IOPscience

Fitting methods for baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9

Fitting methods for baryon acoustic oscillations in the Lyman-α forest fluctuations in BOSS data release 9

David Kirkby,[a] **Daniel Margala,**[a] **Anže Slosar,**[b] **Stephen Bailey,**[c]
Nicolás G. Busca,[d] **Timothée Delubac,**[e] **James Rich,**[e]
Julian E. Bautista,[d] **Michael Blomqvist,**[a] **Joel R. Brownstein,**[f]
Bill Carithers,[c] **Rupert A.C. Croft,**[g] **Kyle S. Dawson,**[f]
Andreu Font-Ribera,[h,c] **Jordi Miralda-Escudé,**[i,j] **Adam D. Myers,**[k]
Robert C. Nichol,[l] **Nathalie Palanque-Delabrouille,**[e]
Isabelle Pâris,[m,n] **Patrick Petitjean,**[m] **Graziano Rossi,**[e]
David J. Schlegel,[c] **Donald P. Schneider,**[o,p] **Matteo Viel,**[q,r]
David H. Weinberg[s] **and Christophe Yèche**[e]

[a]Department of Physics and Astronomy, University of California,
 Irvine, 92697, U.S.A.

[b]Brookhaven National Laboratory,
 Blgd 510, Upton NY 11375, U.S.A.

[c]Lawrence Berkeley National Laboratory,
 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.

[d]APC, Université Paris Diderot-Paris 7, CNRS/IN2P3, CEA, Observatoire de Paris,
 10, rue A. Domon & L. Duquet, Paris, France

[e]CEA, Centre de Saclay,
 IRFU, F-91191 Gif-sur-Yvette, France

[f]Department of Physics and Astronomy, University of Utah,
 115 S 1400 E, Salt Lake City, UT 84112, U.S.A.

[g]Bruce and Astrid McWilliams Center for Cosmology, Carnegie Mellon University,
 Pittsburgh, PA 15213, U.S.A.

[h]Institute of Theoretical Physics, University of Zurich,
 8057 Zurich, Switzerland

[i]Institució Catalana de Recerca i Estudis Avançats,
 Barcelona, Catalonia

[j]Institut de Ciències del Cosmos, Universitat de Barcelona/IEEC,
 Barcelona 08028, Catalonia

[k]Department of Physics and Astronomy, University of Wyoming,
 Laramie, WY 82071, U.S.A.

[l]Institute of Cosmology and Gravitation, Dennis Sciama Building, University of Portsmouth,
 Portsmouth, PO1 3FX, U.K.

^{m}Université Paris 6 et CNRS, Institut d'Astrophysique de Paris,
  98bis blvd. Arago, 75014 Paris, France

^{n}Departamento de Astronomía, Universidad de Chile,
  Casilla 36-D, Santiago, Chile

^{o}Department of Astronomy and Astrophysics, The Pennsylvania State University,
  University Park, PA 16802, U.S.A.

^{p}Institute for Gravitation and the Cosmos, The Pennsylvania State University,
  University Park, PA 16802, U.S.A.

^{q}INAF, Osservatorio Astronomico di Trieste,
  Via G. B. Tiepolo 11, 34131 Trieste, Italy

^{r}INFN/National Institute for Nuclear Physics,
  Via Valerio 2, I-34127 Trieste, Italy

^{s}Department of Astronomy, Ohio State University,
  140 West 18th Avenue, Columbus, OH 43210, U.S.A.

  E-mail: dkirkby@uci.edu

Abstract. We describe fitting methods developed to analyze fluctuations in the Lyman-α forest and measure the parameters of baryon acoustic oscillations (BAO). We apply our methods to BOSS Data Release 9. Our method is based on models of the three-dimensional correlation function in physical coordinate space, and includes the effects of redshift-space distortions, anisotropic non-linear broadening, and broadband distortions. We allow for independent scale factors along and perpendicular to the line of sight to minimize the dependence on our assumed fiducial cosmology and to obtain separate measurements of the BAO angular and relative velocity scales. Our fitting software and the input files needed to reproduce our main BOSS Data Release 9 results are publicly available.

Contents

1 Introduction

The surprising discovery [1, 2] of accelerating expansion in the current universe reveals that either some form of dark energy is driving the expansion or else that our theory of gravity is incomplete on the largest scales (see reference [3] for a recent review). The length scale of baryon acoustic oscillations (BAO) imprinted at the moment of baryon-photon decoupling provides a standard ruler that has been well measured in the temperature anisotropies of the cosmic microwave background [4, 5] (CMB) and in the number-density fluctuations of galaxies at $z < 1$ [6–24]. The Baryon Oscillation Spectroscopic Survey [25] (BOSS) of the third generation of the Sloan Digital Sky Survey [26] (SDSS-III) recently announced its ninth data release [27] (DR9), including an unprecedented number of high-redshift quasar spectra. The BAO feature is imprinted in the correlated fluctuations of intergalactic Lyman-α absorption of the light from these quasars, enabling us to measure the BAO standard ruler at redshifts $z \simeq 2.4$, during the predicted era of matter-dominated deceleration.

An optimal extraction of the BAO signal from a large sample of Lyman-α forest pixels requires new fitting techniques beyond those previously developed to fit the large-scale clustering of galaxies [28]. For example, the non-uniform sampling of the absorption field, large expected redshift-space distortions, and relatively large depth suggest that estimates of the correlation function or power spectrum should be fit to three-dimensional models (rather than low-order multipoles at fixed redshift) expressed directly in terms of the physical absorption wavelengths and angular separations between lines of sight. We describe here new

fitting techniques developed specifically for a near-optimal analysis of the DR9 correlation-function estimates described in a companion paper [29], and highlight the challenges and lessons learned. We are also making our fit input files and fitting code publicly available as a companion to this paper, so that readers may reproduce the main results presented here and in ref. [29].

The outline of our paper is as follows. In section 2 we define our constituent models for linear theory with redshift-space distortions, non-linear effects, and redshift evolution. We also describe our parametrizations of possible deviations between our assumed fiducial cosmology[1] and the cosmology preferred by our data, and of broadband distortion of our correlation-function estimate introduced by analysis systematics. Finally, we introduce two data-reduction techniques using interpolated models. In section 3, we describe the DR9 fitting inputs which consist of $N = 1512$ correlation-function estimates on a three-dimensional physical coordinate grid, accompanied by an initial estimate of their covariance in a block diagonal form that reduces the number of non-zero elements from $N(N+1)/2 \simeq 1,114$K to $\simeq 64$K. Next, we describe a novel method for internally validating and refining our initial covariance estimate, necessitated by a lack of available simulated mock statistics. In section 4, we present our results on the fitting method and what it teaches us about the BOSS DR9 Lyman-α forest dataset. We discuss the expected parameter sensitivities and relative contributions from different regions of the three-dimensional separation space, and present model-independent data reduction results. Cosmological fitting results for DR9 are presented in the companion paper ref. [29]. We conclude in section 5 with a discussion of the main lessons learned and our plans for future development. Appendix A provides details on public access to our fit inputs and fitting code.

2 Models and parameters

We model a measurement of the correlation function

$$\xi(r, \mu, z) \equiv \langle \delta(\mathbf{s}_1)\delta(\mathbf{s}_2) \rangle - \langle \delta(\mathbf{s}_1) \rangle \langle \delta(\mathbf{s}_2) \rangle \tag{2.1}$$

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[2]

$$|\mathbf{s}_2 - \mathbf{s}_1| = r \quad , \quad |\mathbf{s}_2| - |\mathbf{s}_1| = \mu \cdot r \quad , \quad \frac{1}{2}|\mathbf{s}_1 + \mathbf{s}_2| = c\int_0^z \frac{dz'}{H_{\text{fid}}(z')} \tag{2.2}$$

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

2.1 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

---

[1]We assume a flat ΛCDM universe with $\Omega_\Lambda = 0.73$, $h = 0.7$, $\Omega_{\text{b}}h^2 = 0.0227$, and $n_s = 0.97$ throughout.

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

λ_i and λ_j. We convert the observed wavelengths to a relative velocity for the absorption systems[3] [30]

$$\Delta v_{ij} = c \log(\lambda_j/\lambda_i) \tag{2.3}$$

and an average absorption redshift

$$z_{ij} = \frac{\sqrt{\lambda_i \lambda_j}}{\lambda_\alpha} - 1 \tag{2.4}$$

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

$$r_\parallel = D_{C,\text{fid}}(z_j) - D_{C,\text{fid}}(z_i) = \frac{1+z_{ij}}{H_{\text{fid}}(z_{ij})} \cdot \Delta v_{ij} \left[1 + \mathcal{O}\left(\Delta v_{ij}/c\right)^2\right] \tag{2.5}$$

$$r_\perp = D_{A,\text{fid}}(z_{ij}) \cdot \Delta\theta_{ij} = c \int_0^{z_{ij}} \frac{dz'}{H_{\text{fid}}(z')} \cdot \Delta\theta_{ij} \tag{2.6}$$

where $D_{C,\text{fid}}(z)$ and $D_{A,\text{fid}}(z)$ are the co-moving line of sight and angular distance functions, respectively, for the assumed fidicual cosmology $H_{\text{fid}}(z)$. We introduce separate scale factors $\alpha_\parallel$ and $\alpha_\perp$ in section 2.2.4 to allow for small discrepancies between the true cosmology and our assumed fiducial model. Figure 1 shows that the BOSS blue camera wavelength limit of $\lambda \gtrsim 3600$ Å limits pixel pairs contributing to a BAO peak feature near 110 Mpc/h to $\Delta\theta_{ij} \lesssim 100$ arcmin. Similarly, the BOSS redshift coverage $z_{ij} \lesssim 3.25$ limits the BAO peak to $\Delta v_{ij}/c = \log(\lambda_2/\lambda_1) \lesssim 0.04$ and observed wavelength differences $|\lambda_j - \lambda_i| \lesssim 200$ Å.

2.2 Cosmological models

We build the cosmological model starting from an isotropic linear power spectrum prediction $\tilde{P}(k, z_0)$ 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 a plane-parallel redshift-space distortion $r \to (r, \mu)$ we have [31]:

$$\xi_{\text{cosmo}}(r, \mu, z_0) = \sum_{\ell \text{ even}} L_\ell(\mu) \xi_{\ell,\text{cosmo}}(r, z_0) \tag{2.7}$$

with

$$\xi_{\ell,\text{cosmo}}(r, z_0) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr) P_\ell(k, z_0) dk \tag{2.8}$$

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 $\mu_k \equiv \hat{z} \cdot \hat{k}$:

$$P_\ell(k, z_0) = \frac{2\ell + 1}{2} \int_{-1}^{+1} P(k, \mu_k, z_0) L_\ell(\mu_k) d\mu_k . \tag{2.9}$$

Specializing to linear theory and the distant observer approximation [32], the infinite series of eq. (2.7) is truncated at $\ell = 4$, with

$$\tilde{P}_\ell(k, z_0) = b^2(z_0) C_\ell(\beta(z_0)) \tilde{P}(k, z_0) \tag{2.10}$$

---

[3]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.
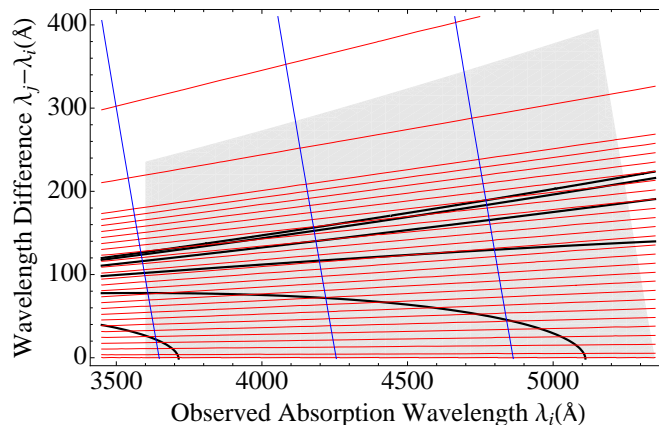
Figure 1. Physical coordinates for pixel pairs with observed Lyman-α absorption wavelengths $\lambda_i \leq \lambda_j$. Grid lines of z_{ij} (vertical blue, values left to right are 2.25, 2.75, 3.25), $\Delta v_{ij}/c$ (horizontal red, values bottom to top cover 0.001-0.049 with 0.002 spacing, with additional contours at 0, 0.059, and 0.083) represent the nominal sampling grid used in a fit. The shaded gray region shows the pixel pairs contributing to a typical BAO fit, bounded by $\lambda_1 > 3600$ Å, $r_\parallel < 170$ Mpc/h, and $z_{ij} < 3.25$. Contours of $\Delta\theta$ (thick black curves) at which the 3D separation is 110 Mpc/h (values from bottom left corner out are 100, 80, 60, 40, 20, 0 arcmins) identify pixel pairs contributing to the BAO peak region at different angular separations.

and

$$C_\ell(\beta) \equiv \frac{2\ell+1}{2} \int_{-1}^{+1} \left(1 + \beta\mu_k^2\right)^2 L_\ell(\mu_k)\, d\mu_k = \begin{cases} 1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2 & \ell = 0 \\ \frac{4}{3}\beta + \frac{4}{7}\beta^2 & \ell = 2 \\ \frac{8}{35}\beta^2 & \ell = 4 \end{cases} , \qquad (2.11)$$

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

$$\xi_{\text{cosmo}}(r, \mu, z_0) = b^2(z_0) \sum_{\ell=0,2,4} C_\ell(\beta(z_0)) L_\ell(\mu)\, \tilde{\xi}_{\ell,\text{cosmo}}(r, z_0) \qquad (2.12)$$

in terms of the undistorted linear-theory multipoles

$$\tilde{\xi}_{\ell,\text{cosmo}}(r, z_0) = \frac{i^\ell}{2\pi^2} \int_0^\infty k^2 j_\ell(kr)\, \tilde{P}(k, z_0)\, dk \ . \qquad (2.13)$$

Figure 2 shows examples of the linear models we use for Lyman-α fitting in this paper, calculated with $z_0 = 2.25$. Note that the undistorted multipoles $\tilde{\xi}_\ell$ are not independent since they derive from the same underlying power spectrum via eq. (2.13). Explicitly, we find that $\tilde{\xi}_2(r)$ and $\tilde{\xi}_4(r)$ can be calculated directly from $\tilde{\xi}_0(r')$ specified on an interval $r_0 \leq r' \leq r$ via (we have dropped z_0 here for clarity):

$$\begin{aligned}
\tilde{\xi}_2(r) &= \tilde{\xi}_0(r) + \left(\frac{r_0}{r}\right)^3 \left[\tilde{\xi}_2(r_0) - \tilde{\xi}_0(r_0)\right] - \frac{3}{r^3} \int_{r_0}^r \tilde{\xi}_0(r') r'^2 dr' \\
\tilde{\xi}_4(r) &= \tilde{\xi}_0(r) + \left(\frac{r_0}{r}\right)^5 \left[\tilde{\xi}_4(r_0) - \tilde{\xi}_0(r_0)\right] - \frac{5}{r^5} \int_{r_0}^r \left[\tilde{\xi}_0(r') + \tilde{\xi}_2(r')\right] r'^4 dr' ,
\end{aligned} \qquad (2.14)$$

where r_0 is an arbitrarily chosen scale and the influence of scales beyond $[r_0, r]$ is fully specified by the constants of integration $\tilde{\xi}_\ell(r_0)$.
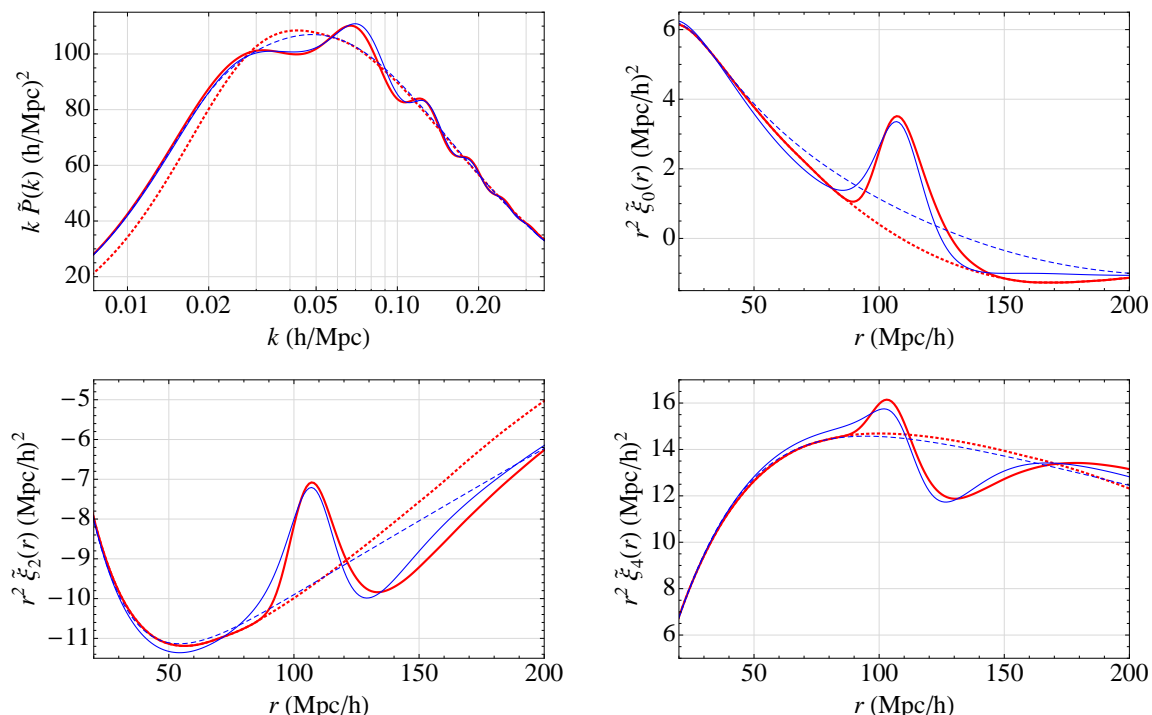
Figure 2. Cosmological linear models calculated for $z_0 = 2.25$ and assuming a flat universe with $\Omega_\Lambda = 0.73$, $h = 0.7$, $\Omega_b h^2 = 0.0227$, and $n_s = 0.97$. Panels show the (k-weighted) power spectrum (top-left) and the (r^2-weighted) correlation function monopole (top-right), quadrupole (bottom-left), and hexadecapole (bottom-right). Curves are calculated with CAMB [33] (thick,red) and using ref. [34] (light,blue) with solid curves showing the full cosmological model and dotted (dashed) curves showing the corresponding CAMB "sideband" ("no-wiggles" of ref. [34]) smooth model.

2.2.1 Peak decomposition

We expect a Lyman-α analysis to distort the measured broadband shape of the correlation function, so our goal is to only use information from the localized peak near $r \simeq 110$ Mpc/h to measure BAO parameters. To achieve this goal, it is useful to decompose ξ_{cosmo} into separate "peak" and "smooth" (or "no-wiggles") components. Following ref. [34], we can isolate the oscillations in a multiplicative term of the effective baryon transfer function $T_b(k)$ (see equation (16) in ref. [34]), which leads to a rather complicated decomposition for the correlation function. Instead, we adopt the unphysical but more tractable decomposition

$$\tilde{\xi}_{\ell,cosmo}(r, z_0) = \tilde{\xi}_{\ell,smooth}(r, z_0) + \tilde{\xi}_{\ell,peak}(r, z_0) \tag{2.15}$$

with the understanding that this model is only valid when using parameter values that are sufficiently close to their nominal values to recover a physically plausible $\tilde{\xi}_{\ell,cosmo}$.

