Mon. Not. R. Astron. Soc. **335,** 887–908 (2002)

The power spectrum of galaxies in the 2dF 100k redshift survey

Max Tegmark,[1]⋆† Andrew J. S. Hamilton[2]⋆† and Yongzhong Xu[1]
[1]*Department of Physics, University of Pennsylvania, Philadelphia, PA 19104, USA*
[2]*JILA and Department Astrophysical & Planetary Sciences, Box 440, Univ. Colorado, Boulder, CO 80309, USA*

Accepted 2002 April 26. Received 2002 April 14; in original form 2001 December 7

ABSTRACT

We compute the real-space power spectrum and the redshift-space distortions of galaxies in the 2dF 100k galaxy redshift survey using pseudo-Karhunen-Loève eigenmodes and the stochastic bias formalism. Our results agree well with those published by the 2dFGRS team, and have the added advantage of producing easy-to-interpret uncorrelated minimum-variance measurements of the galaxy–galaxy, galaxy–velocity and velocity–velocity power spectra in 27 k-bands, with narrow and well-behaved window functions in the range $0.01\,h\,\mathrm{Mpc}^{-1} < k < 0.8\,h\,\mathrm{Mpc}^{-1}$. We find no significant detection of baryonic wiggles, although our results are consistent with a standard flat $\Omega_\Lambda = 0.7$ 'concordance' model and previous tantalizing hints of baryonic oscillations. We measure the galaxy-matter correlation coefficient $r > 0.4$ and the redshift-distortion parameter $\beta = 0.49 \pm 0.16$ for $r = 1$ ($\beta = 0.47 \pm 0.16$ without finger-of-god compression). Since this is an apparent-magnitude limited sample, luminosity-dependent bias may cause a slight red-tilt in the power spectrum. A battery of systematic error tests indicate that the survey is not only impressive in size, but also unusually clean, free of systematic errors at the level to which our tests are sensitive. Our measurements and window functions are available at http://www.hep.upenn.edu/∼max/2df.html together with the survey mask, radial selection function and uniform subsample of the survey that we have constructed.

Key words: methods: data analysis – galaxies: distances and redshifts – galaxies: statistics – large-scale structure of Universe.

1 INTRODUCTION

Three-dimensional maps of the Universe provided by galaxy redshift surveys place powerful constraints on cosmological models, which has motivated ever more ambitious observational efforts such as the the Updated Zwicky Catalog (CfA/UZC; Huchra et al. 1990; Falco et al. 1999), Las Campanas Redshift Survey (Schectman et al. 1996) and *IRAS* Point Source Catalogue Redshift Survey (PSCz; Saunders et al. 2000) surveys, each well in excess of 10^4 galaxies. This has been an exciting year in this regard, with early results released from two even more ambitious projects; the Anglo-Australian Telescope (AAT) two-degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001) and the Sloan Digital Sky Survey (SDSS; York et al. 2000), which aim for 250 000 and 1 million galaxies, respectively.

Analysis of the first 147 000 2dFGRS galaxies (Peacock et al. 2001; Percival et al. 2001; Norberg et al. 2001a; Madgwick et al. 2002) and the first 29 000 SDSS galaxies (Zehavi et al. 2002) have supported a flat dark-energy dominated cosmology, as have angular clustering analyses of the parent catalogs underlying the 2dFGRS (Efstathiou & Moody 2001) and SDSS (Scranton et al. 2002; Connolly et al. 2001; Tegmark et al. 2002; Szalay et al. 2002; Dodelson et al. 2002). Tantalizing evidence for baryonic wiggles in the galaxy power spectrum has been discussed (Percival et al. 2001; Miller, Nichol & Batuski 2001), and cosmological models have been constrained in conjunction with cosmic microwave background (CMB) data (Efstathiou et al. 2002).

The 2dFGRS team has kindly made their first 102 000 redshifts publicly available. Given the huge effort involved in creating this state-of-the-art sample, it is clearly worthwhile to subject it to an independent power spectrum analysis. This is the purpose of the present paper, focusing on large ($k \lesssim 0.3\,h\,\mathrm{Mpc}^{-1}$) scales. Since the cosmological constraints from galaxy surveys are only as accurate as our modelling of bias, extinction, integral constraints, geometry-induced power smearing and other real-world effects, we will employ a number of recently developed techniques for tackling these issues. Compared with the solid and thorough analysis by the 2dFGRS team in Peacock et al. (2001) and Percival et al. (2001), our main improvements will be in the following areas.

(i) By using an approach based on information theory, involving pseudo-Karhunen-Loève eigenmodes, quadratic estimators and Fisher matrix decorrelation, we are able to produce uncorrelated measurements of the linear power spectrum with minimal error bars

---

⋆E-mail: max@physics.upenn.edu (MT); Andrew.Hamilton@colorado.edu (AJSH)
†http://www.hep.upenn.edu/∼max/ (MT); http://casa.colorado.edu/∼ajsh/ (AJSH)

and quite narrow window functions. This allows the power spectrum to be plotted in an easy-to-interpret model-independent way and, because of the narrow windows, minimizes aliasing from non-linear scales when fitting to linear models.

(ii) Using the stochastic bias formalism, we measure independently not one power spectrum but three, encoding clustering anisotropy. On large scales where redshift distortions are linear (Kaiser 1987), these three curves are the real-space power spectra of the galaxies, their velocity divergence (related to the matter density) and the cross-correlation between the two. On smaller scales, the information they encode can be extracted using simulations.

The rest of this paper is organized as follows. In Section 2, we describe the 2dFGRS data used and construct an easy-to-interpret subsample that is strictly magnitude limited after taking various real-world complications into account. We perform our basic analysis in Section 3 and report the results in Section 4. In Section 5, we test for a variety of systematic errors and quantify the effect of non-linearity and non-Gaussianity on our measurements. In Section 6, we discuss our results, fit to cosmological models and compare our results with those in the literature.

2 DATA MODELLING

The 2dFGRS is described in detail in Colless et al. (2001, hereafter C01). The publicly released 2dFGRS sample consists of 102 426 unique objects (excluding duplicates), of which 93 843 have survey quality redshifts (quality factor $\geqslant 3$). Of these 5131 objects have heliocentric redshifts $z \leqslant 0.002$ and are therefore probable stars, while a further 240 galaxies lie outside the defined angular boundaries of the survey (usually inside a hole in one of the parent UK Schmidt Telescope (UKST) fields, occasionally marginally outside one of the 381 surveyed 2° fields). This leaves a sample of 88 472 galaxies with survey quality redshifts. To do full justice to the quality of this data set in a power spectrum analysis, it is crucial to model accurately the three-dimensional selection function $\bar{n}(\boldsymbol{r})$, which gives the expected (not observed) number density of galaxies as a function of 3D position. This is the goal of the present section.

As will be described in Section 3, our method for measuring the power spectrum requires, in its current implementation, that the selection function be separable into the product of an angular part and a radial part:

$$\bar{n}(\boldsymbol{r}) = \bar{n}(\hat{\boldsymbol{r}})\bar{n}(r), \qquad (1)$$

where $\boldsymbol{r} = r\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{r}}$ is a unit vector. The angular part $\bar{n}(\hat{\boldsymbol{r}})$ may take any value between 0 and 1, and gives the completeness as a function of position, i.e. the fraction of all survey-selected galaxies for which survey quality redshifts are actually obtained, while $\bar{n}(r)$ gives the radial selection function. Although it would be possible to generalize the method to a non-separable selection function (by breaking up the selection function into a sum of piece-wise separable parts), we have chosen to stick to the simple case of a separable selection function, for two reasons. First, although the selection function of the 2dF 100k release is not separable, it is nearly so (the survey was originally designed so that it would be), and the gain from allowing a non-separable selection function has seemed insufficient to justify the extra complexity. Secondly, as described in Sections 5.3.1 and 5.3.2, we wish to be able to test for possible systematic effects arising from a misestimate of extinction, which would cause a purely angular modulation of density fluctuations, or from a misestimate of the radial selection function, which would cause a purely radial modulation of the density. Such tests are facilitated if the selection function is separable.

There are two complications that cause slight departures from such separability (C01).

(i) The magnitude limit varies slightly across the sky, because both the photometric calibration of the parent UKST fields and the extinction correction at each angular position was improved after the survey had begun.

(ii) Seeing issues lead to lower completeness for faint galaxies, and weather variations therefore cause the magnitude-dependent completeness fraction to vary in different 2° fields.

Below we will eliminate both of these complications with appropriate cuts on the data set, obtaining a uniform subsample with a separable selection function as in equation (1).

2.1 The basic angular mask

In this subsection, we describe our modelling of the angular mask $\bar{n}(\hat{\boldsymbol{r}})$ for the full sample. In subsequent subsections, we will shrink and re-weight this mask slightly to eliminate the above-mentioned complications, obtaining the final result shown in Fig. 1.

Once the 2dFGRS is complete, it will contain a total of 1192 circular 2° fields, including 450 fields in a 75° × 10° strip near the North Galactic Pole, 643 fields in an 85° × 15° strip near the South Galactic Pole, and a further 99 fields distributed randomly around the Southern strip. The various intersections of these fields with each other yield 7189 non-overlapping intersection regions, referred to as sectors. Parts of sectors are excluded if they fall outside the boundaries of the 314 rectangular UKST plates of the parent APM survey, or inside one of the holes excised from the plates in order to eliminate e.g. bright stars and satellite trails. The data release specifies 2024 holes, of which 1670 lie within, or overlap, those parts of the UKST plates designated as part of the 2dF survey.

The 100k release is a subset of the survey, containing data from 381 circular 2° fields, including 39 random fields. Eventually, when the survey is done, the observed region will be complete, but in the interim the released fields are variably incomplete, with a different completeness fraction $\bar{n}(\hat{\boldsymbol{r}})$ in each sector, as described in C01.

