Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **425,** 116–134 (2012) doi:10.1111/j.1365-2966.2012.21402.x

The WiggleZ Dark Energy Survey: the transition to large-scale cosmic homogeneity

Morag I. Scrimgeour,[1,2⋆] Tamara Davis,[3] Chris Blake,[4] J. Berian James,[5,6] Gregory B. Poole,[4] Lister Staveley-Smith,[1,2] Sarah Brough,[7] Matthew Colless,[7] Carlos Contreras,[4] Warrick Couch,[4] Scott Croom,[8] Darren Croton,[4] Michael J. Drinkwater,[3] Karl Forster,[9] David Gilbank,[10] Mike Gladders,[11] Karl Glazebrook,[4] Ben Jelliffe,[8] Russell J. Jurek,[12] I-hui Li,[4] Barry Madore,[13] D. Christopher Martin,[9] Kevin Pimbblet,[14] Michael Pracy,[8] Rob Sharp,[7,15] Emily Wisnioski,[4] David Woods,[16] Ted K. Wyder[9] and H. K. C. Yee[17]

[1] *International Centre for Radio Astronomy Research, M468, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[2] *ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*
[3] *School of Mathematics and Physics, University of Queensland, QLD 4072, Australia*
[4] *Centre for Astrophysics & Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*
[5] *Dark Cosmology Centre, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark*
[6] *Astronomy Department, 601 Campbell Hall, Berkeley, CA 94720, USA*
[7] *Australian Astronomical Observatory, PO Box 296, Epping, NSW 1710, Australia*
[8] *Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia*
[9] *California Institute of Technology, MC 278-17, 1200 East California Boulevard, Pasadena, CA 91125, USA*
[10] *Astrophysics and Gravitation Group, Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[11] *Department of Astronomy and Astrophysics, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637, USA*
[12] *Australia Telescope National Facility, CSIRO, Epping, NSW 1710, Australia*
[13] *Observatories of the Carnegie Institute of Washington, 813 Santa Barbara Street, Pasadena, CA 91101, USA*
[14] *School of Physics, Monash University, Clayton, VIC 3800, Australia*
[15] *Research School of Astronomy & Astrophysics, Australian National University, Weston Creek, ACT 2600, Australia*
[16] *Department of Physics & Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada*
[17] *Department of Astronomy and Astrophysics, University of Toronto, 50 S. George Street, Toronto, ON M5S 3H4, Canada*

Accepted 2012 May 28. Received 2012 May 25; in original form 2011 December 13

ABSTRACT

We have made the largest volume measurement to date of the transition to large-scale homogeneity in the distribution of galaxies. We use the WiggleZ survey, a spectroscopic survey of over 200 000 blue galaxies in a cosmic volume of $\sim 1\,h^{-3}\,\mathrm{Gpc}^3$. A new method of defining the 'homogeneity scale' is presented, which is more robust than methods previously used in the literature, and which can be easily compared between different surveys. Due to the large cosmic depth of WiggleZ (up to $z = 1$), we are able to make the first measurement of the transition to homogeneity over a range of cosmic epochs. The mean number of galaxies $N(< r)$ in spheres of comoving radius r is proportional to r^3 within 1 per cent, or equivalently the fractal dimension of the sample is within 1 per cent of $D_2 = 3$, at radii larger than $71 \pm 8\,h^{-1}$Mpc at $z \sim 0.2$, $70 \pm 5\,h^{-1}$ Mpc at $z \sim 0.4$, $81 \pm 5\,h^{-1}$ Mpc at $z \sim 0.6$ and $75 \pm 4\,h^{-1}$ Mpc at $z \sim 0.8$. We demonstrate the robustness of our results against selection function effects, using a Λ cold dark matter (ΛCDM) N-body simulation and a suite of inhomogeneous fractal distributions. The results are in excellent agreement with both the ΛCDM N-body simulation and an analytical ΛCDM prediction. We can exclude a fractal distribution with fractal dimension below $D_2 = 2.97$ on scales from $\sim 80\,h^{-1}$ Mpc up to the largest scales probed by our measurement, $\sim 300\,h^{-1}$ Mpc, at 99.99 per cent confidence.

Key words: surveys – galaxies: statistics – cosmology: observations – large-scale structure of Universe.

[⋆] E-mail: morag.scrimgeour@icrar.org

1 INTRODUCTION

One of the main assumptions of the standard theory of cosmology, Λ cold dark matter (ΛCDM; based on CDM and a cosmological constant), is that the Universe is homogeneous and isotropic on large scales, and hence can be described by the Friedmann–Robertson–Walker (FRW) metric. 'Homogeneous' means that its statistical properties (such as density) are translationally invariant; 'isotropic' means it should be rotationally invariant. The Universe clearly deviates from this on small scales, where galaxies are clustered, but on large enough scales ($\gtrsim 100\,h^{-1}$ Mpc in ΛCDM), the distribution of matter is assumed to be 'statistically homogeneous' – i.e. the small-scale inhomogeneities can be considered as perturbations, which have a statistical distribution that is independent of position. However, this is merely an assumption, and it is important for it to be accurately verified by observation. Over the last decade, there has been a debate in the literature as to whether the Universe really is homogeneous, or whether it has a fractal-like structure extending to large scales. It is important to resolve this contention if we are to be justified in assuming the FRW metric.

In fact, although ΛCDM is based on the assumption of large-scale homogeneity and an FRW metric, inflation (which ΛCDM incorporates) actually predicts a certain level of density fluctuations on all scales. Inflation predicts that the primordial density power spectrum was close to scale invariant. In the standard model, the scalar index n_s, which quantifies the scale dependence of the primordial power spectrum, is close to 0.96 (Baumann 2009), while a scale-invariant power spectrum has $n_s = 1$. In this case, these density fluctuations induce fluctuations in the metric, $\delta\Phi$, which are virtually independent of scale, and are of the order of $\delta\Phi/c^2 \sim 10^{-5}$ (Peacock 1999). Since these perturbations are small, the FRW metric is still valid, but it means that we expect the Universe to have a gradual approach to large-scale homogeneity rather than a sudden transition.

The most important implication of inhomogeneity is the so-called 'averaging problem' in general relativity (GR). This arises when we measure 'average' quantities (such as the correlation function and power spectrum, and parameters such as the Hubble constant) over a spatial volume. In doing so, we assume the volume is homogeneous and smooth, when it may not be. Since the Einstein equations are non-linear, density fluctuations can affect the evolution of the average properties of the volume – this is known as the 'backreaction mechanism' (e.g. Buchert 2000; Ellis & Buchert 2005; Li & Schwarz 2007; see Räsänen 2011, for a summary). If we observe a quantity within such a volume, we need to take averaging into account to compare it with theory. It is therefore important to know how much inhomogeneity is present, in order to obtain meaningful results from averaged measurements (and hence most, if not all, cosmological measurements).

Backreaction has also been proposed as an explanation of dark energy, which is believed to be a negative-pressure component of the Universe that drives the accelerated expansion. Some authors have suggested that instead of introducing exotic new forms of dark energy, or modifications to GR, we should revisit the fundamental assumptions of the ΛCDM model, such as homogeneity. If we assume that GR holds, but take inhomogeneities into account, it can be shown that backreaction can cause a global cosmic acceleration, without any additional dark energy component (see e.g. Schwarz 2002; Kolb et al. 2005; Räsänen 2006, 2011; Wiltshire 2007a; Buchert 2008). This effect appears to be too small to explain the observed acceleration, but highlights the importance of understanding the amount of inhomogeneity in the Universe.

Another, related, consequence of inhomogeneity is that it can affect the path travelled by light rays, and the calibration of clocks and rods of observers. It can therefore affect distance measurements, such as redshifts and luminosity distances (Wiltshire 2009; Meures & Bruni 2011). This has also been proposed as a possible explanation of the observed cosmic acceleration, although the effect appears to be of the order of only a few per cent at $z \sim 1$ (Brouzakis, Tetradis & Tzavara 2007).

Homogeneity is required by several important statistical probes of cosmology, such as the galaxy power spectrum and n-point correlation functions, in order for them to be meaningful. Applying these to a galaxy sample below the scale of homogeneity would be problematic, since, if a distribution has no transition to homogeneity, it does not have a defined mean density, which is required to calculate and interpret these statistics. It is also not possible to model its cosmic variance, so the error in the measurements would be ill defined, making it impossible to relate these statistics to a theoretical model. It is therefore important to quantify the scale on which the Universe becomes close enough to homogeneous to justify their use.

Large-scale homogeneity is already well supported by a number of different observations. In particular, the high degree of isotropy of the cosmic microwave background (CMB; Fixsen et al. 1996) gives very strong support for large-scale homogeneity in the early Universe, at redshift $z \sim 1100$. The isotropy of the CMB also indicates the Universe has remained homogeneous, since there are no significant integrated Sachs–Wolfe (ISW) effects distorting our view of the isotropic CMB (Wu, Lahav & Rees 1999). Other high-redshift evidence for homogeneity includes the isotropy of the X-ray background (Peebles 1993; Scharf et al. 2000), believed to be emitted by high-redshift sources, and the isotropy of radio sources at $z \sim 1$ (Blake & Wall 2002).

However, these measurements of high-redshift isotropy do not necessarily imply homogeneity of the present Universe. If every point in the Universe is isotropic, then this implies the Universe is homogeneous; so if we accept the Copernican principle that our location is non-special, then the observed isotropy should imply homogeneity (Peacock 1999).[1] However, most of these measurements (except the ISW effect) only tell us about the high-redshift Universe. We know that it has evolved to a clustered distribution since then, and it is possible that it could also have become anisotropic. It is also possible for the matter distribution to be homogeneous, while the galaxy distribution is not, since the galaxy distribution is biased relative to the matter field (Kaiser 1984) – although since galaxy bias is known to be linear on large scales (Coles 1993; Scherrer & Weinberg 1998), this seems unlikely. The ISW effect (Sachs & Wolfe 1967) gives information about the low-redshift Universe, since it mostly probes the dark energy dominated era, $z \lesssim 1$ (Afshordi 2004), but it is an integral over the line of sight and so does not give full 3D information.

Galaxy surveys are the only 3D probe of homogeneity in the nearby Universe, and a number of homogeneity analyses have been carried out with different surveys, with seemingly conflicting results. Most statistical methods used to measure homogeneity have been based on the simple 'counts-in-spheres' measurement, i.e. the number of galaxies $N(< r)$ in spheres of radius r centred on galaxies, averaged over a large number of such spheres. This quantity scales in proportion to the volume (r^3) for a homogeneous distribution, and homogeneity is said to be reached at the scale above which this holds. Hogg et al. (2005) applied this to the Sloan Digital Sky Survey (SDSS) luminous red galaxy (LRG) sample at $z \sim 0.3$, and found that the data became consistent with homogeneity at $\sim 70\,h^{-1}$ Mpc for this galaxy population (using a different method of determining the homogeneity scale than we do).

---

[1] We note that the Copernican principle is not incompatible with an inhomogeneous Universe. It assumes only that our location is non-special, not that every location is the same (Joyce et al. 2000; Clifton, Ferreira & Land 2008; Sylos Labini, Vasilyev & Baryshev 2009).

This measurement can also be extended to a fractal analysis. Fractal dimensions can be used to quantify clustering; they quantify the scaling of different moments of galaxy counts-in-spheres, which in turn are related to the n-point correlation functions. The most commonly used is the correlation dimension $D_2(r)$, which quantifies the scaling of the two-point correlation function, and is based on the counts-in-spheres, which scale as $N(<r) \sim r^{D_2}$. One can also consider the more general dimensions D_q, where q are different moments of the counts-in-spheres. Using fractal analyses, some researchers have found a transition from $D_2 < 3$ to $D_2 = 3$, at around 70–150 h^{-1} Mpc (Martínez & Coles 1994; Guzzo 1997; Martínez et al. 1998; Scaramella et al. 1998; Amendola & Palladino 1999; Pan & Coles 2000; Kurokawa, Morikawa & Mouri 2001; Yadav et al. 2005; Sarkar et al. 2009), whereas other authors have found no such transition (Coleman & Pietronero 1992; Pietronero, Montuori & Sylos Labini 1997; Sylos Labini, Montuori & Pietronero 1998; Joyce, Montuori & Sylos Labini 1999; Sylos Labini et al. 2009; Sylos Labini 2011). However, many of the galaxy redshift surveys used in the above-mentioned works are too shallow, sparse, or have survey geometries too complicated, to give conclusive results.

In this work, we use the WiggleZ Dark Energy Survey (Drinkwater et al. 2010) to make a new measurement of the counts-in-spheres and correlation dimension, to test for the transition to homogeneity. WiggleZ provides a larger volume than previous surveys, making it ideal for a homogeneity measurement, and it covers a higher redshift range, allowing us to also investigate how homogeneity changes with cosmic epoch. It is not volume limited, but we show that this does not significantly affect our measurement. The transition to homogeneity can be used as a test of a particular cosmological model, since we would expect it to differ for different cosmologies. In this work, we test a ΛCDM model with best-fitting parameters from the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data (Komatsu et al. 2011), which we refer to as ΛCDM+*WMAP*. We demonstrate the robustness of our measurement against systematic effects of the survey geometry and selection function, by repeating our analysis on both the GiggleZ N-body simulation and on a suite of inhomogeneous, fractal distributions.

Before we make any meaningful test of homogeneity, however, it is crucial to properly define what we mean by the so-called 'scale' of homogeneity. Since there is only a gradual approach to homogeneity, such a definition may be arbitrary. In the past, authors have defined the 'scale of homogeneity' as the scale where the data become consistent with homogeneity within 1σ (e.g. Hogg et al. 2005; Bagla, Yadav & Seshadri 2008; Yadav, Bagla & Khandai 2010). However, this method has several disadvantages. It depends on the size of the error bars on the data, and hence on the survey size. A larger survey should have smaller error bars, and so will automatically measure a larger scale of homogeneity. It also depends on the bin spacing, and is susceptible to noise between data points. We therefore introduce a different, and more robust, method for determining homogeneity: we fit a smooth, model-independent polynomial curve to all the data points, and find where this intercepts chosen values close to homogeneity.

Certain parts of our analysis require the assumption of a cosmological model and, implicitly, homogeneity (i.e. for converting WiggleZ redshifts to distances, correcting for the selection function, calculating the uncertainties using lognormal realizations and finding the best-fitting bias). In these cases, we use an input ΛCDM cosmology with $h = 0.71$, $\Omega_{\rm m} = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_{\rm b} = 0.044\,82$, $\sigma_8 = 0.8$ and $n_{\rm s} = 0.96$. Here, the Hubble constant is $H_0 = 100\,h\,{\rm km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{\rm m}$ is the mass density, Ω_Λ is the dark energy density, $\Omega_{\rm b}$ is the baryon density, σ_8 is the rms mass variation within spheres of $8\,h^{-1}$ Mpc radius and $n_{\rm s}$ is the spectral index of the primordial power spectrum. This is the same fiducial cosmology used by Blake et al. (2011a), and we use this for consistency. We discuss the implications of assuming a ΛCDM model on the results of our homogeneity measurement in Section 7.

This paper is organized as follows. In Section 2, we describe the WiggleZ survey and our data set. In Section 3, we describe our methodology. We also explain our definition of the 'scale of homogeneity' and present a new model-independent method for measuring this from data. In Section 4, we describe our analytic ΛCDM+*WMAP* model. We present our results in Section 5. We test the robustness of our method using fractal distributions and a ΛCDM N-body simulation in Section 6. We discuss our results in Section 7 and conclude in Section 8.

