFGS), Beutler et al. [2011] (MGS), Alam et al. [2017] (BOSS galaxies), Bautista et al. [2017] (BOSS LyaF auto-correlation) and Font-Ribera et al. [2014] (BOSS LyaF cross-correlation). Note that the BOSS galaxy and LyaF auto-correlation results are based on DR12, however, the LyaF-QSO cross-correlation and combined LyaF results are from DR11.

with

$$D_V(z)/r_d = \alpha D_{V,\mathrm{fid}}(z)/r_{d,\mathrm{fid}} \quad . \tag{6.6}$$

6.2 BAO constraints

A summary of BAO measurements from various surveys is presented in Table 6.1.

6.2.1 LyaF auto-correlation

The BAO scale was first measured at higher redshift ($z \sim 2.4$) from the auto-correlation of the Lyα forest fluctuations in the spectra of high-redshift quasars from BOSS DR9 [Busca et al., 2013, Kirkby et al., 2013, Slosar et al., 2013] following the pioneering work of measuring 3D fluctuations in the forest [Slosar et al., 2011]. Improved measurements using roughly twice as many quasar spectra from BOSS DR11 are presented in Delubac et al. [2015] and the full BOSS DR12 analysis is presented in Bautista et al. [2017].

6.2.2 LyaF-QSO cross-correlation

BAO have also been measured in the cross-correlation of the Lyα forest with the density of quasars in BOSS DR11 [Font-Ribera et al., 2014]. While the number of quasar-pixel pairs is much lower than the number of pixel-pixel pairs in the auto-correlation function, the clustering signal is much stronger because of the high bias factor of quasars.

6.2.3 Joint LyaF auto and cross correlation

Combine the LyaF auto-correlation and LyaF-QSO cross-correlation measurements. Even though these results are derived from the same volume, we can consider them as independent because their uncertainties are not dominated by cosmic variance. They are dominated instead by the combination of noise in the spectra and sparse sampling of the structure in the survey volume, both of which affect the auto-correlation and cross-correlation almost independently. A number of tests using mock catalogs and several analysis procedures are presented in Delubac et al. [2015], finding good agreement between error estimates from the likelihood function and from the variance in mock catalogs. This independence allows us to add the χ^2 surfaces from both publications, which are publicly available at `http://darkmatter.ps.uci.edu/baofit/`.

6.3 Visualizing BAO constraints

Figure 6.1 shows $H(z)/(1+z)$, which is the proper velocity between two objects with a constant comoving separation of 1 Mpc. This quantity is declining in a decelerating universe and increasing in an accelerating universe. The x-axis is set to be $\sqrt{1+z}$, which makes $H(z)/(1+z)$ a straight line of slope H_0 in an Einstein-de-Sitter ($\Omega_m = 1$) model. Transverse

BAO measurements are shown in Figure 6.2. These are plotted as $c\ln(1+z)/D_M(z)$ so that a constant (horizontal) line in the $H(z)/(1+z)$ plot would produce the same constant line in this panel, assuming a flat universe. This quantity would decrease monotonically in a non-accelerating flat cosmology.

The quantities in both Figure 6.1 and Figure 6.2 approach H_0 as z approaches zero, independent of other cosmological parameters. Note that the $H(z)$ and $D_M(z)$ measurements from a given data set at a particular redshift are covariant. In these figures, the points are anti-correlated. For example, if $H(z)$ at $z = 2.34$ were scattered upward by a statistical fluctuation, then the corresponding $z = 2.34$ point in Figure 6.2 would, labeled "Margala", be scattered downward. Note that the uncertainty in measurement of $D_M(z)$ presented in this work is much larger than published BOSS result. This is not surprising as there are several differences between the methods and the BOSS analysis significantly more mature.

Both figures show predictions for a flat ΛCDM model with $\Omega_m = 0.3$ and $H_0 = 67.3 \pm 1.1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ as determined from a combination of SNIa and BAO results from Aubourg et al. [2015]. It is also interesting to note that the value of $D_M(1090)$ from the angular scale in the CMB, transformed to $c\ln(1+z)/D_M(z) = 151\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ at $z = 1090$ with percent level accuracy, is a factor of two larger than any of the low-redshift values in Figure 6.2.

6.4 Constraints on Dark Energy Models

The combination of BAO measurements with precise CMB measurements from the Plank and WMAP satellites yields tight constraints on the parameters of the ΛCDM cosmological model and on one-parameter extensions of this model. Figure 6.3 and Figure 6.4 compare the predictions of models constrained by CMB data to the BOSS BAO constraints on D_M and D_H at $z = 0.57$ and $z = 2.34$, from CMASS galaxies and the LyaF. Figure 6.3 shows