As part of the 2dFGRS data release, Peder Norberg and Shaun Cole provide software that evaluates $\bar{n}(\hat{\boldsymbol{r}})$ in each of approximately 2.5 million 3×3 arcmin^2 pixels, taking all the various complications into account. However, we wish to adopt a different angular mask that admits a separable selection function, and we also wish to be able to compute the spherical harmonics of the angular mask using the fast, analytic method described in Appendix A of Hamilton (1993). We therefore use a more explicit geometric (not pixelized) specification of the mask, described immediately below.

All field, plate and hole boundaries are simple arcs on the celestial sphere, corresponding to the intersection of the sphere with some appropriate plane. This means that any spherical polygon (a field, plate, hole, sector, etc.) can be defined as the intersection of a set of *caps*, where a cap is the set of directions $\hat{\boldsymbol{r}}$ satisfying $\hat{\boldsymbol{a}} \cdot \hat{\boldsymbol{r}} > b$ for some unit vector $\hat{\boldsymbol{a}}$ and some constant $b \in [-1, 1]$. For instance, a 2° field is a single cap, and a rectangular plate is the intersection of four caps. We define masks such as that the one plotted in Fig. 1 as a list of non-overlapping polygons such that $\bar{n}(\hat{\boldsymbol{r}})$ is constant in each one. We construct the basic 2dFGRS mask as follows.

(i) We generate a list of 8903 polygons comprised of 7189 sectors and 1670 holes, plus 44 polygons defining boundaries of UKST plates.

(ii) Whenever two polygons intersect, we split them into non-intersecting parts, thereby obtaining a longer list of 12 066

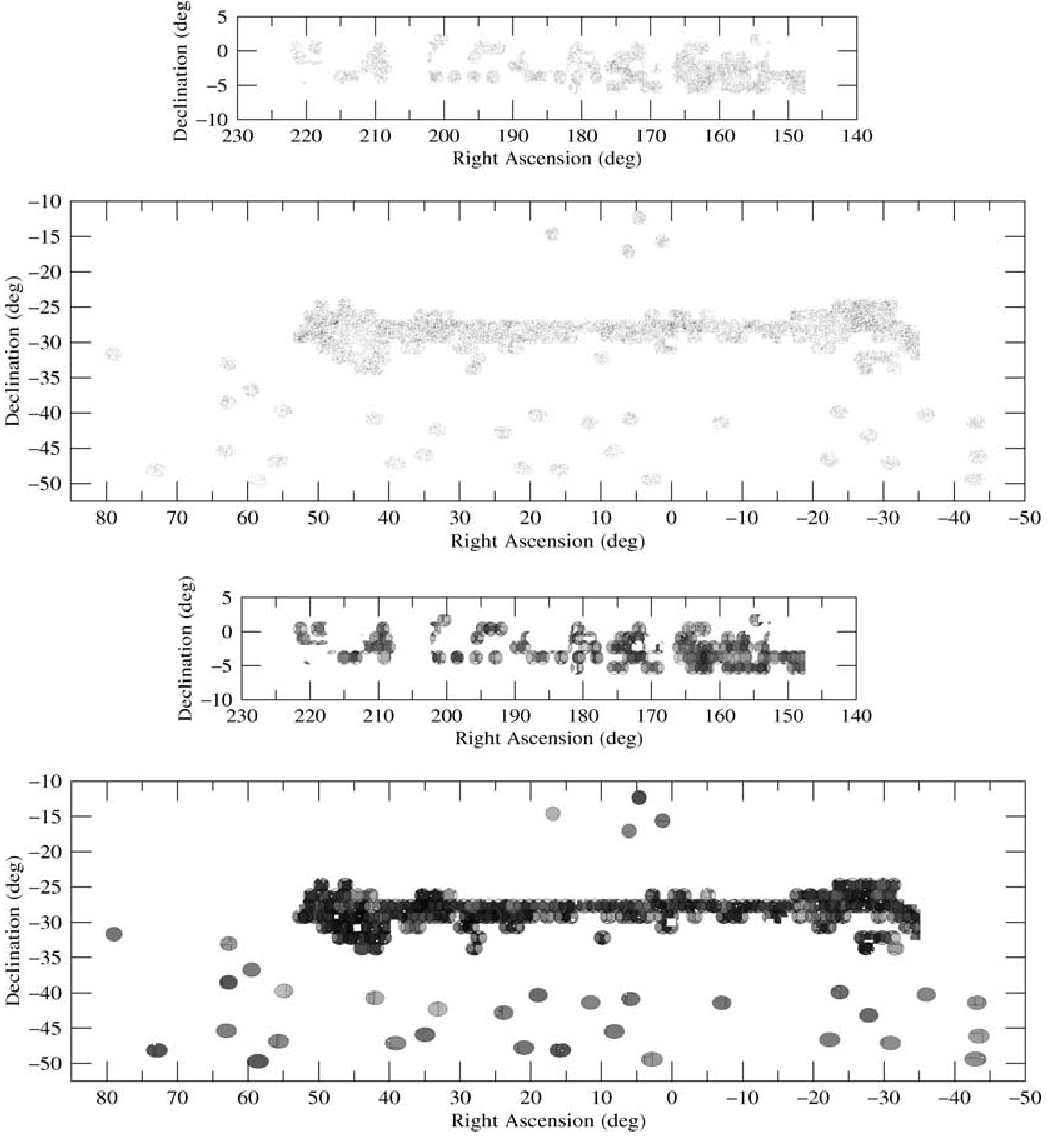

Figure 1. The upper half shows the 59 832 2dF galaxies in our baseline sample, in equatorial 1950 coordinates. The lower half shows the corresponding angular mask, the relative probabilities that galaxies in various directions get included.

non-overlapping polygons. Although slightly tricky in practice, such an algorithm is easy to visualize: if you draw all boundary lines on a sphere and give it to your child as a colouring exercise, using four crayons and not allowing identically coloured neighbours, you would soon be looking at such a list of non-overlapping polygons.

(iii) We compute the completeness $n(\hat{\mathbf{r}})$ for each of these new polygons, originally using the Norberg-Cole software, but subsequently using our own computations, described in the following subsections.

(iv) We simplify the result by omitting polygons with zero weight and merging adjacent polygons that have identical weight.

With the original Norberg and Cole completenesses, the result is a list of 3765 polygons, with a total (unweighted) area of 983 deg^2, and an effective (weighted) area $\int \bar{n}(\hat{\mathbf{r}})\,\mathrm{d}\Omega$ of of 537 deg^2.

With the revised completenesses described in Section 2.2, there are 3614 polygons, with an (unweighted) area of 711 deg^2, and an effective (weighted) area $\int \bar{n}(\hat{\boldsymbol{r}})\,\mathrm{d}\Omega$ of 431 deg^2. This angular mask, and the polygons into which it resolves, are illustrated in Fig. 1.

Section 2.2 explains how we eliminate the two above-mentioned complications, the variations in the magnitude limit, and the variations in the weather, so as to produce an angular mask with the same radial selection function at all points. The reader uninterested in such details can safely skip all this, jumping straight to Section 2.3, remembering only the simple bottom line: we create a uniform sample with 64 844 galaxies over 711 deg^2 that is complete down to b_J magnitude 19.27.

2.1.1 *Cutting to a uniform magnitude limit*

The 2dFGRS aimed to be complete to a limiting b_J magnitude $m = 19.45$ after correction for extinction. However, the actual limiting magnitude varies slightly across the sky as described in C01. This is because after the survey began, there have been improvements in both the photometric calibrations of the underlying parent catalogue (Maddox et al., in preparation) and in the extinction corrections (Schlegel, Finkbeiner & Davis 1998). We eliminate this complication by creating a subsample that is complete down to a slightly brighter limiting magnitude m_*, applying the following two cuts.

(i) Reject all galaxies with extinction-corrected magnitude b_J fainter than m_*.

(ii) Reject all sectors with extinction-corrected magnitude brighter than m_*. The magnitude limit of a sector is defined in the most conservative possible fashion: it is the brightest among all the magnitude limits at the position of each galaxy and of each Norberg-Cole pixel within the sector. The extinction at each position is evaluated using the extinction map of Schlegel et al. (1998).

Fig. 2 shows the number of surviving galaxies as a function of m_*. As we increase m_*, the first cut eliminates fewer galaxies whereas the second cut eliminates more galaxies. The result is seen to be a rather sharply peaked curve, taking its maximum for $m_* = 19.27$, for which 66 050, or 75 per cent, of the 88 472 galaxies survive.

The choice $m_* = 19.27$ turns out to maximize not only the number of galaxies, but the effective survey volume as well. As the flux cut F_* is made fainter, the depth of the survey ($\propto F_*^{-1/2}$ in the Euclidean limit) increases, but the area decreases because there are fewer sectors complete down to F_*. Therefore the survey volume $\overset{\sim}{\propto}$ (area) $\times F_*^{-3/2}$, and this also happens to peak for F_* corresponding to magnitude 19.27.

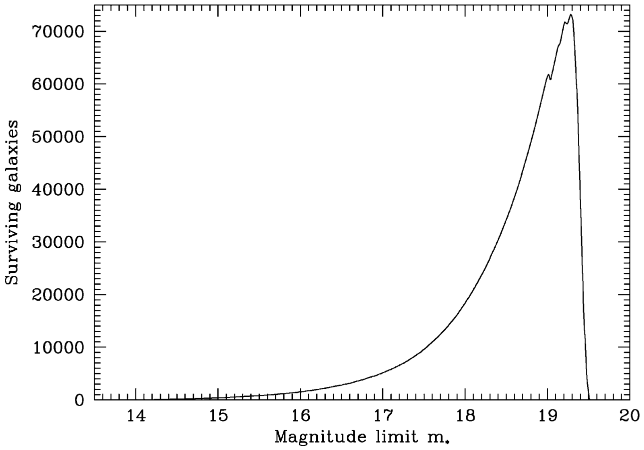

Figure 2. Number of galaxies surviving as a function of uniform magnitude cut.