2 THE WiggleZ SURVEY

The WiggleZ Dark Energy Survey (Drinkwater et al. 2010) is a large-scale spectroscopic galaxy redshift survey conducted at the 3.9-m Anglo-Australian Telescope, and was completed in 2011 January. It maps a cosmic volume of $\sim 1\,{\rm Gpc}^3$ up to redshift $z = 1$, and has obtained 239 000 redshifts for ultraviolet-selected emission-line galaxies with a median redshift of $z_{\rm med} = 0.6$. Of these, 179 599 are in regions contiguous enough to be used for our analysis. It covers $\sim 1000\,{\rm deg}^2$ of equatorial sky in seven regions, shown in Drinkwater et al. (2010, their fig.1).

The observing strategy and galaxy selection criteria of the WiggleZ survey are described in Blake et al. (2009) and Drinkwater et al. (2010). The selection function we use is described in Blake et al. (2010). The targets were selected from *Galaxy Evolution Explorer* (*GALEX*) satellite observations matched with ground-based optical measurements, and magnitude and colour cuts were applied to preferentially select blue, extremely luminous high-redshift star-forming galaxies with bright emission lines.

The WiggleZ survey offers several advantages for a new study of the scale of homogeneity. Its very large volume allows homogeneity to be probed on scales that have not previously been possible, and at a higher redshift; it probes a volume at $z > 0.5$ comparable to the SDSS LRG catalogue at $z < 0.5$. This allows us to make the first measurement of the change in the homogeneity scale over a range of cosmic epochs. We divide our sample into four redshift slices, $0.1 < z < 0.3$, $0.3 < z < 0.5$, $0.5 < z < 0.7$ and $0.7 < z < 0.9$. The sizes of the WiggleZ regions in each redshift slice are listed in Table 1; they sample scales well above the expected scale of homogeneity. The numbers of galaxies in each redshift slice are listed in Table 2.

We also benefit from having seven regions distributed across the equatorial sky, which reduce the effect of cosmic variance. In addition, WiggleZ probes blue galaxies, whereas SDSS (which has obtained the largest scale measurements of homogeneity to date) used LRGs, and so it can also constrain any systematic effects introduced by the choice of tracer galaxy population (Drinkwater et al. 2010). Since blue galaxies are less biased, they are also more representative of the underlying matter distribution.

Table 1. Comoving dimensions of WiggleZ regions in each redshift slice, in units of h^{-1} Mpc. The dimensions correspond to the line of sight, RA and Dec. directions, respectively.

Region	$0.1 < z < 0.3$	$0.3 < z < 0.5$	$0.5 < z < 0.7$	$0.7 < z < 0.9$
00-hr	551 × 146 × 222	505 × 232 × 353	459 × 309 × 471	417 × 378 × 575
01-hr	549 × 127 × 132	499 × 202 × 209	450 × 269 × 279	405 × 329 × 341
03-hr	549 × 133 × 129	499 × 211 × 206	450 × 281 × 274	405 × 343 × 335
09-hr	551 × 221 × 133	504 × 351 × 212	458 × 467 × 283	416 × 571 × 346
11-hr	553 × 280 × 145	509 × 445 × 231	466 × 592 × 308	426 × 724 × 376
15-hr	553 × 295 × 150	510 × 468 × 238	468 × 623 × 317	429 × 762 × 387
22-hr	550 × 142 × 143	500 × 225 × 228	452 × 300 × 303	408 × 367 × 371

Table 2. Number of WiggleZ galaxies in each redshift slice.

Redshift	Number of galaxies
$0.1 < z < 0.3$	25 187
$0.3 < z < 0.5$	45 698
$0.5 < z < 0.7$	70 191
$0.7 < z < 0.9$	38 523

There are, however, several aspects of the survey that could potentially be detrimental to a homogeneity measurement. WiggleZ has a complex window function, with a complicated edge geometry including holes in the angular coverage, and the spectroscopic completeness varies across the sky. In addition, the population properties of the galaxies are known to vary with redshift, due to the effects of Malmquist bias (since WiggleZ is a flux-limited survey), downsizing (the observed fact that the size of the most actively star-forming galaxies decreases with time; Cowie et al. 1996; Glazebrook et al. 2004) and the colour and magnitude selection cuts (Blake et al. 2010). This means that WiggleZ preferentially selects larger mass, higher luminosity galaxies at higher redshift. A consequence of this is that it is not possible to define volume-limited subsamples of WiggleZ. However, we can correct for these effects by using random catalogues (Section 3.1), which account for the survey selection function (Blake et al. 2010). We also divide the survey into four redshift slices, reducing the amount of galaxy population evolution in any region. In addition, we show that our results are not biased by the assumption of homogeneity in the selection function corrections, using an N-body simulation and a suite of inhomogeneous fractal distributions, described in Section 6. We are therefore confident that our result is not distorted by any features of the survey.

We convert the redshifts of the WiggleZ galaxies to comoving distances d_c, using

$$d_c(z) = \frac{c}{H_0}\int_0^z \frac{dz'}{E(z')}, \tag{1}$$

where

$$E(z) = \frac{H(z)}{H_0} = [\Omega_{m,0}(1+z)^3 + \Omega_{\Lambda,0}]^{1/2}, \tag{2}$$

and we use the fiducial ΛCDM parameter values listed in Section 1. To do this, we assume the FRW metric and ΛCDM. This is necessary for any homogeneity measurement, since we must always assume a metric in order to interpret redshifts. Therefore, in the strictest sense this can only be used as a consistency test of ΛCDM. However, if we find the trend towards homogeneity matches the trend predicted by ΛCDM, then this is a strong consistency check for the model and one that an inhomogeneous distribution would find difficult to mimic. We discuss this further in Section 7.

3 METHODOLOGY

Here we describe our methodology for measuring the transition to homogeneity. We first calculate the mean counts-in-spheres $N(<r)$, then we find the fractal correlation dimension $D_2(r)$, from the slope of $N(r)$. Although they are closely related, it is interesting to consider both, since the counts-in-spheres is the simplest measurement of homogeneity, whilst the correlation dimension provides direct information about the fractal properties of the distribution. We also describe our method of determining uncertainties, and our method of defining the 'homogeneity scale' R_H.

3.1 Scaled counts-in-spheres $\mathcal{N}(<r)$

The simplest test of homogeneity of a set of points is to find the average number $N(<r)$ of neighbouring points from any given point, up to a maximum distance r; if the distribution is homogeneous, then (for large enough r)

$$N(<r) \propto r^D, \tag{3}$$

where D is the ambient dimension (the number of dimensions of the space; for a homogeneous volume, $D = 3$).

We find $N(<r)$ for spheres centred on each of the WiggleZ galaxies, and correct for incompleteness by dividing by the number expected for a homogeneous distribution with the same level of completeness. (We show that this does not bias our final results.) This is done by finding the mean $N(<r)$ about the coordinate position of the WiggleZ galaxy from 100 random catalogues, each with the same number density, window function and redshift distribution of the WiggleZ survey. The method of generating the random catalogues is described in Blake et al. (2010). We then take the average over all the galaxies, to obtain the mean, scaled counts-in-spheres measurement $\mathcal{N}(<r)$:

$$\mathcal{N}(<r) = \frac{1}{G}\sum_{i=1}^{G} \frac{N^i(<r)}{\frac{1}{R}\sum_{j=1}^{R} \rho_j N_R^{i,j}(<r)}, \tag{4}$$

where G is the number of WiggleZ galaxies used as sphere centres, R is the number of random catalogues, $N(<r)$ is the counts for WiggleZ galaxies, $N_R(<r)$ is the counts for random galaxies (centred on the position of the ith WiggleZ galaxy) and $\rho_j \equiv n_W/n_{\mathrm{rand},j}$ is the ratio of the total number of WiggleZ galaxies (n_W) to the number of random galaxies in the jth random catalogue ($n_{\mathrm{rand},j}$). In our analysis, we have $G = n_W$, but this would not be the case if, for example, we excluded spheres near the survey edges.

The random catalogue correction has the effect of reducing the scaling by the number of dimensions (i.e. $D = 3$), so that for a homogeneous distribution $\mathcal{N}(<r)$ scales as

$$\mathcal{N}(<r) \propto r^{3-3} = 1. \tag{5}$$

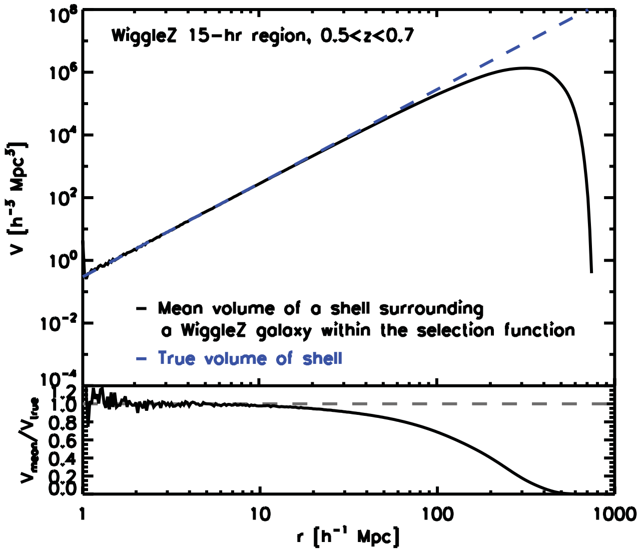

Figure 1. An illustration of the scales for which survey edge effects become important for the homogeneity measurement. Top panel: the mean volume $\overline{V}(r)$ in a thin shell of mean radius r surrounding a WiggleZ galaxy within the selection function (black curve). We show this for the 15-hr $0.5 < z < 0.7$ region. The true volume of the shells is shown as a blue dashed line. The black curve deviates from the blue at large scales, where an increasing proportion of the shells extends outside the survey. Bottom panel: the ratio of the mean volume to the true volume, as a function of r. The grey dashed line indicates a ratio of 1.

In each WiggleZ region, we make $\mathcal{N}(<r)$ measurements in spheres with 12–15 logarithmically spaced radial bins (depending on the size of the region). We determine the large-scale cut-off in each region by calculating the mean volume in the selection function, $\overline{V}(r)$, in a thin shell of mean radius r enclosing a central galaxy. We illustrate this for the 15-hr $0.5 < z < 0.7$ region in Fig. 1, and compare it with the 'true' volume of the shells. Below $\sim 2\,h^{-1}$ Mpc, there is noise due to low resolution, but above this the two curves are very close. On large scales, however, an increasing proportion of shells surrounding galaxies goes off the edge of the survey, and their volume within the survey decreases. On such scales, corrections for edge effects will become important. We take the large-scale cut-off of our homogeneity measurement at the radius of the maximum-volume shell. For the region in Fig. 1, this corresponds to shells that are ∼20 per cent complete. We show later that edge effects up to this scale do not impact our homogeneity measurement.

Our correction method using random catalogues maximizes the use of the data, and accounts for incompleteness and the fact that WiggleZ is not volume limited. However, it can potentially bias our result towards detecting homogeneity, since it assumes homogeneity on the largest scales of the survey. It is equivalent to weighting each measurement by the volume of the sphere included within the survey, multiplied by an arbitrary mean density. Therefore, while the counts-in-spheres measurement should have the advantage of not assuming a mean density (Hogg et al. 2005), our correction method means that we do. Therefore, $\mathcal{N}(<r)$ should tend to 1 on the largest scales of the survey, regardless of whether homogeneity has been reached. However, for a distribution with homogeneity size smaller than the survey, $\mathcal{N}(<r)$ should reach 1, and remain at 1, for a range of scales smaller than the survey scale. We check the robustness of our method against effects of the selection function and correction method in Section 6 and show that our analysis is robust against this potential source of systematic error out to scales far greater than the homogeneity scale we measure.

There are other correction methods used in the literature. There is the so-called 'exclusion' or 'deflation' method (e.g. Coleman & Pietronero 1992; Pan & Coles 2000, 2002; Sylos Labini et al. 2009), which only considers central points that are surrounded by complete spheres within the survey. This therefore excludes as central points any galaxies within a certain distance from the survey edges. However, this does not make the best use of the data, since it excludes data and so reduces the volume of the sample.

There is also the so-called 'angular correction' model, which has been shown to be more optimal, by using all the available data without introducing a bias due to edge corrections (Pan & Coles 2002). This corrects measurements in a sphere by the solid angle subtended by regions in the sphere that are outside the survey boundary. However, as they point out, this method is difficult to apply to surveys with a complicated geometry, especially if they contain holes, as WiggleZ does. Finally, another correction that minimizes bias at the survey edges, but wastes little data, is the 'Ripley' estimator (Ripley 1977; Martínez et al. 1998), which corrects the measurement in a sphere by the *area* of the sphere contained within the survey. Both the angular correction and the Ripley estimator assume isotropy of the samples. Due to the geometry of the WiggleZ survey, we choose to use a random catalogue correction, but make robustness tests to quantify any bias it may introduce to the results.

To demonstrate the robustness of our measurement against the method of correcting for the selection function, we compare our method to an analysis using only complete spheres, with and without correcting for incompleteness, in Section 6.3, and show we obtain consistent results.

3.2 Correlation dimension $D_2(r)$

Fractal dimensions can be used to describe the clustering of a point distribution. There exists a general family of dimensions D_q, the Minkowski–Bouligand dimensions, which describe the scaling of counts-in-spheres centred on points (see, e.g., Borgani 1995; Martínez & Saar 2002, for a review). To completely characterize the clustering of our galaxy distribution, we would need to consider all the moments q of the distribution (corresponding to combinations of n-point correlation functions). However, to identify the scale of homogeneity we consider only D_2, the 'correlation dimension', which quantifies the scaling behaviour of the two-point correlation function $\xi(r)$.

If we take a galaxy and count the number of other galaxies, $N(<r)$, within a distance r, then this quantity scales as

$$N(<r) \propto r^{D_2}, \tag{6}$$

where D_2 is the fractal dimension of the distribution.

From this, the correlation dimension is defined as

$$D_2(r) \equiv \frac{\mathrm{d}\ln N(<r)}{\mathrm{d}\ln r}. \tag{7}$$

Since we must correct each WiggleZ $N(<r)$ measurement for completeness, obtaining the scaled quantity $\mathcal{N}(<r) \propto r^{D_2-3}$, we must calculate $D_2(r)$ via

$$D_2(r) = \frac{\mathrm{d}\ln \mathcal{N}(<r)}{\mathrm{d}\ln r} + 3. \tag{8}$$

For a homogeneous distribution, $D_2 = 3$. If $D_2 < 3$ then the distribution has a scale invariant, fractal (and so inhomogeneous) clustering pattern. If $D_2 > 3$ the distribution is said to be 'superhomogeneous' and corresponds to a lattice-like distribution (Gabrielli, Joyce & Sylos Labini 2002). A power law in $1 + \xi(r) \sim r^{-\gamma}$ has $D_2 = 3 - \gamma$ for $\xi \gg 1$.

Some previous works have found $D_2(r)$ by simply fitting a straight line to a log–log plot of $\mathcal{N}(<r)$. This method can give a false indication of a fractal (Martínez et al. 1998); calculating $D_2(r)$ explicitly gives a more reliable measurement.

We note that our estimator for $\mathcal{N}(<r)$ (equation 4) is essentially equivalent to $1 + \bar{\xi}_g(r)$, where (Hamilton 1992)

$$\bar{\xi}(r) = \frac{3}{r^3}\int_0^r x^2 \xi(x)\,\mathrm{d}x. \tag{9}$$

[This can be seen more clearly by rearrangement of the theoretical expression for $\mathcal{N}(<r)$ given by equation 19.] Many measurements of $\xi_g(r)$ have been made with different galaxy surveys (e.g. Hawkins et al. 2003; Zehavi et al. 2005; Beutler et al. 2011; Blake et al. 2011b). On small scales, the correlation function is well described by a power law:

$$\xi(r) = \left(\frac{r_0}{r}\right)^{\gamma}, \tag{10}$$

where $r_0 \approx 5\,h^{-1}$ Mpc is the so-called clustering length, and $\gamma \approx 1.8$, depending on the galaxy population. On scales $\gtrsim 20\,h^{-1}$ Mpc, however, the correlation function is observed to turn over, consistent with large-scale homogeneity.