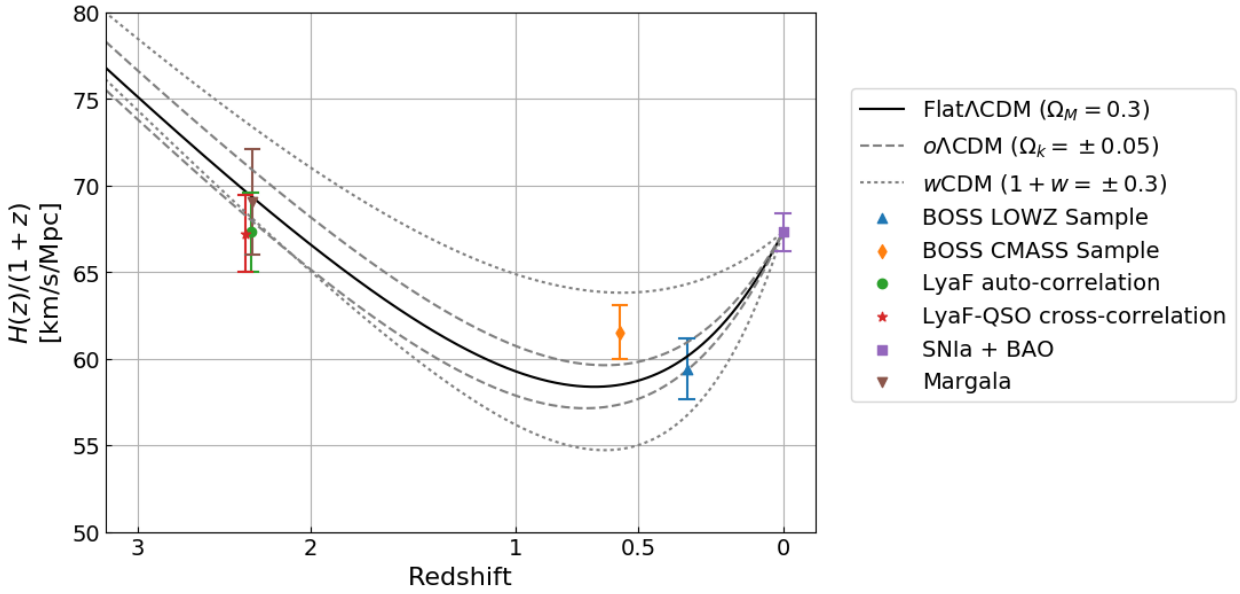

Figure 6.1: BAO measurements and model predictions of $H(z)$ as a function of redshift, with physically informative scalings. The quantity plotted is $H(z)/(1+z)$, the proper velocity between objects 1 comoving Mpc apart. The x-axis is set to $\sqrt{1+z}$ so that a pure matter universe ($\Omega_m = 1$) appears as a decreasing straight line

results for $o\Lambda$CDM models (`base_omegak/planck_lowl_lowLike_highL`), which assume a constant dark energy density but allow non-zero space curvature. Each point in the left and middle panels represents a model from the Planck Collaboration MCMC chains, based on a combination of Planck, WMAP polarization, and ACT/SPT data. The right panel shows the one-dimensional PDF for the curvature parameter Ω_k and the equation of state parameter w in Figure 6.3 and Figure 6.4, respectively. Figure 6.4 shows results for wCDM models (`base_w/planck_lowl_lowLike_highL`), which assume a flat universe but allow a constant equation-of-state parameter $w \equiv p/\rho \neq -1$ for dark energy. The CMB data alone are consistent with a wide range of w values, and they are generally better fit with $w < -1$. However, the combination with CMASS BAO data sharply limits the acceptable range of w, favoring values close to -1.0 (a cosmological constant). Improving the fit to LyaF BAO results by going to $w < -1.3$ is in opposition with CMASS measurements.

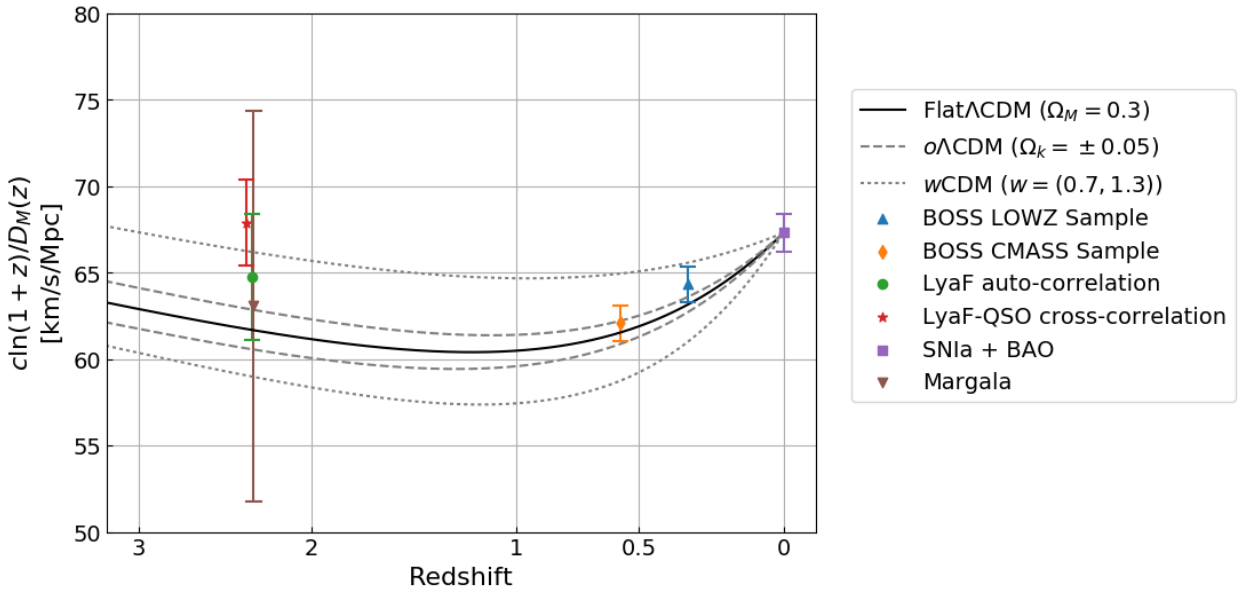

Figure 6.2: BAO measurements and model predictions of $D_M(z)$ as a function of redshift, with physically informative scalings. The quantity plotted is $c\ln(1+z)/D_M(z)$, a scaling that matches a constant line $H(Z) = (1+z)H_0$ in Figure 6.1 to the same constant line in this figure for a flat universe. The x-axis is set to $\sqrt{1+z}$ as in Figure 6.1.

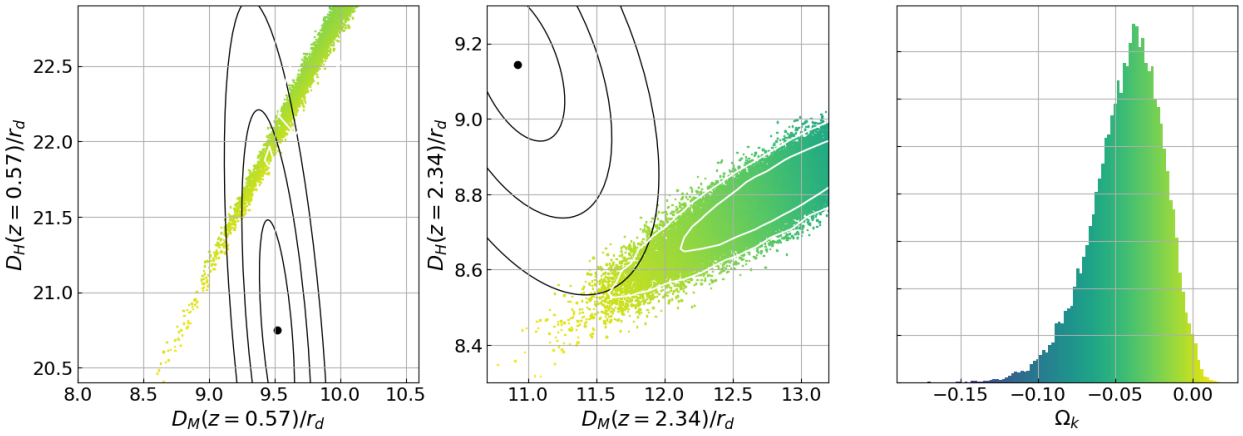

Figure 6.3: BAO constraints in the $D_M - D_H$ planes at $z = 0.57$ (left) and $z = 2.34$ (middle) compared to predictions of $o\Lambda$CDM constrained by CMB data. Black curves show 68%, 95%, and 99.7% likelihood contours from the CMASS and LyaF BAO measurements, relative to the best-fit values (black dots). Colored points represent individual models from Planck+WP+ACT/SPT MCMC chains, which are color-coded by the value of Ω_k as illustrated in the right panel. White curves show 68% and 95% likelihood contours for the CMB data alone

