

# Exploring cosmic homogeneity with the BOSS DR12 galaxy sample

Pierros Ntelis,<sup>a,1</sup> Jean-Christophe Hamilton,<sup>a</sup> Jean-Marc Le Goff,<sup>b</sup> Etienne Burtin,<sup>b</sup> Pierre Laurent,<sup>b</sup> James Rich,<sup>b</sup> Nicolas Guillermo Busca,<sup>a</sup> Jeremy Tinker,<sup>m</sup> Eric Aubourg,<sup>a</sup> Hélion du Mas des Bourboux,<sup>b</sup> Julian Bautista,<sup>f</sup> Nathalie Palanque Delabrouille,<sup>b</sup> Timothée Delubac,<sup>g</sup> Sarah Eftekharzadeh,<sup>d</sup> David W. Hogg,<sup>c</sup> Adam Myers,<sup>d</sup> Mariana Vargas-Magaña,<sup>j</sup> Isabelle Pâris,<sup>e</sup> Partick Petitjean,<sup>l</sup> Graziano Rossi,<sup>h</sup> Donald P. Schneider,<sup>k</sup> Rita Tojeiro<sup>i</sup> and Christophe Yeche<sup>b</sup>

<sup>a</sup>APC, Université Paris Diderot-Paris 7, CNRS/IN2P3, CEA, Observatoire de Paris, 10, rue A. Domon & L. Duquet, Paris, France

<sup>b</sup>IRFU, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France

<sup>c</sup>Center for Cosmology and Particle Physics, New York University, 4 Washington Place, Meyer Hall of Physics, New York, NY 10003, USA

<sup>d</sup>Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

<sup>e</sup>Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France

<sup>f</sup>Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA.

<sup>g</sup>Laboratoire d'astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, CH-1290 Versoix, Switzerland

<sup>h</sup>Department of Astronomy and Space Science, Sejong University, Seoul, 143-747, Korea

<sup>i</sup>School of Physics and Astronomy, University of St Andrews, St Andrews, KY16 9SS, UK

<sup>j</sup>Instituto de Física, Universidad Nacional Autónoma de México, Apdo. Postal 20-364, México

<sup>k</sup>Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802

<sup>l</sup>Institut d'Astrophysique de Paris, CNRS-UPMC, UMR7095, 98bis bd Arago, Paris, 75014 France

<sup>m</sup>Department of Physics and Center for Cosmology and Particle Physics, New York University

E-mail: [pntelis <at> apc <point> in2p3 <point> fr](mailto:pntelis@apc.in2p3.fr)

**Abstract.** In this study, we probe the transition to cosmic homogeneity in the Large Scale Structure (LSS) of the Universe using the CMASS galaxy sample of BOSS spectroscopic survey which covers the largest effective volume to date,  $3 h^{-3} \text{ Gpc}^3$  at  $0.43 \leq z \leq 0.7$ . We study the scaled counts-in-spheres,  $\mathcal{N}(< r)$ , and the fractal correlation dimension,  $\mathcal{D}_2(r)$ , to assess the homogeneity scale of the universe using a *Landy & Szalay* inspired estimator.

Defining the scale of transition to homogeneity as the scale at which  $\mathcal{D}_2(r)$  reaches 3 within 1%, i.e.  $\mathcal{D}_2(r) > 2.97$  for  $r > \mathcal{R}_H$ , we find  $\mathcal{R}_H = (63.3 \pm 0.7) h^{-1} \text{ Mpc}$ , in agreement at the percentage level with the predictions of the  $\Lambda\text{CDM}$  model  $\mathcal{R}_H = 62.0 h^{-1} \text{ Mpc}$ . Thanks to the large cosmic depth of the survey, we investigate the redshift evolution of the transition to homogeneity scale and find agreement with the  $\Lambda\text{CDM}$  prediction. Finally, we find that  $\mathcal{D}_2$  is compatible with 3 at scales larger than  $300 h^{-1} \text{ Mpc}$  in all redshift bins.

These results consolidate the Cosmological Principle and represent a precise consistency test of the  $\Lambda\text{CDM}$  model.

DRAFT

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Dataset</b>	<b>4</b>
2.1	The BOSS survey	4
2.2	Data sample	4
2.3	Mock-Catalogues	5
<b>3</b>	<b>Methodology</b>	<b>6</b>
3.1	Observables	6
3.2	Estimators	6
3.3	Homogeneity scale definition	7
<b>4</b>	<b>Theoretical Model</b>	<b>8</b>
4.1	Prediction for $\xi(r)$ , $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$	8
4.2	Correction for bias and RSD	9
<b>5</b>	<b>Results</b>	<b>10</b>
5.1	Analysis range	10
5.2	Covariance matrices	10
5.3	Homogeneity scale for the galaxy distribution	11
5.4	Determining the bias and the velocity dispersion	12
5.5	Homogeneity scale for the matter distribution	13
5.6	Consistency with $\Lambda$ CDM	14
5.7	Constraints on fractal correlation dimension	14
<b>6</b>	<b>Analysis Robustness</b>	<b>17</b>
6.1	Bias in the fractal correlation dimension	17
6.2	Sensitivity to RSD model	18
<b>7</b>	<b>Conclusions</b>	<b>19</b>
<b>A</b>	<b>From <math>\xi(r)</math> to <math>\mathcal{N}(&lt; r)</math></b>	<b>21</b>
<b>B</b>	<b>Choice of estimator for <math>\mathcal{N}(&lt; r)</math></b>	<b>21</b>
<b>C</b>	<b>Tuning errors in RSD Analysis</b>	<b>21</b>
<b>D</b>	<b>Test of spline fit on QPM mock catalogues</b>	<b>22</b>
<b>E</b>	<b>Homogeneity scale at 0.1%</b>	<b>22</b>

---

# 1 Introduction

Most models of modern Cosmology are based on solutions of General Relativity for an isotropic and homogeneous universe. The standard model, known as  $\Lambda$ CDM, is mainly composed of a Cold Dark Matter (CDM) and  $\Lambda$  corresponds to a cosmological constant. This model shows excellent agreement with current data, be it from Type Ia supernovae [1, 2], temperature and polarisation anisotropies in the Cosmic Microwave Background [3] or Large Scale Structure [4–6]. The two main assumptions of this model are the validity of General Relativity and the *Cosmological Principle* [7] that states that the Universe is isotropic and homogeneous on large enough scales, or equivalently that the Universe statistical properties are both rotationally and translationally invariant on large scales.

Isotropy is well tested through several probes at various cosmic epochs: at  $z \approx 1100$ , Cosmic Microwave Background temperature anisotropies, corresponding to density fluctuations in the young Universe, have been shown to be of order  $10^{-5}$  [8]. In the more recent Universe, the distribution of sources in X-ray [9] and radio [10] surveys strongly supports isotropy. Large spectroscopic galaxy surveys, such as the baryon oscillation spectroscopic survey (BOSS) of the third Sloan digital sky survey (SDSS-III), show no evidence for anisotropies in the projected galaxy distribution in volumes of a few  $\text{Gpc}^3$  [11].

We should stress, however, that these observations test isotropy at a given redshift or after projection over a given range of redshifts. This "projected" isotropy is a weaker assumption than "spatial" isotropy, which is an isotropy at each redshift<sup>1</sup>. Combining spatial isotropy and the Copernican principle, which states that our position in the Universe is not privileged, implies that the Universe is homogeneous [7, 12–14]. However, as shown by Durrer et al. [15], this implication is not true if we only have projected isotropy. So CMB isotropy, for instance, cannot be combined with the Copernican principle to prove homogeneity.

It is therefore important to test homogeneity. Large three dimensional spectroscopic surveys offer an excellent occasion to strengthen this aspect of the cosmological model with accurate observations. Most of the studies conducted so far found a transition to homogeneity in the galaxy distributions at scale between 70 and 150  $h^{-1}$  Mpc [16–27] therefore strengthening the cosmological principle. Some studies did not find a transition to homogeneity [28–34]. However none of those studies reached scale larger than 200  $h^{-1}$  Mpc, where homogeneity becomes clear.

However, these 3D surveys investigate the statistical properties on the observer past light-cone and not inside of it, so they do not observe galaxies at the same epoch. An attempt to overcome this limitation was to use star formation history in order to study the homogeneity inside the past light-cone [35], but this is model dependent. Another possibility [36] is to use a combination of secondary CMB probes, including integrated Sachs-Wolfe, kinetic Sunyaev-Zel'dovich and Rees-Sciama effects.

In this article, we use the Data Release 12 of the SDSS-III/BOSS spectroscopic galaxy sample to search for a transition to homogeneity [37]. The BOSS CMASS galaxy catalogue covers a volume of 5.1  $h^{-3}$   $\text{Gpc}^3$  in the redshift range  $0.43 < z < 0.7$  and contains nearly one million galaxies in 8500 square degrees. It represents the largest effective volume sampled as of today<sup>2</sup> and therefore offers a unique opportunity to observe the transition to homogeneity.

---

<sup>1</sup>More precisely, spatial isotropy is the assumption that  $\rho(r, \theta_1) = \rho(r, \theta_2)$  for any given  $(r, \theta_1, \theta_2)$ , while projected isotropy is the assumption that  $\rho(\theta_1) = \rho(\theta_2)$  for any  $(\theta_1, \theta_2)$ , where  $\rho(\theta) = \int \rho(r, \theta)W(r)dr$  and  $W(r)$  is the window function.

<sup>2</sup>The effective volume of a survey is given by  $V_{\text{eff}} = V \frac{nP}{1+nP}$ , which describes the statistical power of the

A similar study of homogeneity has been performed with the DR12 quasar sample of BOSS [38]. We mostly follow the method introduced in [26] to measure the "fractal correlation dimension" of the distribution,  $\mathcal{D}_2(r)$ , as an indicator of the transition to homogeneity [37].

Due to the possible evolution of the tracer with redshift, a redshift survey is blind to strictly radial variation of the density [39]. So strictly speaking we can only test homogeneity up to a radial variation, that is we can test spatial isotropy (and then combining spatial isotropy with the Copernican principle we can prove homogeneity). On the other hand, as discussed by Laurent et al. [38], we can demonstrate spatial isotropy independently of a fiducial cosmology<sup>3</sup>.

Throughout this study, unless stated otherwise, we use a fiducial flat  $\Lambda$ CDM cosmological model with the following parameters:

$$\mathbf{p}_{\text{cosmo}} = (\omega_{\text{cdm}}, \omega_b, h, n_s, \ln [10^{10} A_s]) = (0.1198, 0.02225, 0.6727, 0.9645, 3.094), \quad (1.1)$$

where  $\omega_b = \Omega_b h^2$  and  $\omega_{\text{cdm}} = \Omega_{\text{cdm}} h^2$  are the reduced fractional density of baryons and cold dark matter, respectively;  $h = H_0 / [100 \text{ km s}^{-1} \text{ Mpc}^{-1}]$ , with  $H_0$  the Hubble constant; and finally  $n_s$  and  $A_s$  are, respectively, the spectral index and the amplitude of the primordial scalar power spectrum. The numerical values in Eq. 1.1 are from Planck 2015  $TT, TE, EE + lowP$  analysis [40].