However, the correlation function cannot be used to test large-scale homogeneity, since the way it is determined from surveys depends on the mean galaxy density. Determinations of $\xi(r)$ commonly use the Landy–Szalay (Landy & Szalay 1993) or Hamilton estimators (Hamilton 1993), which compare the galaxy clustering to that of random catalogues of the same mean density as the survey.

Our $\mathcal{N}(<r)$ estimator also compares the data to random catalogues, since we must correct for the selection function. Therefore, $\mathcal{N}(<r)$, like $\xi(r)$, does assume a mean density on the scale of the survey. However, our estimator is slightly different, as we correct each object separately rather than an average pair count.

The correlation dimension $D_2(r)$, on the other hand, measures the *scaling* of $\mathcal{N}(<r)$, which is *not* affected by the assumption of the mean density. [This only affects the amplitude of $\mathcal{N}(<r)$.] Deviations from a volume-limited sample, which require random catalogue corrections, only cause second-order changes to $D_2(r)$, whilst they would be leading order in the raw correlation function. Therefore, $D_2(r)$ is much more robust to both the assumed mean density and details of the selection function, making it the most reliable measure of homogeneity.

3.3 Lognormal realizations and covariance matrix

Lognormal realizations (Coles & Jones 1991) are an important tool for determining uncertainties in galaxy surveys. A lognormal random field is a type of non-Gaussian random field, which can be used to model the statistical properties of the galaxy distribution and simulate data sets with an input power spectrum. We have used 100 such realizations, generated using an input ΛCDM power spectrum with the fiducial parameters listed in Section 1. These are sampled with the survey selection function, to create 100 mock catalogues for each of the WiggleZ regions. We use these to calculate the full covariance and errors of our measurement. Jackknife resampling does not permit enough independent regions within the survey volume to give a reliable estimate of the uncertainties.

We obtain $\mathcal{N}(<r)$ and $D_2(r)$ for each of the lognormal realizations in the same way as for the WiggleZ data. The covariance matrix between radial bins i and j is given by

$$C_{ij} = \frac{1}{n-1}\sum_{l=1}^{n}[x_l(r_i) - \overline{x(r_i)}][x_l(r_j) - \overline{x(r_j)}], \tag{11}$$

where $x(r)$ is $\mathcal{N}(<r)$ or $D_2(r)$, the sum is over lognormal realizations l, n is the total number of lognormal realizations and $\overline{x(r)} = \frac{1}{n}\sum_{l=1}^{n} x_l(r)$. The diagonal values $j = k$ give the variance, σ^2.

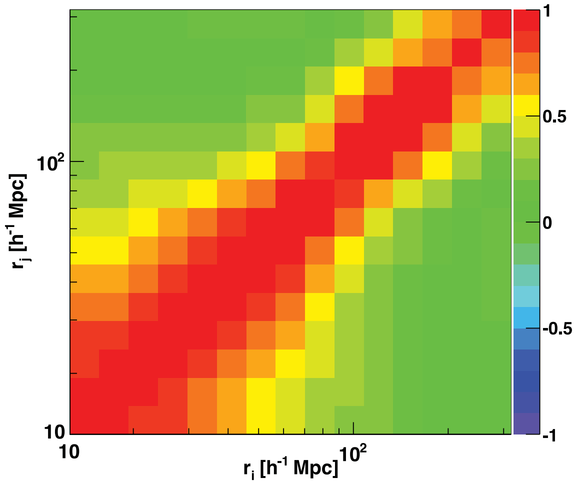

Figure 2. The correlation matrix $C_{ij}/\sqrt{C_{ii}C_{jj}}$ for the $\mathcal{N}(r)$ measurement, obtained from lognormal realizations, for the $0.5 < z < 0.7$ redshift slice.

The correlation coefficient between bins i and j is given by

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}}. \tag{12}$$

The correlation matrices for $\mathcal{N}(<r)$ and $D_2(r)$ are shown in Figs 2 and 3, respectively, for the combined regions (see the next section) in the $0.5 < z < 0.7$ redshift slice.

It is notable that the $\mathcal{N}(<r)$ measurement is more correlated than $D_2(r)$. This is because $\mathcal{N}(<r)$ is effectively an integral of $D_2(r)$, so its covariance is effectively a cumulative sum of that of $D_2(r)$. It can also be explained by the fact that $D_2(r)$ is the logarithmic slope of $\mathcal{N}(<r)$, so it does not depend on the correlations between

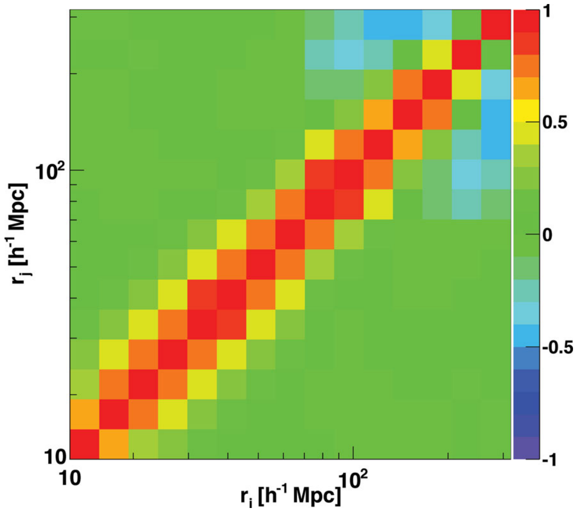

Figure 3. The correlation matrix $C_{ij}/\sqrt{C_{ii}C_{jj}}$ for the $D_2(r)$ measurement, obtained from lognormal realizations, for the $0.5 < z < 0.7$ redshift slice.

widely separated $\mathcal{N}(<r)$ bins but rather the variations between neighbouring bins.

We note that the uncertainties calculated using lognormal realizations assume ΛCDM, and represent the variance we would expect to measure in a ΛCDM universe. It would not be possible to calculate uncertainties for a fractal universe, since a fractal has no defined cosmic variance as it has no defined mean density. However, calculating uncertainties this way gives a valid consistency check of ΛCDM.

3.4 Combining WiggleZ regions

In each of the four redshift slices, we combine the measurements from each of the seven WiggleZ regions using inverse-variance weighting. The combined measurements $x_{\rm comb}$ are given by

$$x_{\rm comb}(i) = \frac{\sum_n w_n(i) x_n(i)}{\sum_n w_n(i)}, \qquad (13)$$

where n are the seven WiggleZ regions, $x_n(i)$ is the measurement in the ith radial bin in the nth region and $w_n(r) = 1/C_n(i, i)$ is a weight function. $C_n(i, j)$ is the covariance matrix of the nth region, and the combined covariance matrix is calculated by

$$C_{\rm comb}(i, j) = \frac{\sum_n C_n(i, j) w_n(i) w_n(j)}{\sum_n w_n(i) \sum_n w_n(j)}. \qquad (14)$$

We show $C_{\rm comb}(i, j)$ for the $\mathcal{N}(<r)$ and $D_2(r)$ measurements in the $0.5 < z < 0.7$ redshift slice, in terms of the correlation matrix $C_{ij}/\sqrt{C_{ii}C_{jj}}$, in Figs 2 and 3.

3.5 Likelihood (χ^2) and parameter fitting

The χ^2 of our model fit to the data is given by

$$\chi^2 \equiv \sum_{i=1}^{n_{\rm bins}} \sum_{j=1}^{n_{\rm bins}} [x_{\rm th}(r_i) - x_{\rm obs}(r_i)] C_{ij}^{-1} [x_{\rm th}(r_j) - x_{\rm obs}(r_j)], \qquad (15)$$

where $x(r)$ is $\mathcal{N}(<r)$ or $D_2(r)$, $n_{\rm bins}$ is the number of radial bins, $x_{\rm th}$ is the theoretical model and $x_{\rm obs}$ is the measured value. C_{ij}^{-1} is the inverse of the covariance matrix.

3.6 A new method of defining the 'homogeneity scale' $R_{\rm H}$

As previously mentioned, the definition of a 'homogeneity scale' is somewhat arbitrary since we expect the Universe to have a gradual approach to homogeneity. Recently, Bagla et al. (2008) and Yadav et al. (2010) proposed that the homogeneity scale be defined where the measured fractal dimension D_q becomes consistent with the ambient dimension within the 1σ statistical uncertainty, $\sigma_{\Delta D_q}$. They derived an approximation for $D_q(r)$ and $\sigma_{\Delta D_q}$, in the limit of weak clustering, for a given correlation function, and showed that both scale the same way, and so the homogeneity scale stays constant, with bias and epoch. This definition is therefore beneficial as it is not arbitrary, and is robust to the tracer galaxy population. However, in deriving $\sigma_{\Delta D_q}$ they considered only shot noise and cosmic variance from variance in the correlation function, while ignoring contributions from the survey geometry and selection function. Any real survey will have these contributions to the statistical uncertainty, which cannot be separated from the variance due to the correlation function alone. The value derived from this definition in a real survey is therefore difficult to interpret in a meaningful way, i.e. one that allows comparison with theory or with different surveys of differing volume and selection function.

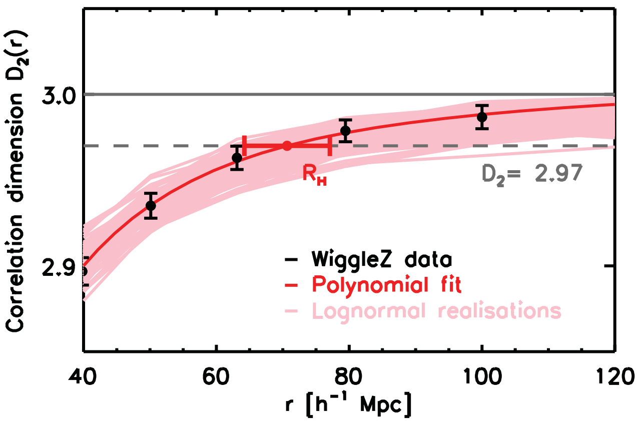

Figure 4. Illustration of our method of defining the homogeneity scale, $R_{\rm H}$, shown here for the $D_2(r)$ measurement. We first fit a model-independent polynomial (red curve) to the data (black data points). We then find where this intercepts a chosen value close to homogeneity, e.g. 1 per cent from homogeneity, $D_2 = 2.97$ (dotted grey line). This gives us $R_{\rm H}$. We find the uncertainty in $R_{\rm H}$ from the rms variance of 100 lognormal realizations (pink curves).

We therefore introduce a different method for determining a 'homogeneity scale' $R_{\rm H}$, which is easier to compare with theory and between surveys. Our method is to fit a smooth, model-independent polynomial to the data, and find the scale at which this intercepts a chosen value, or 'threshold', close to homogeneity. This scale is then defined as the homogeneity scale $R_{\rm H}$. For example, $R_{\rm H}$ could be the value of r at which the polynomial intercepts a line 1 per cent from $\mathcal{N}(<r) = 1$ or $D_2(r) = 3$ (see Fig. 4 for an illustration). The uncertainty is found using the 100 lognormal realizations. The homogeneity scale measured this way does not depend directly on the survey errors (although the *uncertainties* on $R_{\rm H}$ do), and is less susceptible to noise in the data, making this a preferable method that allows comparisons between different surveys. It also allows easy comparison between the data and a given model, e.g. ΛCDM, and we can check that the data converge to $\mathcal{N} = 1$ or $D_2 = 3$ as expected for a homogeneous distribution, by choosing a range of thresholds approaching homogeneity (see Fig. 9).

We can also take this further and construct a likelihood distribution for the homogeneity scale, as described in Section 5.3.

Although our choice of D_2 threshold is arbitrary, by choosing the same threshold for different surveys we obtain an $R_{\rm H}$ value that can be meaningfully compared, and that can be easily compared to a theoretical model. Our choice of threshold may be limited by the amount of noise in the data, however. For instance, we can measure an intercept 1 per cent away from homogeneity for the WiggleZ data, but cannot measure 0.1 per cent in two of the redshift slices, due to noise (the data do not come this close to homogeneity, although they are consistent with it within the uncertainties; see Fig. 9). None the less, we can easily choose a threshold that is possible given our data, and use this to compare with a model. The baryon acoustic oscillation (BAO) feature can also potentially affect the appropriate choice of intercept value as it causes a small distortion in $D_2(r)$ – we discuss this in Section 7.

For the main results in this paper, we have chosen a threshold of 1 per cent away from homogeneity, since this is about the closest threshold to homogeneity we can measure, considering the noisiness of the data.

4 ΛCDM MODEL PREDICTION OF $\mathcal{N}(<R)$ AND $D_2(R)$

In this section, we derive theoretical ΛCDM predictions for the counts-in-spheres and correlation dimension. This allows us to compare our measurements of the transition to homogeneity to the predictions of a ΛCDM model that fits *WMAP* data.

4.1 $\mathcal{N}(<r)$ and $D_2(r)$

For a particular galaxy population, we can calculate the mean counts-in-spheres $N(<r)$ from the two-point matter correlation function predicted by ΛCDM. The two-point correlation function $\xi(r)$ is defined as the excess probability above random of finding two objects in volumes $\mathrm{d}V_1$ and $\mathrm{d}V_2$, separated by distance r (Peebles 1980):

$$P(r) = \bar{\rho}^2[1 + \xi(r)]\,\mathrm{d}V_1\,\mathrm{d}V_2, \tag{16}$$

where $\bar{\rho}$ is the mean number density.

The mean number of galaxies surrounding a random galaxy up to a distance r is found by integrating the correlation function:

$$N(<r)_{\mathrm{g}} = \bar{\rho}\int_0^r [1 + b^2\xi(r')]4\pi r'^2\,\mathrm{d}r', \tag{17}$$

where b is the galaxy bias, relating the clustering of a particular galaxy population to the underlying dark matter distribution. Note that ΛCDM assumes large-scale homogeneity, and indeed we must assume large-scale homogeneity in order for a mean density $\bar{\rho}$ to be defined.

We obtain our model correlation function by transforming a ΛCDM matter power spectrum $P_{\delta\delta}(k)$ generated using CAMB (Lewis, Challinor & Lasenby 2000). Since we make our measurements in redshift space, we first convert $P_{\delta\delta}(k)$ to the redshift-space galaxy power spectrum $P_{\mathrm{g}}^{\mathrm{s}}(k)$ (described in the next section), then convert this to the redshift-space galaxy correlation function $\xi_{\mathrm{g}}(s)$, where s denotes distance in redshift space. Since we use the angle-averaged power spectrum (assuming the power spectrum is isotropic), we do not need to integrate the angular part of the k-space integral, and so use a spherical Hankel transform rather than a Fourier transform to obtain $\xi_{\mathrm{g}}(s)$:

$$\xi_{\mathrm{g}}(s) = \frac{1}{2\pi^2}\int P_{\mathrm{g}}^{\mathrm{s}}(k)\frac{\sin ks}{ks}k^2\,\mathrm{d}k. \tag{18}$$

To compare with our WiggleZ measurement (equation 4), where we correct for incompleteness, we divide our counts-in-spheres prediction by the number that would be expected for a random distribution, i.e. $\bar{\rho}\frac{4}{3}\pi r^3$:

$$\mathcal{N}(<r) = \frac{3}{4\pi r^3}\int_0^r [1 + \xi_{\mathrm{g}}(s)]4\pi s^2\,\mathrm{d}s. \tag{19}$$

We calculate the model $D_2(r)$ by simply applying equation (7) to our model $\mathcal{N}(<r)$.