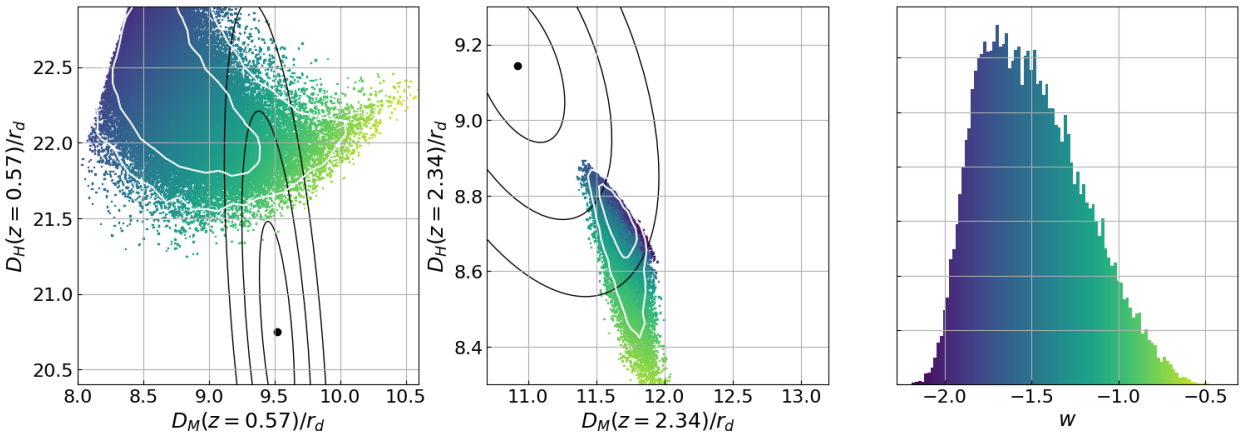

Figure 6.4: BAO constraints in the $D_M - D_H$ planes at $z = 0.57$ (left) and $z = 2.34$ (middle) compared to predictions of $w\Lambda$CDM constrained by CMB data. Black curves show 68%, 95%, and 99.7% likelihood contours from the CMASS and LyaF BAO measurements, relative to the best-fit values (black dots). Colored points represent individual models from Planck+WP+ACT/SPT MCMC chains, which are color-coded by the value of w as illustrated in the right panel. White curves show 68% and 95% likelihood contours for the CMB data alone

Chapter 7

Conclusion

As discussed in Chapter 1, observations of the baryon acoustic oscillation scale provide a unique probe for better understanding our universe. Combined with Cosmic Microwave Background data, BAO measurements yield powerful constraints on the cosmological models and the dark energy equation of state.

In Chapter 2, I presented corrections to spectrophotometric calibration process that provide a significant improvement over the uncorrected quasar spectra in DR12. Other improvements over the method described here include using a more realistic description for the PSF, using the actual guiding adjustments for determining centroid offsets, and applying corrections to individual exposures prior to co-addition. This is especially relevant for the extended Baryon Oscillation Spectroscopic Survey (eBOSS) [1] of SDSS-IV, which uses similar focal plane offsets for quasar targets. We are currently working with the eBOSS data processing team to implement the throughput corrections described for individual exposures into the data processing pipeline. The proposed Dark Energy Spectroscopic Instrument [DESI; Levi et al., 2013] includes an atmospheric dispersion compensator in its design so the effects of ADR should be much reduced for that survey.

[1] http://www.sdss.org/surveys/eboss/

A new procedure for simultaneously fitting a universal quasar continuum and observing frame systematics was described in Chapter 3. So far, preliminary results using the technique on DR12 observations are comparable to some of the early work done by the BOSS collaboration but there is a lot of potential left to investigate. In Chapter 4, I presented the process for estimating the two-point correlation function and its covariance using various sub-sampling and binning schemes. The software developed for this is publicly available and will be useful for future surveys such as eBOSS and DESI. In Chapter 5, I provided a brief overview of the technique used for fitting cosmological models to the two-point correlation function of the observed Lyman-α flux density tracer field of high-redshift quasars. Finally, in Chapter 6, I summarized the BAO measurements and provided a glimpse into how they impact constraints of cosmological models, specifically for dark energy.

Soon, the eBOSS, using the same telescope but an alternative class of targets selected for observation, will extend the range of BAO measurement via galaxy clustering to redshifts $0.7 < z < 2.0$. DESI will go above and beyond the observing prowess of the SDSS 2.5m telescope with a 4m telescope capable of observing 5,000 objects simultanesouly. This will enable measurements of BAO using various probes covering redshifts $0.1 < z < 4$. Since the dawn of time, humans have looked to the sky for answers regarding existence and origins. These measurements continue this great quest to understand our universe.

Bibliography

S. Alam, F. D. Albareti, C. Allende Prieto, F. Anders, S. F. Anderson, B. H. Andrews, E. Armengaud, É. Aubourg, S. Bailey, J. E. Bautista, and et al. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. *ArXiv e-prints*, Jan. 2015.

