Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 1 February 2008 (MN LATEX style file v2.2)

The Galaxy Power Spectrum: Precision Cosmology from Large Scale Structure?

Ariel G. Sánchez[1,2⋆], Shaun Cole[2].

[1] *Instituto de Astronomía Teórica y Experimental (IATE), OAC, UNC, Laprida 854, X5000BGR, Córdoba, Argentina.*
[2] *The Institute for Computational Cosmology, Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK.*

Submitted to MNRAS

ABSTRACT

Published galaxy power spectra from the 2dFGRS and SDSS are not in good agreement. We revisit this issue by analyzing both the 2dFGRS and SDSS DR5 catalogues using essentially identical techniques. We confirm that the 2dFGRS exhibits relatively more large scale power than the SDSS, or, equivalently, SDSS has more small scale power. We demonstrate that this difference is due to the r-band selected SDSS catalogue being dominated by more strongly clustered red galaxies, which have a stronger scale dependent bias. The power spectra of galaxies of the same rest frame colours from the two surveys match well. If not accounted for, the difference between the SDSS and 2dFGRS power spectra causes a bias in the obtained constraints on cosmological parameters which is larger than the uncertainty with which they are determined. We also found that the correction developed by Cole et al. (2005) to model the distortion in the shape of the power spectrum due to non-linear evolution and scale dependent bias is not able to reconcile the constraints obtained from the 2dFGRS and SDSS power spectra. Intriguingly, the model is able to describe the differences between the 2dFGRS and the much more strongly clustered LRG sample, which exhibits greater nonlinearities. This shows that more work is needed to understand the relation between the galaxy power spectrum and the linear perturbation theory prediction for the power spectrum of matter fluctuations. It is therefore important to accurately model these effects to get precise estimates of cosmological parameters from these power spectra and from future galaxy surveys like Pan-STARRS, or the Dark Energy Survey, which will use selection criteria similar to the one of SDSS.

Key words: cosmological parameters, large scale structure of the universe

1 INTRODUCTION

In the last decade, the advent of new precise cosmological observations coming mainly from measurements of fluctuations in the temperature of the cosmic microwave background radiation (CMB) and the large scale structure of the Universe (LSS), have shown a dramatic improvement. This has revolutionized our knowledge of the values of the basic cosmological parameters which are constrained to an accuracy of around 10% (Sánchez et al. 2006; Spergel et al. 2007).

In particular, the measurements of large scale galaxy clustering place important constraints on cosmological parameters that complement those from the analysis of fluctuations in the CMB. Measurements of the galaxy power spectrum constrain the parameter combinations $\Omega_{\rm m}h$ and $\Omega_{\rm b}/\Omega_{\rm m}$. The constraint on $\Omega_{\rm m}h$ is particularly important as, for instance, it breaks a degeneracy in the CMB data and allows accurate determination of $\Omega_{\rm m}$ (e.g. Efstathiou et al. 2002).

In view of these rapid improvements in the amount and quality of the observations the control of the systematic effects introduced in the analysis pipeline becomes increasingly important. New experiments and surveys are being planned which will push the level of precision of the constraints on cosmological parameters even further. Thus it is necessary to establish how robust the constraints really are with respect to the details of the hypothesis implemented to establish the link between theory and observations.

The tension found between the constraints coming from the published analysis of the two-degree field galaxy redshift survey (2dFGRS) of Cole et al. (2005) and the Sloan Dig-

⋆ E-mail: arielsan@oac.uncor.edu

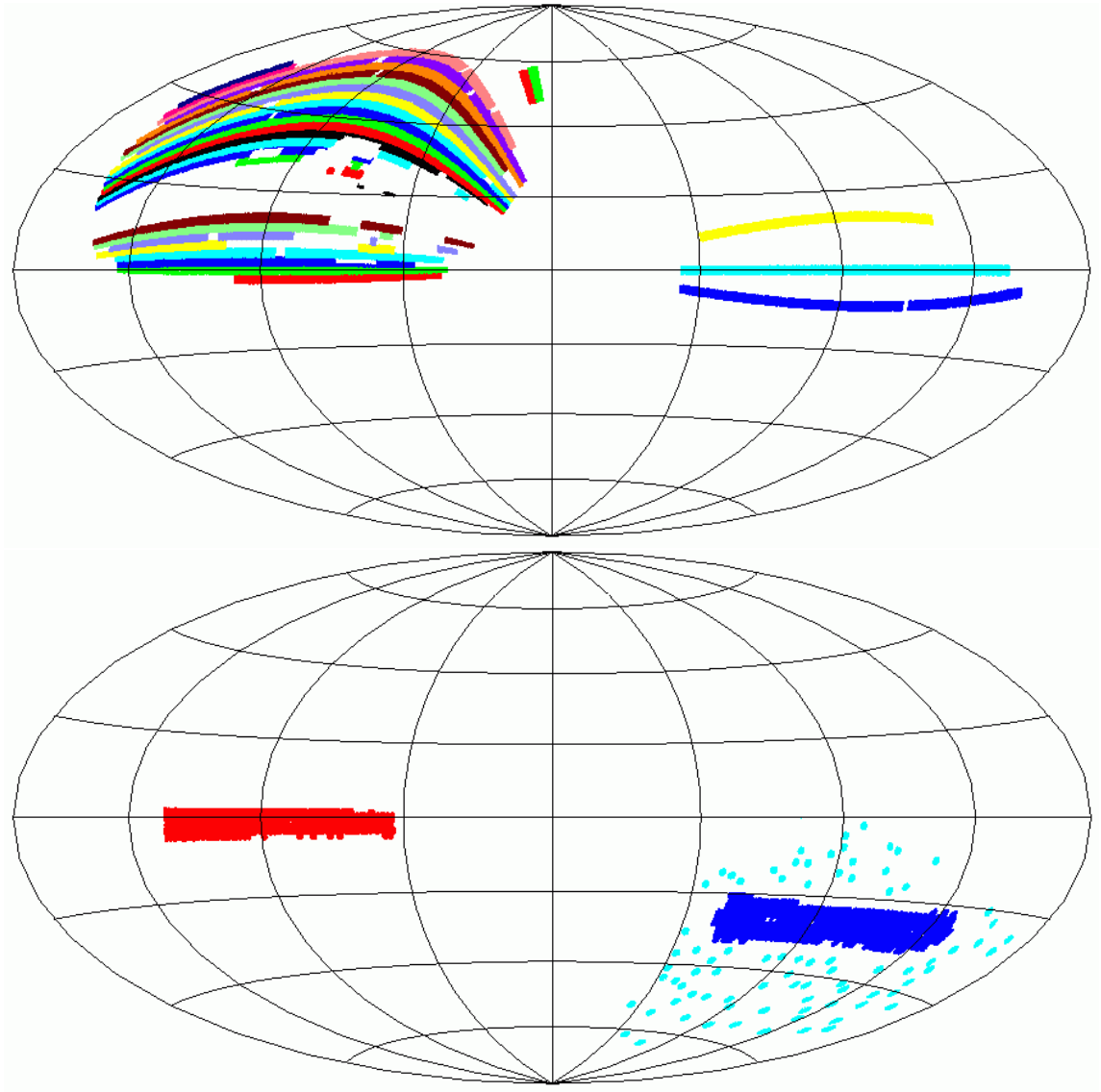

Figure 1. Equal area, all sky, Aitoff projections of all the galaxies accepted by our SDSS (upper) and 2dFGRS (lower) masks used in our analysis.

ital Sky Survey (SDSS) of Tegmark et al. (2004) might be an indication of possible systematic effects. In figure 16a of Cole et al. (2005), which compares the published estimates of the galaxy power spectra, there is evidence of more large scale power in the 2dFGRS than in SDSS. This folds through and results in the headline value of $\Omega_{\rm m}h = 0.168 \pm 0.016$ from 2dFGRS (Cole et al. 2005) being lower than that of Tegmark et al. (2004) SDSS, $\Omega_{\rm m}h = 0.213 \pm 0.023$. Here, one should be cautious as different priors have been assumed, but in the analysis of Sánchez et al. (2006), which treats each dataset on an equal footing and separately combines each with CMB data, one sees in their figure 18 that the SDSS prefers a substantially higher value of $\Omega_{\rm m}$ than does the 2dFGRS. In fact, while the 2dFGRS estimate is in good agreement with that from the CMB data alone (it significantly tightens the constraint without shifting the best fitting value) the SDSS data pull $\Omega_{\rm m}$ to higher values than preferred by the CMB.

The 2dFGRS and SDSS are the largest redshift surveys available and give the most detailed description of the large scale structure of the Universe as traced by galaxies. Hence to be sure that systematic errors are not biasing the parameters it is very useful to see each data set subjected to analysis by a variety of algorithms and codes, as happened following the Percival et al. (2001) 2dFGRS power-spectrum analysis (Tegmark et al. 2001).