4.2 Redshift-space distortions and non-linear velocity damping

Here we describe how we implement redshift-space distortions in our analytical model. In practice, we measure the positions of galaxies in redshift space, which are affected by redshift-space distortions. These are due to the peculiar velocities of galaxies along the line of sight, which add to the measured redshifts and perturb the inferred galaxy positions. This anisotropic effect creates anisotropy in the observed redshift-space galaxy power spectrum $P_{\mathrm{g}}^{\mathrm{s}}(k, \mu)$, and can be modelled by multiplying (convolving in configuration space) the real-space matter power spectrum by an angle-dependent function $F(k, \mu)$:

$$P_{\mathrm{g}}^{\mathrm{s}}(k, \mu) = b^2F(k, \mu)P_{\delta\delta}(k). \tag{20}$$

There are two forms of redshift-space distortion of relevance to our measurement, which we find are necessary for a good fit to the data.

(i) On large, linear scales ($\gtrsim 20\,h^{-1}$ Mpc), the bulk infall of galaxies towards overdensities creates an enhancement in the observed power spectrum along the line of sight. This can be modelled by the linear Kaiser formula (Kaiser 1987):

$$\begin{aligned} P_{\mathrm{g}}^{\mathrm{s}}(k, \mu) &= b^2\left(1 + \frac{f\mu^2}{b}\right)^2 P_{\delta\delta,\mathrm{L}}(k) \\ &= b^2(1 + \beta\mu^2)^2P_{\delta\delta,\mathrm{L}}(k), \end{aligned} \tag{21}$$

where $P_{\delta\delta,\mathrm{L}}(k)$ is the linear matter power spectrum, f is the growth rate of structure and $\beta = f/b$ is the redshift-space distortion parameter.

Note that this formula assumes a perturbed FRW universe with small real-space density perturbations $|\delta(r)| \ll 1$.

(ii) On quasi-linear scales ($10 \lesssim s \lesssim 20\,h^{-1}$ Mpc), the peculiar velocities resulting from the scale-dependent growth of structure distort the shape of the power spectrum, via a scale-dependent damping effect. A common way of modelling this is the 'streaming model' (Peebles 1980; Fisher 1995; Hatton & Cole 1998), which combines the linear theory Kaiser formula with a velocity streaming term. We choose to use a Lorentzian term, $F = [1 + (k\sigma_{\mathrm{p}}\mu)^2]^{-1}$, for an exponential velocity probability distribution function (PDF), since Blake et al. (2011a) found this to be a good fit to the WiggleZ power spectrum for $k < 0.1\,h^{-1}$ Mpc. This gives us the so-called 'dispersion model' (Peacock & Dodds 1994) for the full redshift-space power spectrum:

$$P_{\mathrm{g}}^{\mathrm{s}}(k, \mu) = b^2\frac{(1 + \beta\mu^2)^2}{1 + (k\sigma_{\mathrm{p}}\mu)^2}P_{\delta\delta}(k). \tag{22}$$

Here, σ_{p} is the pairwise velocity dispersion along the line of sight. Both σ_{p} and β are parameters that must be fitted to the data. We use the values obtained by Blake et al. (2011a) in each redshift slice from fits to the WiggleZ 2D power spectrum – these are listed in Table 3.

We note that the streaming model is motivated by virialized motions of particles within haloes, on much smaller scales – the so-called 'Finger of God' effect at $\lesssim 2\,h^{-1}$ Mpc. However, it is heuristic in nature and can also describe physical scales of tens of h^{-1} Mpc. Blake et al. (2011a) apply it this way by fitting for σ_{p} on these

Table 3. Values of the redshift-space distortion parameter β and the pairwise velocity dispersion σ_{p}, used in our modelling of non-linear redshift-space distortion effects. These values were obtained by Blake et al. (2011a) from fits to the WiggleZ 2D power spectrum.

Redshift	β	σ_{p} ($h\,\mathrm{km\,s^{-1}}$)
$0.1 < z < 0.3$	0.69	346
$0.3 < z < 0.5$	0.73	275
$0.5 < z < 0.7$	0.60	275
$0.7 < z < 0.9$	0.51	86

scales, rather than on Finger-of-God scales. We find that including it gives a significant improvement of our model fit to data.

To obtain the angle-averaged redshift-space galaxy power spectrum $P^s_g(k)$, we need to convert the full $F(k, \mu)$ to $F(k)$, which we do by integrating over μ:

$$F(k) = \int_{\mu=0}^{1} \frac{(1+\beta\mu^2)^2}{[1+(k\sigma_v\mu)^2]}\,d\mu. \qquad (23)$$

The angle-averaged redshift-space galaxy power spectrum is then

$$P^s_g(k) = b^2 F(k) P_{\delta\delta}(k). \qquad (24)$$

Equation (22) normally assumes a linear matter power spectrum; however, we choose to use a non-linear $P_{\delta\delta}(k)$, calculated by CAMB using the HALOFIT code (Smith et al. 2003), since Blake et al. (2011a) found that this gave a better fit to the data.

4.3 Correcting the WiggleZ data for galaxy bias

The amplitude of the galaxy correlation function is affected by galaxy bias and redshift-space distortions, and the shape is affected by non-linear damping. Therefore, these also affect the amplitude and shape of $\mathcal{N}(<r)$, as well as the measured scale of homogeneity. It is possible to correct the data for bias, and hence to determine the homogeneity scale for the underlying matter distribution, by assuming a particular model, i.e. our ΛCDM+*WMAP* model, fitting for bias by minimizing the χ^2 value, and correcting the data for this.

In our analysis, we only consider linear galaxy bias. This relates the galaxy correlation function $\xi_g(r)$ to the matter correlation function $\xi_m(r)$ through $\xi_g(r) = b^2\xi_m(r)$. Since we are only interested in large scales, we do not consider scale-dependent bias that may occur on small scales.

We fix the redshift-space distortion parameter β and velocity dispersion σ_p to the values listed in Table 3. We then obtain our corrected measurements $\mathcal{N}_{\text{bias free}}(<r)$:

$$\mathcal{N}_{\text{bias free}}(<r) = \frac{\mathcal{N}(<r) - 1}{b^2} + 1, \qquad (25)$$

and we calculate the bias-corrected correlation dimension $D_{2,\text{bias free}}(r)$ from this, using equation (7).

Since we assume a ΛCDM model to fit for bias, we cannot make a model-independent measurement of transition to homogeneity of the underlying matter distribution. However, it is still interesting to look at the variation of the homogeneity scale of the matter distribution with redshift, assuming ΛCDM+*WMAP*.

We note that our measurement of the homogeneity scale of the WiggleZ galaxies is independent of the ΛCDM modelling shown in this section. This modelling is only done so that we can show that the measurement is consistent with that expected from ΛCDM.

5 RESULTS

5.1 $\mathcal{N}(<r)$ and $D_2(r)$

The $\mathcal{N}(<r)$ measurements in each of the four redshift slices are shown in Fig. 5. The data are compared with a ΛCDM+*WMAP* model (described in Section 4). For each successive redshift slice, the reduced χ^2 values are 0.57, 0.91, 0.69 and 1.1. The first two redshift slices have 14 data bins from 12.5 to 251 h^{-1} Mpc, while the last two have 15 data bins from 12.5 to 316 h^{-1} Mpc. The data are consistent with a monotonically decreasing function, so we can fit a polynomial and find where this intercepts a chosen threshold, as per our definition of homogeneity.

The intercepts of the polynomial fit with $\mathcal{N} = 1.01$ (1 per cent away from homogeneity), which we define as the homogeneity scale R_H, are shown as red error bars. The errors were determined from the 100 lognormal realizations, and correspond to the square roots of the diagonal elements of their covariance matrix. The R_H values and their errors are shown in Fig. 5 and listed in Table 4, along with the values for the ΛCDM model, which are in good agreement.

The $D_2(r)$ measurements in each of the redshift slices are shown in Fig. 6, along with a ΛCDM+*WMAP* model with best-fitting bias. In each redshift slice, the data range and degrees of freedom are the same as for the $\mathcal{N}(<r)$ measurement. The reduced χ^2 values in each redshift slice are 0.83, 0.90, 0.74 and 0.98. In each case, a polynomial is fitted to the data. The homogeneity scales measured where these intercept 1 per cent away from homogeneity, $D_2 = 2.97$, are listed in Table 4 and are also in excellent agreement with the ΛCDM values.

5.2 Effect of bias and $\sigma_8(z)$ on R_H

The homogeneity scale of the model galaxy distribution, measured at 1 per cent from homogeneity, will depend on the amplitude and shape of the correlation function, and so on galaxy bias b and $\sigma_8(z)$. We would also expect it to depend on redshift, since $\sigma_8(z)$ increases over time. These two parameters are in fact completely degenerate in the $\mathcal{N}(<r)$ and $D_2(r)$ measurements.

Fig. 7 shows how our ΛCDM $\mathcal{N}(<r)$ and $D_2(r)$ models vary with bias, at $z = 0.2$, for fixed $\sigma_8(z = 0) = 0.8$. Larger bias means a larger amplitude of clustering, so that both curves are steeper on small scales. This means that the models reach homogeneity at larger radii for higher bias. This can be understood qualitatively, since highly biased galaxies are more clustered together than less biased galaxies, so we must go to larger scales before we reach a homogeneous distribution.

Fig. 8 shows how the homogeneity scale R_H of the galaxy distribution varies with bias, for different intercept values approaching homogeneity, for the $\mathcal{N}(<r)$ and $D_2(r)$ models at $z = 0.2$ with fixed $\sigma_8(z = 0) = 0.8$. For a particular intercept value, larger bias again gives a larger homogeneity scale. Since $\mathcal{N}(<r)$ and $D_2(r)$ approach homogeneity asymptotically, the intercept jumps to higher values when we consider an intercept value closer to homogeneity. This image also shows the mapping between $\mathcal{N}(<r)$ and $D_2(r)$ homogeneity values for intercepts at 1, 0.1 and 0.001 per cent away from homogeneity. They are not identical since they are slightly different methods, but give similar results. This plot illustrates that there are many potential ways to define homogeneity, and so it is important to make consistent measurements between surveys in order for them to be comparable with each other and with theory.

Our bias-corrected $\mathcal{N}(< r)$ and $D_2(r)$ measurements are listed in Table 4. The errors on the bias-corrected data were determined by applying a bias correction to each of the 100 lognormal realizations individually, and recalculating the covariance matrix. Since the bias correction aims to set the bias of all the realizations to $b = 1$, it lowers the overall variance, and so the error bars are slightly smaller than for the pure data. These measurements give our measured homogeneity scale for the matter distribution, assuming ΛCDM. We find that this scale increases with redshift, as expected in ΛCDM. However, since we are assuming ΛCDM, which has $\sigma_8(z)$ increasing with time, this is not a model-independent result.

As already mentioned, the effect of bias on the homogeneity scale of galaxies is degenerate with the amplitude of the correlation function, $\sigma_8(z)$, since the correlation function at redshift z depends on a combination of these, $\xi(r, z) \propto b^2\sigma_8(z)^2$. So far we have

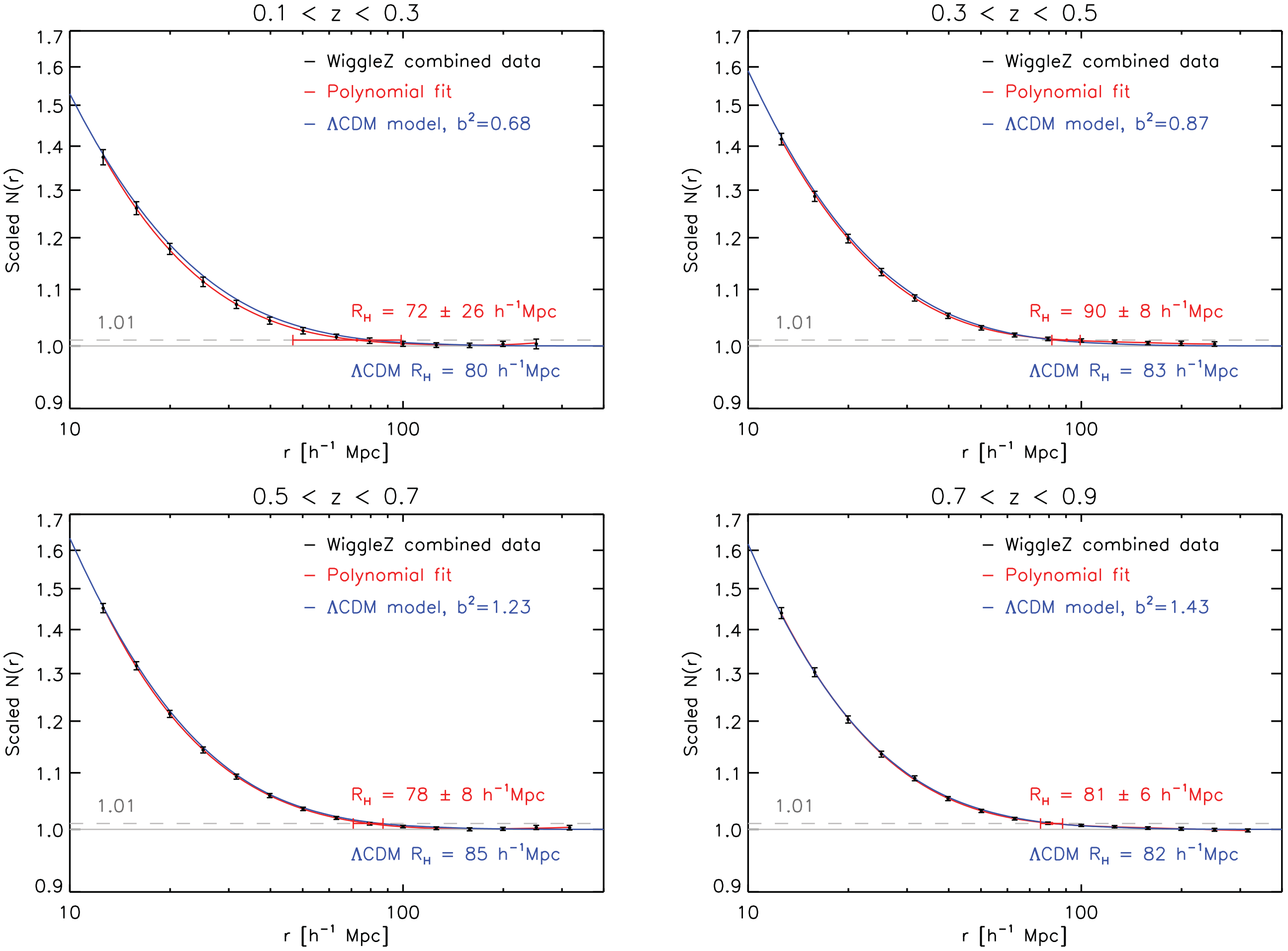

Figure 5. Scaled counts-in-spheres $\mathcal{N}(<r)$ for the combined WiggleZ data in each of the four redshift slices (black error bars). A ΛCDM model with best-fitting bias b^2 is shown in blue. A fifth-degree polynomial fit to the data is shown in red. The red error bar and label show the homogeneity scale $R_{\rm H}$ for the galaxy distribution, measured by the intercept of the polynomial fit with 1.01 (1 per cent away from homogeneity), with the error given by lognormal realizations. This scale is consistent with the ΛCDM intercept with 1.01, labelled in blue.

Table 4. Measured values of the homogeneity scale $R_{\rm H}$ (where data intercept 1 per cent of the homogeneity value). The $R_{\rm H}$ values shown are all for the galaxy distribution, except the values for the bias-corrected data, which are for the underlying matter distribution, i.e. $b^2 = 1$. The bias-corrected values directly assume a ΛCDM+*WMAP* model.