Figure 2 shows the smooth $\tilde{P}(k, z_0)$ and corresponding correlation multipoles suggested in ref. [34]: the oscillations are effectively removed, but the corresponding "peak" feature, defined as the difference between the full and "no-wiggles" models, is not well localized in any of the multipoles, with deviations from the dashed curves extending far from the peak. To remedy this problem, we construct an alternate peak model that is explicitly localized, but only for a single linear combination of multipoles (or, equivalently, a single value of μ) because the integrals in eq. (2.14) effectively spread the peak to all scales above and/or below
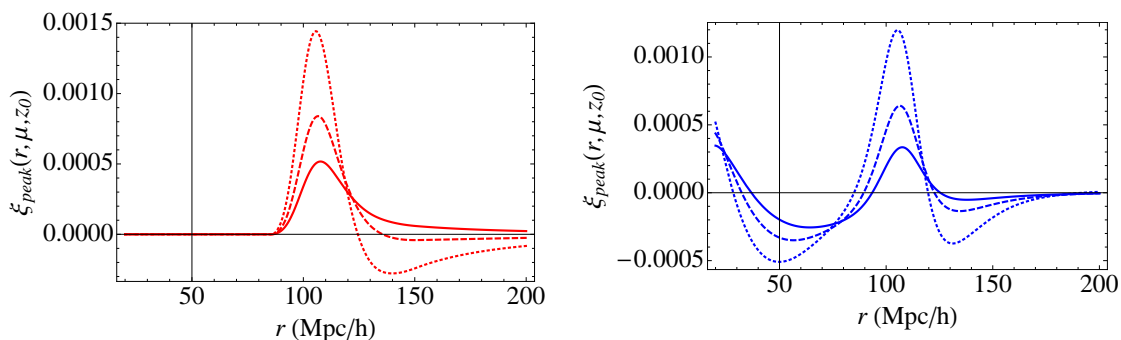
Figure 3. Cosmological peak models calculated using the CAMB "sideband" method described in the text (left, red) and the "no-wiggles" method of ref. [34] (right, blue) described in the text. Curves show $\mu = 0.4$ (solid), 0.7 (dashed), and 1.0 (dotted). There is no r^2 weighting applied here.

(depending on the choice of r_0) for other linear combinations. Our CAMB "sideband" smooth model is constructed as follows: we first isolate a localized peak in the CAMB prediction [33] of figure 2 by simultaneously fitting the regions 50-86 and 150-190 Mpc/h of the monopole to the form

$$\xi_{0,\text{fit}}(r, z_0) = \sum_{j=-3}^{+1} c_j r^j \ . \tag{2.16}$$

Next, we replace $\tilde{\xi}_{0,\text{cosmo}}$ with $\xi_{0,\text{fit}}$ in the region 86-150 Mpc/h to obtain $\tilde{\xi}_{0,\text{smooth}}$. We then calculate $\tilde{\xi}_{2,\text{peak}}$ and $\tilde{\xi}_{4,\text{peak}}$ using eq. (2.14) with $\tilde{\xi}_{\ell,\text{peak}}(r_0) = 0$ at $r_0 = 0$. Finally, we calculate

$$\tilde{\xi}_{\ell,\text{smooth}}(r, z_0) = \tilde{\xi}_{\ell,\text{cosmo}}(r, z_0) - \tilde{\xi}_{\ell,\text{peak}}(r, z_0) \ . \tag{2.17}$$

The resulting CAMB "sideband" smooth model is shown in figure 2.

By construction, our CAMB peak multipoles are exactly zero below 86 Mpc/h, but the peak spreads to large scales for $\ell = 2, 4$, as required by eq. (2.14). Note that although it would be possible to specify independent localized peaks for each multipole, this is unphysical (within the framework of ref. [32]) and does not, in general, reproduce the evolution of the peak shape in $\xi(r, \mu, z_0)$ with μ implied by the constraint of a single underlying $\tilde{P}(k, z_0)$, as illustrated in figure 3. For example, the constrained evolution leads to a percent-level shift in the peak position as a function of μ shown in figure 4 for the two models used here. Note that this shift is essentially the same for both models, despite the rather different peak shapes shown in figure 3.

2.2.2 Nonlinear 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 [35]:

$$\tilde{P}_{\text{NL}}(k, \mu_k, z_0) = \exp(-k^2 \Sigma^2(\mu_k)/2) \cdot \tilde{P}(k, z_0) \tag{2.18}$$

where

$$\Sigma^2(\mu_k) = \mu_k^2 \Sigma_\parallel^2 + (1 - \mu_k^2) \Sigma_\perp^2 \ . \tag{2.19}$$

In general, this approach breaks the decomposition, eq. (2.10), of $\tilde{P}_\ell(k, z_0)$ into separate β- and k-dependent factors and requires that the integrals of eqs. (2.8)–(2.9) be re-evaluated for
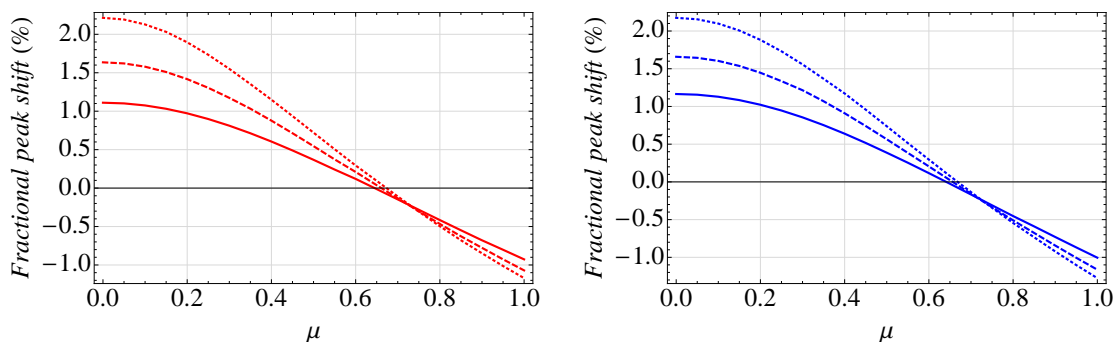
Figure 4. Evolution of the peak position with μ for the cosmological peak models shown in figure 3. Curves show $\beta = 1.0$ (solid), 1.4 (dashed), and 1.8 (dotted). Fractional shifts are measured relative to the position of the monopole peak for each model.

each value of β. However, since we expect $\Sigma \simeq 5$ Mpc/h, compared with an expected peak full-width half-maximum of $\simeq 25$ Mpc/h, we can approximate for $\beta \simeq \beta_0$:

$$P_{\ell,NL}(k, z_0) \simeq \exp(-k^2 \Sigma_\ell^2(\beta_0)/2) \cdot \tilde{P}_\ell(k, z_0) \tag{2.20}$$

with

$$\Sigma_\ell^2(\beta) \equiv f_\ell(\beta) \cdot \Sigma_\parallel^2 + (1 - f_\ell(\beta)) \cdot \Sigma_\perp^2 \tag{2.21}$$

and

$$f_\ell(\beta) \equiv \frac{\int_{-1}^{+1} \mu_k^2 \left(1 + \beta \mu_k^2\right)^2 L_\ell(\mu_k)\, d\mu_k}{\int_{-1}^{+1} \left(1 + \beta \mu_k^2\right)^2 L_\ell(\mu_k)\, d\mu_k} = \begin{cases} \frac{35 + 42\beta + 15\beta^2}{105 + 70\beta + 21\beta^2} & \ell = 0 \\ \frac{7 + 12\beta + 5\beta^2}{14\beta + 6\beta^2} & \ell = 2 \\ \frac{15}{11} + \frac{2}{\beta} & \ell = 4 \end{cases} \tag{2.22}$$

This approximation effectively models anisotropic broadening using different amounts of isotropic broadening for each multipole. The resulting correlation function multipoles $\xi_{\ell,NL}$, calculated with eq. (2.8), can then be substituted in eq. (2.12).

Taking a fiducial value of $\beta_0 = 1.4$, we calculate $f_0 = 0.505$, $f_2 = 1.07$, and $f_4 = 2.79$. At a redshift $z = 2.4$, we expect $\Sigma_\parallel \simeq 6.41$ Mpc/h and $\Sigma_\perp \simeq 3.26$ Mpc/h, so that $\Sigma_0 = 5.10$ Mpc/h, $\Sigma_2 = 6.58$ Mpc/h, and $\Sigma_4 = 9.79$ Mpc/h. Figure 5 compares the resulting approximate models with exact calculations for a range of β values. Note that we are neglecting the non-zero even multipoles $\ell = 6, 8, \ldots$ that are introduced by anisotropic non-linear broadening with this approximation, but Figure 5 shows that most of the BAO peak feature has already been smoothed out by $\ell = 4$. We also neglect any redshift evolution of the non-linear broadening parameters, since this evolution would be a second order effect on what is already a small correction to the linear theory for the purposes of measuring the BAO feature.

In order to achieve a self-consistent decomposition of each correlation-function multipole into peak + smooth components with non-linear broadening effects included, we first broaden each $\xi_{\ell,\text{peak}}(r, z_0)$ by Σ_ℓ to obtain $\xi_{\ell,\text{peak},NL}(r, z_0)$. Broadening by Σ in this context implies

$$P(k) \rightarrow P(k) \exp(-k^2 \Sigma^2/2) \tag{2.23}$$

which transforms to

$$\xi_0(r) \rightarrow \int_0^\infty ds\, \frac{s}{r} \left[G(r - s, \Sigma) - G(r + s, \Sigma)\right] \xi_0(s)\,, \tag{2.24}$$
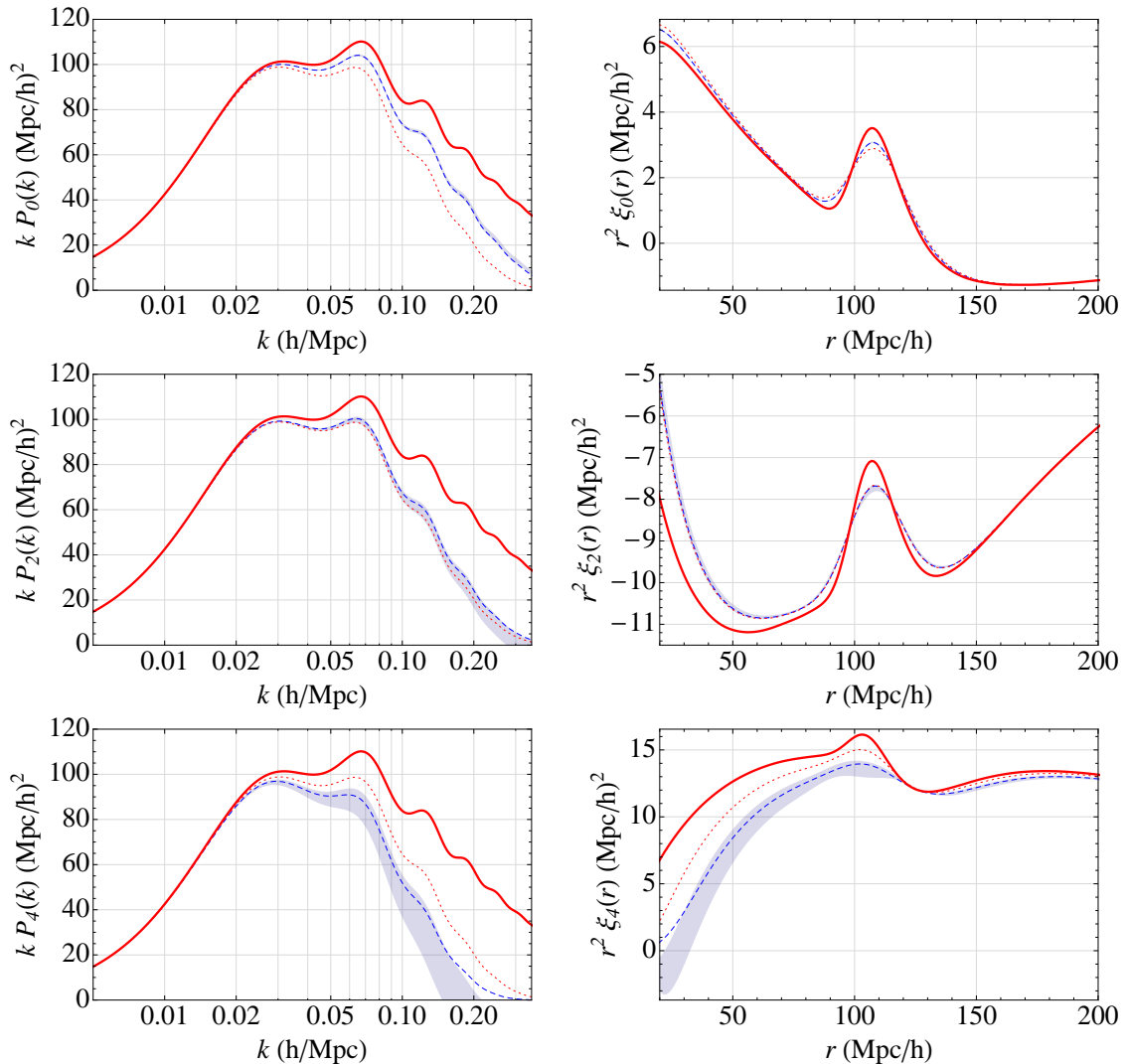
Figure 5. Effects of anisotropic non-linear broadening implemented with eq. (2.18) and applied to $z_0 = 2.25$ linear CAMB predictions [33] using $\Sigma_\parallel = 6.41$ Mpc/h and $\Sigma_\perp = 3.26$ Mpc/h. Curves show no broadening (thick red, same as curves in figure 2), isotropic broadening (dotted red) by $(\Sigma_\parallel^2 + \Sigma_\perp^2)^{1/2}/2 = 5.09$ Mpc/h, the approximate anisotropic model described in the text (dashed blue) with $\beta_0 = 1.4$, and the envelope of full anisotropic calculations (light blue shaded) for $\beta = 0.5$-2.5. Left-hand panels show the k-weighted multipoles $P_{\ell,NL}(k, z_0)$ with $b^2 C_\ell(\beta)$ divided out, for $\ell = 0$ (top), $\ell = 2$ (middle), and $\ell = 4$ (bottom). Right-hand panels show the corresponding r^2-weighted correlation function multipoles $\xi_\ell(r, z_0)$.

where G is the normalized one-dimensional Gaussian

$$G(t, \Sigma) \equiv \frac{1}{\sqrt{2\pi}\Sigma} \exp\left(-\frac{t^2}{2\Sigma^2}\right) . \tag{2.25}$$

Note that when $r \gg \Sigma$, we recover the expected convolution

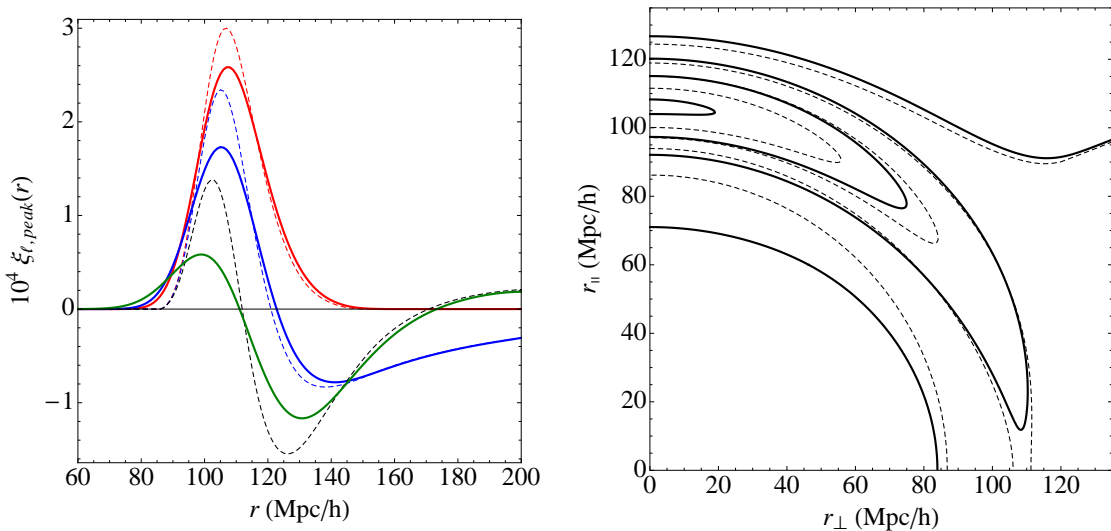$$\xi_0(r) \to \int_0^\infty ds\, G(r - s, \Sigma)\, \xi_0(s) \tag{2.26}$$

Figure 6. Effects of anisotropic non-linear broadening on peak templates, implemented using the approximations described in the text and applied to $z_0 = 2.25$ linear CAMB predictions [33]. Solid (dashed) curves show peak templates with (without) non-linear effects. From top to bottom, curves in the left-hand panel show the $\ell = 0$ (red), 2 (blue), and 4 (green) correlation multipoles, with $b^2 C_\ell(\beta)$ divided out, without any r-weighting and scaled by 10^4. All templates are identically zero for $r < 60$ Mpc/h, by construction. The right-hand panel shows equally spaced contours of $\xi_{\rm peak}(r, \mu)$ calculated with $\beta = 1.4$, with the outer contour corresponding to zero correlation.

to a good approximation. Figure 6 shows the broadened multipole peaks based on our linear CAMB templates. To derive the corresponding smooth templates, we either use the linear smooth templates, or else we calculate non-linear smooth templates by subtracting the broadened peak templates from the broadened multipoles:

$$\xi_{\ell,\mathrm{smooth},NL} = \xi_{\ell,NL} - \tilde{\xi}_{\ell,\mathrm{peak},NL} \ . \tag{2.27}$$

In the first case, we are only applying non-linear broadening to the peak feature, which introduces an unphysical distinction between peak and smooth components. In the second case we are broadening the full correlation function, which removes the distinction but applies unphysical filtering of small scale structure. A better model is probably somewhere in between so we consider both alternatives, and expect that any extracted peak parameters will not depend on this choice. Figure 7 compares both approaches.