The paper is organized as follows. Section 2 describes the galaxy data sample used in this analysis and section 3 the methodology to quantify cosmic homogeneity. In section 4 we introduce the model that is adjusted to the observed data, while in section 5 we present our results. In section 6 we test the robustness of our method, and we conclude in section 7.

---

sample. The CMASS sample has  $V_{\text{eff}}^{\text{CMASS}} \simeq 2.9h^{-3} \text{ Gpc}^3$ , while the QSO sample has  $V_{\text{eff}}^{\text{QSO}} \simeq 0.21h^{-3} \text{ Gpc}^3$ .

<sup>3</sup> Indeed let's assume a homogeneous 3D spectroscopic survey with a constant density  $\rho(r, \theta, \phi) = \text{Cst}$ . This density is the product of the Jacobian  $J(z) = H(z)/cD^2(z)$  with the observed density of sources  $dN(\theta, \phi, z)/d\Omega dz$  which is a function of angles and redshift. As a result the observed density is necessarily a function of redshift only following:  $dN(\theta, \phi, z)/d\Omega dz \propto J^{-1}(z)$ , which is the definition of spatial isotropy.

## 2 Dataset

### 2.1 The BOSS survey

BOSS[41] is dedicated to studying the 3D distribution of  $\sim 1.4 \times 10^6$  galaxies and  $\sim 10^5$  quasars and their Lyman- $\alpha$  forests within an effective area of  $\sim 10,000 \text{ deg}^2$  [42, 43]. Aluminum plates are set on the focal plane of the 2.5m telescope[44] and drilled with 1000 holes corresponding to targets to be observed. An optical fiber is fixed on each hole, allowing 1000 spectra to be measured per exposure [45] with a typical redshift uncertainty of a few tens of  $\text{km} \cdot \text{s}^{-1}$  [46]. Several color cuts are used to select massive galaxy in the redshift range  $0.43 < z < 0.7$  in the imaging data from SDSS-I-II and SDSS-III/BOSS in  $(u, g, r, i, z)$  bands [47]. These cuts are designed to result in a stellar mass limit that is constant with redshift, according to a passively evolving model [48]. The large masses ensure a strong bias with respect to the underlying dark matter field, providing a high signal-to-noise ratio on the two-point correlation function. The resulting sample of galaxies is called CMASS and consists in  $\sim 10^6$  galaxies with accurate spectroscopic redshifts. The CMASS sample has been used to measure the Baryonic Acoustic Oscillation feature (BAO) with unprecedented accuracy, putting strong constraints on cosmological parameters [49]. We use the same galaxy sample to perform a measurement of the transition to homogeneity benefiting from the signal-to-noise enhancement due to the large bias of this galaxy sample.

We weight galaxies with [50]:

$$w_{gal} = (w_{cp} + w_{noz} - 1) \times w_{star} \times w_{see} \times w_{FKP} , \quad (2.1)$$

Here, the close-pair weight,  $w_{cp}$ , accounts for the fact that, due to fiber coating, one cannot assign optical fibers on the same plate to two targets that are closer than  $62''$ . The  $w_{noz}$  weight accounts for targets for which the pipeline failed to measure the redshift. The  $w_{star}$  and  $w_{see}$  weights correct for the dependance of the observed galaxy number density with the stellar density and with seeing, respectively. Finally, we use the FKP weight,  $w_{FKP}$ , [51] in order to reduce the variance of the two-point correlation function estimator.

$z$	NGC	SGC
0.430-0.484	101,383	40,170
0.484-0.538	174,468	63,518
0.538-0.592	151,084	56,805
0.592-0.646	97,155	37,179
0.646-0.700	47,289	17,899
0.430-0.700	571,379	215,571

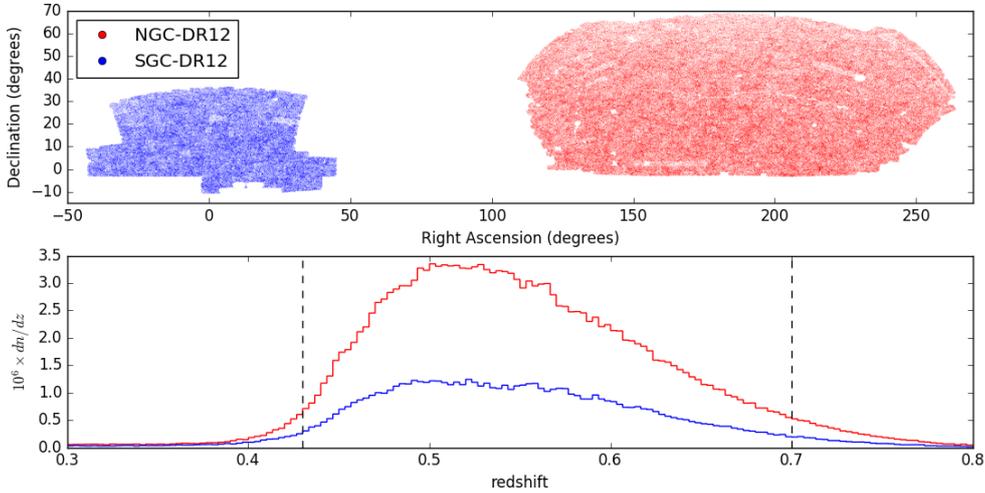
**Table 1.** DR12 data sample in 5 redshift intervals in north (NGC) and south (SGC) galactic caps.

### 2.2 Data sample

We divide our data sample into 5 redshift intervals, as defined in table 1, to study the evolution of the clustering of the CMASS galaxy sample. The angular and redshift distributions of the sample are shown in figure 1. We use a flat  $\Lambda$ CDM model with parameters defined by equation 1.1 to convert the redshift measurements to comoving distances:

$$d_{\text{comov}}(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')} , \quad (2.2)$$

where  $E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$ . A final sample with a total effective volume of  $3.8 h^{-3} \text{Gpc}^3$  [50] is obtained, which is significantly larger than the effective volume used in previous studies, such as  $0.6 h^{-3} \text{Gpc}^3$  for WiggleZ [26],  $1 h^{-3} \text{Gpc}^3$  for a DR7 LRG galaxy sample study [27] or  $0.2 h^{-3} \text{Gpc}^3$  for the SDSS II LRG analysis [16].



**Figure 1.** Right ascension and declination profiles (top) and redshift distribution (bottom) of the DR12 sample for the North (NGC) and South (SGC) Galactic Caps. The vertical dashed lines define the redshift interval used in our analysis.

### 2.3 Mock-Catalogues

Mock catalogues are an important tool for determining uncertainties in galaxy surveys and tuning data analysis pipelines. In our analysis we use 1000 mock catalogues constructed with the quick-particle-mesh (QPM) algorithm [52] for BOSS. These are expected to be more realistic than mocks based on second-order Lagrangian perturbation theory, although not as much as the highly time-consuming N-body simulations. These mock catalogues use a different flat  $\Lambda$ CDM cosmology than the one we use in our analysis, but this difference is accounted for in our analysis:

$$\mathbf{p}_{\text{qpm}} = (\Omega_{\text{cdm}}, \Omega_b, h, n_s, \sigma_8) = (0.274, 0.046, 0.7, 0.95, 0.8) \quad (2.3)$$

where  $\sigma_8$  is the variance of the matter power spectrum computed in spheres of  $8 h^{-1} \text{Mpc}$  radius.

### 3 Methodology

In this section we present the method we have used to study the transition to homogeneity in the distribution of the BOSS CMASS galaxy catalogue. We follow Scrimgeour et al. [26] in assessing the homogeneity of the catalogue through the fractal dimension,  $D_2(r)$ , but use significantly different methods.

#### 3.1 Observables

The ‘‘counts-in-spheres’’,  $N(< r)$ , is defined as the average number of objects in a sphere of radius  $r$ , it should obviously scale as  $r^3$  for an homogeneous distribution. This scaling is the basis for the definition of the fractal dimension:

$$D_2(r) \equiv \frac{d \ln N(< r)}{d \ln r}. \quad (3.1)$$

For a homogeneous distribution one gets  $D_2(r) = 3$ , while for a fractal distribution  $D_2 < 3$  is a measure of the fractal dimension. These simple expressions are no longer valid in the case of a survey that has a peculiar geometry and a non-uniform completeness. We correct for these geometrical effects by using catalogues of random points that uniformly fill the survey volume. We use random samples that are five times larger than the data sample in order to ensure that statistical fluctuations due to the random points are significantly smaller than those due to the data. We therefore define a ‘‘scaled counts-in-spheres’’

$$\mathcal{N}(< r) = \frac{N_{gal}(< r)}{N_{rand}(< r)} \quad (3.2)$$

which is the ratio between the counts-in-spheres of the galaxy distribution,  $N_{gal}(< r)$ , and the counts-in-spheres of the random point distribution inside the same survey geometry,  $N_{rand}(< r)$ . This quantity is now expected to be independent of  $r$  (ratio of two quantities  $\propto r^3$ ) for a homogeneous distribution and therefore the fractal dimension is redefined as:

$$\mathcal{D}_2(r) \equiv \frac{d \ln \mathcal{N}(< r)}{d \ln r} + 3, \quad (3.3)$$

in order to be equal to 3 for a homogeneous distribution.

$\mathcal{N}(< r)$  and  $\mathcal{D}_2(r)$  are the two observables we will consider in the analysis, and we show in the next section that there are different ways to obtain them from the data.

#### 3.2 Estimators

We define the following quantities from our data and random catalogues:

- $dd(r) = \frac{DD(r)}{n_g(n_g - 1)/2}$ : the normalized number of galaxy pairs distant by  $r$ ,
- $rr(r) = \frac{RR(r)}{n_r(n_r - 1)/2}$ : the normalized number of random-point pairs distant by  $r$ ,
- $dr(r) = \frac{DR(r)}{n_g n_r}$ : the normalized number of galaxy random-point pairs distant by  $r$ ,

where  $n_g$  and  $n_r$  are the total number of galaxies and random points, respectively. The definition of  $\mathcal{N}(< r)$  in Eq. 3.2 recalls the Peebles-Hauser estimator,  $\widehat{\xi}(r) = dd(r)/rr(r) - 1$ , for the two-point correlation function [53]. This estimator is known to be less efficient than the more sophisticated Landy-Szalay estimator [54].

$$\widehat{\xi}_{ls}(r) = \frac{dd(r) - 2dr(r) + rr(r)}{rr(r)}, \quad (3.4)$$

which has minimal variance on scales where  $\xi(r) \ll 1$  [54].

Laurent et al. [38] defines a Landy-Szalay inspired estimator for the counts-in-spheres:

$$\widehat{\mathcal{N}}_{lau}(< r) = 1 + \frac{\int_0^r (dd(s) - 2dr(s) + rr(s)) ds}{\int_0^r rr(s) ds}. \quad (3.5)$$

Alternatively, we can directly compute the counts-in-spheres from the Landy-Szalay estimator of the two-point correlation function itself, as explained in appendix A:

$$\widehat{\mathcal{N}}_{cor}(< r) = 1 + \frac{3}{r^3} \int_0^r \widehat{\xi}_{ls}(s) s^2 ds. \quad (3.6)$$

These two estimators are expected to be closer to optimal than the estimator of equation 3.2. They result in two different estimators for the fractal dimension that we name accordingly  $\widehat{\mathcal{D}}_{2,lau}(r)$  and  $\widehat{\mathcal{D}}_{2,cor}(r)$ . In appendix B, we show that the *cor*-estimators,  $\widehat{\mathcal{N}}_{cor}(< r)$  and  $\widehat{\mathcal{D}}_{2,cor}(r)$ , are less biased than the *lau*-estimators. In our final analysis, unless stated otherwise, we use the *cor*-estimators and we drop the *lau* indice and the hat for sake of simplicity.