	WiggleZ data (h^{-1} Mpc)	ΛCDM (h^{-1} Mpc)	Measured bias, b^2	Bias-corrected data (h^{-1} Mpc)	ΛCDM, $b^2 = 1$ (h^{-1} Mpc)	Likelihood analysis (h^{-1} Mpc)	Bias-corrected likelihood analysis (h^{-1} Mpc)
$\mathcal{N}(<r)$							
$0.1 < z < 0.3$	72 ± 26	80	0.68	83 ± 35	93	–	–
$0.3 < z < 0.5$	90 ± 8	83	0.87	99 ± 11	87	–	–
$0.5 < z < 0.7$	78 ± 8	85	1.23	73 ± 6	79	–	–
$0.7 < z < 0.9$	81 ± 6	82	1.43	70 ± 4	72	–	–
$D_2(r)$							
$0.1 < z < 0.3$	70 ± 7	76	0.70	84 ± 11	88	71 ± 8	89 ± 14
$0.3 < z < 0.5$	70 ± 5	78	0.87	74 ± 6	82	70 ± 5	75 ± 5
$0.5 < z < 0.7$	81 ± 4	81	1.25	74 ± 3	73	81 ± 5	76 ± 4
$0.7 < z < 0.9$	74 ± 4	78	1.46	64 ± 2	66	75 ± 4	65 ± 4

assumed a fixed value of $\sigma_8(z = 0)$. However, we can also make predictions independent of σ_8, by finding how $R_{\rm H}$ changes as a function of the combination $b^2\sigma_8(z)^2$. This is shown in Fig. 9. We also show the WiggleZ results, which we have plotted for the best-fitting $b^2\sigma_8(z)^2$ value in each redshift slice. This increases with redshift, so the data points from left to right go from low to high redshift. The WiggleZ results are in very good agreement with the ΛCDM+*WMAP* predictions. We show, for comparison, the values obtained from defining the homogeneity scale as where the data come within 1σ of homogeneity. These values have much greater

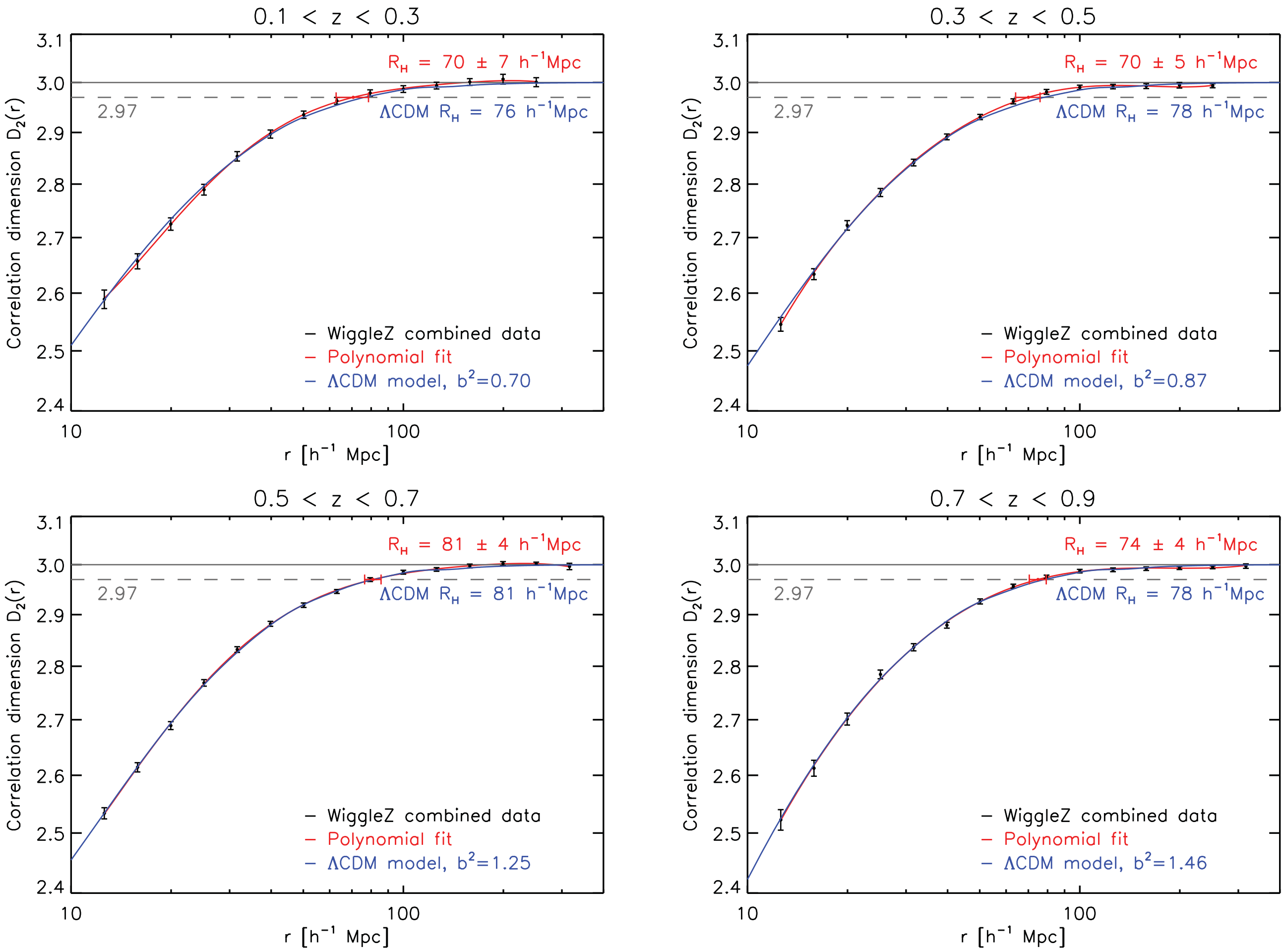

Figure 6. Same as for Fig. 5 but for the correlation dimension $D_2(r)$. The $D_2(r)$ measurements for the combined WiggleZ data in each of the four redshift slices are shown as black error bars. A ΛCDM model with best-fitting bias b^2 is shown in blue. A fifth-degree polynomial fit to the data is shown in red. The red error bar and label show the homogeneity scale R_H measured by the intercept of the polynomial fit with 2.97 (1 per cent away from homogeneity), with the error given by lognormal realizations. This scale is consistent with the ΛCDM intercept with 2.97, labelled in blue.

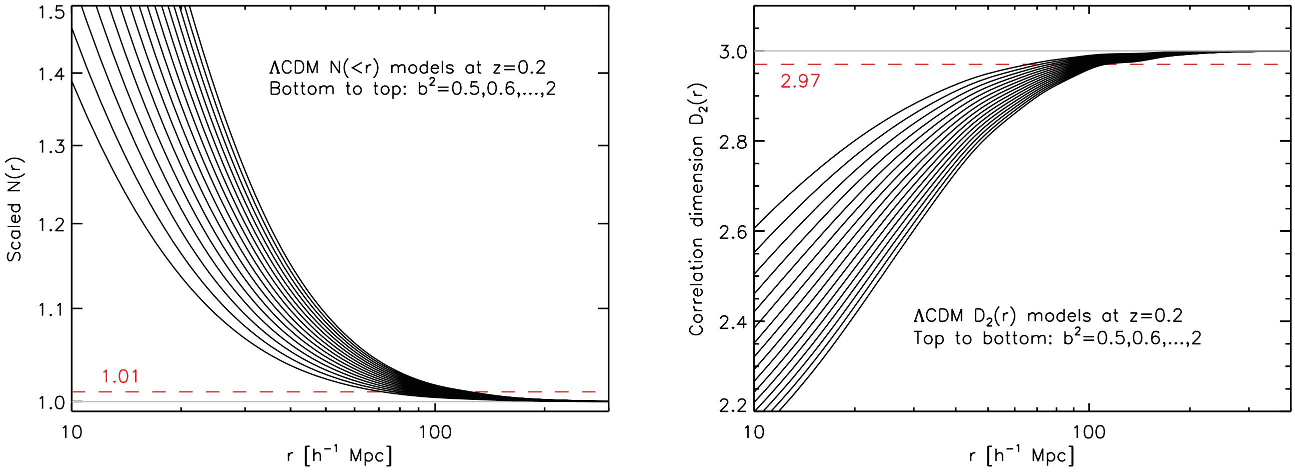

Figure 7. The effect of bias on a ΛCDM $\mathcal{N}(<r)$ model (left) and $D_2(r)$ model (right) at $z = 0.2$. Increasing the bias increases the value of $\mathcal{N}(<r)$ on small scales, and decreases the value of $D_2(r)$ on small scales, and produces a larger homogeneity scale, as seen by the intercepts of the curves with 1 per cent of homogeneity ($\mathcal{N} = 1.01$ and $D_2 = 2.97$, red dotted lines).

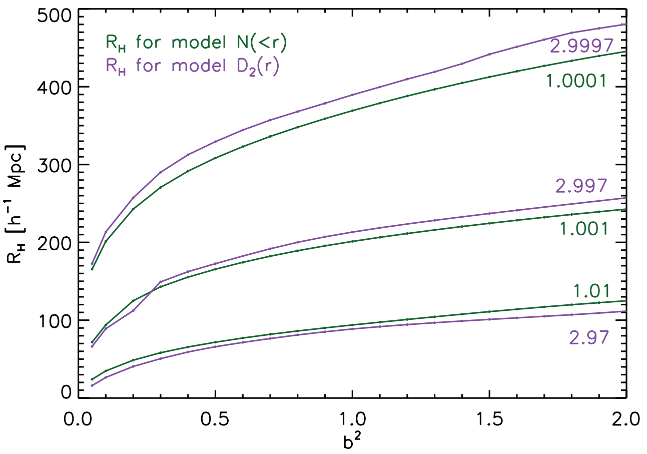

Figure 8. $R_{\rm H}$ for ΛCDM $\mathcal{N}(r)$ (green) and $D_2(r)$ (purple) models at $z = 0.2$ with differing bias b^2. Each curve corresponds to $R_{\rm H}$ evaluated at a different threshold – 1, 0.1 or 0.01 per cent away from homogeneity (from bottom to top, labelled).

stochasticity than those from our method of fitting a smooth curve to many data points, and do not give informative results in this plane.

We see that the model $R_{\rm H}$–$b^2\sigma_8(z)^2$ curves are monotonically increasing. Since we expect $\sigma_8(z)$ in ΛCDM to grow over time due to growth of structure, we would therefore also expect the homogeneity scale to increase over time, for galaxies with fixed bias.

For the WiggleZ data, however, the measured homogeneity scale does not appear to decrease with redshift. This is explained by the fact that the WiggleZ galaxies have increasing bias with redshift, assuming a ΛCDM growth rate. As explained previously, this is understood to be due to the effects of Malmquist bias and downsizing on the selection of the WiggleZ galaxy population, and the colour and magnitude cuts. This counteracts the effect of decreasing σ_8 with redshift. As can be seen in Fig. 9, within the $b^2\sigma_8(z)^2$ range of the data we would not expect a significant change in $R_{\rm H}$, measured at 1 per cent from homogeneity, with redshift.

5.3 Likelihood analysis for homogeneity scale

Rather than trying to measure the scale of homogeneity directly, it can be informative to consider the likelihood that the Universe has reached homogeneity by a certain scale. We can construct a PDF for the homogeneity scale, by combining the likelihood distribution of the data with that of the definition of the homogeneity scale. This gives the probability that homogeneity is reached at a certain scale r.

We could apply a likelihood analysis to either our $\mathcal{N}(<r)$ or $D_2(r)$ measurements. However, it would arguably be invalid for highly correlated data, such as the $\mathcal{N}(<r)$ measurement, since the different contributions to the probability distribution would be correlated, and we do not correct for this. The $D_2(r)$ measurement is much less correlated (see Figs 2 and 3). Also, as we have explained, $D_2(r)$ is the most robust measurement of homogeneity, so we have chosen this for our likelihood analysis.

The likelihood distribution on the data is simply given by the mean and variance of the lognormal realizations in each bin. These give the expected variance of a ΛCDM distribution sampled with the WiggleZ selection function and hence take into account both cosmic variance and shot noise. Here we assume they are Gaussian distributed by virtue of the central limit theorem. However, we also consider the true distributions provided by the lognormal realizations – this gives similar results but with larger errors (see Appendix A).

At each r, the data therefore provide a probability distribution $p_{\rm D}[D_2(r)]$, which we model as Gaussians, as

$$p_{\rm D}[D_2(r)] = \frac{1}{\sigma\sqrt{2\pi}} {\rm e}^{-(D_2(r)-\mu)^2/2\sigma^2}, \qquad (26)$$

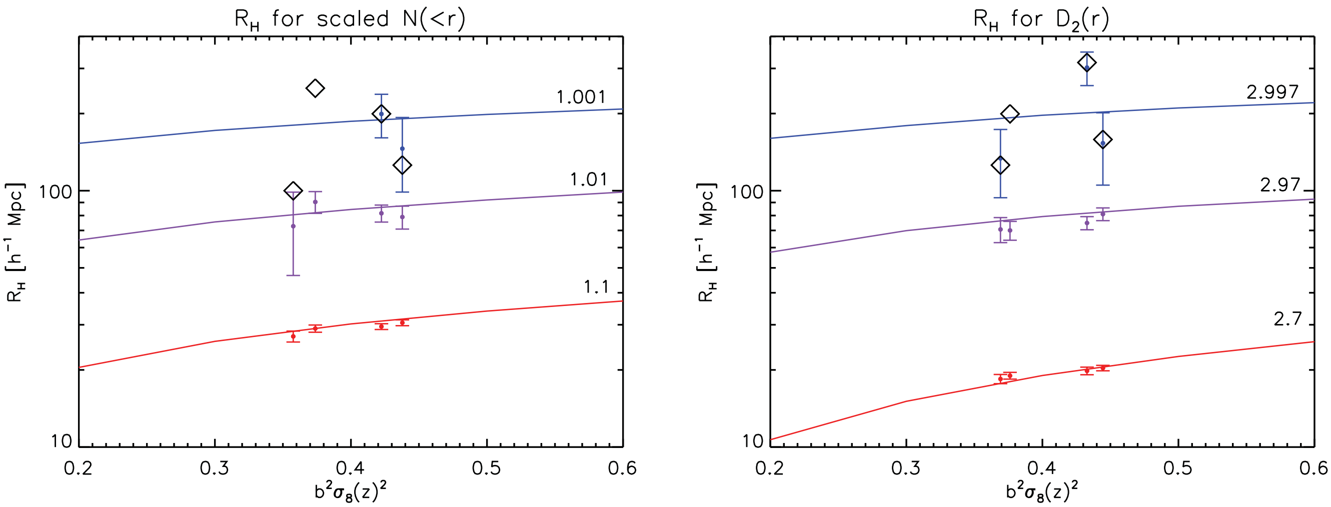

Figure 9. Homogeneity scale $R_{\rm H}$ as a function of $b^2\sigma_8(z)^2$, as predicted by the ΛCDM model, for different thresholds approaching homogeneity (10, 1 and 0.1 per cent from homogeneity, coloured curves from bottom to top), for $\mathcal{N}(<r)$ (left) and $D_2(r)$ (right). The corresponding WiggleZ results are shown as error bars of corresponding colour (the errors are found using lognormal realizations). The $b^2\sigma_8(z)^2$ values of the data increase with redshift slice (so the data points from left to right are from low to high redshift). Not all redshift slices have measurements at 0.1 per cent (blue) since the data do not reach this value in those slices. The definition of homogeneity used by previous authors is where the data come within 1σ of homogeneity – we show these scales as black diamonds. It is clear that this definition has much greater stochasticity than our definition, which fits a smooth curve to many data points. Indeed, this approach gives quite uninformative results in this plane and cannot be compared to the model prediction.