2.2.3 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

$$p(z) = p_0 \left(\frac{1+z}{1+z_0} \right)^{\gamma_p} \ . \tag{2.28}$$

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,

$$C_p = \begin{pmatrix} \sigma_0^2 & \rho \sigma_0 \sigma_\gamma \\ \rho \sigma_0 \sigma_\gamma & \sigma_\gamma^2 \end{pmatrix} \tag{2.29}$$
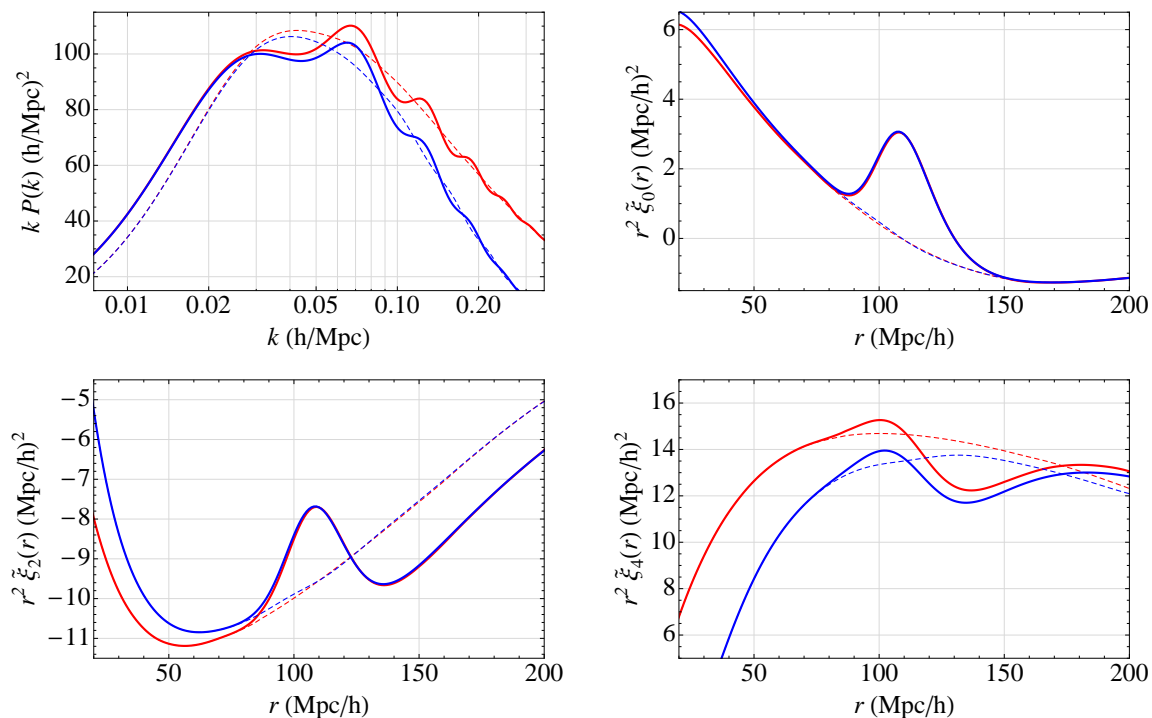
Figure 7. Comparison of two different schemes for incorporating anisotropic non-linear effects into the final templates used for fitting. Panels show the (k-weighted) power spectrum (top-left) and the (r^2-weighted) correlation function monopole (top-right), quadrupole (bottom-left), and hexadecapole (bottom-right). Red curves are calculated with non-linear effects applied to the BAO peak feature only. Blue curves are calculated with non-linear broadening applied at all scales. Both schemes are derived from $z_0 = 2.25$ linear CAMB predictions [33] and use the same broadened peaks shown in figure 6. The combined peak + smooth (smooth only) templates are represented with solid (dashed) curves.

the variance of $p(z)$ with $z \simeq z_0$ is given by

$$\sigma_{p(z)}^2 = J \cdot C_p \cdot J^t \tag{2.30}$$

where

$$J = \left(\frac{\partial p(z)}{\partial p_0}, \frac{\partial p(z)}{\partial \gamma_p} \right) \tag{2.31}$$

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

$$\log \left(\frac{1+z}{1+z_0} \right) = -b - \rho a + \sqrt{b^2 - (1-\rho^2)a^2} \tag{2.32}$$

with

$$a \equiv \frac{\sigma_0}{p_0 \sigma_\gamma} \quad , \quad b \equiv \frac{1}{2\gamma_p} . \tag{2.33}$$

2.2.4 Scale factors

When fitting our model to data, we allow for an overall relative normalization factor $a_{\text{peak}} \simeq 1$ as well as a general coordinate transform, $r' = r'(r, \mu, z)$ and $\mu' = \mu(r, \mu, z)$, that allows for

possible small differences[4] between the fiducial and actual cosmologies:

$$\xi_{\text{cosmo}}(r,\mu,z) \to a_{\text{peak}} \cdot \left[\xi(r',\mu',z) - \xi_{\text{smooth}}(r',\mu',z) \right] + \xi_{\text{smooth}}(r'',\mu'',z) \,. \tag{2.34}$$

We consider two options for the cosmological broadband ξ_{smooth}: we either apply the same transform as for the peak ($r'' = r'$), or else we keep it fixed ($r'' = r$) so that only the peak is transformed. Neither approach is exact when considering variations of the cosmological parameters around our fidicual model, but we find that decoupling the peak from the cosmological broadband ($r'' = r$) better localizes the separations contributing to a BAO measurement to the peak region (compare figures 16 and 19), so is preferred when broadband distortion is not fully under control. The corresponding k-space transforms are defined by

$$k'(k,\mu_k,z) \cdot r'(r,\mu,z) = k \cdot r \quad , \quad \mu_k'(k,\mu_k,z) \cdot \mu(r,\mu,z) = \mu_k \cdot \mu \,. \tag{2.35}$$

Including a μ-dependence in the coordinate transform also enables us to study the constraining power of our data separately along and transverse to the line of sight. Similarly, a z-dependence allows us to determine the redshift at which our results are best measured using eq. (2.32).

For our baseline isotropic model, we use

$$r_{\text{iso}}'(r,\mu,z) = \alpha_{\text{iso}}(z) \cdot r$$
$$\mu_{\text{iso}}'(r,\mu,z) = \mu \,. \tag{2.36}$$

For our baseline anisotropic model, we decouple the line-of-sight ($r_\parallel \to \alpha_\parallel r_\parallel$) and transverse ($r_\perp \to \alpha_\perp r_\perp$) scales using

$$r_{\text{ani}}'(r,\mu,z) = \alpha_{\text{ani}}(\mu,z) \cdot r$$
$$\mu_{\text{ani}}'(r,\mu,z) = \frac{\alpha_\parallel(z)}{\alpha_{\text{ani}}(\mu,z)} \cdot \mu \tag{2.37}$$

with

$$\alpha_{\text{ani}}(\mu,z) = \sqrt{\alpha_\parallel^2(z)\mu^2 + \alpha_\perp^2(z)(1-\mu^2)} \,. \tag{2.38}$$

When the parameters of this anisotropic model are determined by measuring the physical coordinate separation scales $\Delta v_{\text{BAO}}(z)$ and $\Delta\theta_{\text{BAO}}(z)$ corresponding to the comoving BAO scale r_{BAO} at some redshift z, related by

$$\Delta v_{\text{BAO}}(z) \simeq r_{\text{BAO}}\, H(z)/(1+z) \tag{2.39}$$
$$\Delta\theta_{\text{BAO}}(z) = r_{\text{BAO}}/D_A(z) \,, \tag{2.40}$$

then the expected best fit values of $\alpha_\parallel$ and $\alpha_\perp$ satisfy

$$\alpha_\parallel \cdot \frac{1+z}{H_{\text{fid}}(z)}\Delta v_{\text{BAO}}(z) \simeq r_{\text{BAO,fid}} \tag{2.41}$$
$$\alpha_\perp \cdot D_{A,\text{fid}}(z)\Delta\theta_{\text{BAO}}(z) = r_{\text{BAO,fid}} \tag{2.42}$$

where $r_{\text{BAO,fid}}(z)$ is the comoving BAO scale predicted by the assumed fiducial cosmology. Combining these results, the best-fit anisotropic scale factors measure

$$\alpha_\parallel(z) = \frac{r_{\text{BAO,fid}}}{r_{\text{BAO}}} \cdot \frac{H_{\text{fid}}(z)}{H(z)} \quad , \quad \alpha_\perp(z) = \frac{r_{\text{BAO,fid}}}{r_{\text{BAO}}} \cdot \frac{D_A(z)}{D_{A,\text{fid}}(z)} \,. \tag{2.43}$$
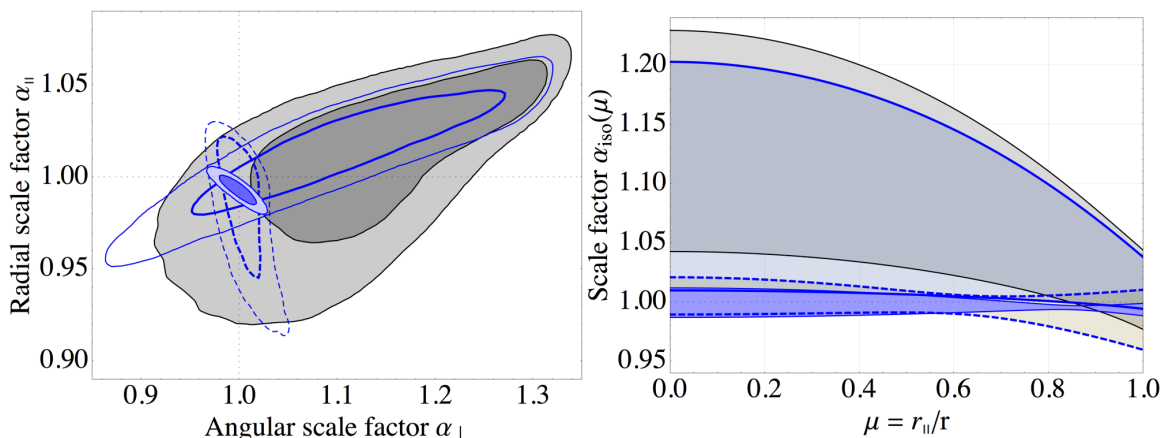
Figure 8. Predictions at $z = 2.4$ for the anisotropic scale factors (left) and $\alpha_{\rm ani}(\mu)$ (right) from WMAP9 data [5]. Filled blue contours show ΛCDM predictions, while solid (dashed) curves show the effects of allowing the curvature Ω_k (dark energy equation of state parameter w_0) to vary in the model. The gray filled contours show the combined effects of varying both Ω_k and w_0. Contours in the left-hand plot enclose 68% and 95% while the bands in the right-hand plot are ± 1 standard deviation. The fiducial cosmology $\alpha_{\|} = \alpha_{\perp} = \alpha(\mu) = 1$ is indicated by dotted lines.

The distributions of these ratios predicted at $z = 2.4$ by WMAP9 [4] observations of the CMB are shown in figure 8.

We can evaluate a transformed correlation function ξ' in the original coordinates (r, μ, z) using eq. (2.12) (with z-dependencies omitted for clarity)

$$
\begin{aligned}
\xi'(r, \mu) &= \xi_{\rm cosmo}(r'(r, \mu), \mu'(r, \mu)) \\
&= b^2 \sum_{\ell=0,2,4} C_\ell(\beta)\, L_\ell(\mu'(r, \mu))\, \tilde{\xi}_{\ell,\rm cosmo}(r'(r, \mu)) \ .
\end{aligned}
\tag{2.44}
$$

The corresponding multipoles in the original coordinates are then given by

$$
\xi'_\ell(r) \equiv \frac{2\ell+1}{2} \int_{-1}^{+1} \xi'(r, \mu)\, L_\ell(\mu) d\mu
\tag{2.45}
$$

$$
= \frac{2\ell+1}{2} b^2 \sum_{\ell'=0,2,4} C_\ell(\beta) \int_{-1}^{+1} \tilde{\xi}_{\ell,\rm cosmo}(r'(r, \mu))\, L_\ell(\mu'(r, \mu)) \cdot L_\ell(\mu) d\mu \ .
\tag{2.46}
$$

In the case of isotropic distortion, we can simply replace r with $\alpha_{\rm iso}(z) \cdot r$ in the multipoles obtained with eq. (2.13). However, anisotropic distortions mix multipoles and lead to more complicated expressions. For small $\epsilon \equiv (\alpha_{\|} - \alpha_{\perp})/(\alpha_{\|} + \alpha_{\perp})$ and arbitrary $\alpha \equiv (\alpha_{\|} + \alpha_{\perp})/2$, we have

$$
\alpha_{\|} = \alpha(1 + \epsilon) \quad , \quad \alpha_{\perp} = \alpha(1 - \epsilon)
\tag{2.47}
$$

and corresponding transforms

$$
\begin{aligned}
r'_{\rm ani}(r, \mu) &= \alpha r \left[1 + 2\epsilon(2\mu^2 - 1) + \epsilon^2\right]^{1/2} &\simeq \alpha r \left[1 + (2\mu^2 - 1) \cdot \epsilon\right] \\
\mu'_{\rm ani}(r, \mu) &= \mu(1 + \epsilon) \left[1 + 2\epsilon(2\mu^2 - 1) + \epsilon^2\right]^{-1/2} &\simeq \mu \left[1 + 2(1 - \mu^2) \cdot \epsilon\right] \ .
\end{aligned}
\tag{2.48}
$$

---

[4]In case there is evidence for large differences, the analysis should be repeated with a fiducial cosmology that better matches the data.

Using these approximations, and

$$\tilde{\xi}_{\ell,\text{cosmo}}(r'(r,\mu)) \simeq \tilde{\xi}_{\ell,\text{cosmo}}(\alpha r) + \epsilon \cdot \alpha r \cdot \partial_r \tilde{\xi}_{\ell',\text{cosmo}}(\alpha r) \cdot (2\mu^2 - 1) \,, \tag{2.49}$$

we obtain distorted multipoles:

$$\xi'_\ell(r) \simeq b^2 \sum_{\ell'=0,2,4} C_{\ell'}(\beta) \left\{ \tilde{\xi}_{\ell',\text{cosmo}}(\alpha r) \cdot A_{\ell,\ell'}(\epsilon) + \epsilon \cdot \alpha r \cdot \partial_r \tilde{\xi}_{\ell',\text{cosmo}}(\alpha r) \cdot B_{\ell,\ell'} \right\} \tag{2.50}$$

where

$$A_{\ell,\ell'}(\epsilon) \equiv \frac{2\ell+1}{2} \int_{-1}^{+1} L_{\ell'}(\mu + 2\epsilon\mu(1-\mu^2)) L_\ell(\mu) d\mu \tag{2.51}$$

$$B_{\ell,\ell'} \equiv \frac{2\ell+1}{2} \int_{-1}^{+1} (2\mu^2 - 1) L_{\ell'}(\mu) L_\ell(\mu) d\mu \,. \tag{2.52}$$

We find the following non-zero coefficients, to first order in ϵ:

$$u = \arctan x \qquad dv = 1\,dx \tag{2.53}$$

$$du = \frac{1}{1+x^2} dx \qquad v = x. \tag{2.54}$$

$$A_{0,0} = 1, \qquad B_{0,0} = -\frac{1}{3}, \qquad A_{0,2} = \frac{4}{5}\epsilon, \qquad B_{0,2} = \frac{4}{15} \tag{2.55}$$

$$B_{2,0} = \frac{4}{3}, \qquad A_{2,2} = 1 + \frac{4}{7}\epsilon, \quad B_{2,2} = \frac{1}{21}, \qquad A_{2,4} = \frac{40}{21}\epsilon, \quad B_{2,4} = \frac{8}{21} \tag{2.56}$$

$$A_{4,2} = -\frac{48}{35}\epsilon, \qquad B_{4,2} = \frac{24}{35}, \qquad A_{4,4} = 1 + \frac{40}{77}\epsilon, \quad B_{4,4} = \frac{1}{77} \tag{2.57}$$

$$A_{6,4} = -\frac{80}{33}\epsilon, \qquad B_{6,4} = \frac{20}{33} \,. \tag{2.58}$$

Note that an $\ell = 6$ term must be included in order to match eq. (2.46) to first order in ϵ, but is numerically negligible. On the other hand, second-derivative terms in the Taylor expansion (2.49) are formally $\mathcal{O}(\epsilon^2)$ but are numerically significant when

$$\epsilon \gtrsim \frac{2|\partial_r \tilde{\xi}_\ell(\alpha r)|}{\alpha r |\partial_r^2 \tilde{\xi}_\ell(\alpha r)|} \,, \tag{2.59}$$

which is guaranteed to occur at the peak of the BAO feature and at the transition to a rising smooth broadband below the peak. Figure 9 shows examples of anisotropic coordinate transforms based on eqs. (2.44) and (2.50). Note that eq. (2.50) is only exact when $\mu^2 = 1/2$ ($r_\parallel = r_\perp$) and grows less accurate when moving away from the diagonal, while eq. (2.44) is exact and therefore preferred when fitting for the anisotropic scale parameters.

2.3 Broadband distortion models

A Lyman-α analysis measures the correlation function

$$\xi_{ij} \equiv \langle \delta_i \delta_j \rangle - \langle \delta_i \rangle \langle \delta_j \rangle = \frac{\langle F_i F_j \rangle - \langle F_i \rangle \langle F_j \rangle}{\overline{F}(\lambda_i) \overline{F}(\lambda_j)} \tag{2.60}$$

for the overdensity proxy

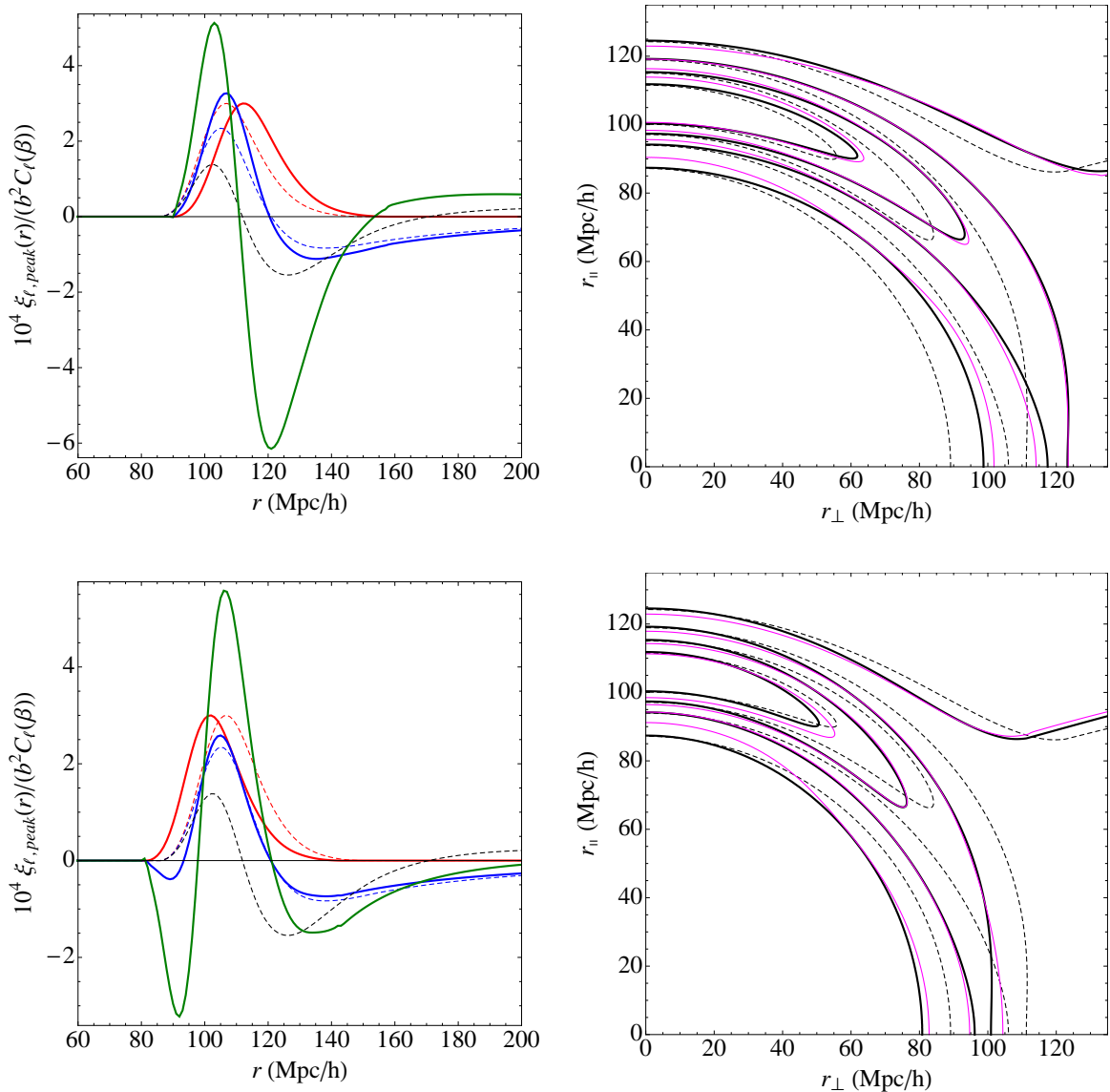$$\delta_i = \frac{F_i - \overline{F}(\lambda_i)}{\overline{F}(\lambda_i)} \tag{2.61}$$