Since we are interested in matter distribution, we must use the galaxy bias,  $b$ , to convert the counts-in-spheres measured for galaxies to the one we would get if we were measuring the whole matter (mostly dark matter). We also have to account for redshift space distortions and we explain in details how we do that in section 5.4.

### 3.3 Homogeneity scale definition

Following Scrimgeour et al. [26], we define an homogeneity scale  $R_H$  for both  $\mathcal{N}$  and  $\mathcal{D}_2$  observables, as the scale for which the observable approaches its homogeneous value within 1%. We then have two a-priori different transition-to-homogeneity scales defined by:

$$\mathcal{D}_2(R_H^{\mathcal{D}_2=2.97}) = 2.97 \quad \text{and} \quad \mathcal{N}(R_H^{\mathcal{N}=1.01}) = 1.01. \quad (3.7)$$

These are, of course, arbitrary definitions for the homogeneity scales but they do not depend on the survey and can be used to test cosmological models and compare different survey measurements as long as the same definitions are used in all cases [26].

Other authors have proposed different ways of defining the homogeneity scale. In particular Yadav et al. [18] proposed to define it as the scale at which  $\mathcal{D}_2$  cannot be distinguished from 3 within the survey errors. This is obviously another proper definition because it is independent of the bias and similar conclusions can be drawn as we demonstrate at the end of section 5.7. Furthermore, using this estimator on the mock catalogues (See section 5.7), we find larger variations on the estimated homogeneity scale in respect with the one obtained by Eq. 3.7. For these reasons, we have made the choice in this study to use the arbitrary but universal definition in Eq. 3.7, although we show that beyond 300  $h^{-1}$ Mpc, we do find consistency with  $\mathcal{D}_2 = 3$  within our measurement errors.

## 4 Theoretical Model

### 4.1 Prediction for $\xi(r)$ , $\mathcal{N}(< r)$ and $\mathcal{D}_2(r)$

We use the CLASS software [55] to compute the  $\Lambda$ CDM theoretical prediction for the two-points correlation function and the observables we consider in the analysis. CLASS computes the theoretical matter power spectrum,  $P_{\delta\delta}^{(r)}(k)$ , where the  $(r)$  exponent indicates that it is calculated in real space, by opposition to redshift space. In order to get a prediction for the redshift-space galaxy power-spectrum,  $P_{gg}^{(s)}(k)$  where the  $(s)$  exponent indicates redshift space, we model redshift-space-distortion (RSD):

- On large scales, galaxies are falling into large gravitational potentials, which tends to sharpen their distribution along the line-of-sight in redshift space. This is known as the Kaiser effect [56], which results in:

$$P_{gg}^{(s)}(k, \mu; b) = b^2(1 + \beta\mu^2)^2 P_{\delta\delta}^{(r)}(k), \quad (4.1)$$

where  $b$  is the linear (scale independent) bias between galaxy and matter distributions,  $\beta = f/b$  where  $f$  is the linear growth rate, which can be approximated by  $f \approx \Omega_m^{0.55}(z)$ [57] and  $\mu = \cos\theta$ , with  $\theta$  the angle relative to the line-of-sight.

- On small scales, galaxies have a velocity dispersion whose projection on the line-of-sight in redshift space gives rise to the "finger-of-God" effect (FoG). These distortions can be modeled with a simple Gaussian orientation-dependent and scale-dependent damping model, which takes into account the pairwise peculiar-velocity dispersion of the galaxies  $\sigma_p$  [58]:

$$\ln D(k, \mu; \sigma_p) = -\frac{1}{2} \left( \frac{k\sigma_p\mu}{H_0} \right)^2. \quad (4.2)$$

This damping factor represents well actual data down to scales where  $k\sigma_p \sim H_0$  [58], which corresponds to  $r \sim 15h^{-1}\text{Mpc}$ .

Finally, accounting for both effects and integrating over all orientations relative to the line-of-sight, the redshift-space galaxy power spectrum reads:

$$P_{gg}^{(s)}(k; b, \sigma_p) = b^2 \int_0^1 (1 + \beta\mu^2)^2 D(k, \mu; \sigma_p) d\mu P_{\delta\delta}^{(r)}(k), \quad (4.3)$$

which can be analytically integrated over  $\mu$ , leading to:

$$P_{gg}^{(s)}(k; b, \sigma_p) = P_{\delta\delta}^{(r)}(k) \times \frac{b^2 H_0}{2k^5 \sigma_p^5} \times \left[ -2 \exp\left(-\frac{k^2 \sigma_p^2}{2H_0^2}\right) k\beta\sigma_p H_0 (k^2(2 + \beta)\sigma_p^2 + 3\beta H_0^2) + \sqrt{2\pi} \text{Erf}\left(\frac{k\sigma_p}{\sqrt{2}H_0}\right) (k^4 \sigma_p^4 + 2k^2 \beta \sigma_p^2 H_0^2 + 3\beta^2 H_0^4) \right] \quad (4.4)$$

Applying a Fast Fourier Transform to equation (4.4) results in the two-point correlation function with two parameters  $b$  and  $\sigma_p$ :

$$\xi^{(s)}(r; b, \sigma_p) = \text{FFT} \left[ P_{gg}^{(s)}(k; b, \sigma_p) \right]. \quad (4.5)$$

Then we use equation (3.6) and (3.3) to compute the  $\Lambda$ CDM prediction for  $\mathcal{N}(< r)$  and  $\mathcal{D}_2(r)$ .

## 4.2 Correction for bias and RSD

The estimators presented in section 3.2 measure the clustering properties of the galaxy distribution, we need to convert them into estimators that describe the clustering properties of the total matter distribution. We fit  $\xi(r)$  in the range  $1 h^{-1} < r < 40 h^{-1}$  Mpc with our model of equation 4.5 to obtain  $b$  and  $\sigma_p$ . If we take into account only the bias and the Kaiser effect, the angle-averaged correlation function is just multiplied by a constant factor, a squared effective bias. Since this factor is independent of  $r$ , the same factor applies to  $\mathcal{N}(< r) - 1$ , which is an integral over  $\xi(r')$  (see equation 3.6). Taking into account also finger-of-God effect, the multiplicative factor is no longer independent of  $r$ . To transform the estimated  $\mathcal{N}(< r) - 1$  for galaxies into an estimation for matter, we multiply it by the ratio of our model for  $\mathcal{N}(< r) - 1$  for matter (i.e.  $b = 1$  and  $\sigma_p = 0$ ) to our model for the best fit value of  $b$  and  $\sigma_p$ , as

$$\widehat{\mathcal{N}}(< r) = \frac{\mathcal{N}^{\text{model}}(< r; b = 1, \sigma_p = 0) - 1}{\mathcal{N}^{\text{model}}(< r; b, \sigma_p) - 1} \times [\widehat{\mathcal{N}}_{gal}(< r) - 1] + 1. \quad (4.6)$$

Then the fractal correlation dimension  $\widehat{\mathcal{D}}_2(r)$  is obtained from  $\widehat{\mathcal{N}}(< r)$  using equation 3.3.

Section 6.2 shows that taking into account the finger-of-God effect only contributes in a small change in the measurement of the homogeneity scale by typically 1%. So the error due to the imperfection in the modelling of this effect is negligible.

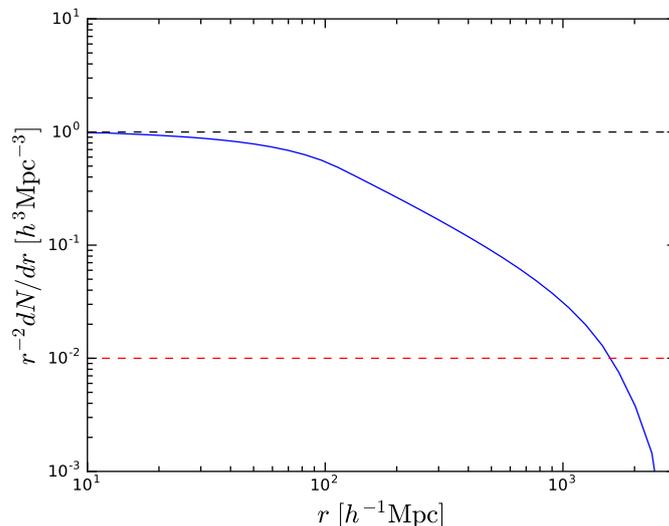
DRAFT

## 5 Results

In this section, we present the results of our analysis. We determine the range in comoving distance for which we can measure  $\mathcal{N}(< r)$  with our sample. We quantify the uncertainties in our measurements using mock catalogues. We measure the homogeneity scale for the galaxy distribution, we fit the redshift-space-distortion parameters, measure the homogeneity scale for matter distribution. Finally we make comparisons with the  $\Lambda$ CDM model and estimate the average of  $\mathcal{D}_2$  on large scales.

### 5.1 Analysis range

Figure 2 shows the number density of random pairs,  $dN/dr$ , divided by  $r^2$  and normalized such that at small scales it is unity. This scaled number density is constant at small scales and then decreases with  $r$  due to the finite size of the survey. We analyse data up to a maximum  $r = 1300 h^{-1}$  Mpc, where the scaled density goes down to 1%.



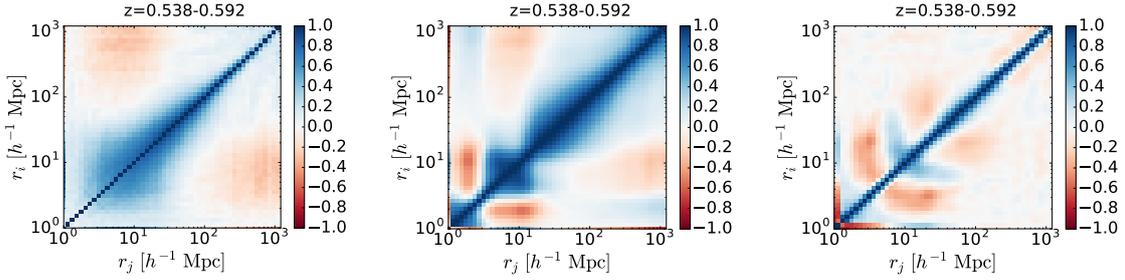
**Figure 2.** Scaled number density of random pairs,  $r^{-2}dN/dr$ , versus the comoving radius of spheres,  $r$ , normalized to be unity at small  $r$ .

### 5.2 Covariance matrices

We use a set of 1000 QPM mock catalogues to estimate the covariance matrices. We compute our observables for each of the catalogues and derive the bin-to-bin covariance matrices for each of the relevant observables,  $\xi(r)$ ,  $\mathcal{N}(< r)$  and  $\mathcal{D}_2(r)$ .