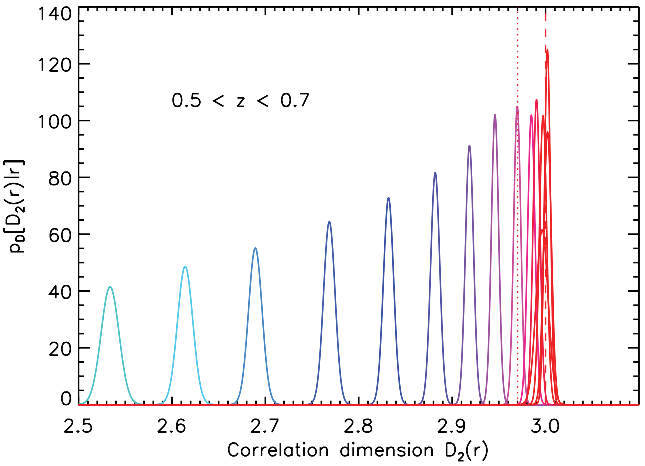

Figure 10. The probability distributions $p_{\rm D}[D_2(r)]$ for each of the r bins in the $0.5 < z < 0.7$ redshift slice. Blue-to-red gradient indicates small to large radius, from 12.5 to 316 h^{-1} Mpc. The area of each distribution above $D_2(r) = 2.97$ (dotted line) gives the probability that homogeneity has been reached within 1 per cent by that value of r.

where μ is the WiggleZ value of D_2 at radius r and σ is the rms variance given by the lognormal realizations. See Fig. 10 for an illustration.

We also expect there to be a likelihood distribution on the $D_2(r)$ value we would measure for a perfectly homogeneous distribution, due to cosmic variance and shot noise caused by our selection function. We can represent this by a likelihood distribution on the homogeneity scale – this would not be a simple delta function at $D_2(r) = 3$, but would have some spread. This distribution should be one-sided – i.e. we do not expect to measure $D_2(r) > 3$, only $D_2(r) \leq 3$, if we have a distribution that approaches homogeneity.[2] We might expect it to be represented by the variance in the random catalogues, which are essentially homogeneous distributions sampled with the WiggleZ selection function. However, the same sources of noise are also present in the lognormal realizations, so we would potentially double count errors if we also used lognormal realizations to determine the likelihood distribution of the data. This means we cannot easily determine the true variance in the value of $D_2(r)$ we would expect to measure for a homogeneous distribution, independently of the variance of the data.

For this reason, we choose to find the likelihood distribution of the data reaching 1 per cent of homogeneity. That is, we assume the likelihood distribution on the homogeneity scale, $p_{\rm H}[D_2(r)]$, is a delta function at $D_2 = 2.97$:

$$p_{\rm H}[D_2(r)] \equiv \delta[D_2(r) - 2.97]. \tag{27}$$

We can then construct the cumulative PDF $P(R_{\rm H} \leq r)$, which gives the probability that the homogeneity scale has been reached at or before scale r, from

$$\begin{aligned} P(R_{\rm H} \leq r) &= \int_{-\infty}^{\infty} p_{\rm D}[D_2(r)] \left(\int_{-\infty}^{D_2(r)} p_{\rm H}(x)\, {\rm d}x \right) {\rm d}D_2(r) \\ &= \int_{2.97}^{\infty} p_{\rm D}[D_2(r)]\, {\rm d}D_2(r). \end{aligned} \tag{28}$$

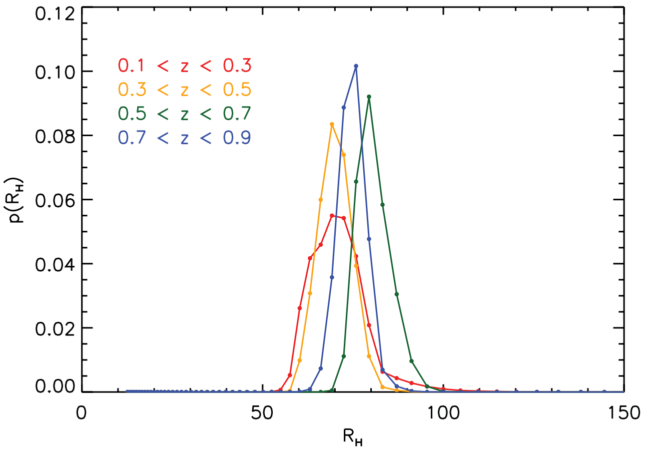

Figure 11. Probability distributions for the scale of homogeneity, $p(R_{\rm H})$, for WiggleZ galaxies in each of the four redshift slices. The homogeneity scale is defined as the scale where the data reach values 1 per cent away from $D_2 = 3$, i.e. $D_2 = 2.97$.

That is, the probability of having reached homogeneity is the area of the likelihood distribution of the data that falls at or above $D_2 = 2.97$.

This cumulative probability is calculated at each scale r. We can then find the PDF for the homogeneity scale, $p(R_{\rm H})$, from

$$p(R_{\rm H}) = \frac{{\rm d}P(R_{\rm H} \leq r)}{{\rm d}r}. \tag{29}$$

The PDFs for the homogeneity scale, for WiggleZ galaxies in each redshift slice, are shown in Fig. 11. We have interpolated between the data bins in order to obtain smoother PDFs. We find the most probable $R_{\rm H}$ values from the mean of the distributions. These are all between 70 and 81 h^{-1} Mpc, and are listed in Table 4. They represent the most probable scale at which the galaxy distribution reaches 1 per cent of homogeneity. We also list the values found for the bias-corrected data, which give the most probable homogeneity scales for the matter distribution, assuming ΛCDM+*WMAP*.

6 ROBUSTNESS OF HOMOGENEITY MEASUREMENT

In this section, we address several issues that could potentially influence our measurement of the homogeneity scale, and perform tests to ensure the robustness of our results. We base our tests on the 15-hr $0.5 < z < 0.7$ region, which is the largest and most populated WiggleZ subregion, but the results are applicable to the entire survey.

6.1 Fractal model test of selection function and boundary effects

A major potential source of bias in our results is the method used to correct for edge effects and the selection function of the survey. We have used 100 random catalogues to correct each individual WiggleZ measurement, as described in Section 3.1. However, this method can potentially bias homogeneity measurements, since it weights measurements by the volume of spheres of radius r included in the survey, and so assumes a homogeneous distribution outside the survey (e.g. Coleman & Pietronero 1992; Sylos Labini et al. 2009). It is therefore important to check that this is not imposing any distortion in our measured correlation dimension, so producing a 'false relaxation' to homogeneity.

---

[2] In some cases, the WiggleZ data do fall below $\mathcal{N} = 1$ or above $D_2 = 3$; this can be explained as the effect of shot noise introduced by the selection function rather than a physical effect, as shown in our comparison with the GiggleZ simulation in Section 6.2.

To test this, we apply our correction method to a range of fractal distributions of known correlation dimension. This has been done previously by a number of works, e.g. Lemson & Sanders (1991), Provenzale, Guzzo & Murante (1994) and Pan & Coles (2002). This allows us to check that our method returns the correct input correlation dimension up to the largest scales we measure in the survey, and to quantify any distortion that may occur.

We generate our fractal distributions using the β-model, a simple self-similar cascading model (see e.g. Castagnoli & Provenzale 1991). This method starts with a cube of side L_0 and splits it up into M smaller cubes of side L_0/n (we take $n = 2$, so $M = 8$). Each subcube is then assigned a probability p of surviving to the next iteration. This is repeated for a certain number of iterations k, and the resultant set of survived points is taken as the final distribution. In the limit of an infinite number of iterations, this produces a monofractal with a correlation dimension given by

$$D_2 = \lim_{k\to\infty} \frac{\log(pM)^k}{\log n^k} = \frac{\log pM}{\log n}. \qquad (30)$$

Our procedure for using β-models to test our analysis method is as follows.

(i) We choose a range of D_2 values (2.7, 2.8, 2.9, 2.95 and 2.97), and for each we generate 100 fractal galaxy distributions, with a boxsize of $(L_0\,h^{-1}\,\mathrm{Gpc})^3$, where L_0 is the length of the longest side of the WiggleZ 15-hr $0.5 < z < 0.7$ selection function grid ($623.5\,h^{-1}$ Mpc).

(ii) We then sample each distribution with the WiggleZ selection function for the 15-hr $0.5 < z < 0.7$ region. We normalize the resulting distribution to give the same number of points as WiggleZ galaxies in this region. This gives us fractal mock catalogues, and we then measure D_2 for these in the same way as for the WiggleZ data, correcting the counts-in-spheres measurements with random catalogues. This gives a result that is influenced by both the WiggleZ selection function, and our correction method.

Fig. 12 shows the mean $\mathcal{N}(<r)$ and $D_2(r)$ results for the different fractal distributions up to $D_2 = 2.95$, with the WiggleZ selection function and correction method applied. Even up to $D_2 = 2.95$, they are clearly distinguishable from ΛCDM and the WiggleZ data. The D_2 values measured are consistent with the input values up to at least $200\,h^{-1}$ Mpc, well above the homogeneity scales measured for the data. These results indicate that the WiggleZ selection function and correction method do not have a significant effect on the measured correlation dimension up to the scales we measure for homogeneity.

It is notable that the size of the error bars gets smaller for models with larger D_2. This is a real effect in the model; for larger D_2, more boxes survive at each iteration, so there are a smaller number of possible configurations for the final distribution, resulting in lower variance for a box of a particular volume.

To quantify how well we can exclude fractal models, we fit a line of constant D_2 to each set of fractal data, over the range [80, 300] h^{-1} Mpc (shown in Fig. 12). This gives us the best-fitting D_2 value we would expect to measure for each fractal distribution over this range, taking into account bias from the selection function. We then find the formal probability of these values fitting the WiggleZ data. Doing this, we find we can exclude a fractal dimension of $D_2(r) = [2.9, 2.95, 2.97]$ at the $[19,6,4]\sigma$ level. In other words, we can exclude fractal distributions with dimension $D_2(r) < 2.97$ at over 99.99 per cent confidence on scales from 80 to $300h^{-1}$ Mpc.

Our results agree with those of Lemson & Sanders (1991), Provenzale et al. (1994) and Pan & Coles (2002), who also find that for samples on scales larger than the homogeneity scale, boundary corrections do not have a significant effect on the analysis. A further check would be to test different types of fractal model other than the β-model, but we leave this for future work. We also note that we still assume an FRW metric in our fractal analysis; an improved consistency check would be to calculate the actual metric in these fractal models (see Section 7), but this is beyond the scope of this paper.

6.2 Comparison with the GiggleZ *N*-body simulation

We have also compared our results with a ΛCDM cosmological *N*-body simulation. This allows us to check that our analytical ΛCDM+*WMAP* model (incorporating Kaiser and streaming models for redshift-space distortions) is consistent, over the relevant

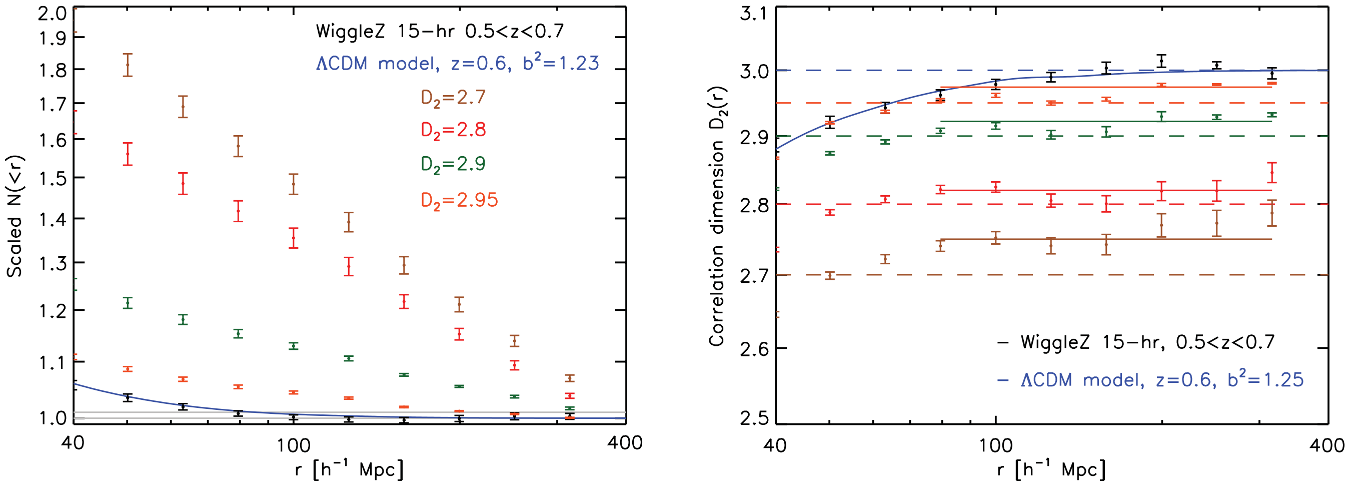

Figure 12. Fractal model comparisons of the measured correlation dimension from the WiggleZ 15-hr region, $0.5 < z < 0.7$ redshift slice. The WiggleZ data (black error bars) and ΛCDM model (blue curve) are compared with several different β-models with different fractal dimension ($D_2 = 2.7$, 2.8, 2.9 and 2.95), which have been sampled with the 15-hr selection function and analysed in the same way as the WiggleZ data (coloured error bars). The uncertainties shown are the error in the mean of 100 fractal realizations. The input fractal dimensions are shown as dotted lines with corresponding colours. The best-fitting $D_2(r)$ value fits over the range $r = [80, 300]\,h^{-1}$ Mpc are shown as solid lines.

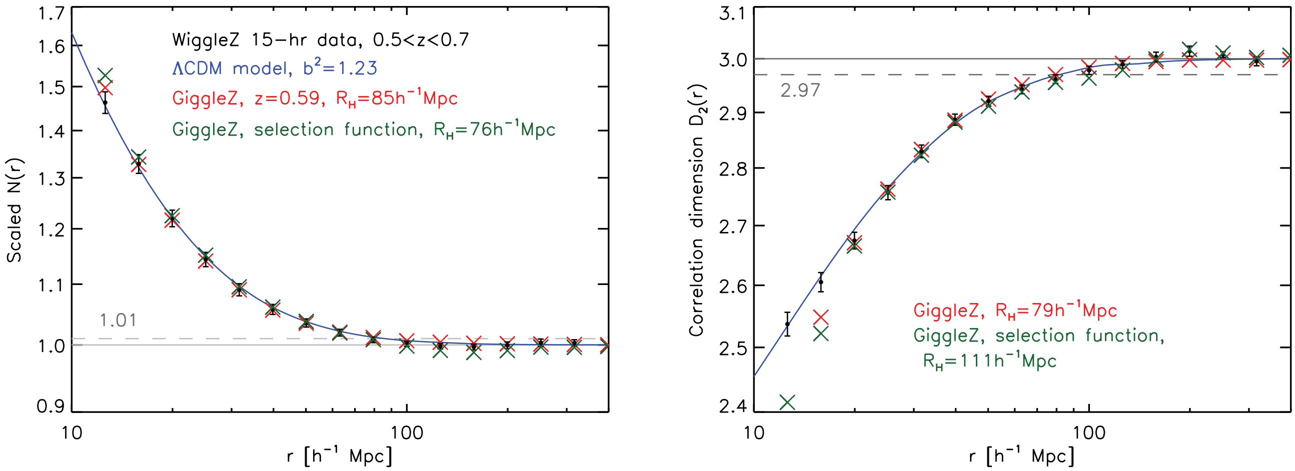

Figure 13. Comparison of the GiggleZ N-body simulation with WiggleZ, for the 15-hr region $0.5 < z < 0.7$ redshift slice. The $\mathcal{N}(<r)$ results are shown on the left, and $D_2(r)$ on the right. The WiggleZ data are shown as black data points, and a ΛCDM model is shown in blue. The results for the full GiggleZ box are shown as the red crosses. The green crosses show the results for the GiggleZ simulation sampled with the WiggleZ 15-hr $0.5 < z < 0.7$ selection function. The measured homogeneity scale R_{H} is indicated for each.

scales, with a full simulation including non-linear effects. It also provides a further test of selection function effects, since we can show that homogeneity measurements of a ΛCDM distribution are not distorted when the WiggleZ selection function is applied.