S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek, A. S. Bolton, J. R. Brownstein, A. Burden, C.-H. Chuang, J. Comparat, A. J. Cuesta, K. S. Dawson, D. J. Eisenstein, S. Escoffier, H. Gil-Marín, J. N. Grieb, N. Hand, S. Ho, K. Kinemuchi, D. Kirkby, F. Kitaura, E. Malanushenko, V. Malanushenko, C. Maraston, C. K. McBride, R. C. Nichol, M. D. Olmstead, D. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, M. Pellejero-Ibanez, W. J. Percival, P. Petitjean, F. Prada, A. M. Price-Whelan, B. A. Reid, S. A. Rodríguez-Torres, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martín, S. Saito, S. Salazar-Albornoz, L. Samushia, A. G. Sánchez, S. Satpathy, D. J. Schlegel, D. P. Schneider, C. G. Scóccola, H.-J. Seo, E. S. Sheldon, A. Simmons, A. Slosar, M. A. Strauss, M. E. C. Swanson, D. Thomas, J. L. Tinker, R. Tojeiro, M. V. Magaña, J. A. Vazquez, L. Verde, D. A. Wake, Y. Wang, D. H. Weinberg, M. White, W. M. Wood-Vasey, C. Yèche, I. Zehavi, Z. Zhai, and G.-B. Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *MNRAS*, 470:2617–2652, Sept. 2017. doi: 10.1093/mnras/stx721.

É. Aubourg, S. Bailey, J. E. Bautista, F. Beutler, V. Bhardwaj, D. Bizyaev, M. Blanton, M. Blomqvist, A. S. Bolton, J. Bovy, H. Brewington, J. Brinkmann, J. R. Brownstein, A. Burden, N. G. Busca, W. Carithers, C.-H. Chuang, J. Comparat, R. A. C. Croft, A. J. Cuesta, K. S. Dawson, T. Delubac, D. J. Eisenstein, A. Font-Ribera, J. Ge, J.-M. Le Goff, S. G. A. Gontcho, J. R. Gott, J. E. Gunn, H. Guo, J. Guy, J.-C. Hamilton, S. Ho, K. Honscheid, C. Howlett, D. Kirkby, F. S. Kitaura, J.-P. Kneib, K.-G. Lee, D. Long, R. H. Lupton, M. V. Magaña, V. Malanushenko, E. Malanushenko, M. Manera, C. Maraston, D. Margala, C. K. McBride, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, P. Noterdaeme, S. E. Nuza, M. D. Olmstead, D. Oravetz, I. Pâris, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, M. Pellejero-Ibanez, W. J. Percival, P. Petitjean, M. M. Pieri, F. Prada, B. Reid, J. Rich, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martín, A. G. Sánchez, L. Samushia, R. T. Génova-Santos, C. G. Scóccola, D. J. Schlegel, D. P. Schneider, H.-J. Seo, E. Sheldon, A. Simmons, R. A. Skibba, A. Slosar, M. A. Strauss, D. Thomas, J. L. Tinker, R. Tojeiro, J. A. Vazquez, M. Viel, D. A. Wake, B. A. Weaver, D. H. Weinberg, W. M. Wood-Vasey, C. Yèche, I. Zehavi, G.-B. Zhao, and BOSS

Collaboration. Cosmological implications of baryon acoustic oscillation measurements. **??***jnlPhys. Rev. D*, 92(12):123516, Dec. 2015. doi: 10.1103/PhysRevD.92.123516.

J. E. Bautista, N. G. Busca, J. Guy, J. Rich, M. Blomqvist, H. du Mas des Bourboux, M. M. Pieri, A. Font-Ribera, S. Bailey, T. Delubac, D. Kirkby, J.-M. Le Goff, D. Margala, A. Slosar, J. A. Vazquez, J. R. Brownstein, K. S. Dawson, D. J. Eisenstein, J. Miralda-Escudé, P. Noterdaeme, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, N. P. Ross, D. P. Schneider, D. H. Weinberg, and C. Yèche. Measurement of baryon acoustic oscillation correlations at z = 2.3 with SDSS DR12 Lyα-Forests. *A&A*, 603:A12, June 2017. doi: 10.1051/0004-6361/201730533.

F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, L. Campbell, Q. Parker, W. Saunders, and F. Watson. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *MNRAS*, 416:3017–3032, Oct. 2011. doi: 10.1111/j.1365-2966.2011.19250.x.

H. Bi and A. F. Davidsen. Evolution of Structure in the Intergalactic Medium and the Nature of the Lyα Forest. *ApJ*, 479:523–542, Apr. 1997.

H. Bi, J. Ge, and L.-Z. Fang. A Simulation of LY alpha Absorption Forests in Linear Approximation of Cold Dark Matter and Cold Plus Hot Dark Matter Models. *ApJ*, 452:90, Oct. 1995. doi: 10.1086/176282.

A. S. Bolton, D. J. Schlegel, É. Aubourg, S. Bailey, V. Bhardwaj, J. R. Brownstein, S. Burles, Y.-M. Chen, K. Dawson, D. J. Eisenstein, J. E. Gunn, G. R. Knapp, C. P. Loomis, R. H. Lupton, C. Maraston, D. Muna, A. D. Myers, M. D. Olmstead, N. Padmanabhan, I. Pâris, W. J. Percival, P. Petitjean, C. M. Rockosi, N. P. Ross, D. P. Schneider, Y. Shu, M. A. Strauss, D. Thomas, C. A. Tremonti, D. A. Wake, B. A. Weaver, and W. M. Wood-Vasey. Spectral Classification and Redshift Measurement for the SDSS-III Baryon Oscillation Spectroscopic Survey. *AJ*, 144:144, Nov. 2012. doi: 10.1088/0004-6256/144/5/144.

N. G. Busca, T. Delubac, J. Rich, S. Bailey, A. Font-Ribera, D. Kirkby, J.-M. Le Goff, M. M. Pieri, A. Slosar, É. Aubourg, J. E. Bautista, D. Bizyaev, M. Blomqvist, A. S. Bolton, J. Bovy, H. Brewington, A. Borde, J. Brinkmann, B. Carithers, R. A. C. Croft, K. S. Dawson, G. Ebelke, D. J. Eisenstein, J.-C. Hamilton, S. Ho, D. W. Hogg, K. Honscheid, K.-G. Lee, B. Lundgren, E. Malanushenko, V. Malanushenko, D. Margala, C. Maraston, K. Mehta, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, P. Noterdaeme, M. D. Olmstead, D. Oravetz, N. Palanque-Delabrouille, K. Pan, I. Pâris, W. J. Percival, P. Petitjean, N. A. Roe, E. Rollinde, N. P. Ross, G. Rossi, D. J. Schlegel, D. P. Schneider, A. Shelden, E. S. Sheldon, A. Simmons, S. Snedden, J. L. Tinker, M. Viel, B. A. Weaver, D. H. Weinberg, M. White, C. Yèche, and D. G. York. Baryon acoustic oscillations in the Lyα forest of BOSS quasars. *A&A*, 552:A96, Apr. 2013. doi: 10.1051/0004-6361/201220724.