Fig. 3 displays the resulting correlation matrices<sup>4</sup> in the redshift bin  $0.538 \leq z \leq 0.592$ . The correlations matrices are similar in the other redshift bins.  $\mathcal{N}(r)$  is more correlated on large scales than  $\xi(r)$  because it is an integral over  $\xi(r)$ . On the other hand  $\mathcal{D}_2(r)$ , which is a derivative of  $\mathcal{N}(r)$ , is not very correlated. Therefore, when studying the redshift evolution in section 5.5, we consider only the homogeneity scale obtained with  $\mathcal{D}_2(r)$ , and not with  $\mathcal{N}(r)$ .

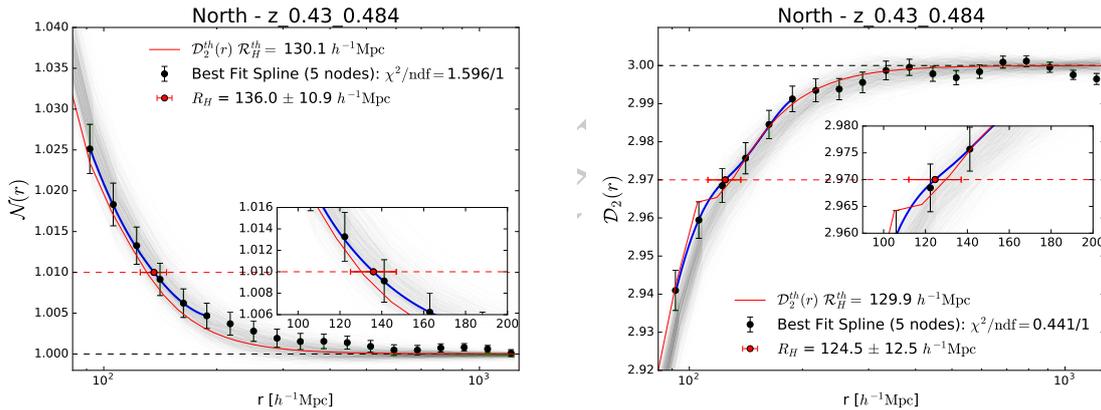
<sup>4</sup>We show the correlation matrices,  $\rho_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}$ , which have less dynamics than the covariance matrices.



**Figure 3.** Correlation matrices for  $\xi(r)$ ,  $\mathcal{N}(<r)$  and  $\mathcal{D}_2(r)$  for the total matter distribution in the redshift range  $0.538 \leq z \leq 0.592$ .

### 5.3 Homogeneity scale for the galaxy distribution

We first compute  $\mathcal{N}^{gal}(<r)$  and  $\mathcal{D}_2^{gal}(r)$  for the CMASS galaxy distribution using equation 3.6 and 3.3, without correcting for bias and redshift space distortions. This provides a measurement that does not rely on  $\Lambda$ CDM model to determine the bias. Results are shown in figure 4 for the  $0.430 \leq z \leq 0.484$  redshift interval in the NGC. As described in section 5.2, the error are obtained from the 1000 QPM mock catalogues.



**Figure 4.** Scaled counts-in-spheres,  $\mathcal{N}(<r)$ , (left) and fractal correlation dimension,  $\mathcal{D}_2(r)$ , (right). The black points with error bars are the result for the galaxy distribution in the NGC in the  $0.430 \leq z \leq 0.484$  redshift bin. The blues lines are the best spline fit of the data. The green lines are the results for the 1000 QPM mock catalogues. The red continuous line is the  $\Lambda$ CDM prediction for  $b = 1.95$  and  $\sigma_p = 252$  km/s.

The horizontal black-dashed lines in figure 4 indicate the value of the corresponding observable for an homogeneous distribution. The two observables reach homogeneity on large scales. The intersection of the data with the red-dashed lines (at  $\mathcal{N} = 1.01$  and  $\mathcal{D}_2 = 2.97$ ) defines the homogeneity scale. In order to determine it, we perform a spline fit over 6 data point around the intersection. The number of nodes of the spline fit is set to 5 in order to get an average  $\chi^2$  equal to the number of degrees of freedom for the 1000 QPM mocks. The error on  $R_H$  was obtained from the errors in the spline factors by error propagation. The results, presented in Table 3 for the case of  $\mathcal{D}_2$ , are consistent with our  $\Lambda$ CDM model predictions. We finally stress the fact that the behaviour of the data is consistent with that of the 1000 QPM mock catalogues.

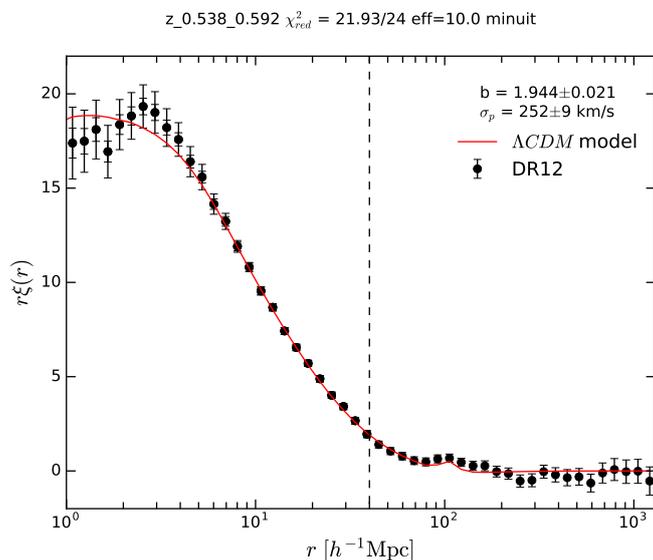
## 5.4 Determining the bias and the velocity dispersion

As illustrated in figure 5, we fit our model (Eq. 4.5) to the measured correlation function for the CMASS galaxy sample, in order to determine the galaxy bias and velocity dispersion. As explained in Appendix C, we account for theoretical uncertainties in the RSD modelling by boosting the covariance matrix on the relevant small scales. We thus achieve a good  $\chi^2/\text{n.d.f.}$  in the range  $[1, 40] h^{-1} \text{ Mpc}$ . Appendix C shows that this change in the covariance matrix at small scales has a negligible effect on the measurement of the homogeneity scale.

Results are given in table 2. The mean precision is 2.6% (NGC) and 4.1% (SGC) for the bias and 12% (NGC) and 23% (SGC) for the velocity dispersion. The bias is in agreement with values obtained by several authors [58–63] and the velocity dispersion is consistent with the typical CMASS-galaxy-sample velocity-dispersion,  $\sigma_p \approx 240 \pm 50 \text{ km/s}$  [64].

$z$	NGC			SGC		
	$b$	$\sigma_p$ [km/s]	$\chi_{red}^2$	$b$	$\sigma_p$ [km/s]	$\chi_{red}^2$
0.430-0.484	$1.879 \pm 0.023$	$243.0 \pm 9.7$	0.44	$1.872 \pm 0.035$	$238.7 \pm 15.2$	0.85
0.484-0.538	$1.846 \pm 0.021$	$247.1 \pm 7.4$	0.67	$1.840 \pm 0.032$	$234.6 \pm 12.1$	0.66
0.538-0.592	$1.944 \pm 0.021$	$252.0 \pm 8.9$	0.91	$1.943 \pm 0.032$	$236.2 \pm 13.7$	0.59
0.592-0.646	$1.995 \pm 0.022$	$234.2 \pm 10.9$	1.47	$2.001 \pm 0.035$	$237.9 \pm 19.6$	0.79
0.646-0.700	$2.153 \pm 0.028$	$235.3 \pm 21.7$	0.79	$2.081 \pm 0.045$	$210.3 \pm 40.1$	0.50

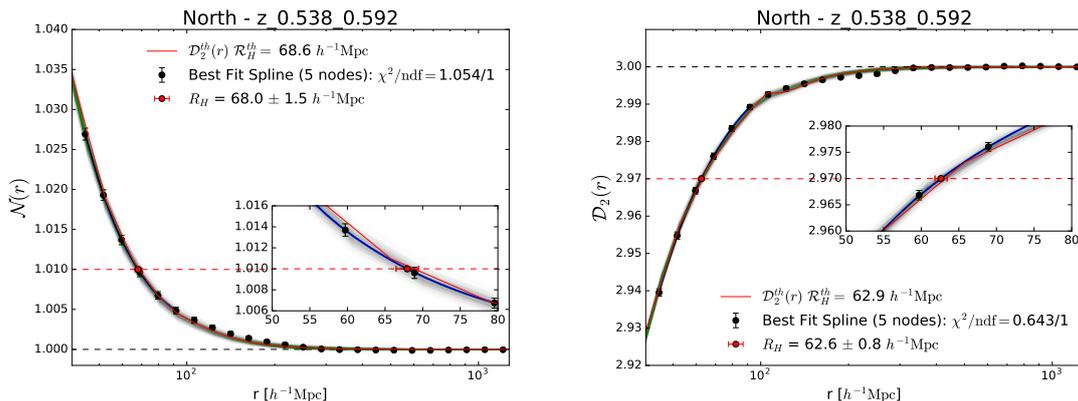
**Table 2.** Fitted values of bias,  $b$ , and velocity dispersion,  $\sigma_p$ , in the different redshift bins, together with the corresponding reduced  $\chi^2$  for 24 degrees of freedom.



**Figure 5.** The correlation function of the CMASS galaxies in the  $[0.538 - 0.592]$  redshift range. Data points have both their original error bars and the error bars enlarged to take into account the uncertainty of the RSD model on small scales (see section 5.4 and appendix C). The red line is the result of the fit performed over the range  $[1, 40] h^{-1} \text{ Mpc}$  (up to the vertical black-dashed line).

## 5.5 Homogeneity scale for the matter distribution

We use equation (4.6) with parameter  $b$  and  $\sigma_p$  obtained in section 5.4 to transform  $\mathcal{N}^{gal}(< r)$  into  $\mathcal{N}(< r)$  for matter. We then use Eq 3.3 to get  $\mathcal{D}_2(r)$  for matter distribution. Results are shown in figure 6 for the redshift interval  $0.538 \leq z \leq 0.592$ . The two observables indicate homogeneity on large scales in this redshift interval, and in the four other intervals as well. As in section 5.3 we fit the data points to determine the homogeneity scales. We stress that the fit range,  $40 < r < 100 h^{-1}$  Mpc, does not overlap with the fit range for determining the bias,  $1 < r < 40 h^{-1}$  Mpc.



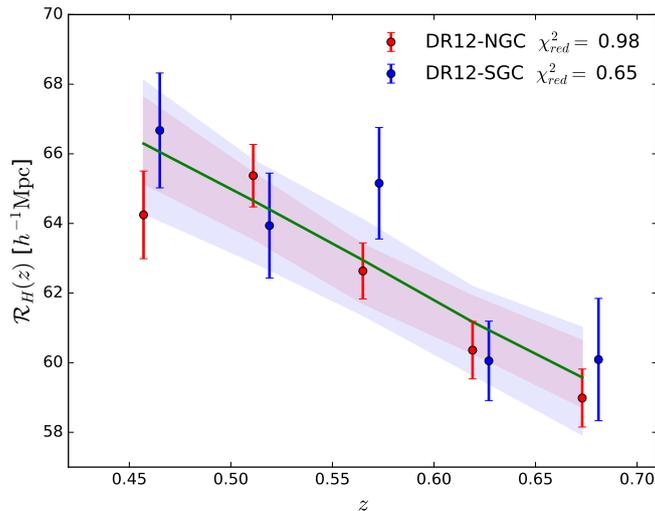
**Figure 6.** Same as figure 4 for matter distribution in the redshift interval  $0.538 \leq z \leq 0.592$ .