The GiggleZ (Giga-parsec WiggleZ) simulation (Poole et al., in preparation) is a suite of dark matter N-body simulations run at Swinburne University of Technology, designed for theoretical analyses of the WiggleZ data set. It was run using a modified version of the N-body code GADGET-2 (Springel, Yoshida & White 2001) using a *WMAP*5 cosmology. We use the main simulation, which has a volume of $(1000\,h^{-1}\,\mathrm{Mpc})^3$ and 2160^3 particles of mass $7.5 \times 10^9\,h^{-1}\,\mathrm{M}_{\odot}$.

Halo finding for GiggleZ was performed using SUBFIND (Springel et al. 2001), which utilizes a friends-of-friends (FoF) algorithm to identify coherent overdensities of particles and a substructure analysis to determine bound overdensities within each FoF halo. For our analysis, we rank-order the resulting SUBFIND substructure catalogues by their maximum circular velocity (V_{max}) and select a contiguous subset of 250 000 haloes (selected to yield a number density comparable to the survey) over a range of V_{max} chosen to yield a bias comparable to that of WiggleZ. We use a catalogue at a redshift of $z = 0.593$, corresponding to the mid-redshift of the WiggleZ $0.5 < z < 0.7$ redshift slice.

We add the effect of redshift-space distortions by shifting the positions of the haloes according to their line-of-sight peculiar velocities. That is, the comoving position of each halo relative to an observer, $\boldsymbol{x}$, is shifted by a vector $\Delta\boldsymbol{x}$,

$$\Delta\boldsymbol{x} = \frac{\boldsymbol{x}}{|\boldsymbol{x}|}\frac{(1+z)}{H(z)}v_{\mathrm{rad}}, \tag{31}$$

where $v_{\mathrm{rad}} = (\boldsymbol{x} \cdot \boldsymbol{v})/|\boldsymbol{x}|$ is the radial velocity of the halo along the observer's line of sight, and we place the observer at the same coordinates relative to the GiggleZ box as for the WiggleZ selection function grid.

We then calculate $\mathcal{N}(<r)$ and $D_2(r)$ using the following two different methods.

(i) Using the full GiggleZ box. We correct the measurement using a random distribution within the same volume box, with 100 times the number of galaxies as the GiggleZ sample.

(ii) Applying the WiggleZ 15-hr $0.5 < z < 0.7$ selection function to GiggleZ. This creates a mock WiggleZ survey containing 10 830 galaxies. We correct the measurement using the random catalogues used for the WiggleZ data. This allows us to see the effects induced purely by the selection function.

The results are shown in Fig. 13. The full GiggleZ data set is very consistent with both the WiggleZ data and the ΛCDM model. The consistency with the model indicates that the implementation of redshift-space distortions in the model, described in Section 4.2, has a good level of accuracy, to scales as small as $\sim 20\,h^{-1}$ Mpc. The deviation of the smallest scale bin is a resolution effect, since the GiggleZ catalogue is sparser than WiggleZ (with 10 830 galaxies, versus 17 928 WiggleZ galaxies in the same region and redshift slice). The GiggleZ results both with and without the WiggleZ selection function are also consistent within the size of the WiggleZ error bars, showing that the selection function and correction method do not have a significant effect. It can be seen that the selection function causes a few data points in the $\mathcal{N}(<r)$ plot to go below 1, and in the $D_2(r)$ plot to go above 3. This shows that adding shot noise can produce this apparent unphysical effect, explaining why this is also seen in some of the WiggleZ results (Figs 5 and 6).

6.3 Comparison of different correction methods

To further demonstrate the robustness of our correction method, we illustrate the results obtained for two alternative correction methods – first, using only complete spheres for the counts-in-spheres measurements (the 'exclusion' method mentioned in Section 3.1), but still correcting for the selection function, and secondly, using no correction for the selection function. By using the exclusion method, we do not have to deal with survey edge effects; however, since WiggleZ is not volume limited, it is still necessary to use random catalogues to correct for the selection function. That is, we recalculate equation (4) but with G equal to the number of WiggleZ galaxies at the centre of complete spheres of radius r. We show this result in Fig. 14, for the 15-hr $0.5 < z < 0.7$ region, as red error bars, where the uncertainty is calculated using our lognormal realizations. We compare this to the result using our correction method (black error bars). The red error bars are consistent with the black error bars, but show more scatter and have higher noise, which increases

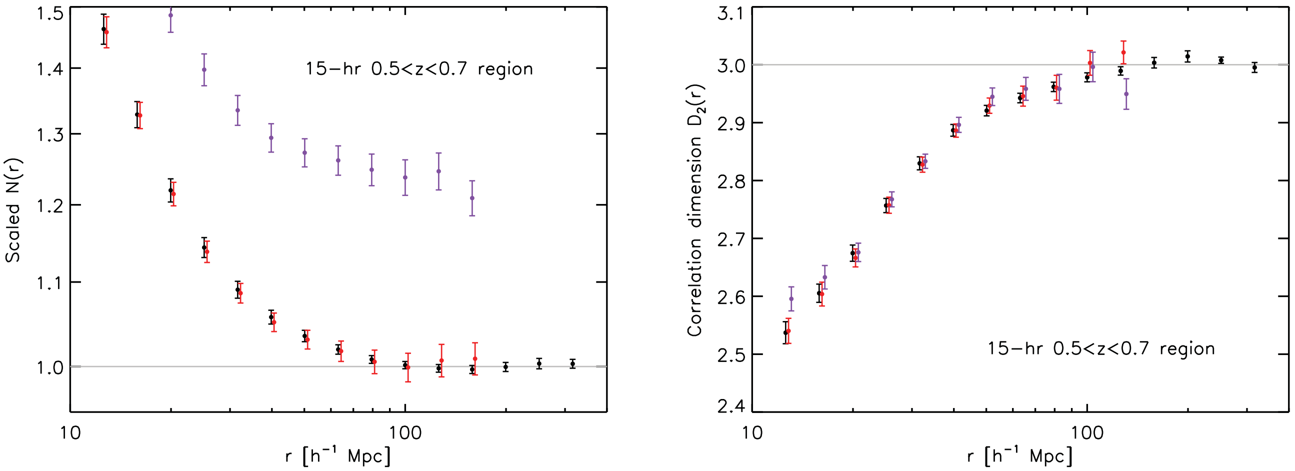

Figure 14. Illustration of the result using different correction methods, for the 15-hr $0.5 < z < 0.7$ region. Black error bars show the result using our correction method with random catalogues. The red error bars are obtained when only complete spheres are used in the counts-in-spheres measurement while still correcting for the selection function. The, purple error bars show the result for complete spheres, with no selection function correction. In each case, the uncertainties are calculated using lognormal realizations. The red and purple error bars are shifted slightly to the right for clarity.

for larger radius, reflecting the lower statistics where fewer spheres contribute to the measurement. The measurements must also be cut off at a lower radius, since not enough larger spheres fit inside the survey region.

We can also compare the result without any selection function correction, which illustrates the necessity of correcting for having a non-volume-limited sample. However, WiggleZ contains holes in the angular coverage, which are independent of assumptions about completion, so we must still take these into account. We therefore normalize each $N^i(r)$ measurement by the volume within the selection function included in that sphere. Therefore, we calculate

$$\mathcal{N}_{\mathrm{no\_corr}}(r) = \frac{1}{G_{\mathrm{complete}}(r)} \sum_{\text{complete spheres } i} \frac{N^i(r)}{V_{\mathrm{SF},i} \times \bar{\rho}}, \tag{32}$$

where $G_{\mathrm{complete}}(r)$ is the number of WiggleZ galaxies at the centre of complete spheres of radius r, $V_{\mathrm{SF},i}$ is the volume within the selection function of the ith galaxy, and $\bar{\rho} = n_{\mathrm{W}}/V_{\mathrm{SF,tot}}$ is the mean density of WiggleZ, i.e. the number of WiggleZ galaxies divided by the total volume in the selection function, excluding holes. We show this in Fig. 14 as purple error bars, where the uncertainty is again calculated using our lognormal realizations (so the uncertainty assumes ΛCDM). Although there is a vertical offset in the $\mathcal{N}(<r)$ plot, caused by the selection function, the $D_2(r)$ results are remarkably similar to those when correcting for the selection function, though again with more noise. There is still a clear transition towards values close to $D_2 = 3$ on large scales. The offset in $\mathcal{N}(<r)$ can be attributed to a selection effect: since we use only complete spheres, we weight the measurement towards the central part of the redshift range, where the completeness is highest and so the number density of WiggleZ galaxies is highest. Since $D_2(r)$ is the slope of $\mathcal{N}(<r)$, it does not depend on the number density itself and so is more robust when summing over a varying selection function.

7 DISCUSSION

Our WiggleZ $\mathcal{N}(<r)$ and $D_2(r)$ results show a very strong agreement with an FRW-based ΛCDM model. However, one of the strongest arguments against previous homogeneity measurements is that the method of correcting for survey selection functions, such as using random catalogues as we do, can distort the data, producing a 'false relaxation' to homogeneity. We have therefore tested this by applying the WiggleZ selection function to a range of inhomogeneous fractal models and a ΛCDM N-body simulation. We have shown that there is no significant impact on our homogeneity measurement, up to at least 200 h^{-1} Mpc. In addition, we have compared the results from our correction method to an analysis using only complete spheres, with and without correcting for the selection function, and have shown that these are consistent, further demonstrating the robustness of our result. We can rule out fractals with fractal dimension up to $D_2 = 2.97$, on scales between 80 and 300 h^{-1} Mpc, at over 99.99 per cent confidence.

We can also be confident that our result is robust to any assumptions in modelling the WiggleZ survey selection function, since Blake et al. (2010) showed that even extreme variations in modelling the angular completeness produce only $\sim 0.5\sigma$ shifts in estimates of the power spectrum. Changes in the parametrization of the redshift distribution were shown to cause larger deviations, but only at scales $>200\,h^{-1}$ Mpc, which are well above the scales on which we measure homogeneity.

Our result is a very good consistency check of ΛCDM. However, it is not independent of the assumption of the FRW metric. A complication for all homogeneity measurements is that we can only observe galaxies on our light cone, and not on spatial surfaces. Maartens (2011) points out that it is therefore not possible to make a homogeneity measurement without making some assumptions, such as the FRW metric (to convert redshifts to distances) and the cosmological principle. Indeed, we must always assume an FRW-based model, ΛCDM in our case, to convert redshifts and angles to distance coordinates. This is also problematic considering that inhomogeneities produce perturbations in the FRW metric, so can potentially distort distance measurements by affecting the paths travelled by light rays (e.g. Wiltshire 2009; Meures & Bruni 2011). However, it seems highly unlikely that our measurements would so closely agree with FRW-based ΛCDM [in both the amplitude and shape of the $\mathcal{N}(<r)$ and $D_2(r)$ curves], if the distribution were, actually, inhomogeneous up to the largest scales probed, and we had incorrectly assumed an FRW metric. We have tested our method

using a fractal model, and shown that this gives a completely different form of these curves. Our assumption of FRW also seems reasonable, since we know that distance measurements from Type Ia supernovae fit the Hubble diagram well to first order up to $z \sim 1.4$ (Conley et al. 2011; Davis et al. 2011), so that any perturbations due to inhomogeneities can only be a second- or higher-order effect. All this means that we can take our results to be a strong consistency check of FRW-based ΛCDM.

It would be possible to test this further by making isotropy measurements in thin redshift shells, over a range of redshifts, without converting them to distances. We could also directly test the assumption of FRW, by calculating the non-FRW metrics of our fractal models and testing our analysis on these. We could also calculate the effects of metric perturbations for ΛCDM, including the effects of backreaction, as in e.g. Buchert (2000), Li & Schwarz (2007), Wiltshire (2007b) and Behrend, Brown & Robbers (2008). This would allow us to test the effect of incorrectly assuming an FRW metric. Alternatively, several possible consistency tests of homogeneity and the Copernican Principle that do not assume FRW have been suggested by e.g., Clarkson et al. (2008); Shafieloo & Clarkson (2010); Maartens (2011); Heavens et al. (2011). However, we leave all these suggestions for future work.

An alternative way of defining the 'homogeneity scale' was recently suggested by Bagla et al. (2008) and Yadav et al. (2010), as where the measured fractal dimension D_q becomes consistent with the ambient dimension within the 1σ statistical uncertainty, $\sigma_{\Delta D_q}$. They make predictions for a ΛCDM model, by deriving an approximation for D_q and $\sigma_{\Delta D_q}$, given a particular correlation function, and showing how these scale with sphere size. Using this, they predict that the 'true' homogeneity scale in ΛCDM is 260 h^{-1} Mpc. They also predict that D_q and $\sigma_{\Delta D_q}$ scale the same way with the correlation function, and so their definition of homogeneity does not change with bias or redshift. Their definition is therefore beneficial, as it is robust to the tracer galaxy population. It is also not arbitrary, and indeed the scale above which the fractal dimension is consistent with homogeneity within cosmic variance is arguably a physically meaningful scale to define as homogeneous, since above this scale the distribution cannot be distinguished from a homogeneous one. However, as we have pointed out (Section 3.6), it is difficult to apply their definition to a real measurement, since their approximation for $\sigma_{\Delta D_q}$ only accounts for the variance of the correlation function (in the limit of weak clustering) and shot noise, but ignores errors due to survey geometry and the selection function. These additional contributions mean that real errors will always be larger, and so will always measure a smaller homogeneity scale than the 'true' one. Since these errors will be different for different surveys, and cannot be separated out from the variance in the underlying correlation function, homogeneity measurements made in this way cannot be easily compared between different surveys, or with theory. The benefit of our method of defining the homogeneity scale, even if we have to make an arbitrary choice about the value of ΔD_q we accept for homogeneity, is that it can be used to easily compare the results from different surveys and with a theoretical model.

A homogeneity scale below 100 h^{-1} Mpc may also seem to contradict the fact that the correlation function has a known feature, the BAO peak at ∼105 h^{-1} Mpc (Eisenstein et al. 2005; Percival et al. 2010; Beutler et al. 2011; Blake et al. 2011c). However, the BAO peak has only a small impact on the counts-in-spheres statistic. It is visible in our ΛCDM prediction for $D_2(r)$ [which is more sensitive to it than $\mathcal{N}(<r)$, since it is a differential measurement], as a small dip at just over 100 h^{-1} Mpc (see Fig. 7). It means that the $D_2(r)$ curve does not increase monotonically around this scale, so we must be careful if attempting to measure an intercept with homogeneity that lies close to this. The magnitude of the distortion due to the BAO peak, ΔD_2, is of order ∼0.01 for $b \sim 1$ so does not affect our homogeneity scale measured at $D_2 = 2.97$. However, it is more significant for more highly biased tracers (as also pointed out by Bagla et al. 2008), and for a highly biased population, such as LRGs, it may be necessary to measure the homogeneity scale at a D_2 value closer to 3, to avoid this region.

It has been pointed out (e.g. Sylos Labini & Pietronero 2010) that measurements of large structures in galaxy surveys are seemingly at odds with a homogeneity scale below 100 h^{-1} Mpc. Previous surveys have detected structures on scales much larger than this (de Lapparent, Geller & Huchra 1989; Geller & Huchra 1989). The largest observed structure in the Universe, the Sloan Great Wall, which is 320 h^{-1} Mpc long (Gott et al. 2005), appears inconsistent with the existence of a homogeneity scale below 100 h^{-1} Mpc. However, this is not incompatible with our results, since we show the scale where the data have $D_2 \geq 2.97$, and it is not impossible to have fluctuations on scales larger than this. Also, these large structures, including the Sloan Great Wall, are usually filamentary, whereas we have measured a volume statistic which averages over fluctuations.