Percival et al. (2007a) computed the power spectrum of a combined Main galaxy and Luminous Red Galaxy (LRG) sample drawn from SDSS Data Release 5 using a similar technique to that of Cole et al. (2005) and found an even higher value of the matter density $\Omega_m h = 0.23 \pm 0.01$, suggesting that the inclusion of the LRGs in the analysis increase the discrepancy between 2dFGRS and SDSS. Their results also show that if the analysis is restricted to large scales ($0.01 < k < 0.06\,h$ Mpc^{-1}) the data favours a lower matter density $\Omega_m h = 0.16 \pm 0.03$. Percival et al. (2007a)

suggested that these differences could be explained by scale-dependent galaxy bias on large scales, but found no significant evidence of it.

Here we seek to investigate whether these differences are as a result of: a) larger than expected cosmic variance, b) systematics due to differences in the analysis technique c) systematics due to problems with galaxy catalogues or d) intrinsic differences in the underlying galaxy clustering. In order to directly compare the 2dFGRS and SDSS, we analyse each dataset using essentially identical methods which we outline in Section 2. In Section 3, we briefly look at the region of overlap between the surveys to study the different selections used and the level of incompleteness. In Section 4, we compare the resulting power spectra from the full catalogues and interpret the differences. Section 5 presents a detailed analysis of the implications of the differences between these datasets in the obtained constraints on cosmological parameters, and the systematic effects that might be introduced by the analysis. Finally, we conclude in Section 6

2 METHODS

2.1 2dFGRS analysis

The 2dFGRS covers approximately 1800 square degrees distributed between two broad strips, one across the South Galactic Pole (SGP) and the other close to the North Galactic Pole (NGP), plus a set of 99 random 2 degree fields spread over the full southern galactic cap. The final catalogue contains reliable redshifts for 221 414 galaxies selected to an extinction-corrected magnitude limit of approximately $b_{\rm J} = 19.45$ (Colless et al. 2001, 2003). The 2dFGRS samples analysed here are the same as described in detail in Cole et al. (2005).

Our method of estimating the galaxy power spectrum and determining statistical errors is essentially identical to that set out in Cole et al. (2005), but with two minor changes. In brief:

We use masks, whose construction is described in Norberg et al. (2002), to describe the angular variation of the survey magnitude limit, redshift completeness and magnitude dependence of the redshift completeness.

Random catalogues are generated by sampling from the luminosity function and viewing through the masks. To generate random catalogues corresponding to red/blue subsets, the luminosity function of only the red/blue galaxies is used.

Close pair incompleteness due to "fibre collisions" is dealt with by redistributing the weights of missed galaxies to their 10 nearest neighbours on the sky.

The power spectrum is estimated using a simple cubic FFT method with the optimal weighting scheme of Percival, Verde & Peacock (2004, hereafter PVP) and then spherically averaged in redshift space. As in Cole et al. (2005), the assumed linear empirical bias factors that are used in this weighting scheme are

$$b_{\rm blue} = 0.9\,(0.85 + 0.15\,L/L_*), \quad (1)$$

for rest frame $b_{\rm J} - r_{\rm F} < 1.07$, and

$$b_{\rm red} = 1.3\,(0.85 + 0.15\,L/L_*), \quad (2)$$

for rest frame $b_{\rm J} - r_{\rm F} > 1.07$. These are based on the scale-independent bias parameter found by Norberg et al. (2001), modified to describe the difference in amplitudes of the power spectra of red and blue galaxies around $k = 0.1\,h\,{\rm Mpc}^{-1}$. This simple description of the luminosity and colour dependence of galaxy bias is in agreement with the results of Swanson et al. (2007).

Cole et al. (2005) showed that the effect on the recovered power spectrum of using Feldman, Kaiser & Peacock (1994, hereafter FKP) rather than the PVP estimator is small (see their figures 17(s) and (t)). The FKP estimator is biased, as it ignores the luminosity dependence of galaxy clustering. Provided the model of luminosity-dependent bias is correct, then the PVP estimator removes this bias. A more complicated recipe for the luminosity and colour dependence of galaxy bias can be applied but, as long as the assumed bias is scale independent (a necessary assumption in our method of analysis) it will hardly change the shape of the estimated power spectrum, as the effect will be significantly smaller than the difference between the FKP and PVP estimates and so entirely negligible.

The covariance matrix describing the errors on the power spectrum measurements and their correlations is estimated using mock catalogues which are constructed from the random catalogues by generating a log-normal density field with a specified power spectrum and using it to modulate the selection of galaxies from the random catalogue. Thus, by construction, these catalogues have a power spectrum very close to the best fitting model and have luminosity and colour dependent clustering consistent with the bias factors of equations (1) and (2).

The survey window function is determined directly from the random catalogue. When fitting models the theoretical model power spectra are convolved with the survey window function.

The two minor changes we have made are:

- We have adopted a simpler binning scheme so that now $P(k)$ is estimated in bins uniformly spaced in $\log_{10} k$, rather than the linearly space bins with different bin widths in different ranges of k that were used in Cole et al. (2005).
- We used new sets of log-normal catalogues in which the modulation of the density field used for galaxies with different bias factors is linear rather than the slightly more complicated scheme that was employed by Cole et al. (2005).

2.2 SDSS analysis

We have analysed the SDSS Data Release 5 (DR5) sample (Adelman-McCarthy et al. 2007), which is considerably larger than the DR3 sample analysed by Tegmark et al. (2004). In most respects our analysis of the SDSS is identical to that of the 2dFGRS. The only differences are a simpler way of generating the survey masks and of populating the corresponding random catalogue.

The sky coverage mask we have adopted for the SDSS-DR5 data is shown in Fig. 1 and compared with corresponding 2dFGRS mask. We constructed this mask by simply noting the angular coverage of each of the stripes from which the SDSS survey is built and by removing a few small regions with poor coverage. Most of the SDSS survey goes to a uniform magnitude limit of $r = 17.77$, but a sub-area, which is easily identified using the target selection date, has

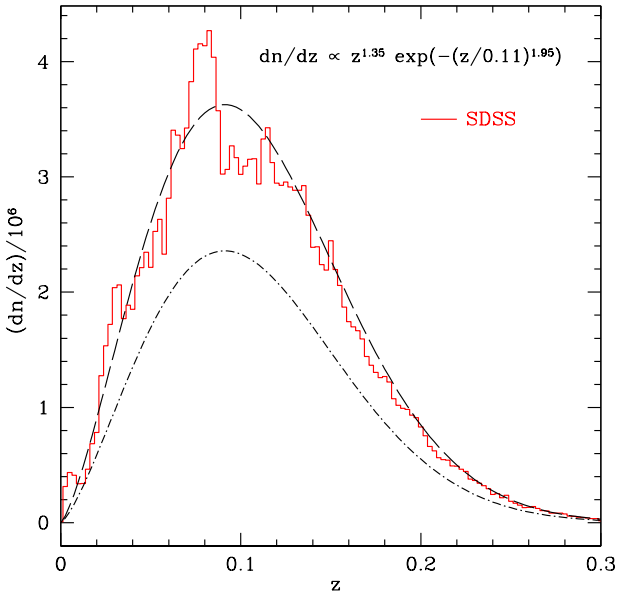

Figure 2. The redshift distribution of the SDSS $r < 17.77$ galaxy sample (solid line). The dashed line shows an analytic fit to the data. The dot-dashed line is this same fit scaled down in amplitude to be always below the redshift histogram. This fit is used to generate the corresponding random catalogue by sampling from the original survey as described in the text.

a variety of different magnitude limits. In this sub-area we simply imposed a fixed magnitude limit of $r = 17.5$ and discarded all galaxies fainter than this limit. The number of galaxies with redshifts that are retained by the mask and magnitude limits and used in our analysis is 443 424. The mask is cruder than the more sophisticated ones employed by Tegmark et al. (2004) and for the 2dFGRS as it ignores the smaller scale variation in the redshift completeness. However in the case of the 2dFGRS, where the incompleteness variation is much larger, Cole et al. (2005) showed that provided this incompleteness is accounted for using our method of redistributing the weights of galaxies without redshifts to neighbours with redshifts, the resulting power spectrum estimates are very accurate (see their figure 17g).

As most of the catalogue goes uniformly to the deeper magnitude limit a simple method can be used to construct the random catalogue. Fig. 2 shows the redshift distribution of this sample (solid line) together with an analytic fit (dashed line) that smoothes away the effect of large scale clustering. Our procedure is:

(i) Select a random direction on the sky.
(ii) Choose at random a genuine galaxy from the region of the survey that goes to $r = 17.77$.
(iii) Keep the galaxy with a probability proportional to the ratio of the fit scaled down in amplitude to be always below the redshift histogram (shown as the dot-dashed line in Fig. 2) to that of the actual height of the redshift histogram at the redshift of the selected galaxy. This effectively adjusts the sampling of the different redshifts to erase the effect of large scale clustering.
(iv) Keep or discard the galaxy based on the sky coverage and magnitude limit of the mask. (Note, for galaxies that fall where the magnitude limit is only 17.5, the fainter galaxies will be discarded and the redshift distribution of the retained galaxies will be appropriately shallower than that of Fig. 2.)