The results, presented in Table 3 for the case of  $\mathcal{D}_2$ , are consistent with our  $\Lambda$ CDM model predictions. In the redshift interval  $0.538 \leq z < 0.592$  we get a precision of 1.6%, which is a factor 5 better than Scrimgeour et al. [26] in spite of their wider redshift range,  $0.5 < z < 0.7$ . The more recent analysis by Sarkar et al. [27] does not give  $R_H$  for matter distribution.

Figure 7 shows that the measured homogeneity scale is compatible with  $\Lambda$ CDM, with reduced  $\chi^2$  smaller than unity both in the NGC and the SGC. The scale is increasing with time, from  $z = 0.7$  to  $z = 0.43$ , as expected when clustering is increasing with time. The reduced  $\chi^2$  remains smaller than unity when adding the result from Laurent et al. [38],  $\mathcal{R}_H(z = 2.4) = 26.2 \pm 0.9 h^{-1}$  Mpc.

$z$	$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Matter		$\mathcal{R}_H[h^{-1}\text{Mpc}]$ Galaxy	
	NGC	SGC	NGC	SGC
0.430 – 0.484	$64.2 \pm 1.3$	$66.7 \pm 1.6$	$124.5 \pm 12.5$	$121.1 \pm 9.8$
0.484 – 0.538	$65.4 \pm 0.9$	$63.9 \pm 1.5$	$111.9 \pm 4.9$	$119.8 \pm 8.8$
0.538 – 0.592	$62.6 \pm 0.8$	$65.2 \pm 1.6$	$116.4 \pm 7.8$	$110.5 \pm 5.1$
0.592 – 0.646	$60.4 \pm 0.8$	$60.1 \pm 1.1$	$108.8 \pm 3.9$	$120.1 \pm 11.7$
0.646 – 0.700	$59.0 \pm 0.8$	$60.1 \pm 1.8$	$125.8 \pm 7.3$	$147.4 \pm 8.4$

**Table 3.** Homogeneity scale,  $R_H^{D_2=2.97}(z)$ , for the galaxy and matter distributions in the north and south galactic caps.



**Figure 7.** The homogeneity scale  $R_H^{\mathcal{D}_2=2.97}(z)$  measured in the NGC (red) and in the SGC (blue) as a function of redshift. The green line is the  $\Lambda$ CDM model prediction. The shaded areas indicate the  $1\sigma$  range for the 1000 QPM mock catalogues.

## 5.6 Consistency with $\Lambda$ CDM

In order to characterize more precisely the agreement with the  $\Lambda$ CDM model, we fit  $\mathcal{D}_2(r)$  from the data and the 1000 mock catalogues in the range  $40 h^{-1} < r < 1300 h^{-1}$  Mpc with

$$\mathcal{D}_2^{\Lambda\text{CDM}} = \mathcal{D}_2(ar; \mathbf{p}_{\text{cosmo}}). \quad (5.1)$$

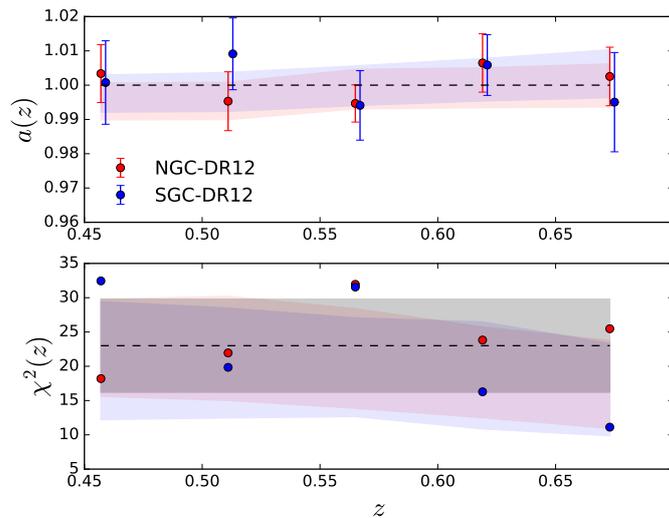
We fix the cosmological parameters  $\mathbf{p}_{\text{cosmo}}$  and leave only free the  $a$ .

Figure 8 shows the results of the fits in different redshift bins for both NGC and SGC. The values of  $a$ , all consistent with 1, demonstrate a good agreement with  $\Lambda$ CDM model. Furthermore, the bottom plot of the figure shows that, for the data, the  $\chi^2$  are consistent with the number of degrees of freedom (23), as illustrated by the grey shaded area that indicates the expected  $1\sigma$  extension,  $\chi^2 = 23 \pm \sqrt{2} \times 23$ . The mock catalogues, however, are a bit at variance with this grey area at large  $z$ .

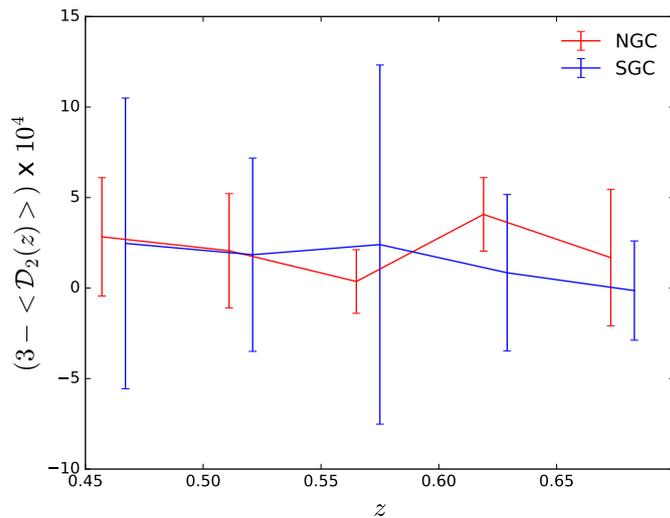
## 5.7 Constraints on fractal correlation dimension

$z$	Matter		Galaxy	
	NGC ( $\times 10^{-4}$ )	SGC ( $\times 10^{-4}$ )	NGC ( $\times 10^{-3}$ )	SGC ( $\times 10^{-3}$ )
0.430 – 0.484	$2.8 \pm 3.3$	$2.5 \pm 8.1$	$1.3 \pm 1.5$	$1.1 \pm 3.7$
0.484 – 0.538	$2.0 \pm 3.1$	$1.9 \pm 5.4$	$0.9 \pm 1.4$	$0.8 \pm 2.4$
0.538 – 0.592	$3.6 \pm 1.7$	$2.4 \pm 9.9$	$0.2 \pm 0.9$	$1.2 \pm 4.9$
0.592 – 0.646	$4.1 \pm 2.0$	$0.9 \pm 4.5$	$2.1 \pm 1.1$	$0.5 \pm 2.3$
0.646 – 0.700	$1.7 \pm 3.8$	$-0.2 \pm 2.8$	$1.0 \pm 2.3$	$0.1 \pm 1.6$

**Table 4.** Fractal correlation dimension,  $3 - \langle \mathcal{D}_2 \rangle_r(z)$ , averaged over  $300 h^{-1} < r < 1300 h^{-1}$  Mpc, with  $1\sigma$  errors, in the NGC and SGC.



**Figure 8.** Top: value of the parameter  $\alpha$  resulting from the fit of the NGC (red) and SGC (blue)  $\mathcal{D}_2(r)$  data with the model of equation 5.1, in different redshift bins. The shaded areas correspond to the  $1\sigma$  region for the 1000 QPM mock catalogues. Bottom: The corresponding  $\chi^2$ .



**Figure 9.** Fractal correlation dimension for the matter distribution,  $3 - \langle \mathcal{D}_2 \rangle$ , averaged over  $300 h^{-1} < r < 1300 h^{-1}$  Mpc, in the different redshift bins.

In order to assess the level of homogeneity of the CMASS DR12 galaxy sample we compute the average of  $\mathcal{D}_2(r)$  over the range  $300 h^{-1} < r < 1300 h^{-1}$  Mpc, accounting for the covariance matrix. The results are presented in table 4 for the different redshift bins in the NGC and SGC. All results are compatible and we average them to get  $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (0.9 \pm 1.2) \times 10^{-3}$  ( $1\sigma$ ).

Strictly speaking, it is not possible to transform the results for the galaxy distribution into results for matter distribution without using  $\Lambda$ CDM prediction to compute the galaxy

bias. However, Laurent et al. [38] show that, with a set of reasonable assumptions independent of  $\Lambda$ CDM, it is possible to obtain a lower limit for the tracer bias. In our case this results in  $b > \sqrt{1.6}$  and  $3 - \langle \mathcal{D}_2 \rangle < 7.12 \pm 9.48 \times 10^{-4}$  for matter distribution.

Alternatively we can assume  $\Lambda$ CDM, use it to get the galaxy bias and compute  $\mathcal{D}_2(r)$ . We average  $\mathcal{D}_2(r)$  over  $300 h^{-1} < r < 1300 h^{-1}$  Mpc and give the results in table 4 and in figure 9. Since the galaxy bias is significantly larger than one, the constraints are tighter for matter distribution than for galaxy distribution. Averaging over redshift and caps we get  $3 - \langle \mathcal{D}_2 \rangle_{r,z} = (1.7 \pm 1.0) \times 10^{-4}$  ( $1\sigma$ ). This is a strong consistency check of  $\Lambda$ CDM.

Using a similar estimator to Yadav et al. [18] on 1000 QPM mock catalogues, we find a minimum  $160 h^{-1}$ Mpc, maximum  $1250 h^{-1}$ Mpc, a spread of  $120 h^{-1}$ Mpc and an average of  $320 h^{-1}$ Mpc corresponding to a precision of about 30% which is less precise than the results, we present in this study. A KS-test [65] for this estimator suggest that data and mocks are drawn from the same distribution at 21% and 83% C.L. for NGC and SGC respectively.

DRAFT

## 6 Analysis Robustness

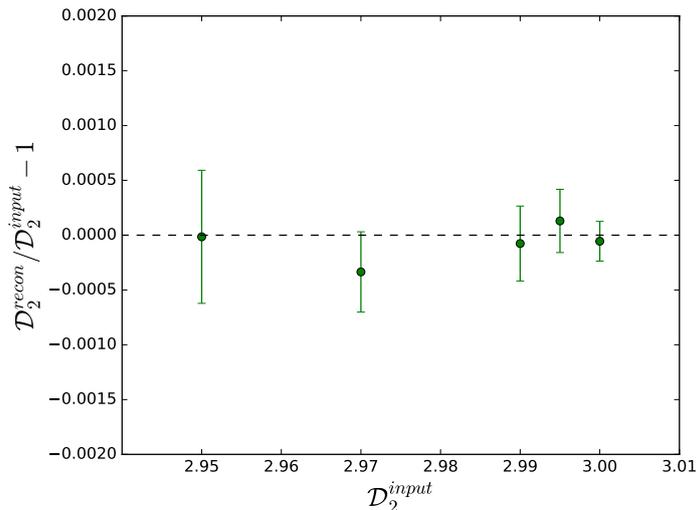
### 6.1 Bias in the fractal correlation dimension

We use homogeneous random catalogues to take into account the inhomogeneity of the survey, so we may wonder if this could bias the resulting  $\mathcal{N}(< r)$  towards homogeneity. To search for such a possible bias, we generate 500 fractal realizations with a given value of the fractal correlation dimension, we pass them through our pipeline analysis and study the resulting  $\mathcal{D}_2$ , as done by Laurent et al. [38].