2.2 Angular selection function

2.2.1 *Modelling the weather*

One of the more time-consuming aspects of our analysis was modelling another departure from uniformity in the 2dFGRS: spatial dependence of the magnitude-dependent incompleteness. As described by C01, although the success-rate P for measuring reliable redshifts (quality $\geqslant 3$) for targeted galaxies is in general quite high, it depends on weather. The poorer the seeing is when a given field is observed, the lower the success rate. Moreover, this weather modulation affects fainter galaxies more than bright ones. C01 found the success rate to be well fitted by an expression of the form

$$P(F) = \gamma\left[1 - (\Phi_f/F)^a\right], \tag{2}$$

where F is the observed flux from the galaxy, $\gamma = 0.99$, $a = 2.5/\ln(10) \approx 1.086$ and Φ_f is a parameter that is fitted for separately for each observed field f, interpretable as the faintest observable flux. Note that since this observational selection effect depends only on magnitude and weather, this issue can be analysed and resolved in terms of apparent magnitudes alone, without explicitly involving redshifts.

The Φ_f-values computed by the 2dFGRS team were not part of the public release, but it is straightforward to generate values from the data provided. Whereas C01 estimated Φ_f from the observed completeness fraction for each field, we performed a maximum-likelihood fit over the fluxes of all objects (galaxies and stars) targeted for observation in each of the 381 field-nights, the likelihood being a product of terms $P(F_i)$ for all successful observations (those yielding a survey-quality redshift), and terms $[1 - P(F_i)]$ for all unsuccessful observations. Maximizing over 382 parameters (Φ_f for each of 381 distinct field-nights f, and a global value of a, with γ fixed equal to 1), we obtain a best-fitting exponent $a = 0.96 \pm 0.04$. Since the exponent is consistent with unity, we set $a = 1$ for simplicity. We repeated the analysis with a permitted to vary separately in each field, but the likelihood is consistent with constant values.

As a cross-check, we repeated the entire analysis sector-by-sector instead of field-by-field, obtaining reassuringly similar results.

2.2.2 *Random sampling to a sharp magnitude limit*

As mentioned, our power spectrum analysis requires a selection function of the separable form of equation (1). Yet the discussion above shows that the shape of the radial selection function $\bar{n}(r)$ varies across the sky, since the success rate $P_f(F)$ is different for each of the 381 field-nights f, as given by equation (2).

We remedy this problem by sparse-sampling the galaxies in such a way that the *shape* of the success rate $P(F)$ (as opposed to its amplitude) becomes the same for all fields. The amplitude variations can then be absorbed into the angular mask $\bar{n}(\hat{\boldsymbol{r}})$, restoring separability. There are clearly infinitely many ways of doing this – we wish to find the way that maximizes the effective volume of the survey for measuring large-scale power.

If we throw away galaxies at random, keeping galaxies in a given field f with a probability p_f that depends on their observed flux F, then the original success rate $P_f(F)$ for the field from equation (2) gets replaced by $P_f(F)p_f(F)$. Our goal then becomes to choose these probabilities $p_f(F)$ such that

$$P_f(F)p_f(F) = w_f P_*(F), \tag{3}$$

where $P_*(F)$ is the desired uniform, global success rate and the weights w_f are are scaling factors that will be absorbed into the angular mask. Since the functions $P_f(F)$ are known, equation (3) immediately specifies how we should choose the probabilities once the function P_* and the weights have been fixed. To maximize the number of surviving galaxies, we want to make p_f and hence w_f as large as possible. Since probabilities cannot exceed unity, this implies that the best weights are

$$w_f = \min_F \left[\frac{P_f(F)}{P_*(F)}\right]. \tag{4}$$

It remains to choose the target success rate $P_*(F)$. Since we are interested in large-scale power, our aim is to maximize not so much the number of galaxies but rather the effective volume of the survey, and we must accomplish this goal by adjusting a function $P_*(F)$ of apparent flux F. The way to do this is to choose $P_*(F)$ so as to retain all galaxies at the faint limit of the survey, and then to make $P_*(F)$ as large as possible at all other fluxes. Given that the original $P(F)$ decreases monotonically to fainter fluxes for all values of the weather parameter Φ_f, and that Φ_f includes cases of perfectly observed fields ($\Phi_f = \infty$), the solution is simply to choose $P_*(F)$ to be a constant, which can be taken to equal 1 without loss of generality, at all values brighter than the flux limit.

This is delightfully simple and convenient: it means that the best choice is a pure magnitude-limited sample with no magnitude incompleteness to keep track of! The corresponding weights are

$$w_f = \min_F P_f(F) = P_f(F_*) = 1 - \Phi_f/F_* \tag{5}$$

where F_* is the flux limit. The scheme thus keeps all galaxies at the flux limit F_*, and discards a progressively larger fraction of the brighter galaxies in each sector so as to cancel exactly the magnitude-dependence of the incompleteness.

The magnitude limit 19.27 arrived at in the previous subsection turns out to maximize the number of galaxies not only before sparse-sampling, but also after sparse-sampling.

The final result is a list of 3614 polygons with associated weights. available at http://www.hep.upenn.edu/~max/2df.html together with the uniform galaxy sample and our power spectrum measurements. The total area is 711 deg^2, and the effective area $\int \bar{n}(\hat{\mathbf{r}})\, \mathrm{d}\Omega$ is 432 deg^2.

2.3 The radial selection function

After the modelling of angular effects above, it remains to measure the radial selection function $\bar{n}(r)$ for the uniform sample. It is important to do this as accurately as possible, since errors in the selection function translate into spurious large-scale power.

The radial selection function $\bar{n}(r)$ that results from the analysis described immediately below is shown in Fig. 3.

In addition to imposing a faint magnitude limit of $b_\mathrm{J} = 19.27$, we follow the advice of the 2dFGRS team (Matthew Colless, private communication) in cutting the survey to a bright limit of $b_\mathrm{J} = 15$. We use a maximum likelihood method based on the C^- method of Lynden-Bell (1971), which assumes that luminosity is uncorrelated with position. We generate an initial approximation to the selection function using a continuous version of the Turner (1979) method, which yields the exact maximum likelihood solution for the case of a survey with a sharp faint flux limit. The Turner method has the merit of being exceedingly fast (less than one CPU second), but it works only if the survey is flux-limited at one end (e.g. the faint end). Starting from the Turner solution, we use an iterative method designed to converge towards the exact maximum likelihood solution for the selection function, which can be shown (Hamilton & Tegmark 2002) to be a step function with steps at the limiting distance of each of the ~60 000 galaxies in the sample. To implement the Bayesian prejudice that the selection function should be smooth, we interpolate the resulting 60 000-point function at ~500 points, through which we pass a cubic spline.

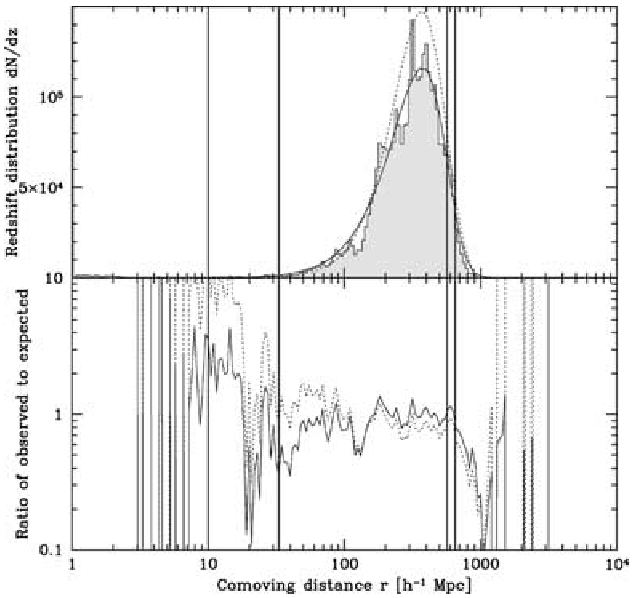

Figure 3. The redshift distribution of the galaxies in our sample is shown both as a histogram (top) and relative to the expected distribution (bottom), in comoving coordinates assuming a flat $\Omega_\mathrm{m} = 0.3$ cosmology. The curves correspond to the the radial selection function $\bar{n}(r)$ employed in our analysis (solid) and by C01 (dotted). The four vertical lines indicate the redshift limits employed in our analysis (10 h^{-1} Mpc $< r <$ 650 h^{-1} Mpc) and where spectral type subsamples are available (33 h^{-1} Mpc $< r <$ 538 h^{-1} Mpc).

We follow the 2dFGRS team in assuming a flat $\Omega_\Lambda = 0.7$ cosmology when converting redshifts to comoving distances r. We transform the galaxy positions into the Local Group frame assuming that the solar motion relative to the Local Group is 306 km s^{-1} toward $l = 99°$, $b = -4°$ (Courteau & van den Bergh 1999). We model k-corrections and luminosity evolution (ε-corrections) together as a power law luminosity evolution $\propto (1+z)^\kappa$ with exponent $\kappa = -0.7$. This exponent was chosen so as to make the comoving density shown in the lower panel in Fig. 3 as flat as possible, i.e. by assuming minimal evolution in the comoving number density of galaxies. Similar results have been reported by Cole (2001), C01, Cross et al. (2001), Madgewick et al. (2001) and Norberg et al. (2001b). The slight differences between our $\bar{n}(r)$ and that of C01 seen Fig. 3 are caused by our different methods for estimating this function from the data, and below we find that they do not have a major impact on the final power spectrum.

We truncate the sample radially by eliminating objects with $r < 10\,h^{-1}$ Mpc (to eliminate stellar contamination) and $r > 650\,h^{-1}$ Mpc (where Fig. 3 shows evidence of incompleteness). This leaves 59 832 galaxies in the sample.