These steps are done repeatedly until a random catalogue containing 100 times more galaxies than the genuine catalogue is built up. Selecting from the genuine catalogue in this way means we automatically have apparent magnitudes and colours for all the galaxies in the random catalogue and so can select sub-samples from it and weight its galaxies in just the same way as the genuine catalogue.

To utilize the Percival, Verde & Peacock (2004) optimum weighting we need to determine bias factors for the galaxies in the genuine, random and mock catalogues. We do this by first converting the SDSS magnitudes to the 2dFGRS b_J and r_F bands using

$$\begin{aligned} b_{\rm J} &= g + 0.15 + 0.13\,(g-r) \\ r_{\rm F} &= r - 0.13 \end{aligned} \qquad (3)$$

and the simple colour dependent k-corrections that were used for the 2dFGRS data (see section 3 of Cole et al. 2005). Then we are able to define the bias factors using equations (1) and (2) just as for the 2dFGRS data.

3 SURVEY OVERLAP

Our main focus is the comparison of the power spectra of the two surveys, but we first directly compare the two surveys in the region of their overlap to get a feel for the different selections used and the level of incompleteness.

In the northern galactic hemisphere there is a contiguous area of overlap between the two surveys, which runs for 74 degrees of RA and is for the most part 5.2 degrees wide in declination. If we select from the SDSS photometric catalogue all galaxies brighter than $b_{\rm J} = 20$ (we do not apply the $r \approx 17.77$ magnitude limit of the SDSS main galaxy survey, but we do apply all the other star-galaxy classification criteria used in that sample (see Strauss et al. 2002)), then in this area there are 53 382 galaxies that are in both catalogues. We find the fraction of SDSS galaxies which are also in the 2dFGRS to be constant at 89% as faint as $b_{\rm J} \approx 18.9$. Fainter than this, SDSS galaxies are missing from the 2dFGRS sample simply due to the (variable) magnitude limit of the 2dFGRS survey and its 0.15 magnitude random photometric errors. This finding is in perfect accord with the estimates made with the SDSS EDR (Stoughton et al. 2002) in Norberg et al. (2002). This 11% incompleteness has been investigated by Cross et al. (2004) as well as Norberg et al. (2002) and has been shown to be predominately due to incorrect star-galaxy classification. The star-galaxy classification parameters based on the APM photometry are noisy and this level of incompleteness is in line with what was expected (Maddox et al. 1990).

The issue here is whether this incompleteness has any influence on estimates of galaxy clustering. We can look at this directly by plotting cone plots (Fig. 3) of the galaxies the two catalogues have in common and those missed by the 2dFGRS. We plot only galaxies with $b_{\rm J} < 19$ to avoid issues with the 2dFGRS magnitude limit. We see that 91% of the SDSS galaxies are in the 2dFGRS. Comparing the two

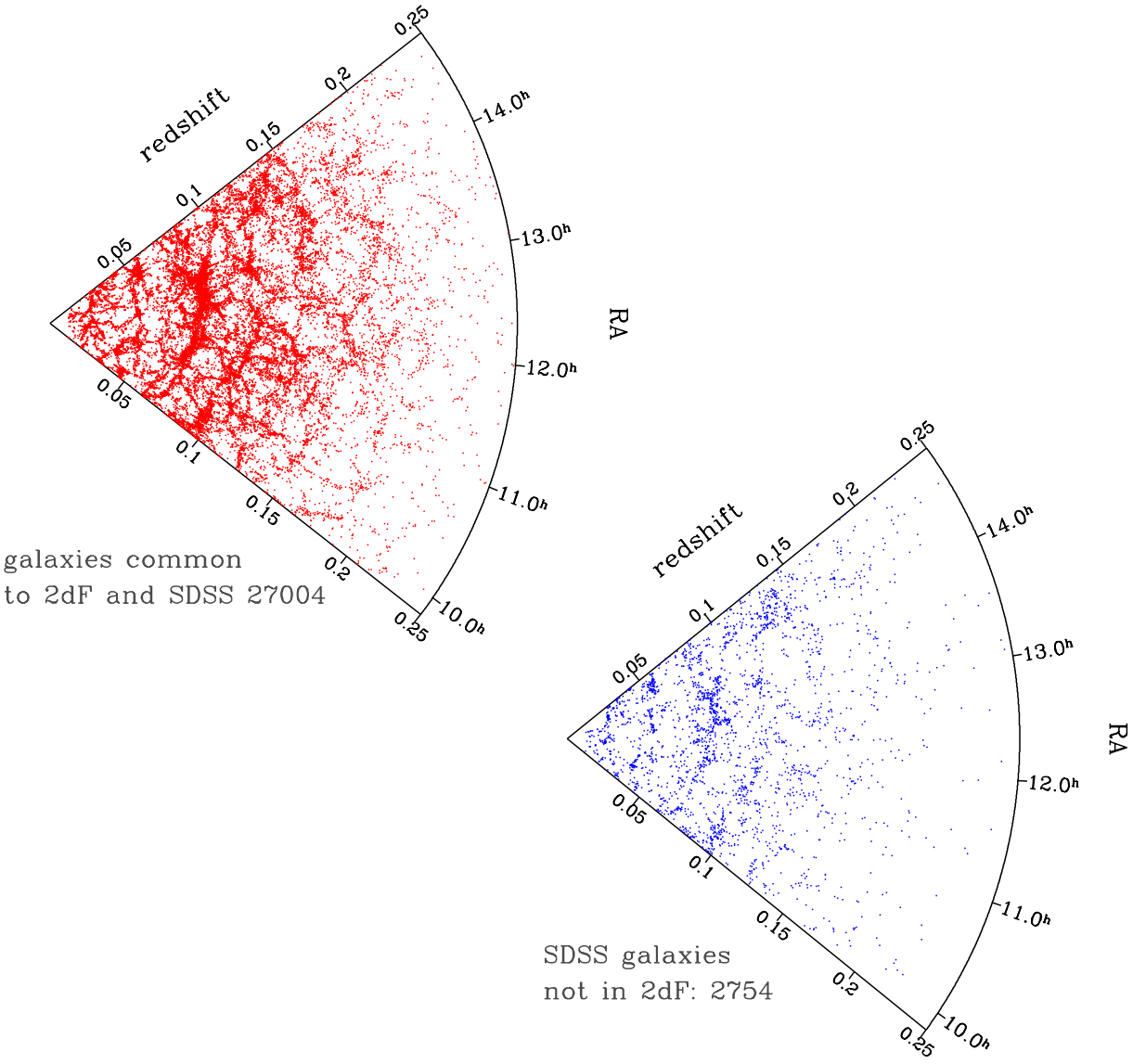

Figure 3. Cone plots showing RA and redshift for galaxies in the region of sky where the 2dFGRS and SDSS surveys overlap. There are 53 382 within this area that are in both surveys and in the upper panel we plot the sub-sample of 27 004 that (estimated from SDSS photometry) have $b_{\rm J} < 19$ and redshift $z > 0.01$. The lower panel shows the 2754 galaxies that are in SDSS and pass the same magnitude and redshift cuts, but are missing from the 2dFGRS catalogue.

cone plots in Fig. 3, it appears that the galaxies missed by 2dFGRS are just a random sparse sampling of the structure seen in the matching sample and so there is no evidence that the incompleteness has a spatial imprint.

4 COMPARISON OF POWER SPECTRA

The power spectra we estimate using the methods outlined in Section 2 are the direct transform of the data, and are thus what CMB researchers would term a pseudo-spectrum. As such, it yields a convolution of the underlying galaxy power spectrum with the modulus squared of the Fourier transform of the window function of either the SDSS or 2dFGRS as appropriate, that is

$$\hat{P}(\mathbf{k}) = P(\mathbf{k}) \otimes W^2(\mathbf{k}). \quad (4)$$

Our random catalogues allow us to accurately estimate $W^2(\mathbf{k})$ and from it determine the matrix of window functions that describe how our spherically averaged band power estimates $\hat{P}(k_j)$ are related to the unconvolved power spectrum $P(k_i)$.

$$\hat{P}(k_j) = \sum_i P(k_i) W_{\rm m}(k_i, k_j). \quad (5)$$

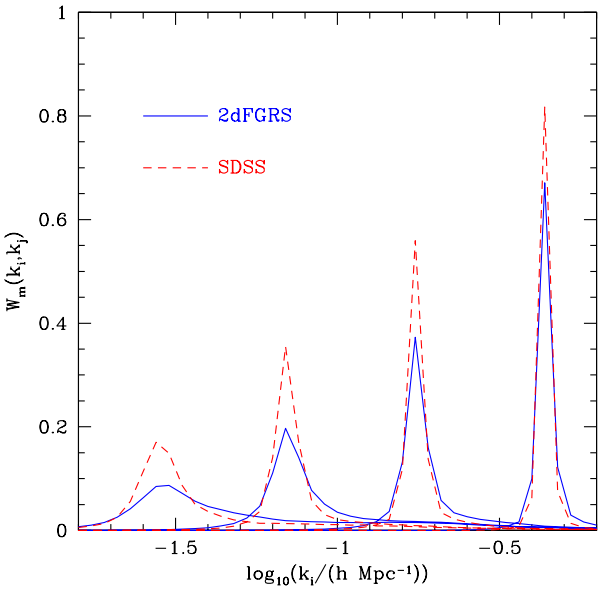

Figure 4. A sample of the normalized window functions for both SDSS (dashed) and 2dFGRS (solid). Each curve shows the relative contribution from the underlying power spectrum $P(k_i)$ as a function of k_i to our individual band power estimates $\hat{P}(k_j)$.