– 13 –

Figure 9. Examples of anisotropic coordinate transforms applied to the correlation function, using $\beta = 1.4$. The top row is calculated with $\alpha = 0.95$, $\epsilon = 0.05$ and the bottom row with $\alpha = 1.05$, $\epsilon = -0.05$, both chosen to match the fiducial cosmology at $\mu = 1$, where the expected sensitivity is largest for $\beta = 1.4$. Left-hand plots show the normalized distorted correlation multipoles $10^4 \xi_\ell(r)/(b^2 C_\ell(\beta))$ calculated with eq. (2.50) (solid curves) or without any coordinate scaling (dashed curves). From top to bottom, curves show $\ell = 0$ (red), 2 (blue), and 4 (green). Right-hand plots show equally spaced contours of $\xi(r, \mu, z_0)$ with the outer contours corresponding to zero correlation. Thick solid curves are calculated with eq. (2.44) and dashed curves are calculated without any coordinate scaling. Thin magenta curves show the results of combining the transformed multipoles (solid curves) of the left-hand plots using eq. (2.8), and demonstrate the level of accuracy provided by the first-order approximations described in the text.

in pixels i measured at rest absorption wavelengths λ_i, where

$$F_i = \frac{f_i - f_{\text{sky},i}}{f_{\text{cont},i}} \tag{2.62}$$

is the transmitted flux fraction, relative to the sky flux level $f_{\mathrm{sky},i}$ and normalized to an assumed quasar continuum flux level $f_{\mathrm{cont},i}$, and $\overline{F}(\lambda_i)$ is an assumed mean transmission fraction at λ_i.

To investigate possible sources of broadband distortion, write

$$f_i = \tilde{f}_i + \epsilon_i$$
$$f_{\mathrm{sky},i} = \tilde{f}_{\mathrm{sky},i} + \epsilon_{\mathrm{sky},i} + s(\lambda_i) \tag{2.63}$$
$$f_{\mathrm{cont},i} = \tilde{f}_{\mathrm{cont},i} + c_i(\mathbf{d})$$

where tilde quantities are true values, ϵ_i and $\epsilon_{\mathrm{sky},i}$ are (zero-mean) noise sources, $s(\lambda_i)$ accounts for any wavelength-dependent residual sky-subtraction bias (as observed in DR9), and $c_i(\mathbf{d})$ describes the continuum modeling error that, in general, depends on the full vector of pixel measurements

$$\mathbf{d} = \{\lambda_j, f_j\}_{i=1}^{N_{\mathrm{pixels}}} \ . \tag{2.64}$$

Combining eqs. (2.62) and (2.63), we find:

$$F_i = \frac{\tilde{f}_i - \tilde{f}_{\mathrm{sky},i} - s(\lambda_i) + \epsilon_i - \epsilon_{\mathrm{sky},i}}{\tilde{f}_{\mathrm{cont},i} + c_i(\mathbf{d})} \tag{2.65}$$
$$\simeq \tilde{F}_i \left[1 - S_i - C_i + E_i \right]$$

with

$$\tilde{F}_i \equiv \frac{\tilde{f}_i - \tilde{f}_{\mathrm{sky},i}}{\tilde{f}_{\mathrm{cont},i}} \quad , \quad S_i \equiv \frac{s(\lambda_i)}{\tilde{f}_i - \tilde{f}_{\mathrm{sky},i}} \quad , \quad C_i \equiv \frac{c_i(\mathbf{d})}{\tilde{f}_{\mathrm{cont},i}} \quad , \quad E_i \equiv \frac{\epsilon_i - \epsilon_{\mathrm{sky},i}}{\tilde{f}_i - \tilde{f}_{\mathrm{sky},i}} \quad , \tag{2.66}$$

where we have assumed that $C_i \ll 1$. Averaging over noise and cosmic realizations, we find:

$$\xi_{ij} \simeq r_{ij} \left[\tilde{\xi}_{ij} \left(1 + A_{ij} \right) + B_{ij} \right] \tag{2.67}$$

where

$$\tilde{\xi}_{ij} \equiv \frac{\langle \tilde{F}_i \tilde{F}_j \rangle - \langle \tilde{F}_i \rangle \langle \tilde{F}_j \rangle}{\langle \tilde{F}_i \rangle \langle \tilde{F}_j \rangle} \tag{2.68}$$

is the undistorted cosmological correlation function we seek to measure, $r_{ij} \simeq 1$ describes the effects of any discrepancy between the assumed and true mean transmitted flux fraction via

$$r_{ij} \equiv \frac{\langle \tilde{F}_i \rangle \langle \tilde{F}_j \rangle}{\overline{F}(\lambda_i) \overline{F}(\lambda_j)} \ , \tag{2.69}$$

A_{ij} is a multiplicative distortion of the true correlation function

$$A_{ij} \equiv \langle S_i S_j \rangle + \langle C_i C_j \rangle + \langle E_i E_j \rangle + \langle S_i \rangle \langle C_j \rangle + \langle S_j \rangle \langle C_i \rangle - \langle S_i \rangle - \langle S_j \rangle - \langle C_i \rangle - \langle C_j \rangle \ , \tag{2.70}$$

and B_{ij} is an additive distortion

$$B_{ij} \equiv \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle + \langle C_i C_j \rangle - \langle C_i \rangle \langle C_j \rangle + \langle E_i E_j \rangle \ , \tag{2.71}$$

where we have used $\langle E_i \rangle = 0$ and assumed that S, C, and E are mutually uncorrelated.

Equation (2.67) demonstrates that, in general, the measured correlation function can be systematically distorted by multiplicative and additive effects. The actual correlation function used for fitting is a weighted sum of pixels in bins of (r, μ, z). Any multiplicative distortion is due to a mismatch between the assumed $\overline{F}(\lambda)$ and the true mean transmitted flux for pixels at the same wavelength λ, whereas additive distortion is due to correlated continuum fit errors.

We model broadband distortion using the following parametrization that combines the multiplicative (B_m) and additive (B_a) effects described above:

$$\xi(r, \mu, z) = \xi_{\rm cosmo}(r, \mu, z) \cdot [1 + B_m(r, \mu, z)] + B_a(r, \mu, z) \cdot \left(\frac{1+z}{1+z_0}\right)^{\gamma_b 2} \tag{2.72}$$

with (B_x represents either B_m or B_a)

$$B_x(r, \mu, z) = \sum_{i=i_{\rm min}}^{i_{\rm max}} \sum_{j=j_{\rm min}}^{j_{\rm max}} \sum_{n=n_{\rm min}}^{n_{\rm max}} b_{i,j,n}^{(x)} \cdot \left(\frac{r}{r_0} - \theta_i\right)^i \cdot L_j(\mu) \cdot \left(\frac{1+z}{1+z_0}\right)^n, \tag{2.73}$$

where, nominally, $r_0 = 100$ Mpc/h and $z_0 = 2.25$, and

$$\theta_i = \begin{cases} 0 & i \leq 0 \\ 1 & i > 0 \end{cases}. \tag{2.74}$$

The integers $i_{\rm min} \leq i_{\rm max}$, $0 \leq j_{\rm min} \leq j_{\rm max}$, and $n_{\rm min} \leq n_{\rm max}$ determine the number of free parameters, and we normally restrict j to even values. The $\ell = 0, 2, 4$ multipoles of the distorted correlation function are given by (arguments of (r, z) are suppressed for clarity):

$$\xi_0 = \xi_{0,{\rm cosmo}} \cdot (1 + B_{0,m}) + \xi_{2,{\rm cosmo}} \cdot \frac{1}{5} B_{2,m} + \xi_{4,{\rm cosmo}} \cdot \frac{1}{9} B_{4,m} + B_{0,a} \tag{2.75}$$

$$\xi_2 = \xi_{0,{\rm cosmo}} \cdot (1 + B_{2,m}) + \xi_{2,{\rm cosmo}} \cdot \left(B_{0,m} + \frac{2}{7} B_{2,m} + \frac{2}{7} B_{4,m}\right) +$$
$$\xi_{4,{\rm cosmo}} \cdot \left(\frac{2}{7} B_{2,m} + \frac{100}{693} B_{4,m} + \frac{25}{143} B_{6,m}\right) + B_{2,a} \tag{2.76}$$

$$\xi_4 = \xi_{0,{\rm cosmo}} \cdot (1 + B_{4,m}) + \xi_{2,{\rm cosmo}} \cdot \left(\frac{18}{35} B_{2,m} + \frac{20}{77} B_{4,m} + \frac{45}{143} B_{6,m}\right) +$$
$$\xi_{4,{\rm cosmo}} \cdot \left(B_{0,m} + \frac{20}{77} B_{2,m} + \frac{162}{1001} B_{4,m} + \frac{20}{143} B_{6,m} + \frac{490}{2431} B_{8,m}\right) + B_{4,a} \tag{2.77}$$

with distortion multipoles given by

$$B_{\ell,x}(r, z) = \sum_{i=i_{\rm min}}^{i_{\rm max}} \sum_{n=n_{\rm min}}^{n_{\rm max}} b_{i,\ell,n}^{(x)} \cdot \left(\frac{r}{r_0} - \theta_i\right)^i \cdot \left(\frac{1+z}{1+z_0} - 1\right)^n. \tag{2.78}$$

For the purposes of fitting DR9, we chose six different parameter configurations for broadband distortion, as summarized in table 1 and plotted in figure 10.

2.4 Interpolated models

For the purposes of visualization and studying non-cosmological signatures in data, it is useful to fit using simple interpolated models with minimal assumptions. We describe here two

	B_m			B_a		
Name	i	j	n	i	j	n
BB1	-	-	-	0,1,2	0,2,4	0
BB2	-	-	-	-2,-1,0	0,2,4	0
BB3	-	-	-	0,1,2	0,2	0,1
BB4	0,1,2	0,2,4	0	-	-	-
BB5	0,1,2	0,2	0,1	-	-	-
BB6	0,1	0,2,4	0	0,1	0,2,4	0

Table 1. Broadband distortion models used to fit DR9. Models are labeled BB1-6 and columns show the range of indices in eq. (2.73) used in the multiplicative (B_m) and additive (B_a) components. Dashes indicate that a component is not used.
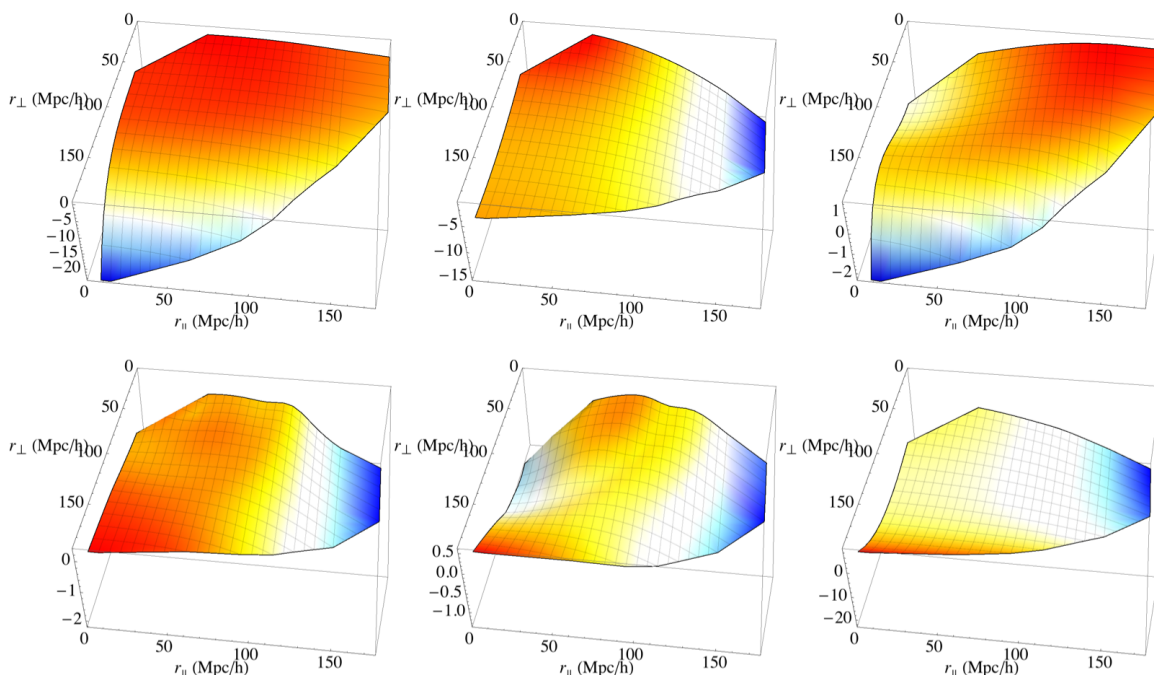
Figure 10. Best fits of the six distortion models BB1-6 described in table 1 to DR9. Plots show the r^2-weighted distortion $\xi(r, \mu, z) - \xi_{\text{cosmo}}(r, \mu, z)$ for $z = 2.5$. Models BB1-3 are on the top row of plots, left to right, and BB4-6 are on the bottom row. The faint negative impression of the BAO feature visible for BB4 and BB5 reflects the fact that they model purely multiplicative distortions.

such models: the first treats the correlation multipoles as (possibly independent) arbitrary interpolations through fixed arbitrary values of the separation, and the second treats corrections to a smoothed power spectrum as (possibly independent) arbitrary interpolations through equally spaced wave numbers covering a limited band.

For our first model, we pick n fixed values of the separation $\{r_j\}$ and model the r^2-weighted correlation multipoles as:

$$r^2 \xi_\ell(r, z) = b_\ell^2(z) \, I_{m,n}(r, z_0; \{r_j\}) \tag{2.79}$$

where $I_{m,n}$ is an m-th order interpolation through the n points $\{r_j\}$ evaluated at separation

r and referenced to the redshift z_0. Comparing with eq. (2.10), we write

$$b_\ell^2(z) = b^2(z) \, C_\ell(\beta(z)) \,, \tag{2.80}$$

so that

$$I_{m,n}(r, z_0) = r^2 \, \tilde{\xi}_{\ell,\text{cosmo}}(r, z_0) \tag{2.81}$$

when $\xi_\ell(r, \mu, z)$ is purely cosmological.

For our second model, we interpolate over a fixed band of k-space and therefore require some assumption about the power outside this band in order to predict r-space correlations. We adopt the following form for the multipoles of the power spectrum:

$$P_\ell(k, z) = b_\ell^2(z) \left[\tilde{P}_{\text{smooth}}(k, z_0) + \Delta P_\ell(k, z_0) \right] \tag{2.82}$$

where $\tilde{P}_{\text{smooth}}(k, z_0)$ is a smoothed fiducial power spectrum, without any BAO features, modulated by functions $\Delta P_\ell(k, z_0)$ that are zero outside some interval $(k_{\text{lo}}, k_{\text{hi}})$, both referenced to z_0. The correlation multipoles corresponding to eq. (2.82) are obtained via eq. (2.8):

$$\xi_\ell(r, z) = b_\ell^2(z) \left[\tilde{\xi}_{\ell,\text{smooth}}(r, z_0) + \Delta \xi_\ell(r, z_0) \right] \tag{2.83}$$

with

$$\Delta \xi_\ell(r, z_0) = \frac{i^\ell}{2\pi^2} \int_{k_{\text{lo}}}^{k_{\text{hi}}} k^2 j_\ell(kr) \Delta P_\ell(k, z_0) \, dk \,. \tag{2.84}$$

We parameterize the k-weighted modulation as a linear combination of m-th order B-spline functions B_m:

$$k \cdot \Delta P_\ell(k, z_0) = \sum_{j=0}^{n-m-2} b_{\ell,j} B_m \left(\frac{k - k_j}{(m+1)\Delta k} \right) \tag{2.85}$$

with n uniformly spaced knots k_j at:

$$k_j = k_{\text{lo}} + j \cdot \Delta k \quad, \quad \Delta k = \frac{k_{\text{hi}} - k_{\text{lo}}}{n - 1} \,. \tag{2.86}$$

Using the property that $B_m(t)$ is only non-zero for $0 \le t \le 1$, we can write

$$\Delta \xi_\ell(r, z_0) = \frac{i^\ell}{2\pi^2} \sum_{j=0}^{n-m-2} b_{\ell,j} \, E_{\ell,m}(r; k_j, \Delta k) \tag{2.87}$$

where

$$E_{\ell,m}(r; k_j, \Delta k) \equiv (m+1)\Delta k \int_0^1 k(t) \, j_\ell(k(t)r) \, B_m(t) dt \tag{2.88}$$

with

$$k(t) = k_j + (m+1)\Delta k \cdot t \tag{2.89}$$

can be solved analytically in terms of trigonometric functions and the sine integral. Figure 11 shows examples of $E_{\ell,m}$ for cubic ($m = 3$) B splines.
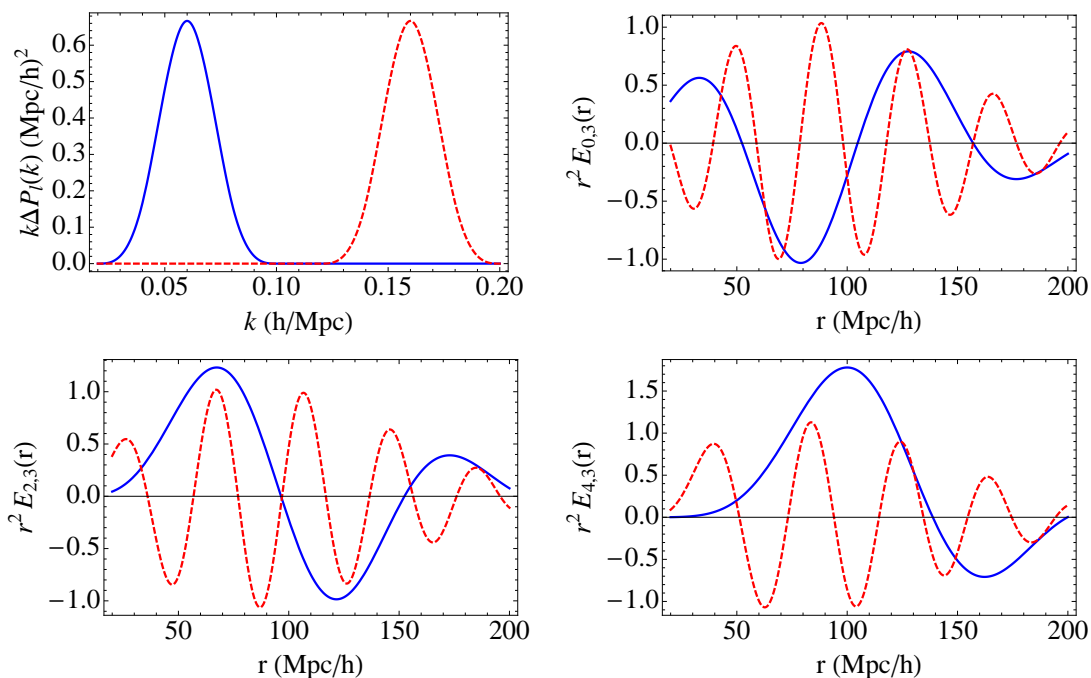
Figure 11. Cubic ($m = 3$) basis spline contributions to the (k-weighted) power multipoles $k\Delta P_\ell(k)$ (top left) with their corresponding (r^2-weighted) contributions $E_{\ell,m}(r; k_j, \Delta k)$ with $\ell = 0, 2, 4$ (clockwise from top right) to the correlation function multipoles $\Delta\xi_\ell(r)$ with $k_{lo}, k_{hi} = (0.02, 0.2)$ h/Mpc, $n = 10$, and $j = 0$ (solid blue curves) and $j = 5$ (dashed red curves).