3 METHOD AND BASIC ANALYSIS

In this section, we analyse the uniform galaxy sample described in the previous section, measuring the power spectrum and redshift space distortions of the galaxy density field. We adopt the matrix-based approach described in Tegmark et al. (1998, hereafter

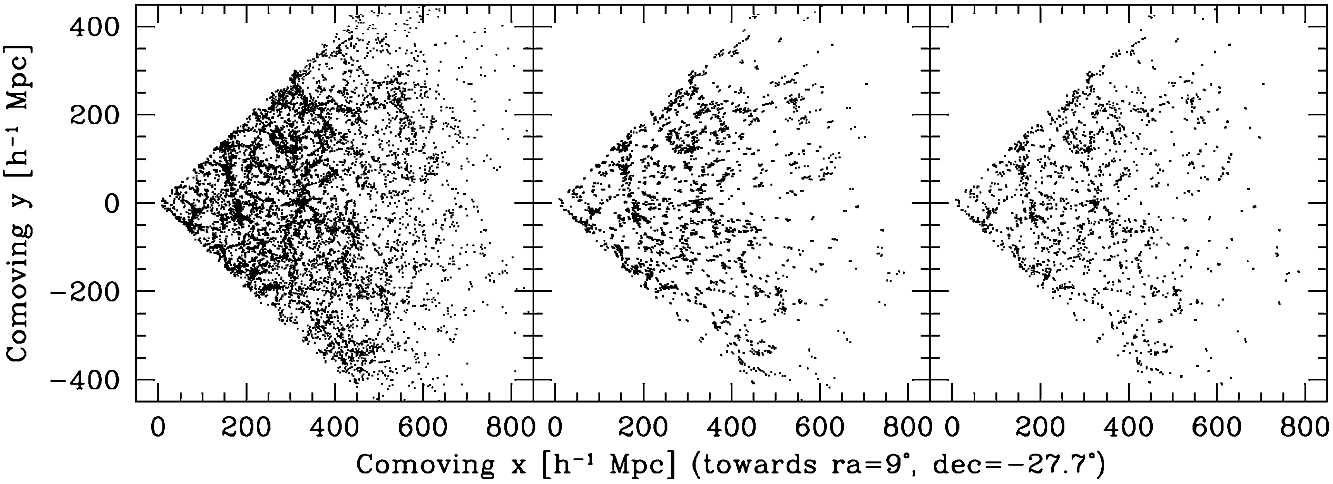

Figure 4. The effect of our fingers-of-God (FOG) removal is shown in the southern slice $\delta = -27.^\circ 7$, $-35^\circ < RA < 53^\circ$. The slice has thickness 2° and has been rotated to lie in the plane of the page. From left to right, the panels show all 15 055 galaxies in the slice, the 6211 that are identified as belonging to FOGs (with density threshold 100) and the same galaxies after FOG compression, respectively.

THSVS98), using the mode expansion of Hamilton & Culhane (1996) and including the stochastic bias formalism. Our analysis consists of the following five steps:

(i) finger-of-God compression;
(ii) pseudo-Karhunen-Loève compression;
(iii) true Karhunen-Loève expansion;
(iv) quadratic band-power estimation;
(v) Fisher decorrelation and flavour disentanglement.

We will now describe these steps in more detail. We will see that step (iii) is not required in practice, and we use it only for systematics tests.

3.1 Step 1: finger-of-God compression

Since our analysis uses the linear Kaiser approximation for redshift-space distortions, it is crucial that we are able to quantify our sensitivity empirically owing to the so-called finger-of-god (FOG) effect whereby radial velocities in virialized clusters make them appear elongated along the line of sight. We therefore start our analysis by compressing (isotropizing) FOGs, as illustrated in Fig. 4. The FOG compression involves a tunable threshold density, and in Section 5.4 below we will study how the final results change as we gradually change this threshold to include lesser or greater numbers of FOGs.

We use a standard friends-of-friends algorithm, in which two galaxies are considered friends, therefore in the same cluster, if the density windowed through an ellipse 10 times longer in the radial than transverse directions, centred on the pair, exceeds a certain overdensity threshold. To avoid linking well-separated galaxies in deep, sparsely sampled parts of the survey, we impose the additional constraint that friends should be closer than $r_{\perp\max} = 5\,h^{-1}$ Mpc in the transverse direction. The two conditions are combined into the following single criterion: two galaxies separated by $r_\parallel$ in the radial direction and by $r_\perp$ in the transverse direction are considered friends if

$$\left[(r_\parallel/10)^2 + r_\perp^2\right]^{1/2} \leqslant \left[(3/4\pi)\pi\bar{n}(1+\delta_c) + r_{\perp\max}^{-3}\right]^{-1/3} \tag{6}$$

where $\bar{n}$ is the selection function (geometrically averaged) at the position of the pair, and δ_c is an overdensity threshold. Note that δ_c represents not the overdensity of the pair as seen in redshift space, but rather the overdensity of the pair after their radial separation has been reduced by a factor of 10. In other words, δ_c is intended to approximate the threshold overdensity of a cluster in real space, not the overdensity of the elongated FOG seen in redshift space. Having identified a cluster by friends-of-friends in this fashion, we measure the dispersion of galaxy positions about the centre of the cluster in both radial and transverse directions. If the one-dimensional radial dispersion exceeds the transverse dispersion, then the cluster is deemed a FOG, and the FOG is then compressed radially so that the cluster becomes round, that is, the transverse dispersion equals the radial dispersion. We perform the entire analysis five times, using progressively more aggressive compression with density cut-offs $1 + \delta_c = \infty$, 200, 100, 50 and 25, respectively. The infinite threshold $1 + \delta_c = \infty$ corresponds to no compression at all.

Fig. 4 illustrates FOG compression with threshold density $1 + \delta_c = 100$, which is the baseline case adopted in this paper. It corresponds to fairly aggressive FOG removal since the overdensity of a cluster is around 200 at virialization and rises as the Universe expands and the background density continues to drop.

3.2 Step 2: Pseudo-KL pixelization

The raw data consists of $N_{\rm gal} = 59\,832$ three-dimensional vectors $\boldsymbol{r}_\alpha$, $\alpha = 1, \ldots, N_{\rm gal}$, giving the measured positions of each galaxy in redshift space. As in THSVS98, we define the density in N_x 'pixels' x_i, $i = 1, \ldots, N_x$ by

$$x_i \equiv \int \frac{n(\boldsymbol{r})}{\bar{n}(\boldsymbol{r})} \psi_i(\boldsymbol{r})\, {\rm d}^3 r \tag{7}$$

for some set of functions ψ_i and work with the N_x-dimensional data vector $\boldsymbol{x}$ instead of the the $3 \times N_{\rm gal}$ numbers $\boldsymbol{r}_\alpha$. Although these are perhaps more aptly termed 'modes' since we will choose quite non-local functions ψ_i, we will keep referring to them as pixels to highlight the useful analogy with CMB map analysis.

Galaxies are (from a cosmological perspective) delta-functions in space, so the integral in equation (7) reduces to a discrete sum over galaxies. We do not rebin the galaxies spatially, which would artificially degrade the resolution. It is convenient to isolate the mean density into a single mode $\psi_1(\boldsymbol{r}) = \bar{n}(\boldsymbol{r})$, with amplitude

$$x_1 = \int n(\boldsymbol{r})\, {\rm d}^3 r = N_{\rm gal}, \tag{8}$$

and to arrange all other modes to have zero mean,

$$\langle x_i \rangle = \int \psi_i(\boldsymbol{r})\,\mathrm{d}^3 r = 0 \quad (i > 1). \tag{9}$$

The covariance matrix of the vector $\boldsymbol{x}$ of amplitudes is a sum of noise and signal terms

$$\langle \Delta\boldsymbol{x}\Delta\boldsymbol{x}^t \rangle = \mathbf{C} \equiv \mathbf{N} + \mathbf{S}, \tag{10}$$

where the shot noise covariance matrix is given by

$$\mathbf{N}_{ij} = \int \frac{\psi_i(\boldsymbol{r})\psi_j(\boldsymbol{r})}{\bar{n}(\boldsymbol{r})}\,\mathrm{d}^3 r \tag{11}$$

and the signal covariance matrix is

$$\mathbf{S}_{ij} = \int \hat{\psi}_i(\boldsymbol{k})\hat{\psi}_j(\boldsymbol{k})^* P(k)\frac{\mathrm{d}^3 k}{(2\pi)^3} \tag{12}$$

in the absence of redshift-space distortions. Here hats denote Fourier transforms and $\bar{n}$ is the three-dimensional selection function described in Section 2, i.e. $\bar{n}(\boldsymbol{r})\,\mathrm{d}V$ is the expected (not the observed) number of galaxies in a volume $\mathrm{d}V$ about $\boldsymbol{r}$. $P(k)$ is the power spectrum, which for a random field of density fluctuations $\delta(\boldsymbol{r})$ is defined by $\langle \hat{\delta}(\boldsymbol{k})^*\hat{\delta}(\boldsymbol{k}')\rangle = (2\pi)^3\delta_{\mathrm{Dirac}}(\boldsymbol{k}-\boldsymbol{k}')$.

As our functions $\psi_i(\boldsymbol{r})$, we use the *pseudo-Karhunen-Loève (PKL) eigenmodes* defined in Hamilton, Tegmark & Padmanabhan (2000, hereafter 'HTP00'). To provide an intuitive feel for the nature of these modes, a sample is plotted in Figs 5 and 6. We use these modes because they have the following desirable properties.

(i) They form a complete set of basis functions probing successively smaller scales, so that a finite number of them (we use the first 4000) allow essentially all information about the density field on large scales to be distilled into the vector $\boldsymbol{x}$.

(ii) They allow the covariance matrices $\mathbf{N}$ and $\mathbf{S}$ to be fairly rapidly computed.

(iii) They are the product of an angular and a radial part, i.e. take the separable form $\psi_i(\boldsymbol{r}) = \psi_i(\hat{\boldsymbol{r}})\psi_i(r)$, which accelerates numerical computations.

(iv) A range of potential sources of systematic problems are isolated into special modes that are orthogonal to all other modes. This means that we can test for the presence of such problems by looking for excess power in these modes, and immunize against their effects by discarding these modes.

We have four types of such special modes.

(i) The very first mode is the mean density, $\psi_1(\boldsymbol{r}) = \bar{n}(\boldsymbol{r})$. The mean mode is used in determining the maximum likelihood normalization of the selection function, but is then discarded from the analysis, since it is impossible to measure the fluctuation of the mean mode. The idea of solving the so-called integral constraint problem by making all modes orthogonal to the mean goes back to Fisher et al. (1993).