8 CONCLUSION

We have measured the large-scale transition to homogeneity in the distribution of galaxies, using the WiggleZ survey, in four redshift bins between $z = 0.1$ and 0.9. We measured the mean, scaled counts-in-spheres $\mathcal{N}(<r)$ and the correlation dimension, $D_2(r)$, and found these to be in excellent agreement with a ΛCDM model with *WMAP* parameters, including redshift-space distortions. We also presented a new, model-independent method for determining the homogeneity scale R_{H} from data. This involves fitting a polynomial curve to the data, and finding where this intercepts chosen values close to homogeneity. This is a more reliable method than finding where the data come within 1σ of homogeneity, since it does not depend directly on the size of the error bars and is less susceptible to noise. It also allows a direct comparison between data and theory, and between different surveys of differing bias and redshift.

We summarize our results as follows.

(i) Our $\mathcal{N}(<r)$ and $D_2(r)$ results show a very strong agreement with an FRW-based ΛCDM+*WMAP* model incorporating redshift-space distortions. They show a clear transition from an inhomogeneous, clustered distribution on small scales, to a homogeneous one on large scales. This transition matches that of the ΛCDM model. We have thereby conducted a very stringent consistency check of ΛCDM.

(ii) If we define the 'homogeneity scale' R_{H} as the scale where the data become consistent with homogeneity within 1 per cent, then from a likelihood analysis of D_2(r), we measure R_{H} to be 71 ± 8 h^{-1}Mpc at $z \sim 0.2$, 70 ± 5 h^{-1}Mpc at $z \sim 0.4$, 81 ± 5 h^{-1}Mpc at $z \sim 0.6$ and 75 ± 4 h^{-1}Mpc at $z \sim 0.8$. These values are consistent with those of the ΛCDM+*WMAP* model with best-fitting bias, of $R_{\mathrm{H}} = 76, 78, 81$ and 78 h^{-1} Mpc.

(iii) We find that the homogeneity scale of our ΛCDM+*WMAP* model increases with clustering amplitude $b(z)\sigma_8(z)$. For a population with fixed bias, we therefore predict the homogeneity scale to grow over time, since $\sigma_8(z)$ increases due to growth of structure. The bias of the WiggleZ galaxies increases with redshift, so the measured R_{H} values do not change with redshift. If we correct our data for bias, assuming ΛCDM, then we measure a homogeneity scale for the matter distribution that increases over time, consistent with our ΛCDM+*WMAP* model.

(iv) The WiggleZ results are in excellent agreement with those of the GiggleZ *N*-body simulation incorporating redshift-space distortions. It is also in excellent agreement with our analytic ΛCDM+*WMAP* model, showing that our model for redshift-space distortions is accurate down to scales as small as $20\,h^{-1}$ Mpc.

(v) We can exclude a fractal with fractal dimension up to $D_2 = 2.97$, on scales between $\sim 80 h^{-1}$ Mpc and the largest scales probed by our measurement, $\sim 300\,h^{-1}$ Mpc, at 99.99 per cent confidence.

(vi) By applying our analysis to the GiggleZ simulation, as well as a suite of fractal distributions of differing fractal dimension, we have shown that our result is not significantly distorted by the WiggleZ selection function and our method of correcting for it. We also show that we obtain consistent results even using different correction methods, i.e. using only complete spheres for the measurement, with and without correcting for incompleteness. This therefore confirms the reliability and robustness of our results.

ACKNOWLEDGMENTS

MIS would like to thank Florian Beutler for many helpful comments on the paper. We also thank Eyal Kazin, David Parkinson, Luigi Guzzo and Andy Taylor for helpful discussions. MIS acknowledges support from a Jean Rogerson Scholarship and a UWA Top-up Scholarship from the University of Western Australia. GBP acknowledges support from two Australian Research Council (ARC) Discovery Projects (DP0772084 and DP1093738). The Centre for All-sky Astrophysics (CAASTRO) is an Australian Research Council Centre of Excellence, funded by grant CE11E0090.

GALEX is a NASA Small Explorer, launched in 2003 April. We gratefully acknowledge NASA's support for construction, operation and science analysis for the *GALEX* mission, developed in cooperation with the Centre National d'Etudes Spatiales of France and the Korean Ministry of Science and Technology.

The WiggleZ survey would not have been possible without the dedicated work of the staff of the Australian Astronomical Observatory in the development and support of the AAOmega spectrograph, and the running of the AAT.

REFERENCES

Afshordi N., 2004, Phys. Rev. D, 70, 083536
Amendola L., Palladino E., 1999, ApJ, 514, L1
Bagla J. S., Yadav J., Seshadri T. R., 2008, MNRAS, 390, 829
Baumann D., 2009, preprint (arXiv:0907.5424)
Behrend J., Brown I. A., Robbers G., 2008, J. Cosmol. Astropart. Phys., 1, 13
Beutler F. et al., 2011, MNRAS, 416, 3017
Blake C., Wall J., 2002, Nat, 416, 150
Blake C. et al., 2009, MNRAS, 395, 240
Blake C. et al., 2010, MNRAS, 406, 803
Blake C. et al., 2011a, MNRAS, 415, 2876
Blake C. et al., 2011b, MNRAS, 415, 2892
Blake C. et al., 2011c, MNRAS, 418, 1707
Borgani S., 1995, Phys. Rep., 251, 1
Brouzakis N., Tetradis N., Tzavara E., 2007, J. Cosmol. Astropart. Phys., 2, 13
Buchert T., 2000, Gen. Relativ. Gravitation, 32, 105
Buchert T., 2008, Gen. Relativ. Gravitation, 40, 467
Castagnoli C., Provenzale A., 1991, A&A, 246, 634
Clifton T., Ferreira P. G., Land K., 2008, Phys. Rev. Lett., 101, 131302
Coleman P. H., Pietronero L., 1992, Phys. Rep., 213, 311
Coles P., 1993, MNRAS, 262, 1065
Coles P., Jones B., 1991, MNRAS, 248, 1
Conley A. et al., 2011, ApJS, 192, 1
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Davis T. M. et al., 2011, ApJ, 741, 67
de Lapparent V., Geller M. J., Huchra J. P., 1989, ApJ, 343, 1
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Eisenstein D. J. et al., 2005, ApJ, 633, 560
Ellis G. F. R., Buchert T., 2005, Phys. Lett. A, 347, 38
Fisher K. B., 1995, ApJ, 448, 494
Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576
Gabrielli A., Joyce M., Sylos Labini F., 2002, Phys. Rev. D, 65, 083523
Geller M. J., Huchra J. P., 1989, Sci, 246, 897
Glazebrook K. et al., 2004, Nat, 430, 181
Gott J. R., III, Jurić M., Schlegel D., Hoyle F., Vogeley M., Tegmark M., Bahcall N., Brinkmann J., 2005, ApJ, 624, 463
Guzzo L., 1997, New Astron., 2, 517
Hamilton A. J. S., 1992, ApJ, 385, L5
Hamilton A. J. S., 1993, ApJ, 417, 19
Hatton S., Cole S., 1998, MNRAS, 296, 10
Hawkins E. et al., 2003, MNRAS, 346, 78
Heavens A. F., Jimenez R., Maartens R., 2011, J. Cosmol. Astropart. Phys., 9, 35
Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, ApJ, 624, 54
Joyce M., Montuori M., Sylos Labini F., 1999, ApJ, 514, L5
Joyce M., Anderson P. W., Montuori M., Pietronero L., Sylos Labini F., 2000, Europhys. Lett., 50, 416
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1987, MNRAS, 227, 1
Kolb E. W., Matarrese S., Notari A., Riotto A., 2005, Phys. Rev. D, 71, 023524
Komatsu E. et al., 2011, ApJS, 192, 18
Kurokawa T., Morikawa M., Mouri H., 2001, A&A, 370, 358
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lemson G., Sanders R. H., 1991, MNRAS, 252, 319
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Li N., Schwarz D. J., 2007, Phys. Rev. D, 76, 083011
Maartens R., 2011, Phil. Trans. R. Soc. London, Ser. A, 369, 5115
Martínez V. J., Coles P., 1994, ApJ, 437, 550
Martínez V. J., Pons-Bordería M.-J., Moyeed R. A., Graham M. J., 1998, MNRAS, 298, 1212
Martínez V., Saar E., 2002, in Starck J.-L., Murtagh F. D., eds, Proc. SPIE, Vol. 4847, Astronomical Data Analysis II. SPIE, Bellingham, p. 86
Meures N., Bruni M., 2011, MNRAS, 419, 1937
Pan J., Coles P., 2000, MNRAS, 318, L51
Pan J., Coles P., 2002, MNRAS, 330, 719
Peacock J. A., 1999, Cosmological Physics. Cambridge Univ. Press, Cambridge, UK
Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Peebles P. J. E., 1993, Principles of Physical Cosmology. Princeton Univ. Press, Princeton, NJ
Percival W. J. et al., 2010, MNRAS, 401, 2148
Pietronero L., Montuori M., Sylos Labini F., 1997, in Turok N., ed., Critical Dialogues in Cosmology. World Scientific Press, Singapore, p. 24
Provenzale A., Guzzo L., Murante G., 1994, MNRAS, 266, 555
Räsänen S., 2006, J. Cosmol. Astropart. Phys., 11, 3
Räsänen S., 2011, Classical Quantum Gravity, 28, 164008
Ripley B. D., 1977, J. R. Statistical Soc. Series B (Methodological), 39, 172
Sachs R. K., Wolfe A. M., 1967, ApJ, 147, 73
Sarkar P., Yadav J., Pandey B., Bharadwaj S., 2009, MNRAS, 399, L128
Scaramella R. et al., 1998, A&A, 334, 404
Scharf C. A., Jahoda K., Treyer M., Lahav O., Boldt E., Piran T., 2000, ApJ, 544, 49
Scherrer R. J., Weinberg D. H., 1998, ApJ, 504, 607
Schwarz D. J., 2002, preprint (astro-ph/0209584)

Shafieloo A., Clarkson C., 2010, Phys. Rev. D., 81, 083537 (Table A1)
Smith R. E. et al., 2003, MNRAS, 341, 1311
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Sylos Labini F., 2011, EPL, 96, 59001
Sylos Labini F., Pietronero L., 2010, J. Statistical Mech.: Theor. Exp., 11, 29
Sylos Labini F., Montuori M., Pietronero L., 1998, Phys. Rep., 293, 61
Sylos Labini F., Vasilyev N. L., Baryshev Y. V., 2009, A&A, 508, 17
Wiltshire D. L., 2007a, New J. Phys., 9, 377
Wiltshire D. L., 2007b, Phys. Rev. Lett., 99, 251101
Wiltshire D. L., 2009, Phys. Rev. D, 80, 123512
Wu K. K. S., Lahav O., Rees M. J., 1999, Nat, 397, 225
Yadav J., Bharadwaj S., Pandey B., Seshadri T. R., 2005, MNRAS, 364, 601
Yadav J. K., Bagla J. S., Khandai N., 2010, MNRAS, 405, 2009
Zehavi I. et al., 2005, ApJ, 630, 1

APPENDIX A: GAUSSIAN VERSUS TRUE PROBABILITY DISTRIBUTIONS $p_{\mathrm{D}}[D_2(r)]$

In our likelihood analysis for the homogeneity scale (Section 5.3), we assume that the distribution of the 100 lognormal realizations in each bin, $p_{\mathrm{D}}[D_2(r)]$, is Gaussian. This allows us to interpolate between the data points and errors, thereby creating more finely spaced $p_{\mathrm{D}}[D_2(r)]$ distributions, in order to determine a smoother PDF for the homogeneity scale, $P(R_{\mathrm{H}} \le r)$. However, it is not obvious that these distributions should be Gaussian. We therefore repeat the analysis, but use the true distributions given by the lognormal realizations.

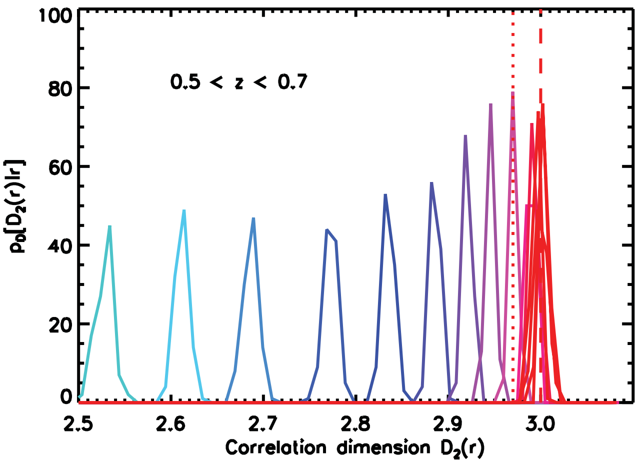

Figure A1. The probability distributions $p_{\mathrm{D}}[D_2(r)]$ for each of the r bins (blue-to-red gradient indicates small to large radius) in the $0.5 < z < 0.7$ redshift slice. Unlike in Fig. 10, we do not assume these are Gaussian distributions; rather, we plot the distributions given by our 100 lognormal realizations.

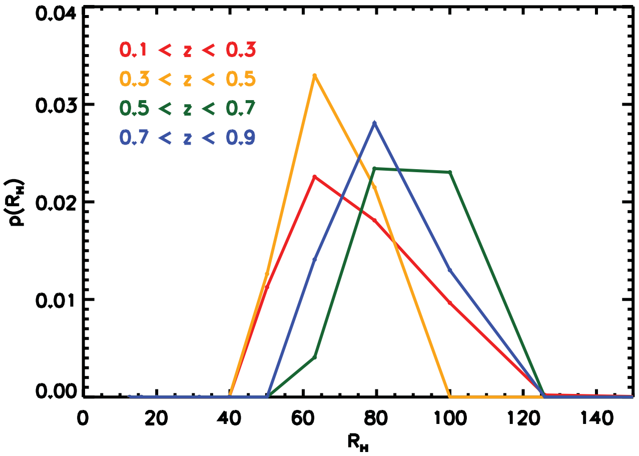

Figure A2. Probability distributions for the homogeneity scale, $p(R_{\mathrm{H}})$, for WiggleZ galaxies in each of the four redshift slices. These are calculated from the probability distributions $p_{\mathrm{D}}[D_2(r)]$ of the 100 lognormal realizations, rather than assuming Gaussians.

Table A1. Comparison of the most probable R_{H} values from a likelihood analysis using the true $p_{\mathrm{D}}[D_2(r)]$ distributions from lognormal realizations and assuming Gaussian distributions.

Redshift	R_{H} for true $p_{\mathrm{D}}[D_2(r)]$ distributions (h^{-1} Mpc)	R_{H} assuming Gaussians (h^{-1} Mpc)
$0.1 < z < 0.3$	79 ± 19	71 ± 8
$0.3 < z < 0.5$	71 ± 13	70 ± 5
$0.5 < z < 0.7$	91 ± 16	81 ± 5
$0.7 < z < 0.9$	84 ± 16	75 ± 4

This gives the $p_{\mathrm{D}}[D_2(r)]$ distributions shown in Fig. A1. They are not smooth Gaussians, although they are close to Gaussian. Their resolution is limited by the number of lognormal realizations, so they could be improved by using more lognormal realizations, although we do not do this here.

We then use these to calculate the PDF for the homogeneity scale, $P(R_{\mathrm{H}} \le r)$, in the same way as we did in Section 5.3 for the Gaussian distributions. This gives the PDFs shown in Fig. A2, for each redshift slice. The mean values and errors are shown in Table A1, along with those from our original analysis. The values are very similar, though the errors are larger. This is because we cannot interpolate between data points as easily, and there are a finite number of lognormal realizations contributing to the distribution for each data point. This means the distribution is effectively smoothed, giving larger uncertainties.

This paper has been typeset from a TEX/LATEX file prepared by the author.