3 Fitting method

We fit the parameterized models of $\xi(r, \mu, z)$ described above to estimates $\xi_{ijk} \equiv \xi(r_{ijk}, \mu_{ijk}, z_k)$ of the correlation function specified on a 3D grid $\Delta v_i \otimes \Delta\theta_j \otimes z_k$ of physical coordinates (see section 2.1), with

$$r_{ijk} = \sqrt{r_\parallel^2(\Delta v_i, z_k) + r_\perp^2(\Delta\theta_j, z_k)} \tag{3.1}$$

$$\mu_{ijk} = r_\parallel(\Delta v_i, z_k)/r_{ijk} . \tag{3.2}$$

Our nominal coordinate grid consists of 28 unequally spaced points in Δv covering $0 \leq \Delta v < 0.083c$ (see figures 1 and 12), 18 equally spaced points in $\Delta\theta$ spanning 5–175 arcminutes, and three redshift values (2, 2.5, 3), for a total of $N = 1512$ grid points. Each set of correlation-function estimates ξ_{ijk} is accompanied by estimated covariances

$$\langle\xi_{ijk}\xi_{i'j'k'}\rangle - \langle\xi_{ijk}\rangle\langle\xi_{i'j'k'}\rangle \tag{3.3}$$

for each physical coordinate pair (ijk) and $(i'j'k')$. In the following, we use the notation **d** to refer to a vector of N correlation function estimates ξ_{ijk}, and write C for the corresponding $N \times N$ covariance matrix. See ref. [29] for details on the estimates **d** and C obtained from BOSS DR9 and accompanying simulated mock data that we use below. Appendix A provides instructions for downloading these estimates as well as the software necessary to reproduce
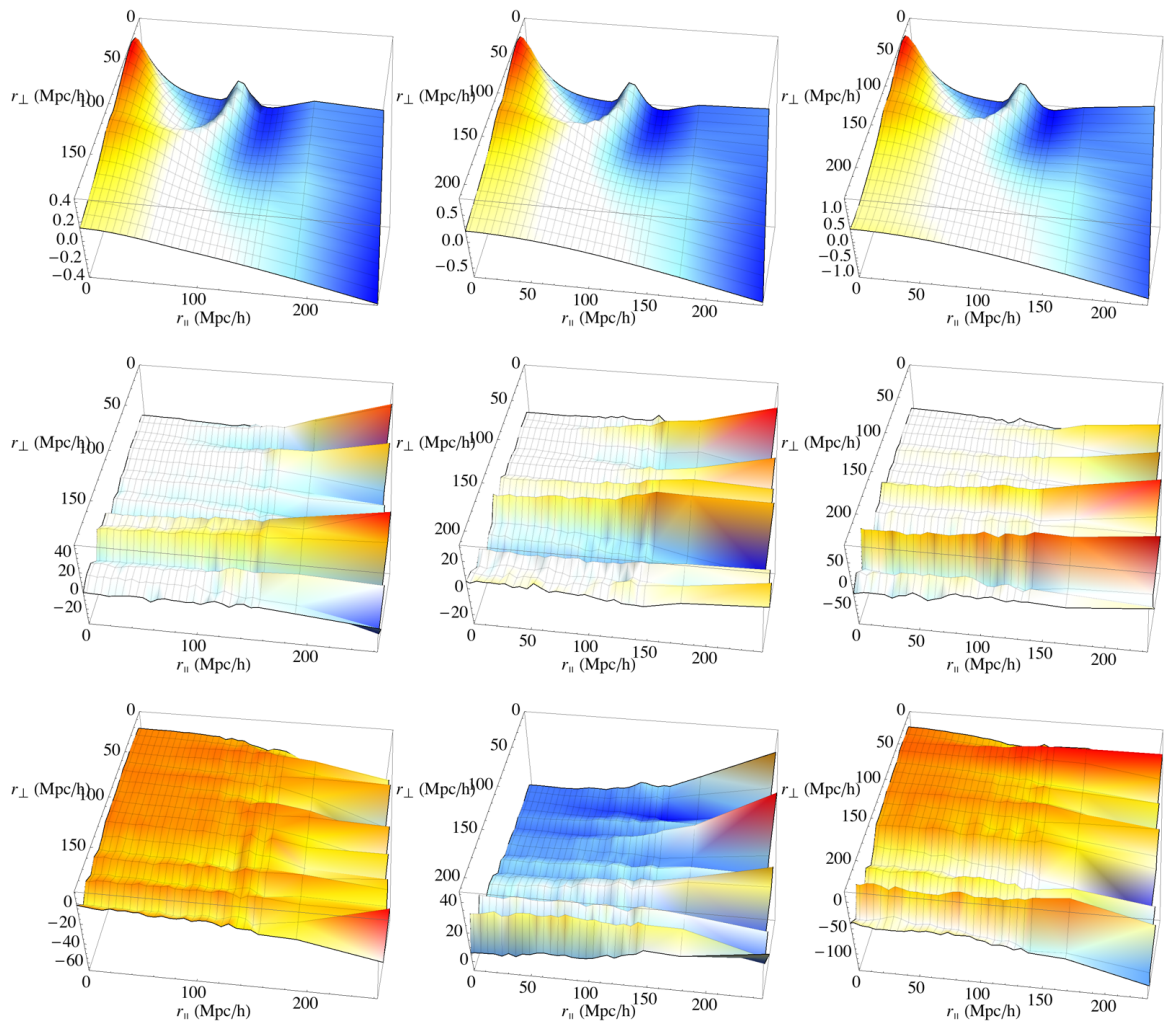
Figure 12. Slices in (r, μ) of the estimated r^2-weighted correlation function for the fiducial cosmology (top row), simulated mocks (middle row), and BOSS data (bottom row). Grid lines show the actual sampling of ξ_{ijk} used. From left to right, columns show redshifts of 2, 2.5, and 3. Physical coordinates are converted to comoving coordinates using the fiducial cosmology.

the main results provided here and in ref. [29]. We define a standard chi-square in terms of a parameter vector $\boldsymbol{\theta}$ and theory prediction $\boldsymbol{\mu}(\boldsymbol{\theta})$ as

$$\chi^2(\boldsymbol{\theta}) \equiv (\boldsymbol{d} - \boldsymbol{\mu}(\boldsymbol{\theta}))^t \, C^{-1} \, (\boldsymbol{d} - \boldsymbol{\mu}(\boldsymbol{\theta})) \ . \tag{3.4}$$

3.1 SDSS-III data release 9 inputs

Figure 12 compares the data vectors for theory, simulated BOSS DR9 data, and actual BOSS DR9 data. The signal is visually obscured in the data by large-amplitude modes at fixed separation angles $\Delta\theta \sim r_\perp$. However, these modes have negligible impact on our ability to measure BAO parameters since they have very small weights (by construction) in the covariance matrix (see figure 13). The value of $\log_{10} |C|/N$ provides a rough measure of the overall (i.e., not BAO specific) signal to noise ratio of a sample. We find values of -8.5 and

-8.3 for data and mocks, respectively, with $N = 1512$, which indicates that the errors in our simulated data are about 30% larger than in real data. The signal to noise ratio is largest at $z = 2.5$, with errors about 40% larger at $z = 2$ and three times larger at $z = 3$.

Figure 13 shows the structure of our covariance estimates, which is essentially the same for simulation and data after accounting for the $\sim 30\%$ normalization difference. The first 36 eigenmodes have artificially large eigenvalues (see figure 13a) due to the template marginalization procedure described in ref. [29], which accounts for the largest-amplitude modes visible in figure 12. We expect correlation function estimates to be uncorrelated between different separation angles and therefore impose this constraint on our estimated inverse covariance matrix reducing both C and C^{-1} to a block diagonal form consisting of 18 sub-matrices, $C_{(j)}$ and $C_{(j)}^{-1}$, with dimensions 84×84. Each of these submatrices has the same structure, shown in figure 13c for $C_{(j)}^{-1}$, but a different normalization, $|C_{(j)}|$ or $|C_{(j)}^{-1}| = |C_{(j)}|^{-1}$, depending on its separation $\Delta\theta_j$ (solid curve in figure 13b). Each of the three 28×28 sub-matrices on the diagonal, corresponding to a fixed separation $\Delta\theta_j$ and redshift z_k, also have the same structure with normalizations shown as the dashed curves in figure 13b. Values of the inverse covariance along the diagonal are roughly independent of Δv_i and only depend on $\Delta\theta_j$ via the factor $|C_{(j)}|$, as plotted in figure 13d where the decreasing values at large Δv_i correspond to increasing errors due to limited statistics.

Our correlation function estimates are actually provided as a set of M quasi-independent data vectors $\mathbf{d_m}$ and covariance matrices C_m, corresponding to angular partition of the BOSS survey footprint into individual observing plates [29]. For fits to DR9, we have $M = 817$ plates. We perform a weighted combination of observations, assumed to be independent, to obtain the final $\mathbf{d}$ and C used for fitting

$$C^{-1} = \sum_{m=1}^{M} C_m^{-1} \quad , \quad \mathbf{d} = C \sum_{m=1}^{M} C_m^{-1} \mathbf{d}_m \ . \tag{3.5}$$

The effective signal to noise ratio of individual observations spans about an order of magnitude (see figure 14) due to varying observing conditions, so it is important that observations are correctly weighted in the combination. However, an overall normalization error in the estimated covariances C_m still leads to a correctly weighted combination, even if the individual C_m have different structures.

3.2 Internal covariance tests and refinement

We exploit the large number of observations available to perform some internal cross checks on the individual C_m and obtain improved estimates, as described below. First, we compare each observation $\mathbf{d}_m$ to the combined observations using the chi-square statistic

$$\chi_m^2 \equiv (\mathbf{d}_m - \mathbf{d})^t (C_m - C)^{-1} (\mathbf{d}_m - \mathbf{d}) \tag{3.6}$$

where $C_m - C$ is the expected covariance of $\mathbf{d}_m - \mathbf{d}$

$$\begin{aligned}
\langle (\mathbf{d}_m - \mathbf{d})(\mathbf{d}_m - \mathbf{d})^t \rangle &= \langle \mathbf{d}_m \mathbf{d}_m^t \rangle - 2\langle \mathbf{d}_m \mathbf{d}^t \rangle + \langle \mathbf{d}\mathbf{d}^t \rangle \\
&= C_m - 2C + \sum_{m'} C C_{m'}^{-1} C \\
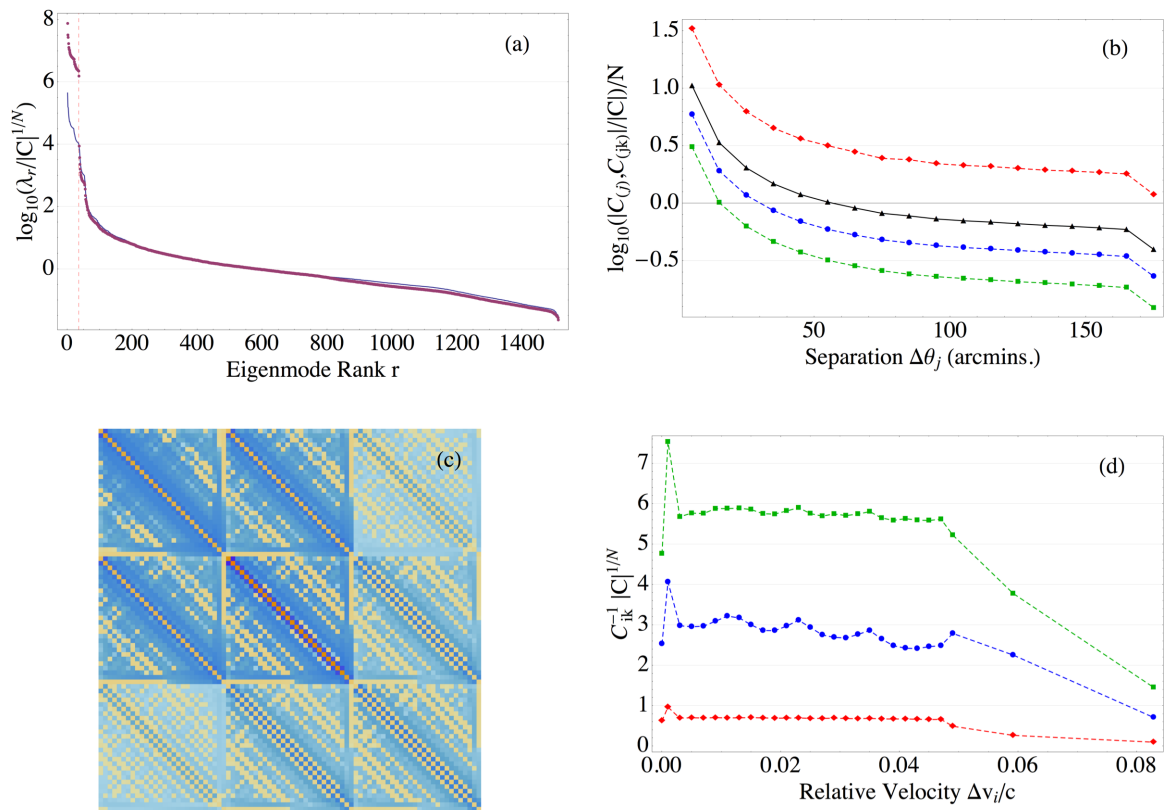&= C_m - C \ ,
\end{aligned} \tag{3.7}$$

Figure 13. Structure of the estimated covariance matrix C. The top-left panel (a) shows the ordered eigenmodes λ_r of C in data (blue curve) and simulation (points). The top-right panel (b) shows the relative normalization of the 18 block-diagonal submatrices $C_{(j)}$ as a function of separation $\Delta\theta_j$ (black), as well as the relative normalization of the three 28×28 sub-matrices $C_{(jk)}$ corresponding to redshifts 2 (blue circles), 2.5 (green squares), and 3 (red diamonds), averaged over relative velocities Δv_i. The bottom-left panel (c) shows the common structure of the 18 C_j^{-1} blocks for each separation $\Delta\theta_j$. The bottom-right panel (d) shows the relative inverse covariance along the diagonal of this common block, as a function of relative velocity $\Delta v_i/c$, for redshifts 2 (blue circles), 2.5 (green squares), and 3 (red diamonds).

and we have assumed that each $\mathbf{d}_m$ is a Gaussian realization of some common true correlation function $\mathbf{d}(\boldsymbol{\theta}_0)$ with additional noise $\mathbf{e}_m$ sampled from C_m

$$\mathbf{d}_m = \mathbf{d}(\boldsymbol{\theta}_0) + \mathbf{e}_m \; . \tag{3.8}$$

Figure 14 shows that, on average, our covariance estimates C_m underestimate the actual variance seen in the data vectors $\mathbf{d}_m$ by about 20%, but that this factor is independent of when the observation was taken. We then search for a correlation between the degree of underestimation and the amount of predicted variance by diagonalizing each $C_m - C$

$$C_m - C = X_m \Lambda_m X_m^t \tag{3.9}$$

where Λ_m is a diagonal matrix of the $N = 1512$ eigenvalues ranked in increasing size (larger predicted variance)

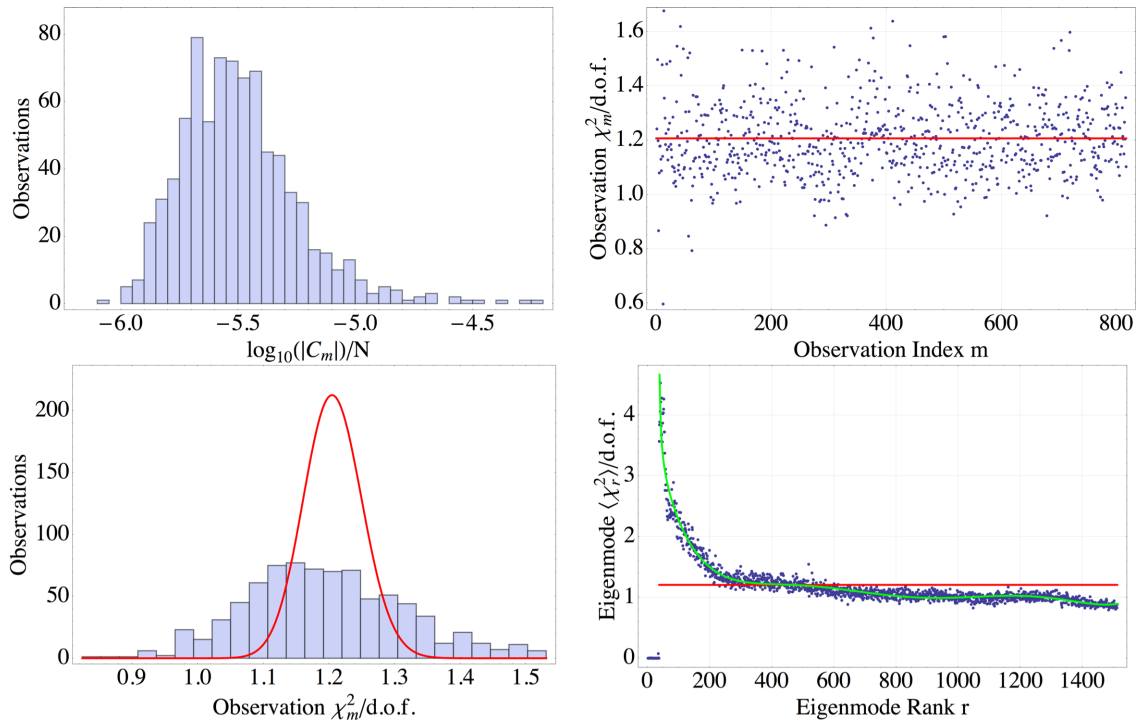$$(\Lambda_m)_{rs} = \delta_{rs}\, \lambda_{m,r} \; , \tag{3.10}$$