Following Castagnoli and Provenzale [66], we create a cubic box of  $L \approx 4 h^{-1}$  Gpc side, containing the whole survey at  $z = 0.538 - 0.592$ . We divide this box in  $M = n^3$  sub-boxes of size  $L/n$ , where  $n = 2$  and we give to each sub-box a survival probability  $p$ . We then repeat the procedure for each surviving sub-box. An infinite number of iterations would give a fractal distribution with:

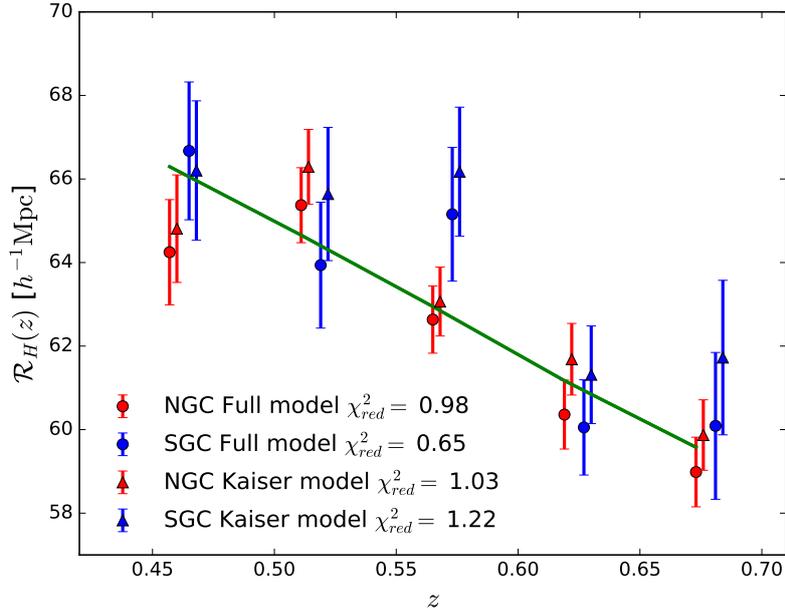
$$D_2 = \frac{\log(pM)}{\log(n)} \quad (6.1)$$

We perform 9 iterations. After the last iteration, we populate each sub-box with random points that follow a Poisson law of mean  $\lambda < 1$ . Then we convert the cartesian coordinates to  $z, RA, DEC$  with the same FLRW metric as used in our analysis. Finally, we apply cuts to simulate the selection function of our galaxy survey. The value of  $\lambda$  is chosen so that after cuts the number of objects in the fractal distribution is approximately the same as in our survey.



**Figure 10.** Reconstructed fractal correlation dimension,  $\mathcal{D}_2$ , in the redshift interval  $z = 0.538-0.592$ , averaged over  $r \in [15, 1300] h^{-1}$  Mpc.

At the last iteration the size of the sub-box is  $15 h^{-1}$  Mpc. So we average the reconstructed  $\mathcal{D}_2$  over  $r \in [15, 1300] h^{-1}$  Mpc. Figure 10 presents this average for different input  $\mathcal{D}_2$ . Fitting these points with a straight line,  $y = \alpha x + \beta$ , results in  $\alpha = 1.00 \pm 0.03$  and  $\beta = (-0.4 \pm 7.9) \times 10^{-2}$  with  $\chi_{\text{red}}^2 = 0.9/3$ . So, in contrast with Scrimgeour et al. [26], we do not observe any bias in the reconstructed fractal correlation dimension.



**Figure 11.** Measured homogeneity scale in different redshift bins, in the NGC and the SGC, with Kaiser (triangle) and full RSD model (circles).

## 6.2 Sensitivity to RSD model

We compare the homogeneity scale obtained with two different modelling of redshift space distortions. Namely, the model with a scale-independent bias and the Kaiser effect, see equation (4.1) :

$$\xi(r; b) = FFT [P_{gg}(k, \mu; b)] , \quad (6.2)$$

and our full model, which includes in addition a modeling of the finger-of-God effect (Eq. 4.3):

$$\xi(r; b, \sigma_p) = FFT [P_{gg}(k, \mu; b, \sigma_p)] \quad (6.3)$$

Figure 11 presents the homogeneity scale for these two models in the five redshift bins for the north and south galactic caps. We note that using the full RSD model lowers the  $\chi^2$ , confirming that this is a better model than the purely linear Kaiser model. The full model tends to lower  $R_H$ , by up to slightly more than  $1\sigma$ , or about 1%. This means that we should not use the purely linear Kaiser model. The remaining error due to the imperfection of our full RSD model is most probably a fraction of the difference between Kaiser and full models, and therefore negligible with respect to our statistical error.

## 7 Conclusions

We use the data release 12 of BOSS CMASS galaxy sample to study the transition to cosmic homogeneity over a volume of  $5.1 h^{-3} \text{ Gpc}^3$ . We do not consider the correlation function,  $\xi(r)$ , to study homogeneity because its definition requires an average density, which is only defined for a homogeneous sample. We rather use the counts-in-spheres,  $N(< r)$ , i.e. the average number of objects around a given object, and its logarithmic derivative, the fractal correlation dimension,  $D_2(r)$ . For a homogeneous sample,  $N(< r) \propto r^3$  and  $D_2 = 3$ . We define a characteristic homogeneity scale,  $\mathcal{R}_H$ , as the value for which  $D_2$  reaches the homogeneous value within 1%, i.e.  $D_2(\mathcal{R}_H) = 2.97$ .

For the distribution of galaxies, we get  $3 - \langle D_2 \rangle = (0.6 \pm 1.3) \times 10^{-3}$  at  $1 \sigma$  over the range  $300 h^{-1} \leq r \leq 1300 h^{-1} \text{ Mpc}$ , consistent with homogeneity and a transition to homogeneity at  $\mathcal{R}_H = 114.2 \pm 5.8 h^{-1} \text{ Mpc}$ . However, our analysis makes use of a random catalog to take into account the geometry and the completeness of the survey. The redshift distribution of this catalogue is taken from the data. We are therefore insensitive to a possible isotropic variation of the density with redshift,  $\rho = \rho(z)$ . In other words, we can only check for spatial isotropy,  $\rho(z, \theta_1) = \rho(z, \theta_2)$ . We stress that the same is true for all galaxy redshift surveys [39]. On the other hand, this spatial isotropy can be obtained without using any fiducial model as discussed by Laurent et al. [38]. So if we assume the Copernican principle our data imply homogeneity of the galaxy sample without any  $\Lambda$ CDM assumption.

Using an estimator similar to that of Yadav et al. [18], we find agreement between the mocks and data, further confirming the transition to homogeneity in the matter and galaxy distributions, despite the lower precision of this estimator (30% compared to our estimator which is near %-level). This estimator provides a qualitative estimate of the homogeneity scale in the range of  $150 h^{-1} \leq r \leq 400 h^{-1} \text{ Mpc}$ .

Alternatively, we can make a cross check of  $\Lambda$ CDM model. Thus, we fit the CMASS galaxy correlation function in the range  $1 h^{-1} < r < 40 h^{-1} \text{ Mpc}$  to obtain the galaxy bias relative to the  $\Lambda$ CDM prediction for the matter correlation function. We correct our measurement of  $N(< r)$  for this bias in order to get the result for the matter distribution, that we finally compare to the  $\Lambda$ CDM prediction.

For the matter distribution, we get  $3 - \langle D_2(r > 300 h^{-1} \text{ Mpc}) \rangle_z = (1.7 \pm 1.0) \times 10^{-4}$  at  $1\sigma$  and a transition to homogeneity at a characteristic scale  $\mathcal{R}_H = 61.9 \pm 0.8 h^{-1} \text{ Mpc}$  at an average  $z = 0.538 - 0.592$ . This measurement of  $\mathcal{R}_H$  is more precise than previous measurement by Scrimgeour et al. [26] by a factor 5, while Sarkar et al. [27], only give a qualitative measurement of  $\mathcal{R}_H$ . We also investigate the redshift evolution of our observables. We find that the homogeneity scale is decreasing with time as expected if clustering is increasing with time. We find accordance with  $\Lambda$ CDM model with a reduced  $\chi^2 = 0.89(0.61)$  for the North(South) Galactic Cap for the 6 redshift bins.

## Acknowledgements

We would like to thank Jean-Philippe Uzan, Cyril Pitrou and Ruth Durrer for useful suggestions and discussions. PN would like to thank James Rich, although a coauthor, for fruitful discussions on the understanding of this work. PN would like also to thank Mikhail Stolpovskiy and Ranajoy Banerji for useful suggestion on the computational side of this project.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University,

Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

MV is partially supported by Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPITT) No IA102516, Proyecto Conacyt Fronteras No 281 and Proyecto DGTIC SC16-1-S-120.

GR acknowledges support from the National Research Foundation of Korea (NRF) through NRF-SGER 2014055950 funded by the Korean Ministry of Education, Science and Technology (MoEST), and from the faculty research fund of Sejong University in 2016

This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

DRAFT

## A From $\xi(r)$ to $\mathcal{N}(< r)$

The probability of finding a galaxy within a volume  $dV$  around another galaxy depends on the two-point correlation function  $\xi(\vec{r})$  [67]:

$$dP = \bar{\rho} [1 + \xi(\vec{r})] dV . \quad (\text{A.1})$$

The counts-in-spheres of the distribution of galaxies is then related to the correlation function:

$$N(< r) = \int dP = \bar{\rho} \int [1 + \xi(\vec{r}')] dV . \quad (\text{A.2})$$

Assuming  $\xi(\vec{r}) = \xi(r)$  we get:

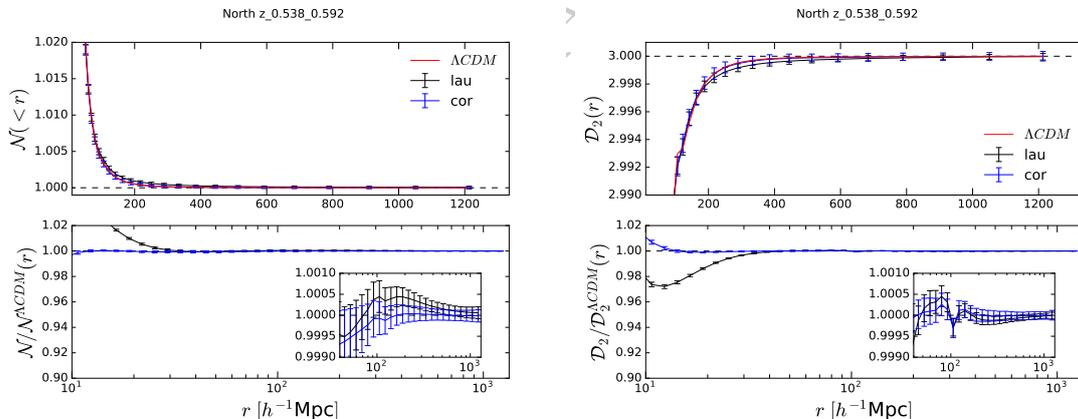
$$N(< r) = 4\pi\bar{\rho} \int_0^r [1 + \xi(r')] r'^2 dr' . \quad (\text{A.3})$$

For the random homogeneous distribution,  $N_R(< r) = \bar{\rho} \frac{4\pi}{3} r^3$ , so

$$\mathcal{N}(< r) = \frac{N(< r)}{N_R(< r)} = \frac{3}{r^3} \int_0^r [1 + \xi(r')] r'^2 dr' = 1 + \frac{3}{r^3} \int_0^r \xi(r') r'^2 dr' . \quad (\text{A.4})$$

## B Choice of estimator for $\mathcal{N}(< r)$

In section 3.2, we consider two estimators for  $\mathcal{N}(< r)$ , defined by equations (3.5) and (3.6). We compute  $\mathcal{N}(< r)$  and the resulting  $\mathcal{D}_2(r)$  with the two estimators for the 1000 QPM mock catalogues. Figure 12 compares the mean of the 1000 mocks to the  $\Lambda$ CDM model. The result with the cor estimator are much closer to the  $\Lambda$ CDM model.