R. Cen, J. Miralda-Escudé, J. P. Ostriker, and M. Rauch. Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest. *ApJL*, 437:L9–L12, Dec. 1994. doi: 10.1086/187670.

K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey, R. H. Barkhouser, J. E. Bautista, A. Beifiori, A. A. Berlind, V. Bhardwaj, D. Bizyaev, C. H. Blake, M. R. Blanton, M. Blomqvist, A. S. Bolton, A. Borde, J. Bovy, W. N. Brandt, H. Brewington, J. Brinkmann, P. J. Brown, J. R. Brownstein, K. Bundy, N. G. Busca, W. Carithers, A. R. Carnero, M. A. Carr, Y. Chen, J. Comparat, N. Connolly, F. Cope, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, T. Delubac, R. de Putter, S. Dhital, A. Ealet, G. L. Ebelke, D. J. Eisenstein, S. Escoffier, X. Fan, N. Filiz Ak, H. Finley, A. Font-Ribera, R. Génova-Santos, J. E. Gunn, H. Guo, D. Haggard, P. B. Hall, J.-C. Hamilton, B. Harris, D. W. Harris, S. Ho, D. W. Hogg, D. Holder, K. Honscheid, J. Huehnerhoff, B. Jordan, W. P. Jordan, G. Kauffmann, E. A. Kazin, D. Kirkby, M. A. Klaene, J.-P. Kneib, J.-M. Le Goff, K.-G. Lee, D. C. Long, C. P. Loomis, B. Lundgren, R. H. Lupton, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. McDonald, I. D. McGreer, R. G. McMahon, O. Mena, J. Miralda-Escudé, A. D. Montero-Dorta, F. Montesano, D. Muna, A. D. Myers, T. Naugle, R. C. Nichol, P. Noterdaeme, S. E. Nuza, M. D. Olmstead, A. Oravetz, D. J. Oravetz, R. Owen, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. K. Parejko, I. Pâris, W. J. Percival, I. Pérez-Fournon, I. Pérez-Ràfols, P. Petitjean, R. Pfaffenberger, J. Pforr, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, R. Rebolo, J. Rich, G. T. Richards, C. M. Rockosi, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martin, L. Samushia, A. G. Sánchez, C. Sayres, S. J. Schmidt, D. P. Schneider, C. G. Scóccola, H.-J. Seo, A. Shelden, E. Sheldon, Y. Shen, Y. Shu, A. Slosar, S. A. Smee, S. A. Snedden, F. Stauffer, O. Steele, M. A. Strauss, A. Streblyanska, N. Suzuki, M. E. C. Swanson, T. Tal, M. Tanaka, D. Thomas, J. L. Tinker, R. Tojeiro, C. A. Tremonti, M. Vargas Magaña, L. Verde, M. Viel, D. A. Wake, M. Watson, B. A. Weaver, D. H. Weinberg, B. J. Weiner, A. A. West, M. White, W. M. Wood-Vasey, C. Yeche, I. Zehavi, G.-B. Zhao, and Z. Zheng. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *AJ*, 145:10, Jan. 2013. doi: 10.1088/0004-6256/145/1/10.

T. Delubac, J. E. Bautista, N. G. Busca, J. Rich, D. Kirkby, S. Bailey, A. Font-Ribera, A. Slosar, K.-G. Lee, M. M. Pieri, J.-C. Hamilton, É. Aubourg, M. Blomqvist, J. Bovy, J. Brinkmann, W. Carithers, K. S. Dawson, D. J. Eisenstein, S. G. A. Gontcho, J.-P. Kneib, J.-M. Le Goff, D. Margala, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, P. Noterdaeme, R. O'Connell, M. D. Olmstead, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, N. P. Ross, G. Rossi, D. J. Schlegel, D. P. Schneider, D. H. Weinberg, C. Yèche, and D. G. York. Baryon acoustic oscillations in the Lyα forest of BOSS DR11 quasars. *A&A*, 574: A59, Feb. 2015. doi: 10.1051/0004-6361/201423969.

S. Dodelson. *Modern cosmology*. Academic Press, 2003.

D. J. Eisenstein, H.-J. Seo, and M. White. On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter. *ApJ*, 664:660–674, Aug. 2007. doi: 10.1086/518755.

D. J. Eisenstein, D. H. Weinberg, E. Agol, H. Aihara, C. Allende Prieto, S. F. Anderson, J. A. Arns, É. Aubourg, S. Bailey, E. Balbinot, and et al. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. *AJ*, 142:72, Sept. 2011. doi: 10.1088/0004-6256/142/3/72.

N. Filiz Ak, W. N. Brandt, P. B. Hall, D. P. Schneider, S. F. Anderson, R. R. Gibson, B. F. Lundgren, A. D. Myers, P. Petitjean, N. P. Ross, Y. Shen, D. G. York, D. Bizyaev, J. Brinkmann, E. Malanushenko, D. J. Oravetz, K. Pan, A. E. Simmons, and B. A. Weaver. Broad Absorption Line Disappearance on Multi-year Timescales in a Large Quasar Sample. *ApJ*, 757:114, Oct. 2012. doi: 10.1088/0004-637X/757/2/114.

N. Filiz Ak, W. N. Brandt, P. B. Hall, D. P. Schneider, S. F. Anderson, F. Hamann, B. F. Lundgren, A. D. Myers, I. Pâris, P. Petitjean, N. P. Ross, Y. Shen, and D. York. Broad Absorption Line Variability on Multi-year Timescales in a Large Quasar Sample. *ApJ*, 777:168, Nov. 2013. doi: 10.1088/0004-637X/777/2/168.

N. Filiz Ak, W. N. Brandt, P. B. Hall, D. P. Schneider, J. R. Trump, S. F. Anderson, F. Hamann, A. D. Myers, I. Pâris, P. Petitjean, N. P. Ross, Y. Shen, and D. York. The Dependence of C IV Broad Absorption Line Properties on Accompanying Si IV and Al III Absorption: Relating Quasar-wind Ionization Levels, Kinematics, and Column Densities. *ApJ*, 791:88, Aug. 2014. doi: 10.1088/0004-637X/791/2/88.