Figure 14. Covariance studies of the 817 independent correlation function estimates in BOSS DR9. The top-left panel shows the distribution of covariance weights $\log_{10}(|C_j|)/N$ for each observation j. The χ_j^2 values for each of the 817 observations are shown in the top-right panel as a function of observation index j, which is roughly ordered by date of observation, compared with their mean of 1.2. The bottom-left panel shows the distribution of normalized χ_j^2, defined in eq. (3.6), for 1512 degrees of freedom, compared with their expected distribution (red curve) for Gaussian statistics with an overall normalization factor of 1.2 applied. The bottom-right panel shows the evolution of $\langle \chi_r^2 \rangle$ with eigenvalue rank r. Figure 15 repeats the bottom two panels after applying the rescaling procedure described in the text.

and X_m is a matrix whose columns are the corresponding eigenvectors of $C_m - C$. We can then rewrite χ_j^2 as

$$\chi_m^2 = \sum_r \chi_{m,r}^2 \tag{3.11}$$

where

$$\chi_{m,r}^2 \equiv \lambda_r^{-1} u_{m,r} \tag{3.12}$$

and $u_{m,r}$ are the components of $\mathbf{d}_m - \mathbf{d}$ in the eigenvector basis for $C_m - C$

$$\mathbf{u}_m = X_m^t (\mathbf{d}_m - \mathbf{d}) . \tag{3.13}$$

Each $\chi_{m,r}^2$ should be chi-square distributed with one degree of freedom (if our assumptions of Gaussianity are valid) so we use their mean over the $M = 817$ observations as a function of eigenvalue rank r

$$\langle \chi_r^2 \rangle \equiv \frac{1}{M} \sum_r \chi_{j,r}^2 \tag{3.14}$$

to provide another internal test, where we expect $\langle \chi_r^2 \rangle \simeq 1$ for an eigenmode whose variance λ_r is correctly estimated. The results show (see figure 14) that most of the 20% underestimation
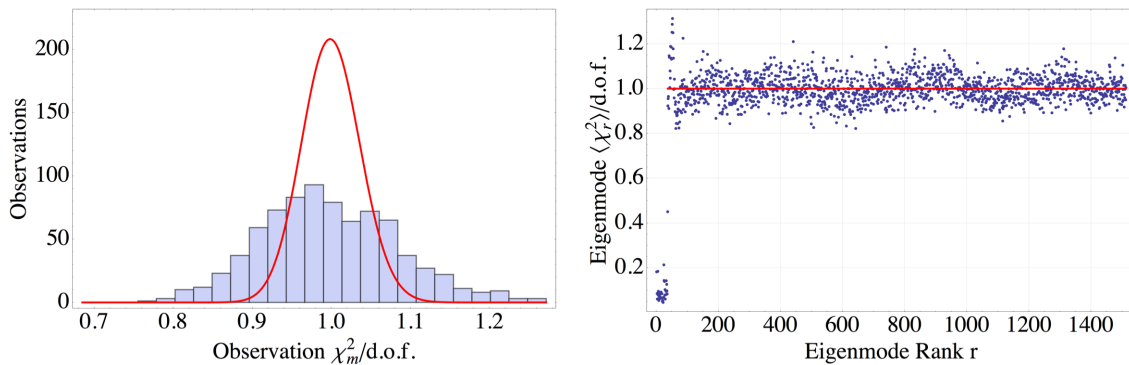
Figure 15. Results of the rescaling procedure described in the text applied to BOSS DR9. Plots shows the same quantities as the bottom row of figure 14 after applying the rescaling procedure described in the text.

is due to the first ~ 250 eigenmodes with the largest predicted variances (but not including the 36 modes with artificially large variance mentioned above, which appear here with very small values of $\langle \chi_r^2 \rangle$).

We can partially correct for the deviations with eigenvalue rank shown in figure 14 by independently rescaling each eigenmode r of C_m by $\langle \chi_r^2 \rangle^{-1}$. This procedure is not exact, however, since $\langle \chi_r^2 \rangle^{-1}$ is derived from $C_m - C$ rather than C_m (but we always have $|C_m| \gg |C|$) and because each C_m generally has different eigenvectors (but we expect similar eigenvectors since the C_m differ primarily in their normalizations). The results of this rescaling procedure are shown in figure 15 and demonstrate that all eigenmodes now have internally consistent $\mathbf{d}_m$ fluctuations and covariance estimates C_m. Two caveats to this procedure are that we do not rescale the 36 previously marginalized modes and that we impose the expected block-diagonal structure on the rescaled C_m, eliminating any off-diagonal contributions introduced by numerical round-off errors. Note that the distribution of the χ_m^2 statistic has improved somewhat, compared with the Gaussian expectation, after rescaling. Finally, we combine the rescaled C_m using eq. (3.5) to obtain the final $\mathbf{d}$ and C used for the fits described below (and in figures 12 and 13).

We use bootstrap sampling in two different ways in our analysis. First, we boostrap our 817 estimates of the data vector $\mathbf{d}_m$ to obtain an estimate of their combined covariance that is independent of the C_m, as a cross-check of our combination using eq. (3.5). With $M = 817$ and $N = 1512$, This approach is only feasible because of the 18×84^2 block-diagonal structure of C: we effectively bootstrap the $84(84 + 1)/2 = 3570$ elements of each block $C_{(j)}$ independently using 817 sub-vectors of length 84. We compare fit results obtained with the calculated combination C and the corresponding bootstrap estimate of C to check the validity of our calculated C.

The second type of bootstrap analysis we perform is to fit a large number of bootstrap samples and then compare the resulting distributions of best-fit parameter values with the parabolic errors calculated from the likelihood surface of the combined fit using $\mathbf{d}$ and C. To the extent that the errors in our fit parameters are not Gaussian, we expect some discrepancy and then prefer the bootstrap estimates. Since our observations span an order of magnitude in weights $|C_m|$, a sampling with replacement generally yields a set of observations whose combined covariance does not match the combined covariance of all observations. Therefore, some care is required in assigning the total covariance used to fit each bootstrap sample since,

although we ignore the resulting fit errors, we still require a correct relative weighting over the physical coordinate grid in eq. (3.4).

We generalize the combination method of eq. (3.5) for bootstrap sampling to

$$\tilde{C}^{-1} = \sum_{m=1}^{M} n_m C_m^{-1} \quad , \quad \mathbf{d} = \tilde{C} \sum_{m=1}^{M} n_m C_m^{-1} \mathbf{d}_m \ , \tag{3.15}$$

where $n_m \geq 0$ is the number of repetitions of observation m and $\sum_{m=1}^{M} n_m = M'$ is the bootstrap sample size (usually $M = M'$). The resulting $\tilde{C}$ is not the covariance of $\mathbf{d}$ since it incorrectly reduces the variance for double-counted observations. Instead, the correct covariance C to use is

$$C = \tilde{C} D^{-1} \tilde{C} \tag{3.16}$$

with

$$D^{-1} \equiv \sum_{m=1}^{M} n_m^2 C_m^{-1} \ . \tag{3.17}$$

Note that eq. (3.15) reduces to eq. (3.5) when each $n_m^2 = n_m$, so that each sample is either not used, $n_m = 0$ (if $M' < M$), or used exactly once, $n_m = 1$. In the limit that all C_m are identical, we find

$$\frac{\langle \tilde{\chi}^2 \rangle}{\langle \chi^2 \rangle} = \frac{M' + M - 1}{M} \ , \tag{3.18}$$

where $\tilde{\chi}^2$ is the chi-squared that would be obtained using $\tilde{C}^{-1}$ instead of C^{-1} in eq. (3.4). In the usual case of $M' = M$, we find a ratio $2 - 1/M \simeq 2$.

Before performing fits, we generally apply some final cuts to select a subset of the N three-dimensional grid points. Cuts can either be in physical or (using the fiducial cosmology) comoving coordinates. We implement cuts by first combining all observations using eq. (3.5), without cuts, and then eliminating elements from the final data vector $\mathbf{d}$, as well as the corresponding rows and columns of C. Note that the resulting C^{-1} is different from what we would have obtained by eliminating rows and columns from C^{-1} directly. Our nominal final cuts on physical coordinates are $0.003 < \Delta v_i / c < 0.083$ and $5 \leq \Delta \theta_j \leq 165$ arcminutes. We also cut on the co-moving separation $50 < r < 190$ Mpc/h. After all final cuts, the size of our data vector is reduced from $N = 3 \times 504 = 1512$ to $341 \, (z = 2) + 310 \, (z = 2.5) + 290 \, (z = 3) = 941$.

4 Results

In this section, we present results related to the fitting methods and what they reveal about the BOSS DR9 Lyman-α forest sample. Cosmological fit results derived from the same sample are presented in the companion paper ref. [29].

4.1 Parameter sensitivities

In order to quantify and visualize how a dataset constrains each parameter p, it is instructive to plot the vector

$$\mathbf{F}_p \equiv \mathbf{d}(\boldsymbol{\theta})_{,p} \circ \left(C^{-1} \mathbf{d}(\boldsymbol{\theta})_{,p} \right) \tag{4.1}$$

where $\mathbf{d}(\boldsymbol{\theta})_{,p}$ is the partial derivative of the theory prediction at some point in parameter space $\boldsymbol{\theta}$, and $\circ$ represents the Hadamard (entrywise) product. The sum of the resulting components
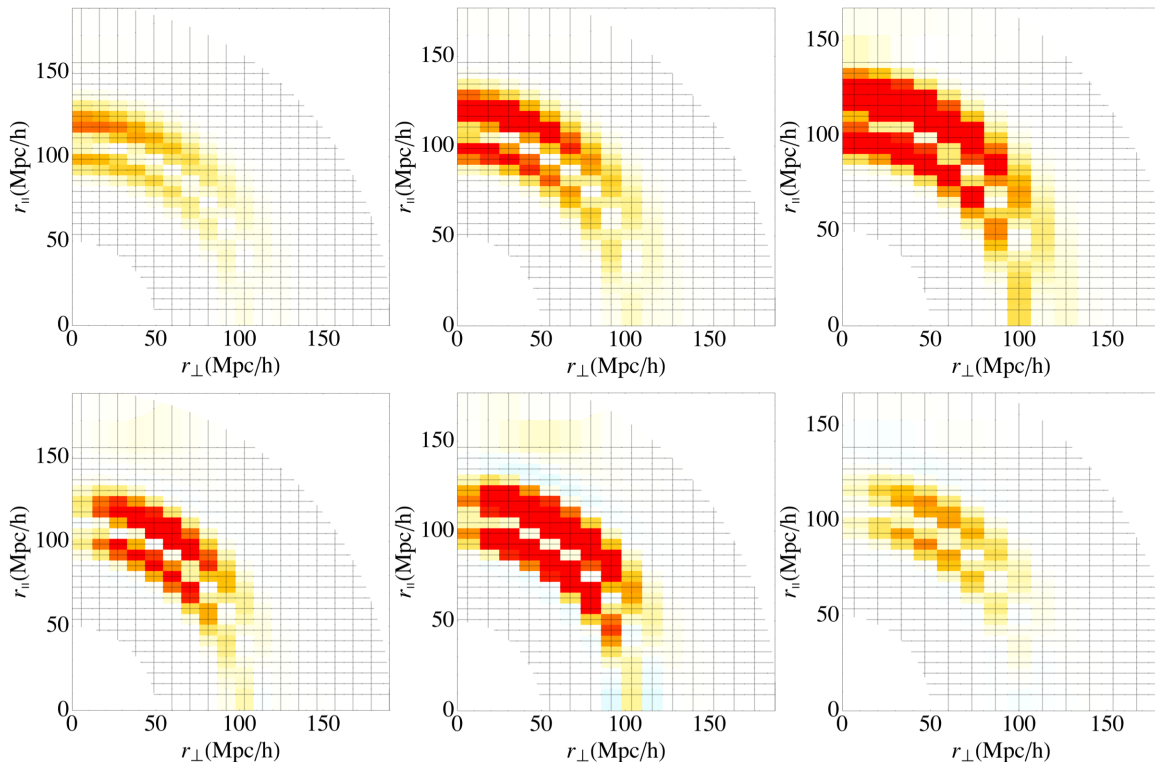
Figure 16. Plots of the components of $\mathbf{F}_p$ for the isotropic scale parameter $\alpha_{\rm iso}$. The top row shows the intrinsic model sensitivity using $C = |C|^{1/n}$, while the bottom row uses the combined covariance matrix for BOSS DR9. The three plots on each row show element values in the $(r_{\parallel}, r_{\perp})$ for $z = 2$ (left), 2.5 (middle), and 3 (right), using the same color scale to preserve the relative magnitudes between redshifts.

is the Fisher information for parameter p [36] that specifies the expected maximum-likelihood error σ_p (in an ensemble-average sense) when all other parameter values are known

$$\sum_{ijk} F_{p,ijk} \simeq \sigma(p)^{-2} \ . \tag{4.2}$$

The theory derivatives $\mathbf{d}(\boldsymbol{\theta})_{,p}$ encode the parameter sensitivity inherent to the model while the covariance C encodes the performance of a particular analysis method. The quantity $\mathbf{F}_p$ is plotted for our baseline fit to BOSS DR9 for the isotropic scale parameter $\alpha_{\rm iso}$ in figure 16 and for the anisotropic scale parameters in figures 17 ($\alpha_{\parallel}$) and 18 ($\alpha_{\perp}$).

The top row of figure 16 uses $C = |C|^{1/n}$ to isolate the intrinsic model sensitivity, resulting in a Fisher error prediction of $\sigma(\alpha_{\rm iso}) = 0.017$, with inverse-variance contributions of 12%, 28%, 60% from $z = 2$, 2.5, 3, respectively, due to the observed large increase in tracer bias $b^2(z)$ with redshift governed by the parameter γ_{b^2}. Information on the BAO scale comes predominantly from the forward direction due to the large observed value of the redshift-distortion parameter β, and is concentrated on two rings in co-moving separation on either side of the BAO peak, where the theory is most sensitive to the peak position. When the combined covariance C is taken into account, the Fisher error prediction increases slightly to $\sigma(\alpha_{\rm iso}) = 0.018$, and the region of peak sensitivity moves from $\mu = 1$ to $\mu \simeq 0.9$, mostly
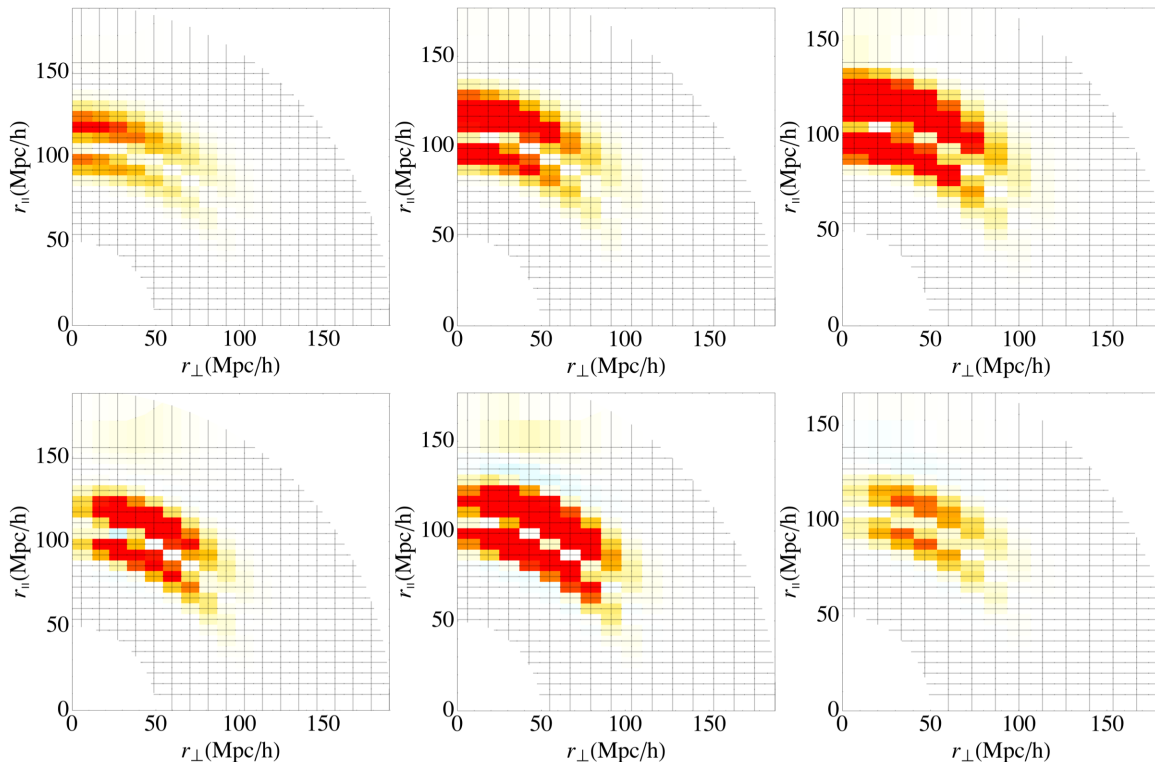
Figure 17. Plots of the components of $\mathbf{F}_p$ for the anisotropic scale parameter $\alpha_\parallel$. See caption to figure 16 for details.

due to the trend shown in figure 13(b). Intermediate redshifts now dominate the statistical power due to the distribution of forest pixel redshifts, with inverse-variance contributions of 29%, 61%, 10% from $z = 2$, 2.5, 3, respectively.

When floating the scales parallel and perpendicular to the line of sight, the Fisher predicted errors are $\sigma(\alpha_\parallel) = 0.024$ (0.021) and $\sigma(\alpha_\perp) = 0.050$ (0.047) for the combined analysis covariance C (using $C = |C|^{1/n}$). Note that our actual covariance actually reduces the expected error on $\sigma(\alpha_\perp)$, relative to $C = |C|^{1/n}$, because of the decreasing errors with separation angle shown in figure 13b. The peak predicted sensitivities are $\mu \simeq 0.95$ for $\alpha_\parallel$ and $\mu \simeq 0.70$ for $\alpha_\perp$.