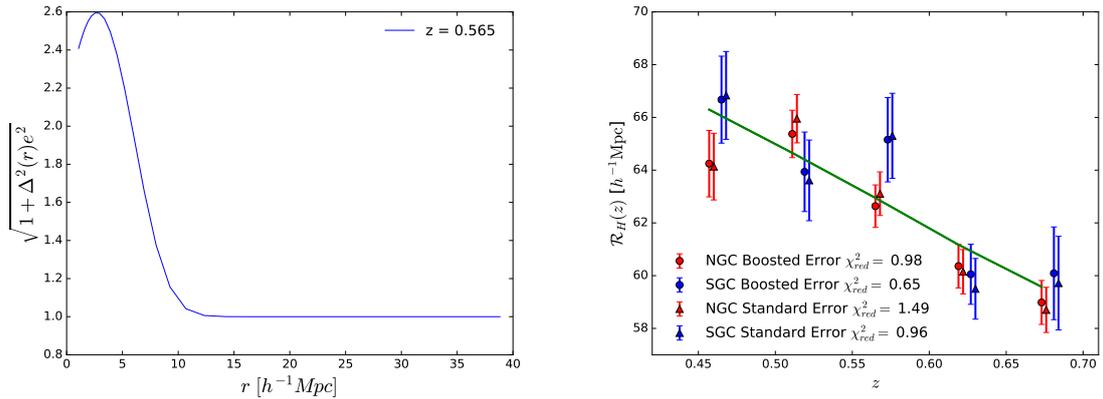


**Figure 12.** Top: the scaled counts-in-spheres,  $\mathcal{N}(< r)$ , (left) and the fractal correlation dimension,  $\mathcal{D}_2(r)$ , (right) for matter distribution, with lau (black) and cor (blue) estimators, compared to  $\Lambda$ CDM model (red). Bottom: the ratio to  $\Lambda$ CDM model for both estimator.

## C Tuning errors in RSD Analysis

The theoretical model for redshift space distortions (Eq. 4.3) is not perfectly accurate at the smallest scales due to the real nonlinear behaviour of gravity at these scales. In order to ensure satisfying  $\chi^2/\text{n.d.f.}$  for the RSD fitting, we boost the error on  $\xi(r)$  at the relevant scales in an empirical way as

$$C'_{ij} = C_{ij}(1 + \delta_{ij} \Delta_i \Delta_j e^2) . \quad (\text{C.1})$$



**Figure 13.** Left: Boosting factor on the error in the correlation function. Right:  $R_H^{D_2=2.97}(z)$  with (circle) and without (triangle) boosting the errors, for NGC (red) and SGC (blue).

Here  $\delta_{ij}$  is the usual Kronecker symbol;  $e$  is a parameter that measures the amount of error boosting we apply;  $\Delta_i$  is the theoretical inaccuracy, estimated as the relative difference between our model and the average of the 1000 QPM mock catalogues. Figure 13 (left) presents the resulting boosting factor on the error in the correlation function. It appears to be significant only on scales smaller than 10  $h^{-1}$  Mpc. Fig. 13 (right) shows that the reconstructed homogeneity scales measured with  $\mathcal{D}_2$  is not significantly modified by the error boosting.

## D Test of spline fit on QPM mock catalogues

We perform a spline fit of  $\mathcal{D}_2$  for the 1000 QPM mock catalogues in order to obtain the homogeneity scale at 1% (Eq. 3.7). The fit is performed in the range  $r \in [40, 100] h^{-1}$  Mpc with 6 data points and 1 degree of freedom. The distribution of the  $\chi^2$  of the mock should therefore follow a  $\chi^2$ -law for 1 degree of freedom. In table (5) we show the mean and the error on the mean of the distribution of the corresponding  $\chi^2$  for the 1000 QPM mock catalogues. The test is successful in both NGC and SGC.

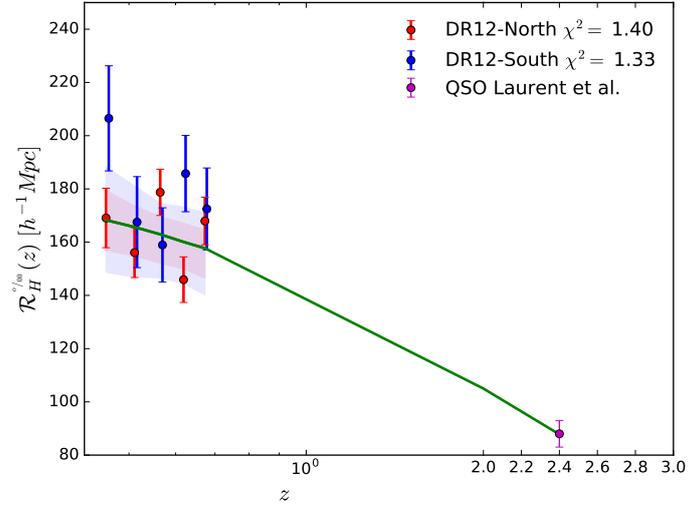
$z$	$\chi_{\text{NGC}}^2$	$\chi_{\text{SGC}}^2$
0.430-0.484	$1.00 \pm 0.05$	$0.99 \pm 0.04$
0.484-0.538	$0.99 \pm 0.05$	$1.00 \pm 0.04$
0.538-0.592	$1.02 \pm 0.05$	$1.00 \pm 0.04$
0.592-0.646	$0.99 \pm 0.04$	$1.00 \pm 0.05$
0.646-0.700	$1.02 \pm 0.05$	$1.00 \pm 0.04$

**Table 5.** Mean and error over the 1000 QPM mock catalogues for the  $\chi^2$  of the spline fit with 1 degree of freedom, in the NGC and SGC for the five redshift bins.

## E Homogeneity scale at 0.1%

The choice of a 1% threshold to define the homogeneity scales is arbitrary. We can define them for instance at 0.1% as:

$$\mathcal{D}_2(R_H^{D_2=2.997}) = 2.997 \quad \text{or} \quad \mathcal{N}(R_H^{\mathcal{N}=1.001}) = 1.001 \quad (\text{E.1})$$



**Figure 14.** The homogeneity scale at 0.1% level,  $R_H^{\mathcal{D}_2=2.997}(z)$ , measured in the NGC (red) and in the SGC (blue) as a function of redshift. The purple point is the result obtained with quasars in the NGC, in the redshift range  $2.2 \leq z \leq 2.8$  by Laurent et al. [38]. The green line is the  $\Lambda\text{CDM}$  model prediction. The shaded areas indicate the  $1\sigma$  range for the 1000 QPM mock catalogues.

Figure 14 shows that the measured homogeneity scale for matter distribution is compatible with  $\Lambda\text{CDM}$ , with  $\chi_{red}^2 = 5.82/6$  in the NGC and  $\chi_{red}^2 = 7.98/6$  in the SGC.

## References

- [1] Perlmutter, S., G. Aldering, G. Goldhaber, et al. Measurements of  $\omega$  and  $\lambda$  from 42 high-redshift supernovae. *The Astrophysical Journal*, 517:565, 1999.
- [2] Riess, A. G., A. V. Filippenko, P. Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116:1009, 1998.
- [3] Adam, R., P. Ade, N. Aghanim, et al. Planck 2015 results. i. overview of products and scientific results. 2015. [arXiv:1502.01582](#).
- [4] Percival, W. et al. Parameter constraints for flat cosmologies from cmb and 2dfgrs power spectra, 2002. *Mon. Not. R. Astron. Soc.*, 337:1068. [arXiv:astro-ph/0206256](#).
- [5] Parkinson, D., S. Riemer-Sørensen, C. Blake, et al. The wigglez dark energy survey: final data release and cosmological results. *Physical Review D*, 86:103518, 2012. [arXiv:1210.2130](#).
- [6] Heymans, C., E. Grocutt, A. Heavens, et al. Cfhtlens tomographic weak lensing cosmological parameter constraints: Mitigating the impact of intrinsic galaxy alignments. *Monthly Notices of the Royal Astronomical Society*, 432:2433–2453, 2013. [arXiv:1303.1808](#).
- [7] Ellis, G. F. Cosmology and verifiability. *Quarterly Journal of the Royal Astronomical Society*, 16:245–264, 1975.
- [8] Fixsen, D., E. Cheng, J. Gales, et al. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *The Astrophysical Journal*, 473:576, 1996. [astro-ph/9605054](#).
- [9] Schärf, C., M. Treyer, K. Jahoda, et al. The 2-10 keV xrb dipole and its cosmological implications. 1999. [arXiv:astro-ph/9908187](#).
- [10] Blake, C. and J. Wall. Detection of the velocity dipole in the radio galaxies of the nrao vla sky survey. 2002. [astro-ph/0203385](#).
- [11] Hamilton, J.-C. What have we learned from observational cosmology? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 46:70–85, 2014.
- [12] Maartens, R. Is the universe homogeneous? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369:5115–5137, 2011. [arXiv:1104.1300](#).
- [13] Clarkson, C. Establishing homogeneity of the universe in the shadow of dark energy. *Comptes Rendus Physique*, 13:682–718, 2012.
- [14] Straumann, N. Minimal assumptions leading to a robertson-walker model of the universe. *Helvetica Physica Acta*, 47:379–383, 1974.
- [15] Durrer, R., J.-P. Eckmann, F. S. Labini, et al. Angular projections of fractal sets. *EPL (Europhysics Letters)*, 40:491, 1997.
- [16] Hogg, D. W., D. J. Eisenstein, M. R. Blanton, et al. Cosmic homogeneity demonstrated with luminous red galaxies. *The Astrophysical Journal*, 624:54, 2005. [arXiv:astro-ph/0411197v1](#).
- [17] Yadav, J., S. Bharadwaj, B. Pandey, et al. Testing homogeneity on large scales in the sloan digital sky survey data release one. *Monthly Notices of the Royal Astronomical Society*, 364: 601–606, 2005. [arXiv:astro-ph/0504315](#).