A. Font-Ribera, D. Kirkby, N. Busca, J. Miralda-Escudé, N. P. Ross, A. Slosar, J. Rich, É. Aubourg, S. Bailey, V. Bhardwaj, J. Bautista, F. Beutler, D. Bizyaev, M. Blomqvist, H. Brewington, J. Brinkmann, J. R. Brownstein, B. Carithers, K. S. Dawson, T. Delubac, G. Ebelke, D. J. Eisenstein, J. Ge, K. Kinemuchi, K.-G. Lee, V. Malanushenko, E. Malanushenko, M. Marchante, D. Margala, D. Muna, A. D. Myers, P. Noterdaeme, D. Oravetz, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, M. M. Pieri, G. Rossi, D. P. Schneider, A. Simmons, M. Viel, C. Yeche, and D. G. York. Quasar-Lyman α forest cross-correlation from BOSS DR11: Baryon Acoustic Oscillations. *JCAP*, 5:027, May 2014. doi: 10.1088/1475-7516/2014/05/027.

D. L. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *J. Opt. Soc. Am.*, 56(10):1372–1379, Oct 1966. doi: 10.1364/JOSA.56.001372. URL `http://www.osapublishing.org/abstract.cfm?URI=josa-56-10-1372`.

M. Fukugita, T. Ichikawa, J. E. Gunn, M. Doi, K. Shimasaku, and D. P. Schneider. The Sloan Digital Sky Survey Photometric System. *AJ*, 111:1748, Apr. 1996. doi: 10.1086/117915.

K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *ApJ*, 622:759–771, Apr. 2005. doi: 10.1086/427976.

J. E. Gunn and B. A. Peterson. On the Density of Neutral Hydrogen in Intergalactic Space. *ApJ*, 142:1633–1641, Nov. 1965. doi: 10.1086/148444.

J. E. Gunn, M. Carr, C. Rockosi, M. Sekiguchi, K. Berry, B. Elms, E. de Haas, Ž. Ivezić, G. Knapp, R. Lupton, G. Pauls, R. Simcoe, R. Hirsch, D. Sanford, S. Wang, D. York, F. Harris, J. Annis, L. Bartozek, W. Boroski, J. Bakken, M. Haldeman, S. Kent, S. Holm, D. Holmgren, D. Petravick, A. Prosapio, R. Rechenmacher, M. Doi, M. Fukugita, K. Shimasaku, N. Okada, C. Hull, W. Siegmund, E. Mannery, M. Blouke, D. Heidtman,

D. Schneider, R. Lucinio, and J. Brinkman. The Sloan Digital Sky Survey Photometric Camera. *AJ*, 116:3040–3081, Dec. 1998. doi: 10.1086/300645.

J. E. Gunn, W. A. Siegmund, E. J. Mannery, R. E. Owen, C. L. Hull, R. F. Leger, L. N. Carey, G. R. Knapp, D. G. York, W. N. Boroski, S. M. Kent, R. H. Lupton, C. M. Rockosi, M. L. Evans, P. Waddell, J. E. Anderson, J. Annis, J. C. Barentine, L. M. Bartoszek, S. Bastian, S. B. Bracker, H. J. Brewington, C. I. Briegel, J. Brinkmann, Y. J. Brown, M. A. Carr, P. C. Czarapata, C. C. Drennan, T. Dombeck, G. R. Federwitz, B. A. Gillespie, C. Gonzales, S. U. Hansen, M. Harvanek, J. Hayes, W. Jordan, E. Kinney, M. Klaene, S. J. Kleinman, R. G. Kron, J. Kresinski, G. Lee, S. Limmongkol, C. W. Lindenmeyer, D. C. Long, C. L. Loomis, P. M. McGehee, P. M. Mantsch, E. H. Neilsen, Jr., R. M. Neswold, P. R. Newman, A. Nitta, J. Peoples, Jr., J. R. Pier, P. S. Prieto, A. Prosapio, C. Rivetta, D. P. Schneider, S. Snedden, and S.-i. Wang. The 2.5 m Telescope of the Sloan Digital Sky Survey. *AJ*, 131:2332–2359, Apr. 2006. doi: 10.1086/500975.

P. B. Hall, W. N. Brandt, P. Petitjean, I. Pâris, N. Filiz Ak, Y. Shen, R. R. Gibson, É. Aubourg, S. F. Anderson, D. P. Schneider, D. Bizyaev, J. Brinkmann, E. Malanushenko, V. Malanushenko, A. D. Myers, D. J. Oravetz, N. P. Ross, A. Shelden, A. E. Simmons, A. Streblyanska, B. A. Weaver, and D. G. York. Broad absorption line quasars with redshifted troughs: high-velocity infall or rotationally dominated outflows? *MNRAS*, 434: 222–256, Sept. 2013. doi: 10.1093/mnras/stt1012.

L. Hernquist, N. Katz, D. H. Weinberg, and J. Miralda-Escudé. The Lyman-Alpha Forest in the Cold Dark Matter Model. *ApJL*, 457:L51, Feb. 1996. doi: 10.1086/309899.

W. Hu and N. Sugiyama. Small-Scale Cosmological Perturbations: an Analytic Approach. *ApJ*, 471:542, Nov. 1996. doi: 10.1086/177989.

L. Hui, N. Y. Gnedin, and Y. Zhang. The Statistics of Density Peaks and the Column Density Distribution of the Lyα Forest. *ApJ*, 486:599–622, Sept. 1997.

Ž. Ivezić, A. Connolly, J. Vanderplas, and A. Gray. *Statistics, Data Mining and Machine Learning in Astronomy*. Princeton University Press, 2014.

D. H. Jones, M. A. Read, W. Saunders, M. Colless, T. Jarrett, Q. A. Parker, A. P. Fairall, T. Mauch, E. M. Sadler, F. G. Watson, D. Burton, L. A. Campbell, P. Cass, S. M. Croom, J. Dawe, K. Fiegert, L. Frankcombe, M. Hartley, J. Huchra, D. James, E. Kirby, O. Lahav, J. Lucey, G. A. Mamon, L. Moore, B. A. Peterson, S. Prior, D. Proust, K. Russell, V. Safouris, K.-I. Wakamatsu, E. Westra, and M. Williams. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. *MNRAS*, 399:683–698, Oct. 2009. doi: 10.1111/j.1365-2966.2009.15338.x.