(ii) Modes 2–5 are associated with the motion of the Local Group through the cosmic microwave background at 622 km s^{-1} towards (B1950 FK4) RA = 162°, Dec. = −27° (Lineweaver et al. 1996; Courteau & van den Bergh 1999). In the angular direction, these Local Group modes are monopole and dipole modes multiplied by the angular mask, while in the radial direction they take the form specified by equation (4.42) of Hamilton (1997c). Mode 2 is a pure monopole mode (multiplied by the angular mask), and is present because the survey is not all-sky. The other three Local Group modes are dipole modes with admixtures of the Local Group monopole mode 2, such as to make them orthogonal to the mean mode 1.

(iii) Purely radial modes (for example mode 104 in Fig. 6) are to first order the only ones affected by misestimates of the radial selection function $\bar{n}(r)$.

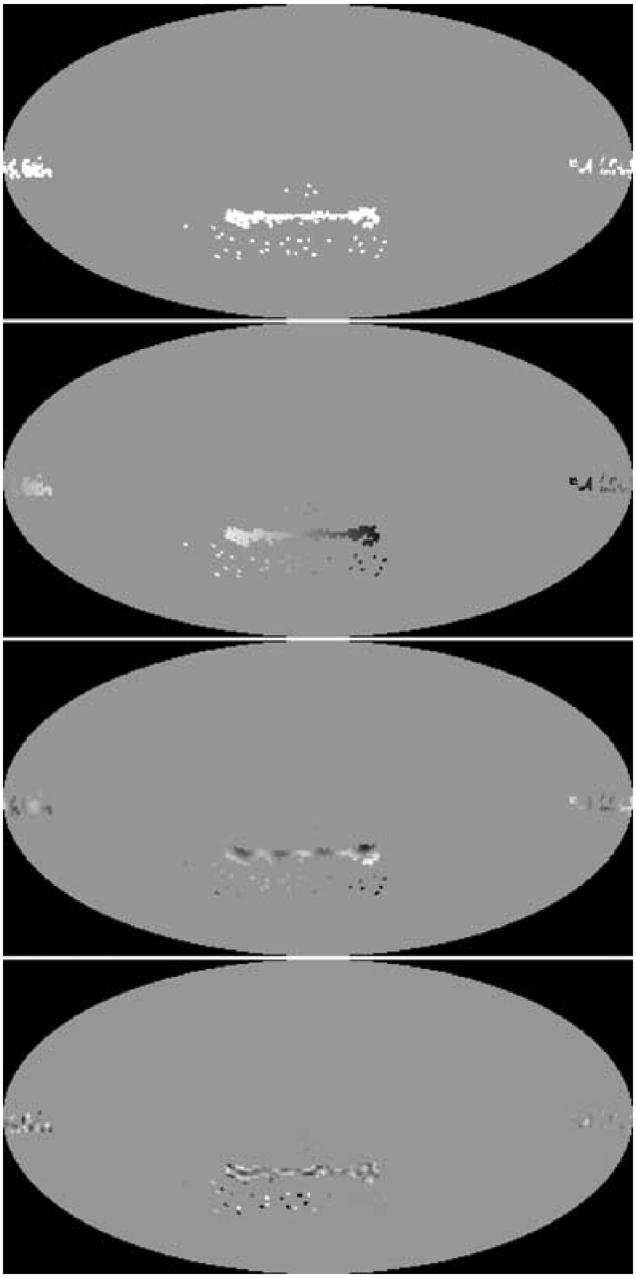

Figure 5. A sample of four angular pseudo-KL (PKL) modes are shown in Hammer-Aitoff projection in equatorial coordinates, with grey representing zero weight, and lighter/darker shades indicating positive/negative weight, respectively. From top to bottom, they are angular modes 1 (the mean mode), 3, 20 and 106, and are seen to probe successively smaller angular scales.

(iv) Purely angular modes (for example mode 148 in Fig. 6) are to first order the only ones affected by misestimates of the angular selection function $\bar{n}(\hat{\boldsymbol{r}})$, as may result from inadequate corrections for, e.g. extinction, the variable magnitude limit, the variable magnitude completeness or photometric zero-point offsets.

As described in HTP00, the modes ψ_i are computed in the logarithmic spherical wave basis (Hamilton & Culhane 1996), which are orthonormal eigenfunctions $Z_{\omega\ell m}(\boldsymbol{r}) = (2\pi)^{-1/2}\mathrm{e}^{-(3/2+\mathrm{i}\omega)\ln r}Y_{\ell m}(\hat{\boldsymbol{r}})$ of the complete set of commuting Hermitian operators

$$\mathrm{i}\left(\frac{\partial}{\partial \ln r} + \frac{3}{2}\right) = -\mathrm{i}\left(\frac{\partial}{\partial \ln k} + \frac{3}{2}\right), \quad L^2, \quad L_z. \tag{13}$$

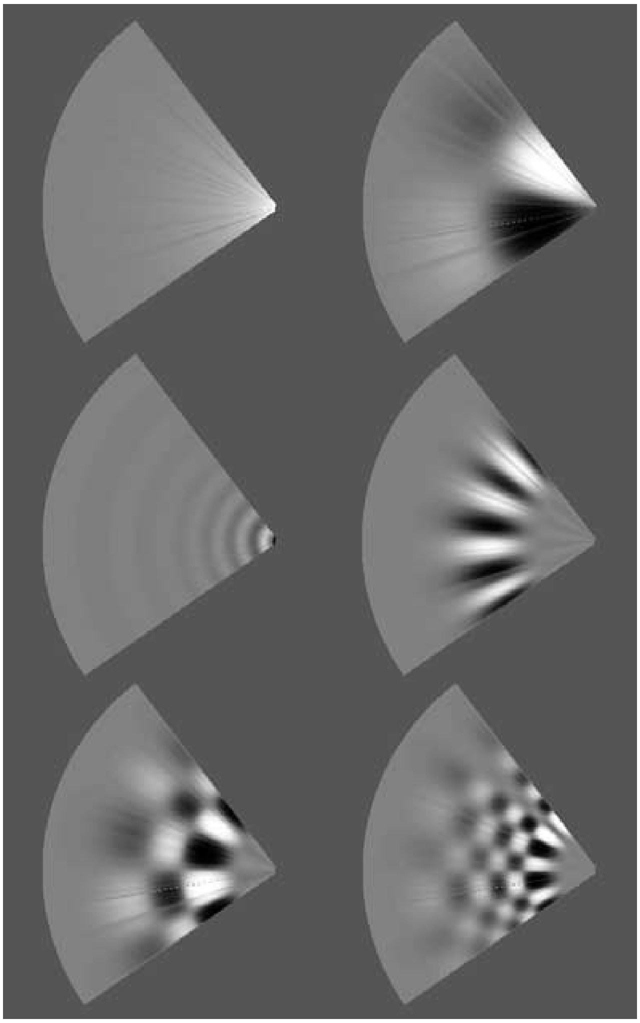

Figure 6. A sample of six pseudo-KL modes are shown in the plane of the southern 2dF slice with $\delta = -27^\circ\!.7$, $-35^\circ < RA < 53^\circ$. Grey represents zero weight, and lighter/darker shades indicate positive/negative weight, respectively. From left to right, top to bottom, these are modes 1 (the mean mode), 14, 104, 148, 58 and 178, and are seen to probe successively smaller scales. Those in the middle panel are examples of purely radial (left) and purely angular (right) modes.

Slightly better numerical behaviour is obtained by expanding not $\psi_i(\boldsymbol{r})$ itself but rather $\psi_i(\boldsymbol{r})/\bar{n}(r)^{1/2}$ (the denominator is the square root of the radial part of the selection function only, not the angular part) in logarithmic spherical waves, since this mitigates some difficulties that arise from the fact that the radial selection function $\bar{n}(r)$ varies by orders of magnitude. The merits of working in a basis of spherical harmonics were first emphasized by Fisher, Scharf & Lahav (1994) and by Heavens & Taylor (1995). The advantages of working with logarithmic radial waves $\mathrm{e}^{-(3/2+\mathrm{i}\omega)\ln r}$, compared for example with spherical Bessel functions, are both numerical and physical.

(i) Numerically, the logarithmic radial wave basis permits rapid transformation between real, ω, and Fourier space using Fast Fourier Transforms. The transformation is mathematically equivalent to the Fast Fourier-Hankel-Bessel Transform FFTLOG described in Appendix B of Hamilton & Tegmark (2000).

(ii) Physically, logarithmic radial waves are well matched to real galaxy surveys like the 2dFGRS, which are finely sampled nearby, and coarsely sampled far away.

(iii) The linear redshift distortion operator is diagonal in this basis (Hamilton & Culhane 1996).

The logarithmic radial wave basis discretizes naturally on to a logarithmically equispaced grid (in both real and Fourier space), and is periodic over a logarithmic interval. To avoid potential problems of aliasing between small and large scales, we embed the survey inside a suitably large logarithmic interval of depths, extending in real space from $10^{-2}\,h^{-1}$ Mpc to $10^4\,h^{-1}$ Mpc. As remarked in Section 2.3, we truncate the survey to radial depths 10–650 h^{-1} Mpc within this interval.

The dimensionless log-frequency ω in the radial eigenmode $\mathrm{e}^{-(3/2+\mathrm{i}\omega)\ln r}$ is a radial analogue (in a precise mathematical sense) of the dimensionless angular harmonic number ℓ. Similar resolution in the radial and angular directions is secured by choosing the maximum log-frequency to be about equal to the the maximum harmonic number, $\omega_{\max} \approx \ell_{\max}$. The maximum log-frequency is related to the radial resolution $\Delta \ln r$ by $\omega_{\max} = \pi/\Delta \ln r$. We adopt a maximum harmonic number of $\ell_{\max} = 40$, and a radial resolution of 32 points per decade, so $\Delta \ln r = (\ln 10)/32$, giving $\omega_{\max} = 43.7$ (the same as in HTP00). These choices ensure comparable effective resolutions in radial and angular directions.