The results above were obtained with $r'' = r$ in eq. (2.34), so that scale factors are applied only to the peak position and not the cosmological broadband. If, instead, we use $r'' = r'$, the resulting Fisher components for an isotropic fit are shown in figure 19, resulting in a smaller predicted error of $\sigma(\alpha_{\rm iso}) = 0.010$ (0.013) for the combined analysis covariance C (using $C = |C|^{1/n}$). Comparing with figure 16, we conclude that a fit using coupled transforms obtains a smaller error (when all other parameter values are fixed) by using information outside of the peak region. However, this benefit is offset by a larger dependence on the broadband distortion model, and the final errors are comparable when marginalizing over distortion parameters. For fits to DR9, we prefer to localize the BAO scale measurement to the peak region and minimize our dependence on the distortion model. With larger data samples and improved analyses, we expect to have better control of distortion effects and this choice should be revisited.
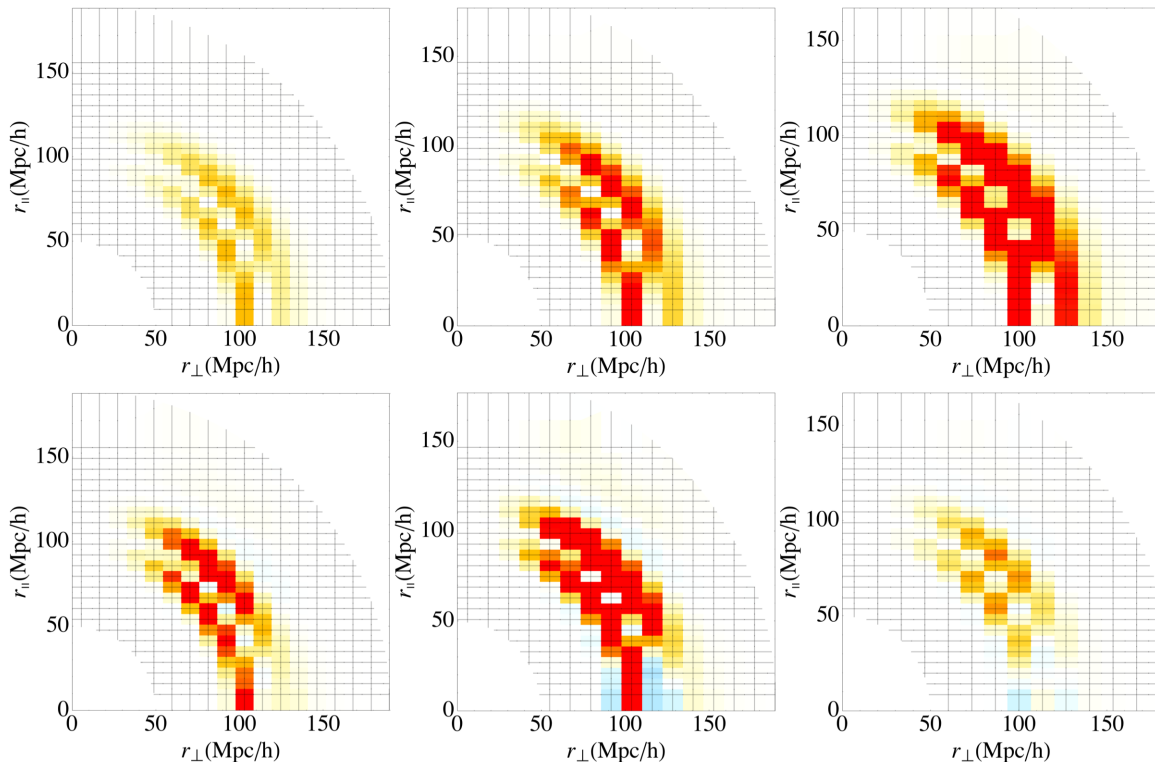
Figure 18. Plots of the components of $\mathbf{F}_p$ for the anisotropic scale parameter $\alpha_\perp$. See caption to figure 16 for details.

4.2 Data reductions

The 1512 fit inputs ξ_{ijk} already represent a substantial reduction of data from the millions of Lyman-α forest pixel pairs from which they are derived, while preserving essentially all of the cosmologically relevant information. We can take this process one step further using the interpolated models of section 2.4 to reduce ξ_{ijk} down to estimates of multipoles of the correlation function or power spectrum. There are two main challenges in this process: first, the input ξ_{ijk} has large correlated errors that will propagate through and, second, the input ξ_{ijk} is expected to include some broadband distortion that prevents direct comparison of reduced data with theory. Therefore, we first test our data reductions on a mock dataset based on our fiducial cosmology and with the same covariance as data, but no distortion, and then compare these to the corresponding results obtained from data.

We use three different types of model-independent fits for data reduction. The first fit estimates the r^2-weighted correlation function multipoles at $z = 2.4$ as a linear interpolation between 22 co-moving separations (5-140 in steps of 5, followed by 150, 160, and 180 Mpc/h) using eq. (2.79), for a total reduction from 1512 to 66 values. The second fit estimates the k-weighted power spectrum multipoles at $z = 2.4$ as a set 20 of equally spaced band powers covering $k = 0.03$-0.33 h/Mpc with $\Delta k = 0.015$ h/Mpc, using eq. (2.82) with 0-th order B-splines, for a total reduction from 1512 to 60 values. Our final model assumes that the $P(k)$ multipoles are identical except for the normalization factor eq. (2.80), which is expected to be a good approximation if broadband distortion is either small or else experiences similar redshift-space distortion as the linear theory, and results in 20 final values. In all of our data
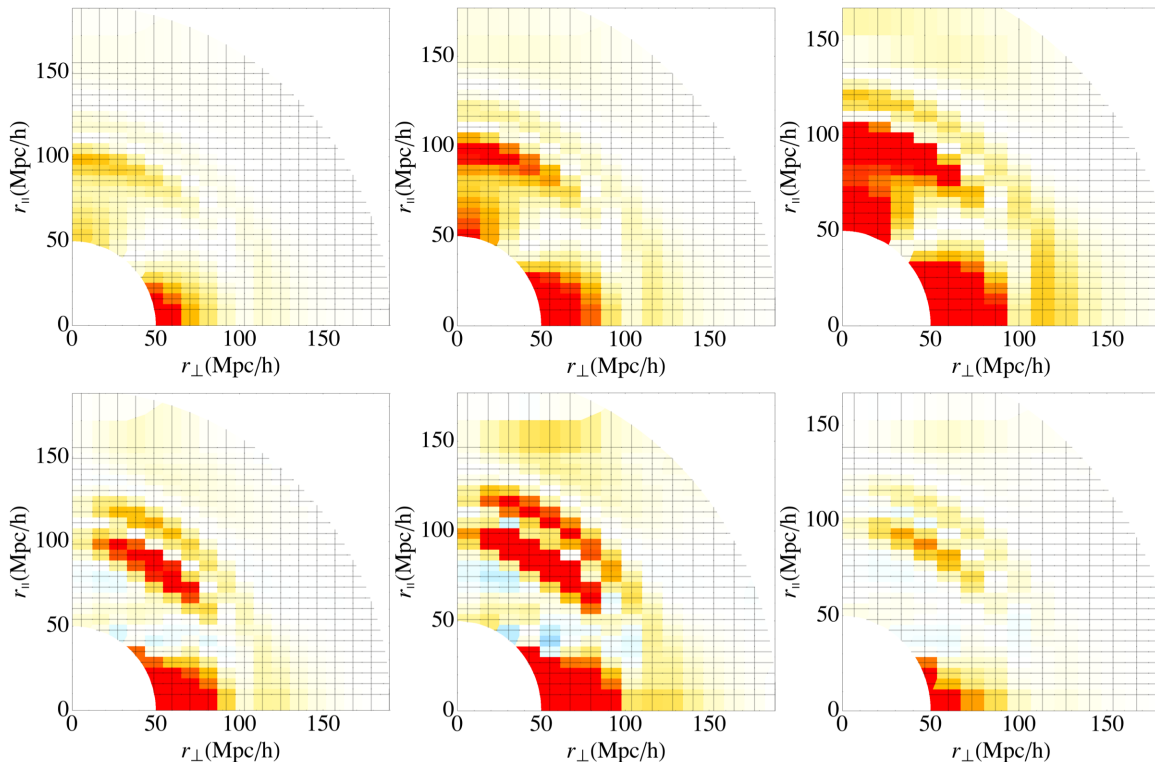
Figure 19. Plots of the components of $\mathbf{F}_p$ for the isotropic scale parameter α_{iso}, using $r'' = r'$ instead of $r'' = r$ in eq. (2.34). See caption to figure 16 for details.

reduction fits, we assume $\beta(z_0) = 1.4$, $\gamma_{b^2} = 3.8$, and $\gamma_\beta = 0$, and our results are normalized using $b(z_0) = -0.183$. We have made the results of these data reductions publicly available (see appendix A).

Figure 20 shows the ranked eigenvalues of the parameter covariance matrices for fits to data and the un-distorted mock dataset, and demonstrates that the data and mock have essentially the same parameter covariance structure. For the purposes of visualization, it is helpful to project out the largest eigenmodes of the parameter covariance matrix, which eliminates the largest sources of correlated errors but can also introduce a projection distortion of the reduced data (in addition to any distortion from the analysis method). The optimum number of modes to project out is a balance between minimizing correlated errors and minimizing projection distortion effects. Figures 21–23 show the results of projecting different numbers of eigenvalues on the monopole of our un-distorted mock, and allows us to identify an appropriate number of modes to filter for each data reduction fit: -14, -36, and -6, respectively.

Figures 24–26 show the same reductions as figures 21–23, but applied to data instead of un-distorted mocks. Comparing these, we conclude that the projection distortion effects are similar for each reduction. We also note that the analysis introduces significant broadband distortion that is most apparent (especially with the r^2 weighting) as a broad excess above 110 Mpc/h in the correlation-function monopole (figure 21 with 14 modes projected out). Since this broadband distortion corresponds to low-k modes, the $P(k)$ data reductions are relatively immune to it. Instead, they show an enhancement of the BAO oscillation signal
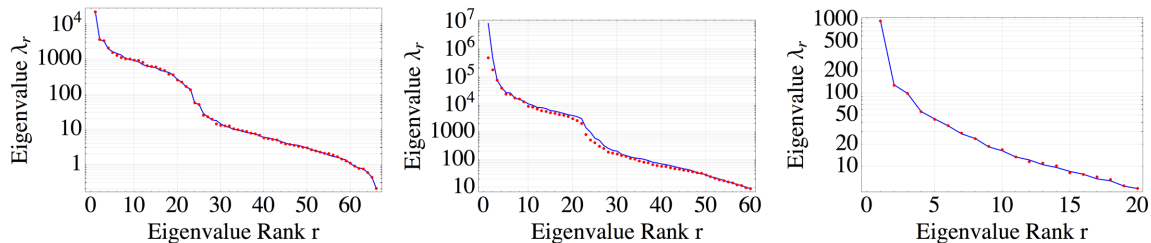
Figure 20. Ranked eigenvalues of parameter covariance matrices for the three data reduction fits described in the text with, left to right, 66, 60, and 20 parameters, respectively. Red curves show eigenvalues for the fits to data and points show fits to a mock dataset without broadband distortion.
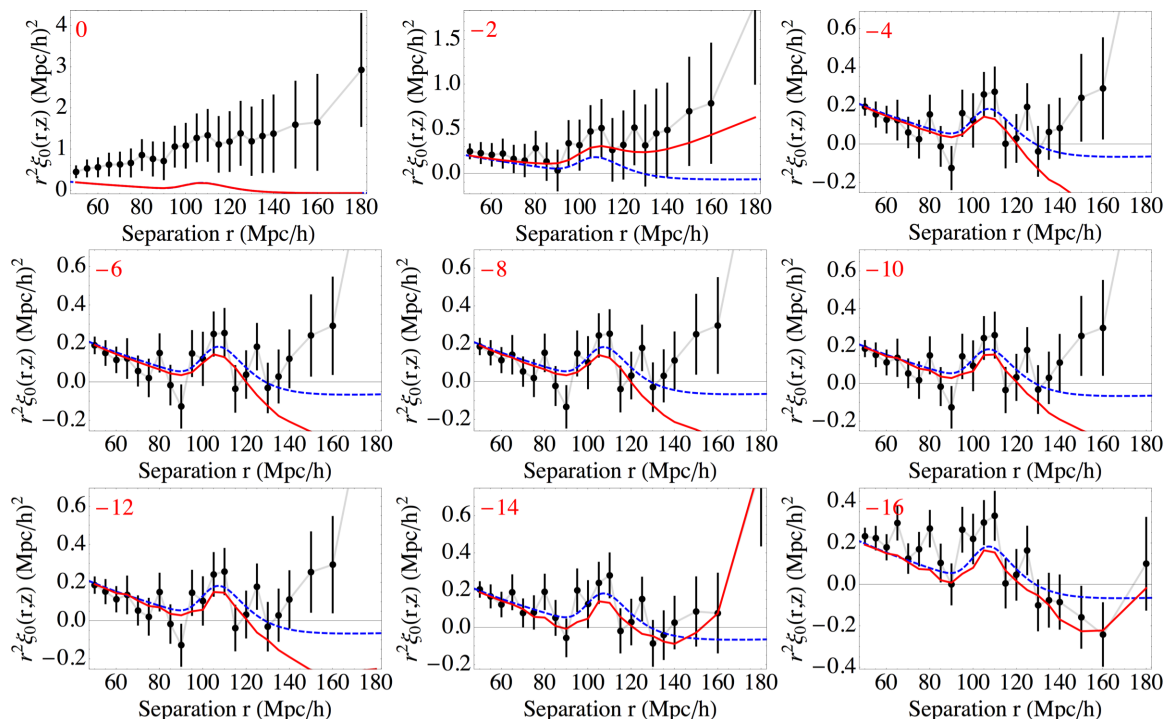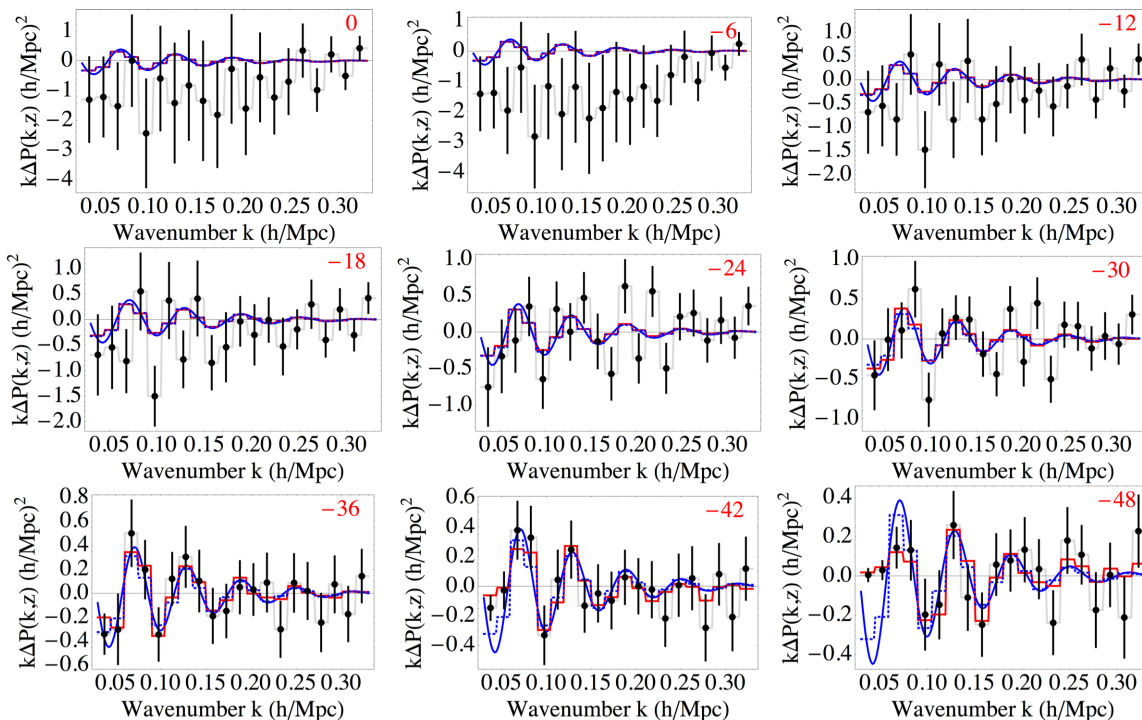

Figure 21. Data reduction fit to an un-distorted mock with an increasing number of the largest eigenmodes (shown in the top-left corner of each plot) projected out of the resulting parameter covariance matrix. Points show the r^2-weighted monopole $\xi_0(r, z)$ at $z = 2.4$ of the reduced data with errors calculated from the corresponding diagonal parameter covariance matrix elements. Curves show the fiducial model at $z = 2.4$ with (red) and without (blue, dashed) the projection distortion. Neither curve is a fit to the data points.

relative to the fiducial model. Further study with additional data is needed to determine if this excess is a fortunate statistical fluctuation or if perhaps the signal is larger than expected.

5   Discussion

In this paper we have described a near-optimal method for fitting correlations observed in the Lyman-α forest of high-redshift quasars to measure the properties of the baryon acoustic oscillation feature. We also explore the properties of the BOSS DR9 correlation estimates

Figure 22. Data reduction fit to an un-distorted mock with an increasing number of the largest eigenmodes (shown in the top-left corner of each plot) projected out of the resulting parameter covariance matrix. Points show the k-weighted deviation in the monopole $\Delta P(k, z)$ at $z = 2.4$ of the reduced data with errors calculated from the corresponding diagonal parameter covariance matrix elements. Curves show the fiducial model at $z = 2.4$ with (red) and without (blue, dashed) the projection distortion. Neither curve is a fit to the data points.

described in ref. [29], and provide the code and input files necessary to reproduce our main results.

Comparing with the Lyman-α fitting methods used in ref. [24], where correlations are modeled in the monopole and quadrupole at a single mean redshift, we have developed a fully three-dimensional fitting technique and included the effects of anisotropic non-linear broadening. These developments do not yield a significant improvement in the cosmological constraints that are possible with DR9, but provide more flexibility to study and quantify systematics and allow direct study the redshift evolution of parameters such as $b(z)$ and $\beta(z)$, and to accurately determine the effective redshift of our cosmological constraints.

The challenge of a three-dimensional analysis is that the resulting large covariance matrix is difficult to estimate and validate, especially with the limited number of mocks available for the DR9 Lyman-α forest. However, we have developed a novel method for internally testing and refining our large covariance matrix, as described in section 3.2. The resulting covariance matrix still has large correlations that visually obscure the BAO signal but have little impact on our ability to measure its properties. We describe data reductions that retain most of the cosmologically relevant information and show how the expected features are visually revealed when large eigenmodes are projected out.