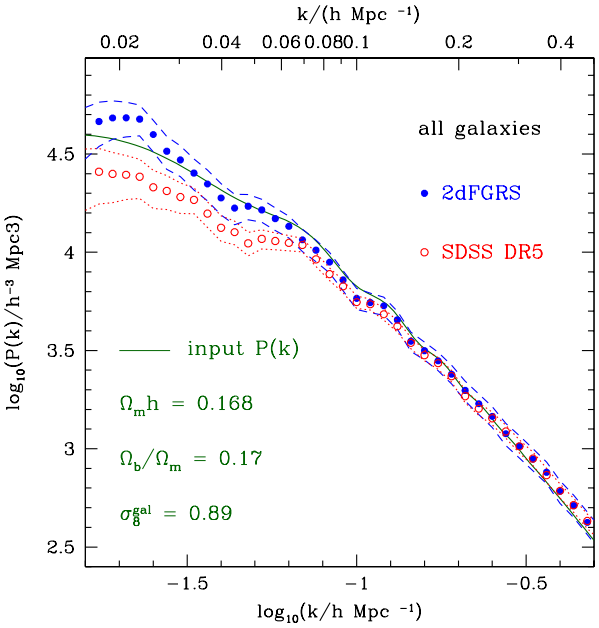

Figure 5. Comparison of the power spectra estimated from the full 2dFGRS and SDSS DR5 samples - corrected for the effect of the window function as described in the text. The solid line shows the input power spectrum of the mock catalogues used to estimate the covariance matrix of the measurements.

Examples of these window functions for the SDSS and 2dFGRS are shown in Fig. 4. For all our quantitative analysis in Section 5 we use these window functions to convolve the model power spectra before comparing with the data. However for the purposes of visually comparing the 2dFGRS and SDSS power spectra we have corrected the convolved estimates by multiplying them through by the ratio of a similar model power spectrum to its convolved counterpart. This 'deconvolution' is accurate provided the power spectra are smooth.

In Fig. 5 we compare the 'deconvolved' power spectra estimated from the full 2dFGRS catalogue and full SDSS-DR5 sample. We note that the SDSS and 2dFGRS galaxy power spectra agree well for $k > 0.07\,h\,\mathrm{Mpc}^{-1}$. The good agreement in amplitude at this wavenumber is due to the bias dependent weights used in the PVP estimator which have successfully modelled the difference in the clustering strength of the red selected SDSS galaxies and blue selected 2dFGRS galaxies. This is by design as the bias factors were normalized empirically by the 2dFGRS red and blue samples at this scale (see figure 15 of Cole et al. 2005). In contrast, on larger scales we see evidence for significantly more large scale power in the 2dFGRS than in SDSS. This clearly shows that the difference between the SDSS and 2dFGRS results that was noted in the introduction and which motivated this analysis is certainly significant and not an artifact of differing analysis techniques.

We now investigate if the discrepant shapes of the galaxy power spectra are due to the difference in the clustering properties of red and blue galaxies. Fig. 6 shows histograms of rest frame $b_\mathrm{J} - r_\mathrm{F}$ colours for both the 2dFGRS and SDSS catalogues. The SDSS magnitudes have been converted to these bands assuming the relations given in equation (3). The colour distributions are clearly bimodal with a natural dividing point at $b_\mathrm{J} - r_\mathrm{F} = 1.07$. The 2dFGRS has roughly equal numbers of red and blue galaxies while the SDSS, being red selected, is naturally dominated by red galaxies.

In Fig. 7 we compare the galaxy power spectra estimated from just the galaxies redder than $b_\mathrm{J} - r_\mathrm{F} = 1.07$ in both samples. Comparing the SDSS $P(k)$ from just the red galaxies with the previous estimate from the full SDSS catalogue reveals them to be in very close agreement. This is to be expected as the SDSS sample is both dominated by red galaxies and the PVP power spectrum estimator that we employ gives them more weight than their less clustered blue counterparts. In contrast, the estimate for just the red 2dFGRS galaxies differs from that from all the 2dFGRS catalogue. In fact it is a much closer match to the result from the SDSS data. The only places where the two estimates are not in excellent agreement is on the very largest scales $k < 0.025\,h\,\mathrm{Mpc}^{-1}$, where the estimates are both noisy and highly correlated, and also around $k \approx 0.05\,h\,\mathrm{Mpc}^{-1}$. In fact this difference is also due to sample variance. Cole et al. (2005) investigated the effect of removing from the 2dFGRS catalogue the two largest super clusters found by the analysis of Baugh et al. (2004). Their figure 17 (panels o and p) shows that this in general has a small effect, but does perturb the power just around $k \approx 0.05\,h\,\mathrm{Mpc}^{-1}$.

5 IMPLICATIONS FOR COSMOLOGICAL PARAMETERS

5.1 The shape of the power spectrum

The power spectrum measured for galaxies differs in a number of ways from the power spectrum for the mass predicted

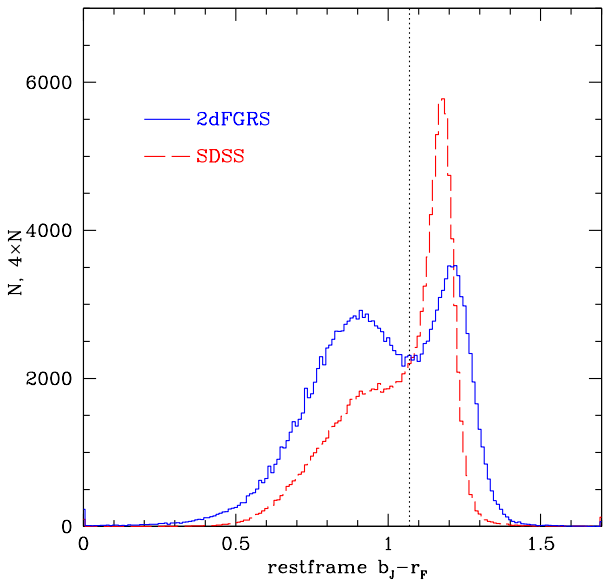

Figure 6. Histograms comparing the distribution of rest frame $b_{\rm J}-r_{\rm F}$ colours in the 2dFGRS (solid line) and SDSS (dashed line) catalogues. The distributions are clearly bimodal with a dividing point at $b_{\rm J}-r_{\rm F}=1.07$. The 2dFGRS has roughly equal numbers of red and blue galaxies while the SDSS is dominated by red galaxies. Note the different units on the y-axis.

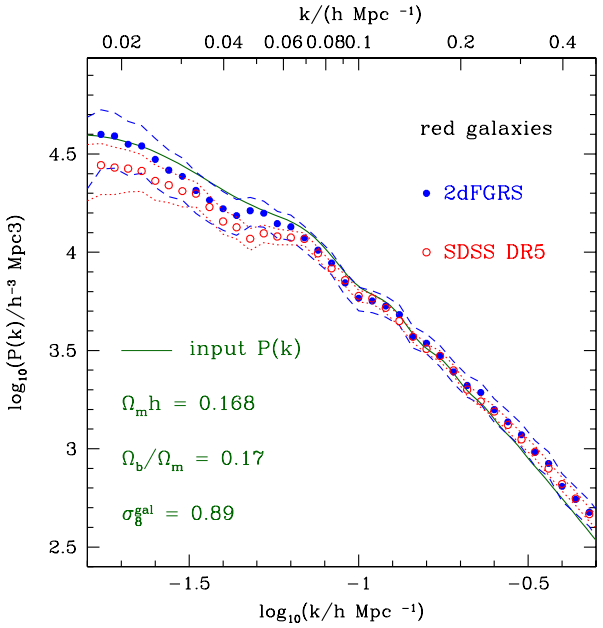

Figure 7. Comparison of the 2dFGRS and SDSS-DR5 power spectra from red subsamples that satisfy the rest frame colour $b_{\rm J}-r_{\rm F}>1.07$. The solid line shows the input power spectrum of the mock catalogues used to estimate the covariance matrix of the measurements.

in linear perturbation theory: (i) Nonlinear evolution of density perturbations leads to coupling between Fourier modes, changing the shape of the power spectrum. (ii) The galaxy power spectrum is distorted by the gravitationally induced peculiar motions of galaxies when a redshift is used to infer the distance to each galaxy. (iii) The power spectrum of the galaxies is a modified version of the power spectrum of the mass. This phenomenon is known as galaxy bias. As we have shown in the previous section these effects change with scale and depend on the galaxy type.