A maximum angular harmonic number of $\ell_{\max} = 40$ gives $(\ell_{\max} + 1)^2 = 1681$ spherical harmonics, while 32 points per radial decade over 6 decades gives 192 radial modes. Thus there is a potential pool of $41^2 \times 192 \approx 320\,000$ modes from which we would like to construct Karhunen-Loève (KL) modes. The usual way to construct such modes would be to diagonalize a $320\,000 \times 320\,000$ matrix, but this is evidently utterly intractable numerically.

How do we build the PKL modes in practice? To make the problem tractable, we instead proceed hierarchically, first constructing angular PKL modes, and then constructing a set of radial PKL modes associated with each angular KL mode. The procedure is possible because we have required the selection function to be separable into angular and radial parts, equation (1). We refer to the resulting modes as pseudo Karhunen-Loève (PKL) modes. The PKL basis contains almost as much information as a true KL basis, but it circumvents the need to diagonalize an impossibly huge matrix. Our procedure is the same as that of HTP00. A different, but similar in spirit, hierarchical approach to the KL problem has been proposed by Taylor et al. (2001).

As we proceed from angular PKL mode to angular PKL mode, extending each into 3D PKL modes by computing associated radial functions, we retain only the $N_x = 4000$ PKL modes with the highest expected signal-to-noise. As detailed below, we make this truncation both to render the various $N_x \times N_x$ matrices numerically tractable and to limit sensitivity to small, non-linear scales. As the signal-to-noise ratio of the angular PKL mode decreases, fewer and fewer of the associated radial PKL modes make the cut into the pool of PKL modes. We stop when 10 successive angular PKL modes have contributed no new PKL mode. In practice only 140 of the angular PKL modes actually contribute to the PKL modes. The reduction from 1681 to 140 angular modes with little information loss is possible because the spherical harmonics are overcomplete and redundant on the modest fraction of the sky actually covered by the 2dFGRS.

The orthogonality of the PKL modes to the mean and the properties of the 'special' modes are enforced in the construction of the modes. We perform the PKL decomposition *after* selecting out the special modes (rather than doing the KL decomposition and then making them orthogonal to the special modes), since we find that this makes better PKL modes. We do this as described in Appendix B of THSVS98, with the complication that we make the non-special

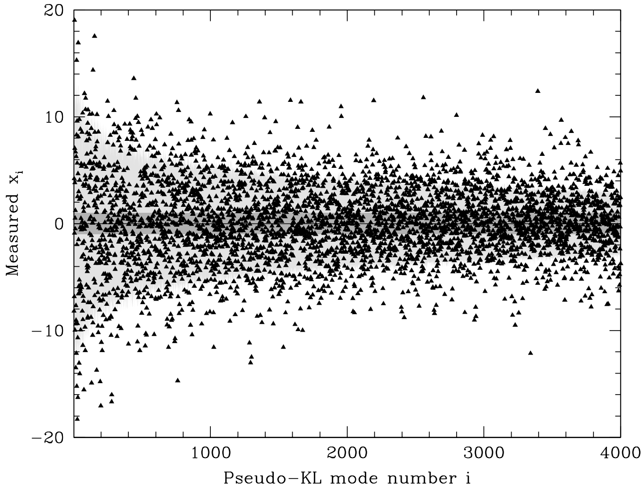

Figure 7. The triangles show the 4000 elements x_i of the data vector $\boldsymbol{x}$ (the pseudo-KL expansion coefficients) for the baseline galaxy sample. If there were no clustering in the survey, merely shot noise, they would have unit variance, and about 68 per cent of them would be expected to lie within the dark grey band. If our prior power spectrum were correct, then the standard deviation would be larger, as indicated by the shaded light grey band.

modes *exactly* orthogonal to the masked mean and the masked LG modes, not merely orthogonal up to the finite order of the discrete matrices.

The pixelized data vector $\boldsymbol{x}$ is shown in Fig. 7. This data compression step has thus distilled the large-scale information about the galaxy density field from $N_{\rm gal} = 59\,832$ galaxy position vectors into 4000 PKL-coefficients. The functions ψ_i are normalized so that $N_{ii} = 1$, i.e. so that the shot noise contribution to their variance is unity. If there were no cosmological density fluctuations in the survey, merely Poisson fluctuations, the PKL-coefficients x_i would thus have unit variance, and about 68 per cent of them would be expected to lie within the blue/dark grey band. Fig. 7 shows that the fluctuations are considerably larger than Poisson, especially for the largest-scale modes (to the left), demonstrating that cosmological density fluctuations are present, as expected.

3.3 Step 3: expansion into true KL modes

Karhunen-Loève (KL) expansion (Karhunen 1947) was first introduced into large-scale structure analysis by Vogeley & Szalay (1996). It has since been applied to the Las Campanas redshift survey (Matsubara et al. 2000), the UZC survey (PTH01) and the SDSS (Szalay et al. 2002; Tegmark et al. 2002) and has been successfully applied to cosmic microwave background data as well, first by Bond (1995) and Bunn (1995).

Given $\boldsymbol{x}$, $\mathbf{N}$ and $\mathbf{S}$ from the previous section, it is straightforward to compute the true Karhunen-Loève (KL) coefficients. They are defined by

$$\boldsymbol{y} \equiv \mathbf{B}^t \boldsymbol{x}, \tag{14}$$

where $\boldsymbol{b}$, the columns of the matrix $\mathbf{B}$, are the N_x eigenvectors of the generalized eigenvalue problem

$$\mathbf{S}\boldsymbol{b} = \lambda \mathbf{N}\boldsymbol{b}, \tag{15}$$

sorted from highest to lowest eigenvalue λ and normalized so that $\boldsymbol{b}^\dagger \mathbf{N} \boldsymbol{b} = \boldsymbol{I}$. This implies that

$$\langle y_i y_j \rangle = \delta_{ij}(1 + \lambda_i), \tag{16}$$

which means that the transformed data values $\boldsymbol{y}$ have the desirable property of being uncorrelated. In the approximation that the distribution function of $\boldsymbol{x}$ is a multivariate Gaussian, this also implies that they are statistically independent – then $\boldsymbol{y}$ is merely a vector of independent Gaussian random variables. Moreover, equation (15) shows that the eigenvalues λ_i can be interpreted as a signal-to-noise ratio S/N. Since the matrix $\mathbf{B}$ is invertible, the final data set $\boldsymbol{y}$ clearly retains all the information that was present in $\boldsymbol{x}$. In summary, the KL transformation partitions the information content of the original data set $\boldsymbol{x}$ into N_x chunks that are mutually exclusive (independent), collectively exhaustive (jointly retaining all the information), and sorted from best to worst in terms of their information content. In most applications, the chief use of KL-coefficients is for data compression, discarding modes containing almost no information and thereby accelerating subsequent calculations. The KL-coefficients for our data set are plotted in Fig. 8, and it is seen that even the worst coefficients still have non-negligible signal-to-noise ratio, bearing numerical testimony to the quality of the PKL modes we have used. This means that KL-compression would not accelerate our particular analysis, and we will indeed work directly with the uncompressed data $\boldsymbol{x}$ in the following subsections.

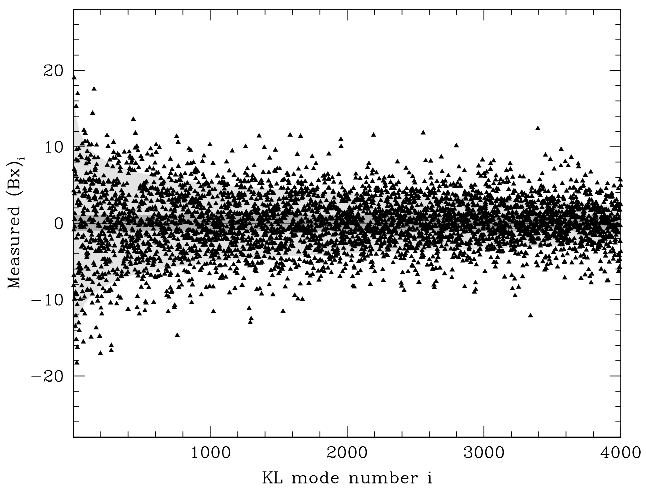

Figure 8. The triangles show the 3999 uncorrelated elements y_i of the transformed data vector $\boldsymbol{y} = \mathbf{B}\boldsymbol{x}$ (the true KL expansion coefficients) for the baseline galaxy sample. If there were no clustering in the survey, merely shot noise, they would have unit variance, and about 68 per cent of them would be expected to lie within the dark grey band. If our prior power spectrum were correct, then the standard deviation would be larger, as indicated by the shaded light grey band. The grey curve is the rms of the data points x_i, averaged in bands of width 25, and is seen to agree better with the light grey band than the dark grey band.

Rather, the reason we have computed KL-coefficients is as an additional check against systematic errors and incorrect assumptions, to verify that we modelled not only the diagonal terms in $\mathbf{C}$ correctly (as seen in Fig. 7), but the off-diagonal correlations as well. As discussed in many of the above-mentioned KL-papers, inspection of the KL-coefficients as in Fig. 8 provides yet another opportunity to detect suspicious outliers and to check whether the variance predicted by the prior power spectrum is consistent with the data. We will provide a detailed test based on the KL-coefficients in Section 5.1.

3.4 What we wish to measure: three power spectra, not one

Before analysing the $\boldsymbol{x}$-vector in the following subsections, let us first discuss precisely what we want to measure. Cosmological constraints based on galaxy power spectrum measurements are only as accurate as our understanding of biasing. We will therefore perform our analysis in a way that avoids making any assumptions about the relation between galaxies and matter, as described in Tegmark (1998) and HTP00.