D. Kirkby, D. Margala, A. Slosar, S. Bailey, N. G. Busca, T. Delubac, J. Rich, J. E. Bautista, M. Blomqvist, J. R. Brownstein, B. Carithers, R. A. C. Croft, K. S. Dawson, A. Font-Ribera, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, G. Rossi, D. J. Schlegel, D. P. Schneider, M. Viel, D. H. Weinberg, and C. Yèche. Fitting methods for baryon acoustic oscillations in the

Lyman-α forest fluctuations in BOSS data release 9. *JCAP*, 3:024, Mar. 2013. doi: 10.1088/1475-7516/2013/03/024.

M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein, B. Flaugher, K. Honscheid, R. Kron, O. Lahav, P. McDonald, N. Roe, D. Schlegel, and representing the DESI collaboration. The DESI Experiment, a whitepaper for Snowmass 2013. *ArXiv e-prints*, Aug. 2013.

I. D. McGreer, L. Jiang, X. Fan, G. T. Richards, M. A. Strauss, N. P. Ross, M. White, Y. Shen, D. P. Schneider, A. D. Myers, W. N. Brandt, C. DeGraf, E. Glikman, J. Ge, and A. Streblyanska. The z = 5 Quasar Luminosity Function from SDSS Stripe 82. *ApJ*, 768: 105, May 2013. doi: 10.1088/0004-637X/768/2/105.

A. A. Meiksin. The physics of the intergalactic medium. *Reviews of Modern Physics*, 81: 1405–1469, Oct. 2009. doi: 10.1103/RevModPhys.81.1405.

J. Miralda-Escudé, R. Cen, J. P. Ostriker, and M. Rauch. The Ly alpha Forest from Gravitational Collapse in the Cold Dark Matter + Lambda Model. *ApJ*, 471:582, Nov. 1996. doi: 10.1086/177992.

J. B. Oke and J. E. Gunn. Secondary standard stars for absolute spectrophotometry. *ApJ*, 266:713–717, Mar. 1983. doi: 10.1086/160817.

N. Palanque-Delabrouille, C. Magneville, C. Yèche, S. Eftekharzadeh, A. D. Myers, P. Petitjean, I. Pâris, E. Aubourg, I. McGreer, X. Fan, A. Dey, D. Schlegel, S. Bailey, D. Bizayev, A. Bolton, K. Dawson, G. Ebelke, J. Ge, E. Malanushenko, V. Malanushenko, D. Oravetz, K. Pan, N. P. Ross, D. P. Schneider, E. Sheldon, A. Simmons, J. Tinker, M. White, and C. Willmer. Luminosity function from dedicated SDSS-III and MMT data of quasars in 0.7 <z <4.0 selected with a new approach. *A&A*, 551:A29, Mar. 2013. doi: 10.1051/0004-6361/201220379.

I. Pâris, P. Petitjean, N. P. Ross, A. D. Myers, É. Aubourg, A. Streblyanska, S. Bailey, É. Armengaud, N. Palanque-Delabrouille, C. Yèche, F. Hamann, M. A. Strauss, F. D. Albareti, J. Bovy, D. Bizyaev, W. N. Brandt, M. Brusa, J. Buchner, J. Comparat, R. A. C. Croft, T. Dwelly, X. Fan, A. Font-Ribera, J. Ge, A. Georgakakis, P. B. Hall, L. Jian, K. Kinemuchi, E. Malanushenko, V. Malanushenko, R. G. McMahon, M.-L. Menzel, A. Merloni, K. Nandra, P. Noterdaeme, D. Oravetz, K. Pan, M. M. Pieri, F. Prada, M. Salvato, D. J. Schlegel, D. P. Schneider, A. Simmons, M. Viel, D. H. Weinberg, and L. Zhu. The Sloan Digital Sky Survey Quasar Catalog: twelfth data release. *ArXiv e-prints*, Aug. 2016.

M. Rauch. The Lyman Alpha Forest in the Spectra of QSOs. *ARA&A*, 36:267–316, 1998. doi: 10.1146/annurev.astro.36.1.267.

N. P. Ross, A. D. Myers, E. S. Sheldon, C. Yèche, M. A. Strauss, J. Bovy, J. A. Kirkpatrick, G. T. Richards, É. Aubourg, M. R. Blanton, W. N. Brandt, W. C. Carithers, R. A. C. Croft, R. da Silva, K. Dawson, D. J. Eisenstein, J. F. Hennawi, S. Ho, D. W. Hogg, K.-G. Lee, B. Lundgren, R. G. McMahon, J. Miralda-Escudé, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, M. M. Pieri, J. Rich, N. A. Roe, D. Schiminovich, D. J. Schlegel,

D. P. Schneider, A. Slosar, N. Suzuki, J. L. Tinker, D. H. Weinberg, A. Weyant, M. White, and W. M. Wood-Vasey. The SDSS-III Baryon Oscillation Spectroscopic Survey: Quasar Target Selection for Data Release Nine. *ApJS*, 199:3, Mar. 2012. doi: 10.1088/0067-0049/199/1/3.

N. P. Ross, I. D. McGreer, M. White, G. T. Richards, A. D. Myers, N. Palanque-Delabrouille, M. A. Strauss, S. F. Anderson, Y. Shen, W. N. Brandt, C. Yèche, M. E. C. Swanson, É. Aubourg, S. Bailey, D. Bizyaev, J. Bovy, H. Brewington, J. Brinkmann, C. DeGraf, T. Di Matteo, G. Ebelke, X. Fan, J. Ge, E. Malanushenko, V. Malanushenko, R. Mandelbaum, C. Maraston, D. Muna, D. Oravetz, K. Pan, I. Pâris, P. Petitjean, K. Schawinski, D. J. Schlegel, D. P. Schneider, J. D. Silverman, A. Simmons, S. Snedden, A. Streblyanska, N. Suzuki, D. H. Weinberg, and D. York. The SDSS-III Baryon Oscillation Spectroscopic Survey: The Quasar Luminosity Function from Data Release Nine. *ApJ*, 773:14, Aug. 2013. doi: 10.1088/0004-637X/773/1/14.

S. Seager, D. D. Sasselov, and D. Scott. A New Calculation of the Recombination Epoch. *ApJL*, 523:L1–L5, Sept. 1999. doi: 10.1086/312250.

A. Slosar, A. Font-Ribera, M. M. Pieri, J. Rich, J.-M. Le Goff, É. Aubourg, J. Brinkmann, B. Carithers, R. Charlassier, M. Cortês, et al. The lyman-α forest in three dimensions: measurements of large scale flux correlations from boss 1st-year data. *Journal of Cosmology and Astroparticle Physics*, 2011(09):001, 2011.