In order to constrain cosmological parameters, these effects need to be modelled. Cole et al. (2005) developed an empirical scheme to correct for these effects by applying a correction for non-linearity and scale-dependent bias to the shape of $P(k)$ of the form

$$P_{\rm gal}(k)=b^2\frac{1+Qk^2}{1+Ak}\,P_{\rm lin}(k), \qquad (6)$$

where $P_{\rm lin}(k)$ is the linear theory power spectrum and b is a constant bias factor. This formula was deduced by comparison with detailed numerical galaxy-formation models: these show that the value of $A=1.4$ is robust, but the exact value of Q depends on galaxy type and also has some uncertainty depending on how the modelling is done. These results were used to determine a range of allowed values for Q. For 2dFGRS the value $Q=4.6$ is preferred.

Although this correction factor was designed and tested for the redshift space power spectrum of 2dFGRS where the scale dependent correction is small, it has also been used to model the real space power spectra for redder and more luminous galaxy samples (Tegmark et al. 2004, 2006; Percival et al. 2007a,b). In these cases larger values of Q were required to reconcile the measured power spectrum with linear theory (e.g $Q=26$ in Percival et al. 2007b). Here we analyse the validity of this correction in general terms, testing if the different shapes of the power spectra for the full and red galaxy samples of 2dFGRS and SDSS can be accounted for by using equation (6) with different values of Q.

Besides the change in the overall shape of $P(k)$, it has been shown that non-linear evolution distorts the acoustic oscillations (Eisenstein et al. 2005; Springel et al. 2005; Eisenstein et al. 2006; Angulo et al. 2005, 2007). These oscillations are damped in a way that erases the higher harmonic peaks. Eisenstein et al. (2006) modelled the damping of the oscillations by computing a 'dewiggled' power spectrum by

$$P_{\rm dw}(k)=P_{\rm lin}(k)G(k)+P_{\rm nw}(k)(1-G(k)), \qquad (7)$$

where $P_{\rm nw}(k)$ is a smooth power spectrum, with the same shape as $P_{\rm lin}(k)$ but without baryonic oscillations, computed using the fitting formulas of Eisenstein & Hu (1999) and $G(k)\equiv\exp\left[-(k/\sqrt{2}k_\star)^2\right]$. This function regulates the transition from the large scales, where $P_{\rm dw}(k)$ follows linear theory to the small scales where the acoustic oscillations are completely damped. As described in Tegmark et al. (2006), the value of the damping scale $k_\star$ is a function of $\Omega_{\rm m}$ and the primordial amplitude of scalar fluctuations $A_{\rm s}$. Here we analyse the effect of this modelling on the obtained constraints on cosmological parameters. When this correction is applied, $P_{\rm dw}(k)$ is used in equation (6) instead of $P_{\rm lin}(k)$.

We use the data from the observed power spectra for $k<0.2\,h\,{\rm Mpc}^{-1}$ and discard measurements with $k<0.02\,h\,{\rm Mpc}^{-1}$ which could be affected by uncertainties in the mean density of galaxies within the survey. We compare the data against a restricted parameter space given by

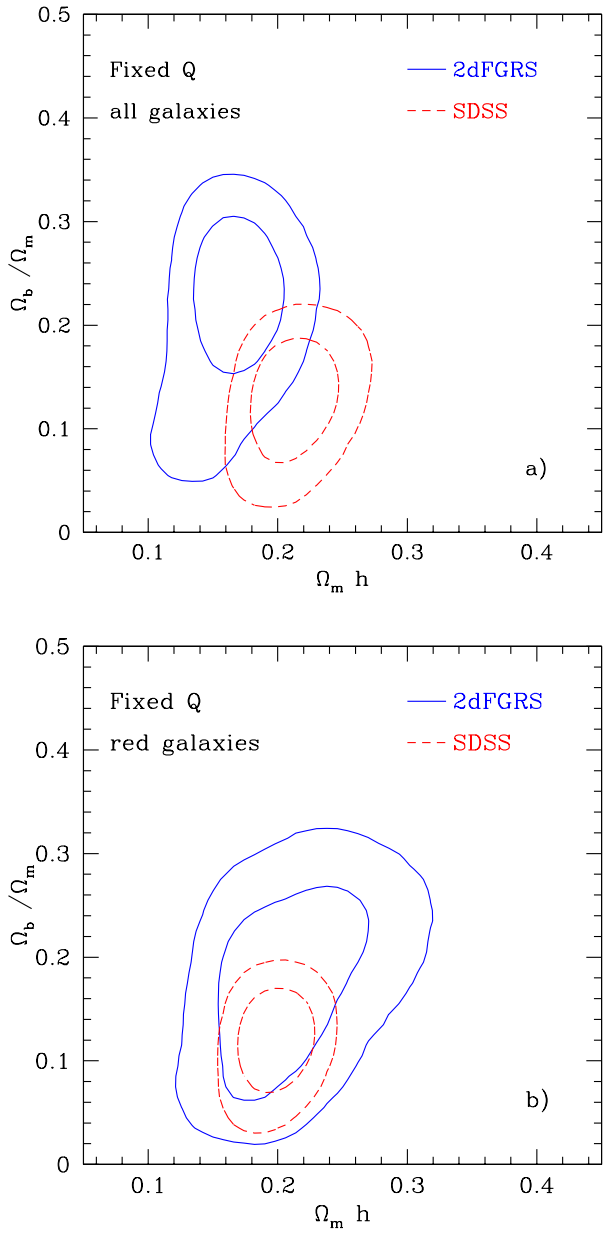

Figure 8. Contours show 68% and 95% joint confidence intervals for the baryon fraction $\Omega_{\rm b}/\Omega_{\rm m}$ and $\Omega_{\rm m}h$ for fits to the SDSS and 2dFGRS data in the range $0.02 < k < 0.2\,h\,{\rm Mpc}^{-1}$. The parameter Q modelling the distortion of power spectrum due to nonlinearity, redshift space distortions and scale dependent bias was kept fixed at $Q = 5$ in these fits.

$$\mathbf{P} \equiv (\Omega_{\rm m}h, \Omega_{\rm b}/\Omega_{\rm m}, Q, A_{\rm s}), \qquad (8)$$

and we fix the values of $h = 0.72$ and $n_{\rm s} = 1$. Eq. (8) gives the most important parameters to characterize the full shape of $P(k)$. When constraining the values of these parameters, we only use information from the shape of the galaxy power spectrum and not from its amplitude. This is why A_s only enters in the constraints marginally when the damping of the acoustic oscillations is included in the modelling through the assumed dependency of $k_\star$ on $\Omega_{\rm m}$ and A_s (Eisenstein et al. 2006; Tegmark et al. 2006).

Our results were generated with a modified version of the publicly available CosmoMC code (Lewis & Bridle 2002) as in Sánchez et al. (2006). CosmoMC uses the CAMB package to compute the linear power spectra for the matter fluctuations (Lewis et al. 2000). We have found that the use of the approximate Eisenstein & Hu (1998) fitting formula to model $P_{\rm lin}(k)$ causes a small shift in the recovered values of $\Omega_{\rm m}h$. Our analysis was carried out in parallel on the Cosmology Machine at Durham University. For each parameter set considered, we ran 10 separate chains using the Message Passing Interface (MPI) convergence criterion to stop the chains when the Gelman & Rubin (1992) statistic $R < 1.01$, which is a significantly more stringent criterion than is usually adopted (Verde et al. 2003; Seljak et al. 2005).

To test the efficiency of the modelling to account for the differences between the datasets we analyse three separate cases: first we fixed the parameters A and Q at fiducial values of $A = 1.4$ and $Q = 5$ and ignore the correction due to the damping of the acoustic oscillations (Section 5.2). In this way the difference in the shape of the power spectra for SDSS and 2dFGRS is completely translated into the recovered constraints on $\Omega_{\rm m}h$ and $\Omega_{\rm b}/\Omega_{\rm m}$. Second we extend the model allowing Q to vary over a wide prior (Section 5.3). In this way we can test if the ansatz of equation (6) can account for the observed differences. Finally, we include the effect of the damping of the acoustic oscillations into our modelling of $P(k)$ (Section 5.4). Table 1 summarizes the constraints obtained with the different datasets and parameter sets analysed.