Unfortunately, bias is complicated. The commonly used assumption that the matter density fluctuations $\delta(\boldsymbol{r})$ and the galaxy number density fluctuations $g(\boldsymbol{r})$ obey

$$g(\boldsymbol{r}) = b\delta(\boldsymbol{r}) \tag{17}$$

for some constant b (the bias factor) appears to be violated in a number of ways. It has been long known (Davis & Geller 1976; Dressler 1980) that b must depend on galaxy type. However, there is also evidence that it depends on scale (see e.g. Mann et al. 1998; Blanton et al. 2000; Hamilton & Tegmark 2002, and references therein) and on time (Fry 1996; Tegmark & Peebles 1998; Giavalisco et al. 1998). Finally, there are good reasons to believe that there is *no* deterministic relation that can replace equation (17), but that bias is inherently somewhat stochastic (Dekel & Lahav 1999) – this has been demonstrated in both simulations (Blanton 2000) and real data (Tegmark & Bromley 1999). The term stochastic does of course not imply any randomness in the galaxy formation process, merely that additional factors besides density may be important (gas temperature, say).

The good news for our present analysis is that, restricting attention to second moments, all the information about stochasticity is contained in a single new function $r(k)$ (Pen 1998; Tegmark & Peebles 1998). Grouping the fluctuations into a two-dimensional vector

$$\boldsymbol{x} \equiv \begin{pmatrix} \delta \\ g \end{pmatrix} \tag{18}$$

and assuming nothing except translational invariance, its Fourier transform $\hat{\boldsymbol{x}}(\boldsymbol{k}) \equiv \int \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \boldsymbol{x}(\boldsymbol{r})\, \mathrm{d}^3 r$ obeys

$$\langle \hat{\boldsymbol{x}}(\boldsymbol{k}) \hat{\boldsymbol{x}}(\boldsymbol{k}')^\dagger \rangle = (2\pi)^3 \delta^D(\boldsymbol{k} - \boldsymbol{k}') \begin{pmatrix} P(\boldsymbol{k}) & P_\times(\boldsymbol{k}) \\ P_\times(\boldsymbol{k}) & P_{\mathrm{gg}}(\boldsymbol{k}) \end{pmatrix} \tag{19}$$

for some 2×2 power spectrum matrix that we will denote $\mathbf{P}(\boldsymbol{k})$. Here P is the conventional power spectrum of the mass distribution, P_{gg} is the power spectrum of the galaxies, and $P_\times$ is the cross spectrum. It is convenient to rewrite this covariance matrix as

$$\mathbf{P}(\boldsymbol{k}) = P(\boldsymbol{k}) \begin{pmatrix} 1 & b(\boldsymbol{k})r(\boldsymbol{k}) \\ b(\boldsymbol{k})r(\boldsymbol{k}) & b(\boldsymbol{k})^2 \end{pmatrix} \tag{20}$$

where $b \equiv (P_{\mathrm{gg}}/P)^{1/2}$ is the bias factor (the ratio of galaxy and total fluctuations) and the new function $r \equiv P_\times/(P P_{\mathrm{gg}})^{1/2}$ is the dimensionless correlation coefficient between galaxies and matter. Note that both b and r generally depend on scale k. The Schwarz inequality shows that the special case $r = 1$ implies the simple deterministic equation (17), and the converse is of course true as well.

On large scales where linear perturbation theory is valid, redshift distortions (Kaiser 1987; Hamilton 1997c) conveniently allow all three of these functions to be measured. Specifically, the correlation between the observed densities at any two points depends linearly on these three power spectra:

$$\begin{aligned} &\text{galaxy–galaxy power}: P_{\mathrm{gg}}(k) = b(k)^2 P(k) \\ &\text{galaxy–velocity power}: P_{\mathrm{gv}}(k) = r(k) b(k) f P(k) \\ &\text{velocity–velocity power}: P_{\mathrm{vv}}(k) = f^2 P(k) \end{aligned} \tag{21}$$

Here $f \approx \Omega_{\mathrm{m}}^{0.6}$ is the dimensionless growth rate for linear density perturbations (see Hamilton 2001). More correctly, the 'velocity' here refers to minus the velocity divergence, which in linear theory is related to the mass (not galaxy) overdensity δ by $f\delta + \nabla \cdot \boldsymbol{v} = 0$, where ∇ denotes the comoving gradient in velocity units. Note that $P_{\mathrm{gv}}(k) = f P_\times(k)$ and that the parameter f is conveniently eliminated by defining $\beta(k) \equiv f/b(k)$, which gives

$$\begin{aligned} P_{\mathrm{gv}}(k) &= \beta(k) r(k) P_{\mathrm{gg}}(k), \\ P_{\mathrm{vv}}(k) &= \beta(k)^2 P_{\mathrm{gg}}(k). \end{aligned} \tag{22}$$

3.5 Step 4: quadratic compression into band powers

In this step, we perform a much more radical data compression by taking certain quadratic combinations of the data vector that can easily be converted into power spectrum measurements.

We parametrize the three power spectra $P_{\mathrm{gg}}(k)$, $P_{\mathrm{gv}}(k)$ and $P_{\mathrm{vv}}(k)$ as piecewise constant functions, each with 49 'steps' of height p_i, which we term the *band powers*. To avoid unnecessarily jagged spectra, we take $k^{1.5}P$ rather than P to be constant within each band. We group these 3×49 numbers into a 147-dimensional vector $\boldsymbol{p}$. We choose our 49 k-bands to be centred on logarithmically equispaced k-values $k_i = 10^{(i-41)/16}\, h\,\mathrm{Mpc}^{-1}$, $i = 1, \ldots, 49$, i.e. ranging from $0.00316\, h\,\mathrm{Mpc}^{-1}$ to $3.16\, h\,\mathrm{Mpc}^{-1}$. For instance, $P_{\mathrm{gg}}(k) = (k/k_i)^{-1.5} p_i$ for $|\lg k - \lg k_i| < 1/32$. This should provide fine enough k-resolution to resolve any baryonic wiggles and other spectral features that may be present in the power spectrum. For instance, baryon wiggles have a characteristic scale of order $\Delta k \sim 0.1$, so we oversample the first one around $k \sim 0.1$ by a factor $\Delta k/(k_{26} - k_{25}) \sim 16/\ln 10 \sim 7$.

This parametrization means that we can write the pixel covariance matrix of equation (10) as

$$\mathbf{C} = \sum_{i=0}^{147} p_i \mathbf{C}_{,i}, \tag{23}$$

where the derivative matrix $\mathbf{C}_{,i} \equiv \partial \mathbf{C}/\partial p_i$ is the contribution from the ith band. For notational convenience, we have included the noise term in equation (23) by defining $\mathbf{C}_{,0} \equiv \mathbf{N}$, corresponding to an extra dummy parameter $p_0 = 1$ giving the shot noise normalization. As in Hamilton & Tegmark (2000) and HTP00, we in practice redefine the parameters p_i to be ratio of the actual band power to the prior band power. As long as the prior agrees fairly well with the measured result, this has the advantage of giving better behaved window functions, as described in Hamilton & Tegmark (2000).

Our quadratic band power estimates are defined by

$$q_i \equiv \frac{1}{2} \boldsymbol{x}^t \mathbf{C}^{-1} \mathbf{C}_{,i} \mathbf{C}^{-1} \boldsymbol{x}, \tag{24}$$

$i = 0, \ldots, 147$. These numbers are shown in Fig. 9, and we group them together in a 148-dimensional vector $\boldsymbol{q}$. Note that whereas $\boldsymbol{x}$ (and therefore $\mathbf{C}$) is dimensionless, $\boldsymbol{p}$ has units of power, i.e. volume. Equation (24) therefore shows that $\boldsymbol{q}$ has units of inverse power, i.e. inverse volume. It is not immediately obvious that the vector $\boldsymbol{q}$ is a useful quantity. It is certainly not the final result (the power spectrum estimates) that we want, since it does not even have the right units. Rather, it is a useful intermediate step. In the approximation that the pixelized data has a Gaussian probability distribution (a good approximation in our case because of the central limit theorem, since N_{gal} is large) $\boldsymbol{q}$ has been shown to retain all the power spectrum information from the original data set (Tegmark, Taylor & Heavens 1997, hereafter 'T97'). The numbers q_i have the additional advantage (as compared with, e.g. maximum-likelihood estimators) that their properties are easy to compute: their mean and covariance are given by

$$\langle \boldsymbol{q} \rangle = \mathbf{F}\boldsymbol{p}, \tag{25}$$

$$\langle \boldsymbol{q}\boldsymbol{q}^t \rangle - \langle \boldsymbol{q} \rangle \langle \boldsymbol{q} \rangle^t = \mathbf{F}, \tag{26}$$

where $\mathbf{F}$ is the *Fisher information matrix* (Tegmark et al. 1997)

$$\mathbf{F}_{ij} = \frac{1}{2} \mathrm{tr}\left[\mathbf{C}^{-1} \mathbf{C}_{,i} \mathbf{C}^{-1} \mathbf{C}_{,j}\right]. \tag{27}$$

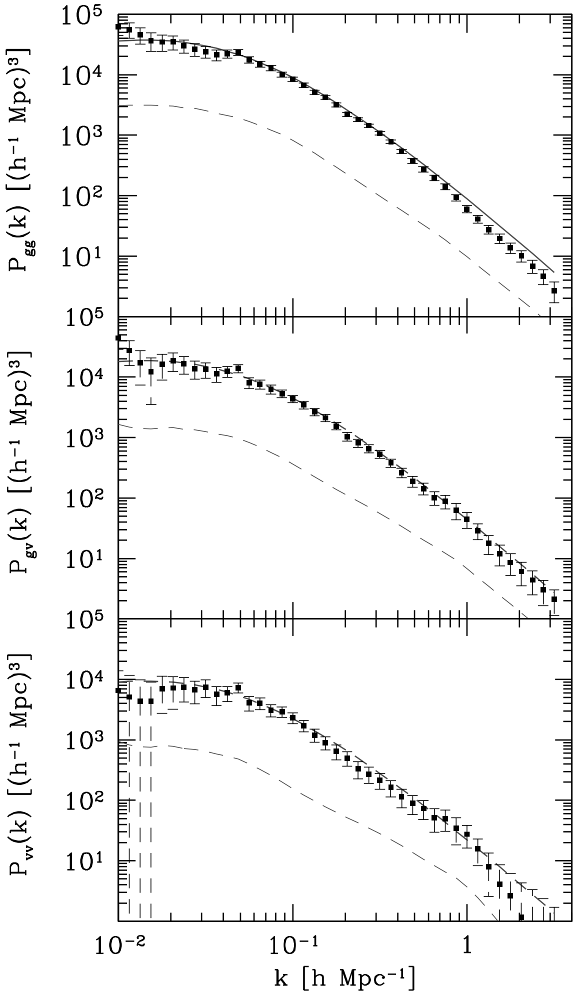

Figure 9. The 147 quadratic estimators q_i, normalized so that their window functions equal unity and with the shot noise contribution f_i (dashed curve) subtracted out. They cannot be directly interpreted as power spectrum measurements, since each point probes a linear combination of all three power spectra over a broad range of scales, typically centred at a k-value different than the nominal k where it is plotted. Moreover, nearby points are strongly correlated, causing this plot to over-representamount of information present in the data. The solid curves show the prior power spectrum used to compute the error bars.