In building models suitable for a three-dimensional analysis, we find that care is needed to isolate the BAO feature in the correlation function and describe an alternative to the "no-wiggles" approach of ref. [34] that is better suited for this purpose. When considering
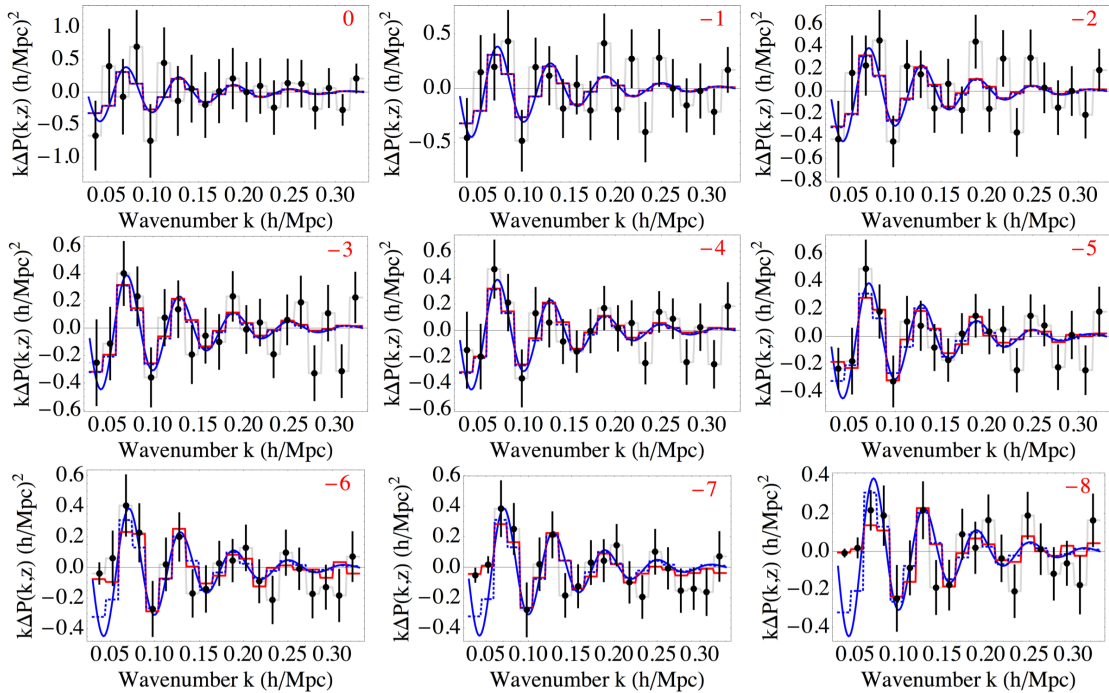
Figure 23. Plots correspond to figure 22 but use a data-reduction fit that assumes identical $P(k)$ multipoles, resulting in 20 final parameters, as described in the text.
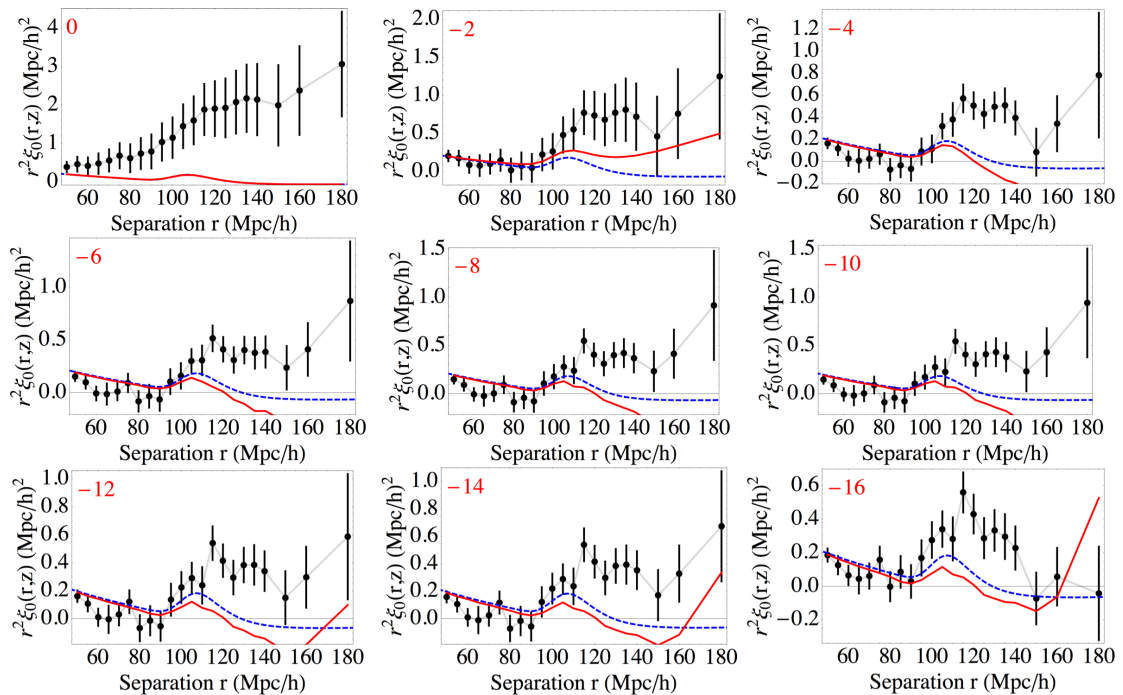

Figure 24. Plots correspond to figure 21 but are based on data instead of the un-distorted mock.

the line of sight ($\alpha_\parallel$) and transverse ($\alpha_\perp$) scale factors independently, we find that a first-order treatment is numerically inaccurate near the BAO peak, and identify the important
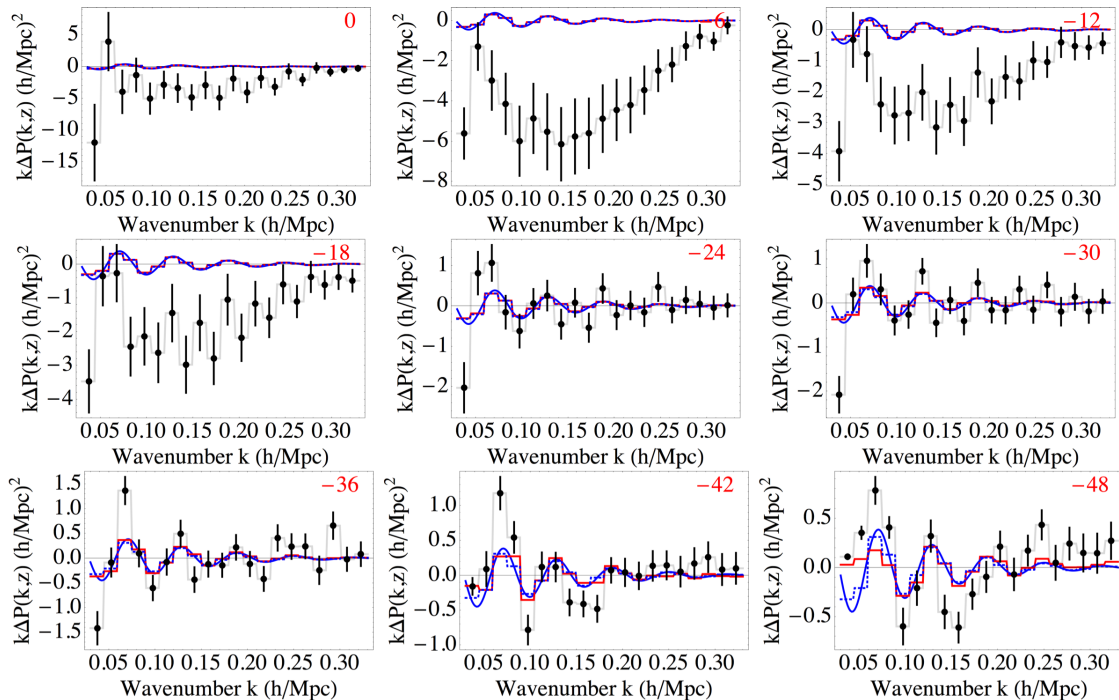
Figure 25. Plots correspond to figure 22 but are based on data instead of the un-distorted mock.
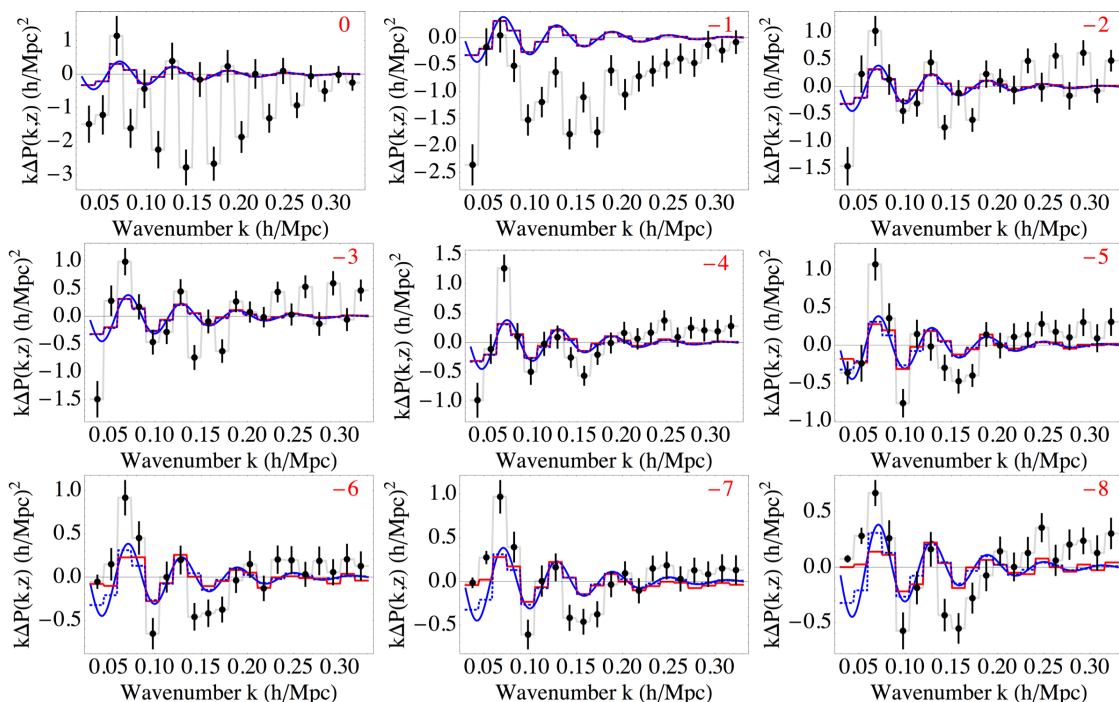
Figure 26. Plots correspond to figure 23 but are based on data instead of the un-distorted mock.

second-order terms. We introduce anisotropic non-linear broadening using an approximation that is valid for a large range of β. We also use a highly flexible parametrization of both

multiplicative and additive broadband distortions to accomodate the expected systematics of the correlation-function estimates due to effects such as a biased continuum estimate. Finally, we demonstrate that information on the BAO feature in the Lyman-α forest arises primarily along the line of sight, as expected due to the large redshift-space distortion, but that the contributions from $H(z)$ and $D_A(z)$ can still be independently measured.

The first measurements of baryon acoustic oscillations in the Lyman-α forest [24, 29] provide a new confirmation of our basic picture of cosmological evolution dominated by dark energy, but do not yet significantly constrain the parameters of the simplest cosmological models. With the SDSS-III high-redshift quasar sample expected to triple the DR9 statistics over the next 18 months, that situation is likely to change and will require near-optimal fitting methods and a flexible framework for careful study of systematics.

Acknowledgments

A Public access to data and code

The software used to generate the results in this paper and ref. [29] are publicly available at http://github.com/baofit/. We also provide instructions to install and run the software, together with the BOSS DR9 correlation-function estimates and configuration files necessary to reproduce our main results, at http://darkmatter.ps.uci.edu/baofit/. The software is written in C++ and uses MINUIT [37] for likelihood minimization. Data and configuration files are in plain text format.

References

[1] SUPERNOVA SEARCH TEAM collaboration, A.G. Riess, et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009 [astro-ph/9805201] [INSPIRE].

[2] SUPERNOVA COSMOLOGY PROJECT collaboration, S. Perlmutter, et al., *Measurements of Omega and Lambda from 42 high redshift supernovae*, *Astrophys. J.* **517** (1999) 565 [astro-ph/9812133] [INSPIRE].

[3] D.H. Weinberg, M.J. Mortonson, D.J. Eisenstein, C. Hirata, A.G. Riess, et al., *Observational Probes of Cosmic Acceleration*, arXiv:1201.2434 [INSPIRE].

[4] WMAP collaboration, E. Komatsu, et al., *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys. J. Suppl.* **192** (2011) 18 [arXiv:1001.4538] [INSPIRE].

[5] G. Hinshaw, D. Larson, E. Komatsu, D. Spergel, C. Bennett, et al., *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results*, arXiv:1212.5226 [INSPIRE].

[6] 2DFGRS collaboration, W.J. Percival, et al., *The 2dF Galaxy Redshift Survey: The Power spectrum and the matter content of the Universe*, *Mon. Not. Roy. Astron. Soc.* **327** (2001) 1297 [astro-ph/0105252] [INSPIRE].

[7] C.J. Miller, R.C. Nichol and D.J. Batuski, *Possible detection of baryonic fluctuations in the large scale structure power spectrum*, *Astrophys. J.* **555** (2001) 68 [astro-ph/0103018] [INSPIRE].

[8] SDSS collaboration, D.J. Eisenstein, et al., *Detection of the baryon acoustic peak in the large-scale correlation function of SDSS luminous red galaxies*, *Astrophys. J.* **633** (2005) 560 [astro-ph/0501171] [INSPIRE].

[9] 2DFGRS collaboration, S. Cole, et al., *The 2dF Galaxy Redshift Survey: Power-spectrum analysis of the final dataset and cosmological implications*, *Mon. Not. Roy. Astron. Soc.* **362** (2005) 505 [astro-ph/0501174] [INSPIRE].

[10] G. Huetsi, *Power spectrum of the sdss luminous red galaxies: constraints on cosmological parameters*, *Astron. Astrophys.* **459** (2006) 375 [astro-ph/0604129] [INSPIRE].

[11] SDSS collaboration, N. Padmanabhan, et al., *The Clustering of Luminous Red Galaxies in the Sloan Digital Sky Survey Imaging Data*, *Mon. Not. Roy. Astron. Soc.* **378** (2007) 852 [astro-ph/0605302] [INSPIRE].

[12] C. Blake, A. Collister, S. Bridle and O. Lahav, *Cosmological baryonic and matter densities from 600,000 SDSS Luminous Red Galaxies with photometric redshifts*, *Mon. Not. Roy. Astron. Soc.* **374** (2007) 1527 [astro-ph/0605303] [INSPIRE].

[13] W.J. Percival, S. Cole, D.J. Eisenstein, R.C. Nichol, J.A. Peacock, et al., *Measuring the Baryon Acoustic Oscillation scale using the SDSS and 2dFGRS*, *Mon. Not. Roy. Astron. Soc.* **381** (2007) 1053 [arXiv:0705.3323] [INSPIRE].

[14] T. Okumura, T. Matsubara, D.J. Eisenstein, I. Kayo, C. Hikage, et al., *Large-Scale Anisotropic Correlation Function of SDSS Luminous Red Galaxies*, *Astrophys. J.* **676** (2008) 889 [arXiv:0711.3640] [INSPIRE].

[15] E. Gaztanaga, A. Cabre and L. Hui, *Clustering of Luminous Red Galaxies IV: Baryon Acoustic Peak in the Line-of-Sight Direction and a Direct Measurement of H(z)*, *Mon. Not. Roy. Astron. Soc.* **399** (2009) 1663 [arXiv:0807.3551] [INSPIRE].

[16] SDSS collaboration, E.A. Kazin, et al., *The Baryonic Acoustic Feature and Large-Scale Clustering in the SDSS LRG Sample*, *Astrophys. J.* **710** (2010) 1444 [arXiv:0908.2598] [INSPIRE].

[17] SDSS collaboration, W.J. Percival, et al., *Baryon Acoustic Oscillations in the Sloan Digital Sky Survey Data Release 7 Galaxy Sample*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 2148 [arXiv:0907.1660] [INSPIRE].

[18] B.A. Reid, W.J. Percival, D.J. Eisenstein, L. Verde, D.N. Spergel, et al., *Cosmological Constraints from the Clustering of the Sloan Digital Sky Survey DR7 Luminous Red Galaxies*, *Mon. Not. Roy. Astron. Soc.* **404** (2010) 60 [arXiv:0907.1659] [INSPIRE].

[19] C. Blake, E. Kazin, F. Beutler, T. Davis, D. Parkinson, et al., *The WiggleZ Dark Energy Survey: mapping the distance-redshift relation with baryon acoustic oscillations*, *Mon. Not. Roy. Astron. Soc.* **418** (2011) 1707 [arXiv:1108.2635] [INSPIRE].

[20] F. Beutler, C. Blake, M. Colless, D.H. Jones, L. Staveley-Smith, et al., *The 6dF Galaxy Survey: Baryon Acoustic Oscillations and the Local Hubble Constant*, *Mon. Not. Roy. Astron. Soc.* **416** (2011) 3017 [arXiv:1106.3366] [INSPIRE].

[21] H.-J. Seo, S. Ho, M. White, A. Cuesta, A. Ross, et al., *Acoustic scale from the angular power spectra of SDSS-III DR8 photometric luminous galaxies*, *Astrophys. J.* **761** (2012) 13 [arXiv:1201.2172] [INSPIRE].

[22] N. Padmanabhan, X. Xu, D.J. Eisenstein, R. Scalzo, A.J. Cuesta, et al., *A 2- I : Methods and Application to the Sloan Digital Sky Survey*, arXiv:1202.0090 [INSPIRE].

[23] L. Anderson, E. Aubourg, S. Bailey, D. Bizyaev, M. Blanton, et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the Data Release 9 Spectroscopic Galaxy Sample*, *Mon. Not. Roy. Astron. Soc.* **428** (2013) 1036 [arXiv:1203.6594] [INSPIRE].

[24] BOSS collaboration, N.G. Busca, et al., *Baryon Acoustic Oscillations in the Ly-α forest of BOSS quasars*, arXiv:1211.2616 [INSPIRE].

[25] BOSS collaboration, K.S. Dawson, et al., *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, arXiv:1208.0022 [INSPIRE].

[26] SDSS collaboration, D.J. Eisenstein, et al., *SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way Galaxy and Extra-Solar Planetary Systems*, *Astron. J.* **142** (2011) 72 [arXiv:1101.1529] [INSPIRE].

[27] SDSS collaboration, C.P. Ahn, et al., *The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey*, *Astrophys. J. Suppl.* **203** (2012) 21 [arXiv:1207.7137] [INSPIRE].

[28] X. Xu, N. Padmanabhan, D.J. Eisenstein, K.T. Mehta and A.J. Cuesta, *A 2- II: Fitting Techniques*, arXiv:1202.0091 [INSPIRE].

[29] A. Slosar, V. Irsic, D. Kirkby, S. Bailey, N.G. Busca, et al., *Measurement of Baryon Acoustic Oscillations in the Lyman-alpha Forest Fluctuations in BOSS Data Release 9*, arXiv:1301.3459 [INSPIRE].

[30] SDSS collaboration, P. McDonald, et al., *The Lyman-alpha forest power spectrum from the Sloan Digital Sky Survey*, *Astrophys. J. Suppl.* **163** (2006) 80 [astro-ph/0405013] [INSPIRE].

[31] A.J.S. Hamilton, *Linear Redshift Distortions: a Review*, Astrophys. Space Sc. L. **231** (1998) 185 [astro-ph/9708102] [INSPIRE].

[32] N. Kaiser, *Clustering in real space and in redshift space*, Mon. Not. Roy. Astron. Soc. **227** (1987) 1 [INSPIRE].

[33] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [astro-ph/9911177] [INSPIRE].

[34] D.J. Eisenstein and W. Hu, *Baryonic features in the matter transfer function*, *Astrophys. J.* **496** (1998) 605 [astro-ph/9709112] [INSPIRE].

[35] D.J. Eisenstein, H.-j. Seo and M.J. White, *On the Robustness of the Acoustic Scale in the Low-Redshift Clustering of Matter*, *Astrophys. J.* **664** (2007) 660 [astro-ph/0604361] [INSPIRE].

[36] M. Tegmark, A. Taylor and A. Heavens, *Karhunen-Loeve eigenvalue problems in cosmology: How should we tackle large data sets?*, *Astrophys. J.* **480** (1997) 22 [astro-ph/9603021] [INSPIRE].

[37] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343 [INSPIRE].