5.2 Fixed Q

First we analyse the constraints on $\Omega_{\rm m}h$ and $\Omega_{\rm b}/\Omega_{\rm m}$ that are obtained by fitting the SDSS and 2dFGRS data with power spectra of the form described in the last section, with the parameters A and Q fixed at fiducial values of $A = 1.4$ and $Q = 5$ and ignoring the correction due to the damping of the acoustic oscillations. In this way these are the only two parameters defining the shape of the model $P(k)$ and the difference in shape of the fitted power spectra is completely characterized by the regions of the explored parameter space selected by each dataset. The parameter constraints obtained in this way are shown in Fig. 8 for the full (upper panel) and red (lower panel) 2dFGRS (solid lines) and SDSS (dashed lines) galaxy power spectra.

Fig. 8(a) shows that the 2dFGRS constraints are in complete agreement with those obtained by CMB data (Sánchez et al. 2006; Spergel et al. 2007). For 2dFGRS data only we obtain $\Omega_m = 0.246 \pm 0.030$ while the WMAP3 analysis of Spergel et al. (2007) gives $\Omega_m = 0.234 \pm 0.035$. The SDSS parameter estimates are completely in accord with those from Tegmark et al. (2004), but with much tighter bounds due to the larger SDSS dataset used here. This shows the consistency of the different analysis techniques applied to this dataset.

We note that the 2dFGRS and SDSS best fit values lie outside each others 95% confidence contours. This clearly shows that, if not accounted for, the difference between the SDSS and 2dFGRS power spectra that was noted in Section 4 causes a bias in the obtained constraints on cosmological parameters which is larger than the uncertainty with which they are determined.

Table 1. Marginalized 68% interval constraints on cosmological parameters obtained using the different datasets and parameter sets analysed.

Data set	Parameter	Fixed Q	Varying Q	Varying Q, dw
2dFGRS all	$\Omega_{\rm m}h$	$0.177^{+0.022}_{-0.025}$	$0.165^{+0.025}_{-0.029}$	$0.191^{+0.064}_{-0.051}$
	$\Omega_{\rm b}/\Omega_{\rm m}$	$0.195^{+0.054}_{-0.057}$	$0.212^{+0.061}_{-0.065}$	$0.250^{+0.081}_{-0.080}$
	Q	-	$7.7^{+3.3}_{-3.4}$	$8.1^{+3.7}_{-3.8}$
2dFGRS red	$\Omega_{\rm m}h$	$0.217^{+0.040}_{-0.040}$	$0.197^{+0.043}_{-0.043}$	$0.023^{+0.094}_{-0.069}$
	$\Omega_{\rm b}/\Omega_{\rm m}$	$0.183^{+0.064}_{-0.068}$	$0.197^{+0.071}_{-0.074}$	$0.238^{+0.086}_{-0.089}$
	Q	-	$8.0^{+4.4}_{-4.5}$	$8.1^{+4.8}_{-4.7}$
SDSS all	$\Omega_{\rm m}h$	$0.199^{+0.014}_{-0.014}$	$0.216^{+0.019}_{-0.020}$	$0.234^{+0.030}_{-0.034}$
	$\Omega_{\rm b}/\Omega_{\rm m}$	$0.128^{+0.029}_{-0.031}$	$0.112^{+0.019}_{-0.033}$	$0.150^{+0.050}_{-0.053}$
	Q	-	$2.8^{+1.5}_{-1.5}$	$3.1^{+1.6}_{-1.6}$
SDSS red	$\Omega_{\rm m}h$	$0.201^{+0.013}_{-0.015}$	$0.221^{+0.017}_{-0.025}$	$0.240^{+0.038}_{-0.039}$
	$\Omega_{\rm b}/\Omega_{\rm m}$	$0.120^{+0.027}_{-0.030}$	$0.115^{+0.031}_{-0.036}$	$0.157^{+0.057}_{-0.058}$
	Q	-	$3.2^{+1.6}_{-1.7}$	$3.5^{+1.7}_{-1.7}$
LRG	$\Omega_{\rm m}h$	-	$0.172^{+0.016}_{-0.016}$	$0.173^{+0.017}_{-0.017}$
	$\Omega_{\rm b}/\Omega_{\rm m}$	-	$0.173^{+0.038}_{-0.036}$	$0.187^{+0.038}_{-0.038}$
	Q	-	$27.1^{+4.7}_{-4.7}$	$28.5^{+4.9}_{-5.0}$

Fig. 8(b) confirms the agreement, within the expected statistical uncertainty, of the power spectra for red galaxies in 2dFGRS and SDSS. Being dominated by red galaxies, the SDSS constraints are almost identical to the ones obtained with the full sample. Instead, the results for 2dFGRS shift towards the same region of the parameter space preferred by SDSS to contain it completely. The best fitting parameters for 2dFGRS lie within the SDSS 65% confidence contour and vice versa. The contours obtained for 2dFGRS broaden as a result of the larger errors in the power spectrum obtained with a smaller sample of galaxies. This shows that the differences in the constraints obtained with SDSS and 2dFGRS are due to the different selection criteria used to construct the samples, with the r-band selected SDSS catalogue being dominated by more strongly clustered red galaxies which are more strongly affected by scale dependent bias.

5.3 Varying Q

In this section we extend the parameter space by allowing Q to vary over a wide prior. This allows us to test if the shape distortion modelled by equation (6) can account for the differences between 2dFGRS and SDSS found in Section 4. The results obtained in this way are shown in Fig. 9 for the full (upper panels) and red (lower panels) 2dFGRS (solid lines) and SDSS (dashed lines) galaxy power spectra. For this case, we have also included in our analysis the power spectrum of the SDSS Luminous Red Galaxies (LRG, dotted lines in Fig. 9) as measured by Tegmark et al. (2006).

From Figs. 9(a) and (b) we note that the correction factor of equation (6) is not able to reconcile the constraints obtained from the 2dFGRS and SDSS power spectra. The constraints on the $\Omega_{\rm m}h-\Omega_{\rm b}/\Omega_{\rm m}$ plane exhibit a similar behaviour to the case where the value of Q was held fixed. This shows that the non-linearities and scale dependent bias of the r-selected SDSS galaxies distort the shape of the power spectrum in a way that can not be described accurately by equation (6). Contrary to what was expected, the constraints on the plane $\Omega_{\rm m}h-Q$ show that SDSS and 2dFGRS select different values of $\Omega_{\rm m}h$ but a similar allowed region for Q. Then, even incorporating this model of the distortion into the methodology for constraining cosmological parameters, the use of a red sample of galaxies introduces strong systematics effects on the resulting parameter constraints.

Intriguingly, the constraints on $\Omega_{\rm m}h$ and $\Omega_{\rm b}/\Omega_{\rm m}$ obtained using the LRG power spectrum lie well within the 2dFGRS contours, in perfect agreement with the results obtained using CMB data. This means that the difference in the shape of the SDSS-LRG and 2dFGRS power spectra is completely consistent with equation (6). Both datasets can be accurately described with the same values of the cosmological parameters, but with a much higher value of $Q=27.1\pm4.7$ for the SDSS-LRG sample, in agreement with the results of Percival et al. (2007b).

As shown in Fig. 9(c) and (d), the situation is similar for the constraints obtained using the power spectra of the red 2dFGRS and SDSS galaxy samples. As in the case of a fixed value of Q, the constraints obtained with the 2dFGRS power spectrum are shifted from its region of agreement with the CMB towards the values preferred by the SDSS. With