A. Slosar, V. Iršič, D. Kirkby, S. Bailey, N. G. Busca, T. Delubac, J. Rich, É. Aubourg, J. E. Bautista, V. Bhardwaj, M. Blomqvist, A. S. Bolton, J. Bovy, J. Brownstein, B. Carithers, R. A. C. Croft, K. S. Dawson, A. Font-Ribera, J.-M. Le Goff, S. Ho, K. Honscheid, K.-G. Lee, D. Margala, P. McDonald, B. Medolin, J. Miralda-Escudé, A. D. Myers, R. C. Nichol, P. Noterdaeme, N. Palanque-Delabrouille, I. Pâris, P. Petitjean, M. M. Pieri, Y. Piškur, N. A. Roe, N. P. Ross, G. Rossi, D. J. Schlegel, D. P. Schneider, N. Suzuki, E. S. Sheldon, U. Seljak, M. Viel, D. H. Weinberg, and C. Yèche. Measurement of baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9. *JCAP*, 4:026, Apr. 2013. doi: 10.1088/1475-7516/2013/04/026.

S. A. Smee, J. E. Gunn, A. Uomoto, N. Roe, D. Schlegel, C. M. Rockosi, M. A. Carr, F. Leger, K. S. Dawson, M. D. Olmstead, J. Brinkmann, R. Owen, R. H. Barkhouser, K. Honscheid, P. Harding, D. Long, R. H. Lupton, C. Loomis, L. Anderson, J. Annis, M. Bernardi, V. Bhardwaj, D. Bizyaev, A. S. Bolton, H. Brewington, J. W. Briggs, S. Burles, J. G. Burns, F. J. Castander, A. Connolly, J. R. A. Davenport, G. Ebelke, H. Epps, P. D. Feldman, S. D. Friedman, J. Frieman, T. Heckman, C. L. Hull, G. R. Knapp, D. M. Lawrence, J. Loveday, E. J. Mannery, E. Malanushenko, V. Malanushenko, A. J. Merrelli, D. Muna, P. R. Newman, R. C. Nichol, D. Oravetz, K. Pan, A. C. Pope, P. G. Ricketts, A. Shelden, D. Sandford, W. Siegmund, A. Simmons, D. S. Smith, S. Snedden, D. P. Schneider, M. SubbaRao, C. Tremonti, P. Waddell, and D. G. York. The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey. *AJ*, 146:32, Aug. 2013. doi: 10.1088/0004-6256/146/2/32.

V. I. Tatarskii. *Wave Propagation in Turbulent Medium*. McGraw-Hill, 1961.

T. Theuns, A. Leonard, G. Efstathiou, F. R. Pearce, and P. A. Thomas. P^3M-SPH simulations of the Lyalpha forest. *MNRAS*, 301:478–502, Dec. 1998. doi: 10.1046/j.1365-8711.1998.02040.x.

S. Vikas, W. M. Wood-Vasey, B. Lundgren, N. P. Ross, A. D. Myers, Y. AlSayyad, D. G. York, D. P. Schneider, J. Brinkmann, D. Bizyaev, H. Brewington, J. Ge, E. Malanushenko, V. Malanushenko, D. Muna, D. Oravetz, K. Pan, I. Pâris, P. Petitjean, S. Snedden, A. Shelden, A. Simmons, and B. A. Weaver. Moderate C IV Absorber Systems Require 10^{12} M $_{\odot}$ Dark Matter Halos at z ˜ 2.3: A Cross-correlation Study of C IV Absorber Systems and Quasars in SDSS-III BOSS DR9. *ApJ*, 768:38, May 2013. doi: 10.1088/0004-637X/768/1/38.

D. H. Weinberg, M. J. Mortonson, D. J. Eisenstein, C. Hirata, A. G. Riess, and E. Rozo. Observational probes of cosmic acceleration. *PhR*, 530:87–255, Sept. 2013. doi: 10.1016/j.physrep.2013.05.001.

M. White, A. D. Myers, N. P. Ross, D. J. Schlegel, J. F. Hennawi, Y. Shen, I. McGreer, M. A. Strauss, A. S. Bolton, J. Bovy, X. Fan, J. Miralda-Escude, N. Palanque-Delabrouille, I. Paris, P. Petitjean, D. P. Schneider, M. Viel, D. H. Weinberg, C. Yeche, I. Zehavi, K. Pan, S. Snedden, D. Bizyaev, H. Brewington, J. Brinkmann, V. Malanushenko, E. Malanushenko, D. Oravetz, A. Simmons, A. Sheldon, and B. A. Weaver. The clustering of intermediate-redshift quasars as measured by the Baryon Oscillation Spectroscopic Survey. *MNRAS*, 424:933–950, Aug. 2012. doi: 10.1111/j.1365-2966.2012.21251.x.

D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda,

and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120: 1579–1587, Sept. 2000. doi: 10.1086/301513.

Y. Zhang, P. Anninos, and M. L. Norman. A Multispecies Model for Hydrogen and Helium Absorbers in Lyman-Alpha Forest Clouds. *ApJL*, 453:L57, Nov. 1995. doi: 10.1086/309752.

Appendix A

Public Access to Data and Code

The data used in this work publicly available at `http://www.sdss.org/dr12/`.

The software developed for this analysis and to generate the results presented here are publicly available. A list of the primary packages is provided here:

- `bossdata`[1]: A python package for working with spectroscopic data from the Sloan Digital Sky Survey.
- `tpcorr`[2]: Throughput correction code for offset fibers in SDSS.
- `lyabao`[3]: Quasar continuum fitting and Lyman-α flux-density tracer extraction.
- `turbo-octo-spice`[4]: A software package for estimating the two-point correlation function of density fields along lines of sight.
- `likely`[5]: A C++ library of numerical utilities to support likelihood analysis.

[1] `https://github.com/dkirkby/bossdata/`
[2] `https://github.com/dmargala/tpcorr/`
[3] `https://github.com/dmargala/lyabao/`
[4] `https://github.com/dmargala/turbo-octo-spice/`
[5] `https://github.com/igmhub/likely/`

- `cosmo`[6]: A C++ library for numerical evaluation of cosmological quantities.
- `baofit`[7]: A software package for analyzing cosmological correlation functions to estimate parameters related to baryon acoustic oscillations and redshift-space distortions.

[6] `https://github.com/igmhub/cosmo/`
[7] `https://github.com/igmhub/baofit/`