- [18] Yadav, J. K., J. Bagla, and N. Khandai. Fractal dimension as a measure of the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 405:2009–2015, 2010. [arXiv:1001.0617](#).
- [19] Sarkar, P., J. Yadav, B. Pandey, et al. The scale of homogeneity of the galaxy distribution in sdss dr6. *Monthly Notices of the Royal Astronomical Society: Letters*, 399:L128–L131, 2009. [arXiv:0906.3431](#).
- [20] Martinez, V. J. and P. Coles. Correlations and scaling in the qdot redshift survey. *The Astrophysical Journal*, 437:550–555, 1994.
- [21] Guzzo, L. Is the universe homogeneous?(on large scales). *New Astronomy*, 2:517–532, 1997. [arXiv:astro-ph/9711206](#).
- [22] Martínez, V. J., M.-J. Pons-Borderia, R. A. Moyeed, et al. Searching for the scale of homogeneity. *Monthly Notices of the Royal Astronomical Society*, 298:1212–1222, 1998. [arXiv:astro-ph/9804073](#).
- [23] Scaramella, R., L. Guzzo, G. Zamorani, et al. The eso slice project [esp] galaxy redshift survey: V. evidence for a  $d=3$  sample dimensionality. 1998. [arXiv:astro-ph/9803022](#).
- [24] Amendola, L. and E. Palladino. The scale of homogeneity in the las campanas redshift survey. *The Astrophysical Journal Letters*, 514:L1, 1999. [arXiv:astro-ph/9901420](#).
- [25] Pan, J. and P. Coles. Large-scale cosmic homogeneity from a multifractal analysis of the psz catalogue. *Monthly Notices of the Royal Astronomical Society*, 318:L51–L54, 2000. [arXiv:astro-ph/0008240v1](#).
- [26] Scrimgeour, M. I., T. Davis, C. Blake, et al. The wigglez dark energy survey: the transition to large-scale cosmic homogeneity. *Monthly Notices of the Royal Astronomical Society*, 425:116–134, 2012. [arXiv:1205.6812](#).
- [27] Sarkar, P., S. Majumdar, B. Pandey, et al. The many scales to cosmic homogeneity: Use of multiple tracers from the sdss. 2016. [arXiv:1611.07915](#).
- [28] Appleby, S. and A. Shafieloo. Testing isotropy in the local universe. *Journal of Cosmology and Astroparticle Physics*, 2014:070, 2014. [arXiv:1405.4595](#).
- [29] Labini, F. S., M. Montuori, and L. Pietronero. Comment on the paper by l. guzzo" is the universe homogeneous?". 1998. [arXiv:astro-ph/9801151](#).
- [30] Coleman, P. H. and L. Pietronero. The fractal structure of the universe. *Physics Reports*, 213:311–389, 1992.
- [31] Labini, F. S., M. Montuori, and L. Pietronero. Scale-invariance of galaxy clustering. *Physics Reports*, 293:61–226, 1998. [arXiv:astro-ph/9711073](#).
- [32] Joyce, M., M. Montuori, and F. S. Labini. Fractal correlations in the cfa2-south redshift survey. *The Astrophysical Journal Letters*, 514:L5, 1999. [arXiv:astro-ph/9901290](#).
- [33] Labini, F. S. Breaking the self-averaging properties of spatial galaxy fluctuations in the sloan digital sky survey–data release six. *Astronomy & Astrophysics*, 508:17–43, 2009. [arXiv:astro-ph/9901290](#).
- [34] Labini, F. S. Very large-scale correlations in the galaxy distribution. *EPL (Europhysics Letters)*, 96:59001, 2011. [arXiv:1110.4041](#).
- [35] Hoyle, B., R. Tojeiro, R. Jimenez, et al. Testing homogeneity with galaxy star formation histories. *The Astrophysical Journal Letters*, 762:L9, 2012.

- [36] Zibin, J., A. Moss, et al. Nowhere to hide: closing in on cosmological homogeneity. 2014. [arXiv:1409.3831](#).
- [37] Wu, K. K., O. Lahav, and M. J. Rees. The large-scale smoothness of the universe. *Nature*, 397: 225–230, 1999.
- [38] Laurent, P., J.-M. Le Goff, E. Burtin, et al. A  $14 \text{ h}^{-3} \text{ Gpc}^3$  study of cosmic homogeneity using BOSS DR12 quasar sample. 2016. [arXiv:1602.09010](#).
- [39] Mustapha, N., C. Hellaby, and G. Ellis. Large-scale inhomogeneity versus source evolution: can we distinguish them observationally? *Monthly Notices of the Royal Astronomical Society*, 292: 817–830, 1997.
- [40] Planck Collaboration and Ade, P. A. R. and Aghanim, N. and Arnaud, M. and Ashdown, M. and Aumont, J. and Baccigalupi, C. and Banday, A. J. and Barreiro, R. B. and Bartlett, J. G. and et al., t. . P. 2016. [arXiv:1502.01589](#).
- [41] Dawson, K. S., D. J. Schlegel, C. P. Ahn, et al. The baryon oscillation spectroscopic survey of sdss-iii. *The Astronomical Journal*, 145:10, 2012.
- [42] Alam, S., F. D. Albareti, C. A. Prieto, et al. The eleventh and twelfth data releases of the sloan digital sky survey: final data from sdss-iii. *The Astrophysical Journal Supplement Series*, 219: 12, 2015. [arXiv:1501.00963v3](#).
- [43] Eisenstein, D. J., D. H. Weinberg, E. Agol, et al. Sdss-iii: Massive spectroscopic surveys of the distant universe, the milky way, and extra-solar planetary systems. *The Astronomical Journal*, 142:72, 2011.
- [44] Gunn, J. E., W. A. Siegmund, E. J. Mannery, et al. The 2.5 m telescope of the sloan digital sky survey. *The Astronomical Journal*, 131:2332, 2006.
- [45] Smee, S. A., J. E. Gunn, A. Uomoto, et al. The multi-object, fiber-fed spectrographs for the sloan digital sky survey and the baryon oscillation spectroscopic survey. *The Astronomical Journal*, 146:32, 2013. [arXiv:1208.2233](#).
- [46] Bolton, A. S., D. J. Schlegel, É. Aubourg, et al. Spectral classification and redshift measurement for the sdss-iii baryon oscillation spectroscopic survey. *The Astronomical Journal*, 144:144, 2012. [arXiv:1207.7326](#).
- [47] Fukugita, M., T. Ichikawa, J. Gunn, et al. The sloan digital sky survey photometric system. *The Astronomical Journal*, 111:1748, 1996.
- [48] Maraston, C., G. Strömbäck, D. Thomas, et al. Modelling the colour evolution of luminous red galaxies—improvements with empirical stellar spectra. *Monthly Notices of the Royal Astronomical Society: Letters*, 394:L107–L111, 2009.
- [49] Aubourg, É., S. Bailey, F. Beutler, et al. Cosmological implications of baryon acoustic oscillation measurements. *Physical Review D*, 92:123516, 2015.
- [50] Reid, B., S. Ho, N. Padmanabhan, et al. Sdss-iii baryon oscillation spectroscopic survey data release 12: galaxy target selection and large-scale structure catalogues. *Monthly Notices of the Royal Astronomical Society*, 455:1553–1573, 2016.
- [51] Feldman, H. A., N. Kaiser, and J. A. Peacock. Power spectrum analysis of three-dimensional redshift surveys. *arXiv preprint astro-ph/9304022*, 1993.
- [52] White, M., J. L. Tinker, and C. K. McBride. Mock galaxy catalogues using the quick particle mesh method. *Monthly Notices of the Royal Astronomical Society*, 437:2594–2606, 2014.

- [53] Peebles, P. J. E. and M. G. Hauser. Statistical analysis of catalogs of extragalactic objects. iii. the shane-wirtanen and zwicky catalogs. *The Astrophysical Journal Supplement Series*, 28:19, 1974.
- [54] Landy, S. D. and A. S. Szalay. Bias and variance of angular correlation functions. *The Astrophysical Journal*, 412:64–71, 1993.
- [55] Blas, D., J. Lesgourgues, and T. Tram. The cosmic linear anisotropy solving system (class). part ii: approximation schemes. *Journal of Cosmology and Astroparticle Physics*, 2011:034, 2011.
- [56] Kaiser, N. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227:1–21, 1987.
- [57] Linder, E. V. and R. N. Cahn. Parameterized beyond-einstein growth. *Astroparticle Physics*, 28: 481–488, 2007. [arXiv:astro-ph/0701317v2](#).
- [58] Ballinger, W., J. Peacock, and A. Heavens. Measuring the cosmological constant with redshift surveys. *arXiv preprint astro-ph/9605017*, 1996. [arXiv:astro-ph/9605017v1](#).
- [59] White, M., M. Blanton, A. Bolton, et al. The clustering of massive galaxies at  $z \approx 0.5$  from the first semester of boss data. *The Astrophysical Journal*, 728:126, 2011. [arXiv:1010.4915v2](#).
- [60] Nuza, S. E., A. G. Sánchez, F. Prada, et al. The clustering of galaxies at  $z \approx 0.5$  in the sdss-iii data release 9 boss-cmass sample: a test for the  $\Lambda$ cdm cosmology. *Monthly Notices of the Royal Astronomical Society*, 432:743–760, 2013. [arXiv:1202.6057](#).
- [61] Ho, S., A. Cuesta, H.-J. Seo, et al. Clustering of sloan digital sky survey iii photometric luminous galaxies: the measurement, systematics, and cosmological implications. *The Astrophysical Journal*, 761:14, 2012.
- [62] Comparat, J., E. Jullo, J.-P. Kneib, et al. Stochastic bias of colour-selected bao tracers by joint clustering–weak lensing analysis. *Monthly Notices of the Royal Astronomical Society*, 433: 1146–1160, 2013. [arXiv:1302.5655v2](#).
- [63] Gil-Marín, H., J. Noreña, L. Verde, et al. The power spectrum and bispectrum of sdss dr11 boss galaxies-i. bias and gravity. *Monthly Notices of the Royal Astronomical Society*, 451:5058–5099, 2015.
- [64] Thomas, D., O. Steele, C. Maraston, et al. Stellar velocity dispersions and emission line properties of sdss-iii/boss galaxies. *Monthly Notices of the Royal Astronomical Society*, page stt261, 2013. [arXiv:1207.6115v2](#).
- [65] Kolmogorov–Smirnov test. Encyclopedia of Mathematics.
- [66] Castagnoli, C. and A. Provenzale. From small-scale fractality to large-scale homogeneity—a family of cascading models for the distribution of galaxies. *Astronomy and Astrophysics*, 246:634–643, 1991.
- [67] Peebles, P. The large-scale structure of the universe (research supported by the national science foundation. princeton, nj, princeton university press, 435 p., 1980). In *ADS*, volume 25, page 82, 1980.