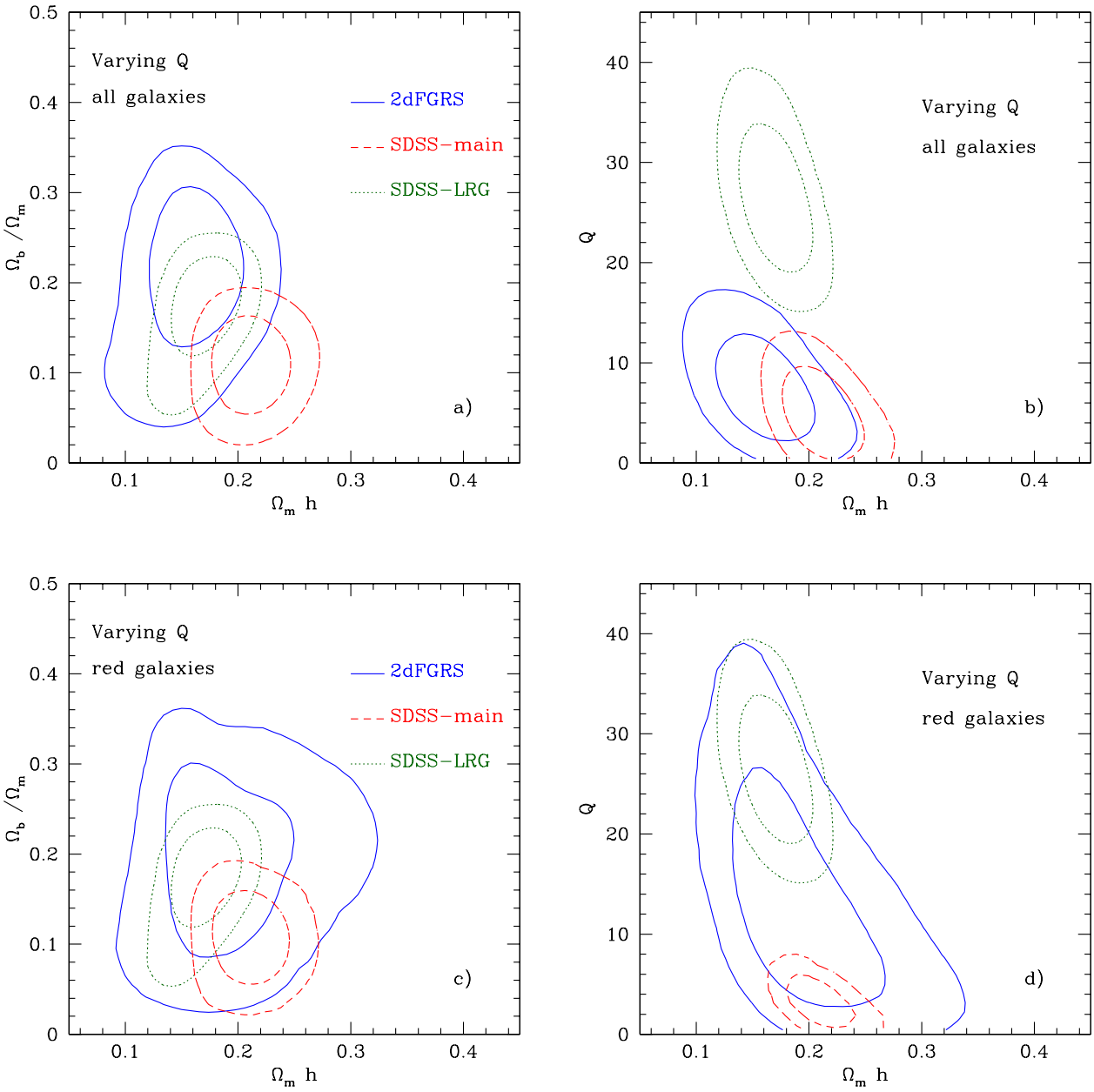

Figure 9. Contours show 68% and 95% joint confidence intervals for the baryon fraction $\Omega_{\rm b}/\Omega_{\rm m}$ and $\Omega_{\rm m}h$ for fits to the 2dFGRS (solid line), SDSS (dashed line) and SDSS-LRG (dotted line) in the range $0.02 < k < 0.2\,h\,{\rm Mpc}^{-1}$. In these fits, the parameter Q modelling the distortion of power spectrum due to nonlinearity, redshift space distortions and scale dependent bias was allowed to vary over a wide prior $(0-45)$.

the addition of Q, the extra degree of freedom gives rise to a degeneracy that involves $\Omega_{\rm m}h$ and Q. These two parameters act in opposite ways which allows them to compensate each others effect on the theoretical power spectrum. This degeneracy can not be broken by the power spectrum of the red 2dFGRS galaxies and results in a considerable increase of the allowed region for the different parameters.

The breakdown of the simple model of equation (6) in describing the differences in the clustering of red and blue galaxies can be more clearly seen in Fig. 10, which shows the different power spectra analysed in this work divided by a reference power spectrum $P_{\rm ref}(k)$ computed using the best fit cosmological parameters from Spergel et al. (2007). Assuming this is the true cosmology, if there were no distortions in the shape of $P(k)$, the different measurements should lie over the dashed line of $P(k)/P_{\rm ref}(k) = 1$. The difference in the shapes of these spectra is a clear demonstration of the fact that scale dependent bias and nonlinearities are a function of galaxy type. The solid lines in Fig. 10 show the shape distortion model of equation (6) for the specified

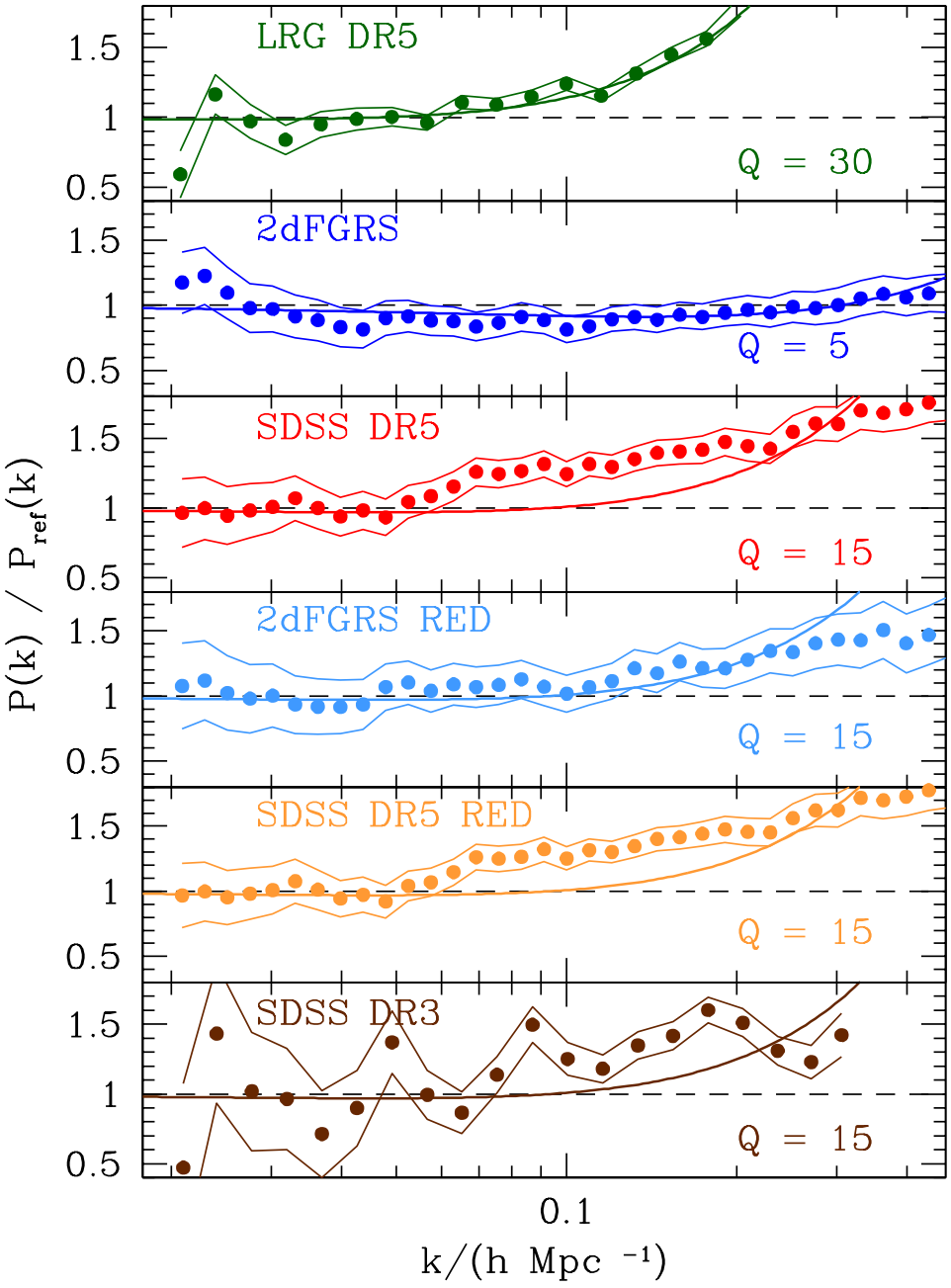

Figure 10. The power spectra analysed in this work divided by a reference power spectrum computed using the best fit cosmological parameters from Spergel et al. (2007). The solid lines show the shape distortion modelled by equation (6) for the specified value of Q.

value of Q. With a value of $Q \approx 5$ this correction gives a good description of the distortions in the power spectrum of the 2dFGRS, with a small decrease in the amplitude of $P(k)$ at intermediate scales and an increase at small scales. The same correction factor can reproduce the effect of the larger non-linearities of the LRG power spectrum, but with a much higher value of $Q \approx 30$.

Despite its success at describing the 2dFGRS and LRG power spectra, equation (6) fails to reproduce the shape of the SDSS power spectrum, which exhibits a different behavior as a function of k and can not be fitted by this simple model. The fact that the power spectrum of the red galaxies from 2dFGRS is distorted in a similar fashion indicates that, as was found in Section 4, red selected galaxy samples exhibit stronger scale dependent galaxy bias. For comparison, Fig. 10 also shows that the power spectrum of SDSS DR3

measured by Tegmark et al. (2004) suffers from the same effect.

This clearly shows that, in order to obtain unbiased constraints on cosmological parameters it is necessary to better understand and constrain the distortion in the shape of the power spectrum caused by nonlinearity and scale dependent bias.

5.4 Damping of acoustic oscillations

In this section we analyse the effect of including the damping of the acoustic oscillations into our modelling of $P(k)$. This case mimics the analysis of Tegmark et al. (2006) of the LRG power spectrum. Although we only use information from the shape of the galaxy power spectrum, when this effect is included in our modelling of $P(k)$ the amplitude A_s enters in the constraints marginally, since the value of the damping scale $k_\star$ depends on this parameter (Eisenstein et al. 2006; Tegmark et al. 2006).

As can be seen in the last column of Table 1 this addition does not change the differences between 2dFGRS and SDSS. The main effect of this is the increase of the allowed regions of the different cosmological parameters. As the oscillations are damped, models with higher values of $\Omega_{\rm b}/\Omega_{\rm m}$, which according to linear theory exhibit larger amplitude oscillations, are now compatible with the data. This causes the signal of the oscillations to effectively loose the power to constrain efficiently the values of the parameters, which in this case are determined almost completely by the overall shape of $P(k)$.

The details of the way in which the oscillations are damped and the correct way to model it are the main focus of several recent studies (Eisenstein et al. 2006; Angulo et al. 2007; Crocce & Scoccimarro 2007). Although substantial progress has been made in this subject, it is not completely understood. This is then another example of how the details of the modelling used to fit theoretical models may introduce systematic effects which may cause significant differences in the obtained constraints.

6 CONCLUSIONS

In this paper we studied in detail the differences between the constraints on cosmological parameters coming from the 2dFGRS and SDSS power spectra. We have analysed both the 2dFGRS and SDSS DR5 catalogues using essentially identical techniques. The main conclusion of this investigation is that a significant difference exists between the shape of the galaxy power spectra measured in the 2dFGRS and SDSS main galaxy sample and that this difference is due to scale dependent bias. Our results extend those of Percival et al. (2007a) who found that the inclusion of the LRGs in the analysis increases the discrepancy between 2dFGRS and SDSS. Their findings also showed that if linear CDM fits are restricted to large scales ($0.01 < k < 0.06\,h$ Mpc^{-1}), SDSS data also favours a lower matter density $\Omega_m h = 0.16 \pm 0.03$. Percival et al. (2007a) suggested that a scale-dependent bias which depends on the r-band luminosity of the galaxies could explain these differences, but found no significant evidence of it.

If a homogeneous sample of red galaxies is selected from each survey then the resulting power spectra agree to within the expected statistical errors. In contrast, when the full 2dFGRS and SDSS catalogues are analyzed the resulting 2dFGRS power spectrum differs in shape to that of SDSS. If normalized on scales around $k \approx 0.1\,h\,{\rm Mpc}^{-1}$, as is done automatically by the scale independent bias factors assumed in the PVP estimator, then the 2dFGRS $P(k)$ exhibits more large scale power than SDSS. However, if one instead normalizes on large scales one finds the equivalent result that SDSS exhibits more small scale power than 2dFGRS. This behaviour is exactly what one expects if the more strongly clustered red galaxies live in denser environments where the effects of nonlinearity are greater.

We have obtained constraints on the parameter combinations $\Omega_{\rm m}h$ and $\Omega_{\rm b}/\Omega_{\rm m}$ which are the most important parameters to characterize the full shape of $P(k)$. The constraints obtained by fitting the SDSS and 2dFGRS data show that the difference between the SDSS and 2dFGRS results which motivated this analysis is significant and is not due to the differing analysis techniques. If not accounted for, the difference between these datasets introduces a systematic effect that bias the obtained constraints on cosmological parameters, whose effect is larger than the uncertainty with which they are determined.

Cole et al. (2005) introduced a first attempt at modelling the distortion in the shape of the power spectrum caused by nonlinearity and scale dependent bias using the Q and A parameters of equation (6). For the mix of red and blue galaxies present in the 2dFGRS survey, the necessary value of Q was reasonably small and hence the scale dependent correction quite modest. The same correction factor has been applied to samples of redder or more luminous and hence more clustered galaxies where one expects greater nonlinearity.

We found that the correction factor of equation (6) is not able to reconcile the constraints obtained from the 2dFGRS and SDSS power spectra. This shows that this simple correction can not describe the effect of non-linearities and scale dependent bias for a general sample of galaxies. Intriguingly, the same model can describe correctly the difference in the shape of the power spectra from 2dFGRS and the much strongly clustered SDSS-LRG sample. Both datasets can be described with the same values of the cosmological parameters, but with a much higher value of $Q = 27.1 \pm 4.7$ for the SDSS-LRG sample.

This comparison has revealed that to get unbiased estimates of the cosmological parameters it is necessary to better understand and constrain the different processes that shape the galaxy power spectrum. Thus to get robust constraints from the main SDSS survey and from future galaxy surveys like Pan-STARRS, or the Dark Energy Survey, which will use selection criteria similar to the one of SDSS, will require more detailed modelling of nonlinearity and scale dependent bias.

ACKNOWLEDGEMENTS

We thank Steve Wilkins for his assistance in preparing preliminary versions of many of our plots. We thank Carlton Baugh for his careful reading of the manuscript and useful discusions. AGS acknowledges the hospitality of the Depart-

ment of Physics at the University of Durham where part of this work was carried out. AGS acknowledges a fellowship from CONICET; This work was supported by PPARC, the European Commission's ALFA-II programme through its funding of the Latin-american European Network for Astrophysics and Cosmology (LENAC), and the Royal Society, through the award of an International Incoming Short Visit grant.

REFERENCES

Adelman-McCarthy J.K. et al., 2007, arXiv:0707.3380
Angulo R., Baugh C.M., Frenk C.S., Bower R.G., Jenkins A., Morris S.L., 2005, MNRAS, 362, L25
Angulo R., Baugh C.M., Frenk C.S., Lacey C.G., 2007, arXiv:astro-ph/0702543
Baugh C.M. et al. (The 2dFGRS Team), 2004, MNRAS, 351, L44
Cole S. et al. (The 2dFGRS Team), 2005, MNRAS, 362, 505
Colless M. et al. (The 2dFGRS Team), 2001, MNRAS, 328, 1039
Colless M. et al. (The 2dFGRS Team), 2003, arXiv:astro-ph/0306581
Crocce M., Scoccimarro R., 2007, arXiv:0704.2783
Cross, N. J. G., Driver, S. P., Liske, J., Lemon, D. J., Peacock, J. A., Cole, S., Norberg, P., Sutherland, W. J., 2004, MNRAS, 349, 576
Efstathiou G. et al., 2002, MNRAS, 330, 29
Eisenstein D.J., Hu, W., 1998, ApJ. 496, 605
Eisenstein D.J., Hu, W., 1999, ApJ. 511, 5
Eisenstein D.J. et al., 2001, AJ, 122, 2267
Eisenstein D.J. et al., 2005, ApJ. 633, 560
Eisenstein D.J., Seo H., Sirko E., Spergel D., 2006, arXiv:astro-ph/0604362
Feldman H. A., Kaiser N., Peacock J. A., 1994, MNRAS, 426, 23
Gelman A., Rubin D., 1992, Stat. Sci., 7, 457
Lewis A., Bridle, S., 2002, Phys. Rev. D, 66, 103511
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Maddox, S. J. Efstathiou, G., Sutherland, W. J., Loveday, J., 1990, MNRAS, 243, 692
Norberg P. et al. (The 2dFGRS Team), 2001, MNRAS, 328, 64
Norberg P. et al. (The 2dFGRS Team), 2002, MNRAS, 336, 907
Percival W.J. et al. (The 2dFGRS Team), 2001, MNRAS, 327, 1297
Percival W.J., Verde L., Peacock J.A., 2004, MNRAS, 347, 645
Percival W.J. et al., 2007a, ApJ, 657, 51
Percival W.J. et al., 2007b, ApJ, 657, 645
Sánchez A.G. et al., 2006, MNRAS, 366, 189
Seljak U. et al., 2005, Phys. Rev. D, 71, 103515
Spergel D. et al., 2007, ApJS, 170, 377
Sringel V. et al., 2005, Nature, 435, 629
Stoughton C. et al., 2002, AJ, 123, 485
Strauss, M. et al., 2002, AJ, 124, 1810
Swanson M.E.C., Tegmark M., Blanton M., Zehavi I., 2007, arXiv:astro-ph/0702584
Tegmark M., Zaldarriaga M., Hamilton A.J.S., 2001, MNRAS,335, 887
Tegmark M. et al., 2004, ApJ, 606, 702
Tegmark M. et al., 2006 Phys Rev. D, 74, 123507
Verde L. et al., 2003, ApJS, 148, 195