Quadratic estimators were first derived for galaxy survey applications (Hamilton 1997a,b). They were accelerated and first applied to CMB analysis (T97; Bond, Jaffe & Knox 2000).

In conclusion, this step takes the vector $\boldsymbol{x}$ and its covariance matrix $\mathbf{C}$ from Fig. 7 and compresses it into the smaller vector $\boldsymbol{q}$ and its covariance matrix $\mathbf{F}$, illustrated in Figs 9 and 10. Although equation (25) shows that we can obtain unbiased estimates of the true powers $\boldsymbol{p}$ by computing $\mathbf{F}^{-1}\boldsymbol{q}$, there are better options, as will be described in the next subsection.

3.6 Step 5: Fisher decorrelation and flavour disentanglement

Let us first eliminate the shot-noise dummy parameter p_0, since we know its value. We define $\boldsymbol{f}$ to be the zeroth column of the Fisher matrix defined above ($f_i \equiv \mathbf{F}_{i0}$) and restrict the indices i and j to run from 1 to 147 from now on, so $\boldsymbol{f}$, $\boldsymbol{q}$ and $\boldsymbol{p}$ are 147-dimensional vectors and $\mathbf{F}$ is a 147×147 matrix. Since $p_0 = 1$, equation (25) then becomes $\langle \boldsymbol{q} \rangle = \mathbf{F}\boldsymbol{p} + \boldsymbol{f}$.

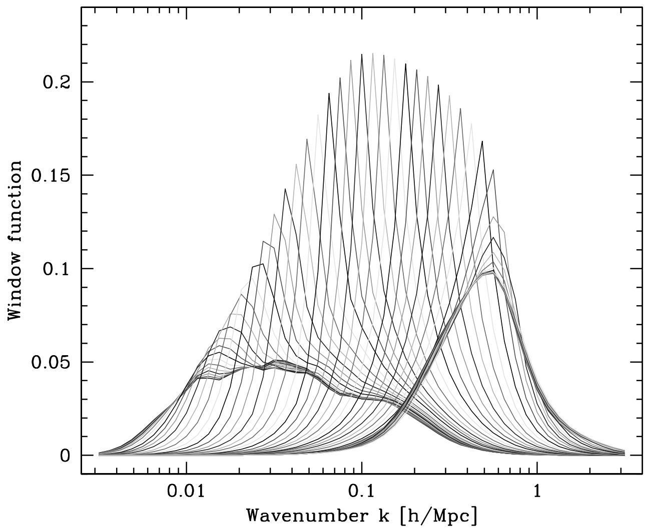

Figure 10. The rows of the gg-portion of the Fisher matrix $\mathbf{F}$. The ith row typically peaks at the ith band, the scale k that the band power estimator q_i was designed to probe. All curves have been renormalized to unit area, so the highest peaks indicate the scales the the window functions obtained are narrowest. The turnover in the envelope at $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ reflects our running out of information due to omission of modes probing smaller scales. For comparison with the next figure, these are the rows of $\mathbf{W}$ when $\mathbf{M}$ is diagonal.

We now define a vector of shot-noise corrected band-power estimates

$$\hat{\boldsymbol{p}} \equiv \mathbf{M}(\boldsymbol{q} - \boldsymbol{f}), \tag{28}$$

where $\mathbf{M}$ is some matrix whose rows are normalized so that the rows of $\mathbf{MF}$ sum to unity. Using equations (25) and (26), this gives the mean and covariance

$$\langle \hat{\boldsymbol{p}} \rangle = \mathbf{W}\boldsymbol{p}, \tag{29}$$

$$\boldsymbol{\Sigma} \equiv \langle \hat{\boldsymbol{p}}\hat{\boldsymbol{p}}^t \rangle - \langle \hat{\boldsymbol{p}} \rangle \langle \hat{\boldsymbol{p}} \rangle^t = \mathbf{MFM}^t, \tag{30}$$

where $\mathbf{W} \equiv \mathbf{MF}$. We will refer to the rows of $\mathbf{W}$ as window functions, since they sum to unity and equation (29) shows that $\hat{p}_i$ probes a weighted average of the true band powers p_j, the ith row of $\mathbf{W}$ giving the weights.

3.6.1 Correlated, anticorrelated and uncorrelated band powers

For the purpose of fitting models $\boldsymbol{p}$ to our measurements $\hat{\boldsymbol{p}}$, we are already done – the last two equations tell us how to compute χ^2 for any given $\boldsymbol{p}$, and the result

$$\chi^2 = (\hat{\boldsymbol{p}} - \langle \hat{\boldsymbol{p}} \rangle)^t \boldsymbol{\Sigma}^{-1} (\hat{\boldsymbol{p}} - \langle \hat{\boldsymbol{p}} \rangle)^t \tag{31}$$

is independent of the choice of $\mathbf{M}$. However, since one of the key goals of our analysis is to provide model-independent measurement of the three power spectra, the choice of $\mathbf{M}$ is crucial. Ideally, we would like both uncorrelated error bars (diagonal $\boldsymbol{\Sigma}$) and well-behaved (narrow, unimodal and non-negative) window functions $\mathbf{W}$ that do not mix the three power spectra.

There are a number of interesting choices of $\mathbf{M}$ that each have their pros and cons (Tegmark & Hamilton 1998; Hamilton & Tegmark 2000). The simple choice where $\mathbf{M}$ is diagonal gives the 'best guess' estimates in the sense of having minimum variance (Hamilton 1997a); (T97; Bond et al. 2000), and also has the advantage of being independent of the number of bands used in the limit of high spectral resolution. It was used for Figs 9 and 10. Here the window

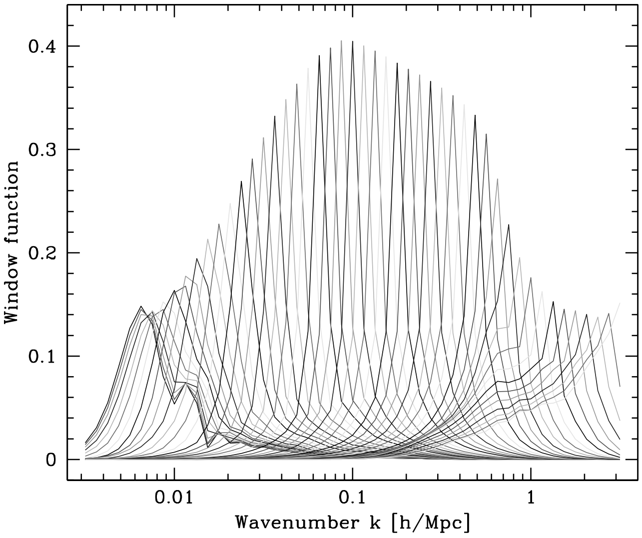

Figure 11. The window functions (rows of the gg-portion of $\mathbf{W}$) are shown using decorrelated estimations. The ith row of $\mathbf{W}$ typically peaks at the ith band, the scale k that the band power estimator $\hat{p}_i$ was designed to probe. Comparison with Fig. 10 shows that decorrelation makes all windows substantially narrower.

functions are simply the rows of the Fisher matrix, and are seen to be rather broad. All entries of $\mathbf{F}$ are guaranteed to be positive as proven in PTH01, which means not only that all windows are positive (which is good) but also that all measurements are positively correlated (which is bad).

Another interesting choice is (T97) $\mathbf{M} = \mathbf{F}^{-1}$, which gives $\mathbf{W} = \mathbf{I}$. In other words, all window functions are Kronecker delta functions, and $\hat{\boldsymbol{p}}$ gives completely unbiased estimates of the band powers, with $\langle \hat{p}_i \rangle = p_i$ regardless of what values the other band powers take. This gives an answer similar to the maximum-likelihood method (THSVS98), and the covariance matrix of equation (30) reduces to $\mathbf{F}^{-1}$. A serious drawback of this choice is that that if we have sampled the power spectrum on a scale finer than the inverse survey size in an attempt to retain all information about wiggles etc. this covariance matrix tends to give substantially larger error bars ($\Delta p_i \equiv \mathbf{M}_{ii}^{1/2} = [(\mathbf{F}^{-1})_{ii}]^{1/2}$) than the first method, anti-correlated between neighbouring bands.

The two above-mentioned choices for $\mathbf{M}$ both tend to produce correlations between the band power error bars. The minimum-variance choice generally gives positive correlations, since the Fisher matrix cannot have negative elements, whereas the unbiased choice tends to give anticorrelation between neighbouring bands. The choice (Tegmark & Hamilton 1998; Hamilton & Tegmark 2000) $\mathbf{M} = \mathbf{F}^{-1/2}$ with the rows renormalized has the attractive property of making the errors uncorrelated, with the covariance matrix of equation (30) diagonal. The corresponding window functions $\mathbf{W}$ are plotted in Fig. 11, and are seen to be quite well-behaved, even narrower than those in Fig. 10 while remaining positive.[1] This choice, which is the one we make in this paper, is a compromise between the two first ones: it narrows the minimum variance window functions at the cost of only a small noise increase, with uncorrelated noise as an extra bonus. The minimum-variance band-power estimators are essentially a smoothed version of the uncorrelated ones, and their lower variance was paid for by correlations which reduced the effective number of independent measurements.

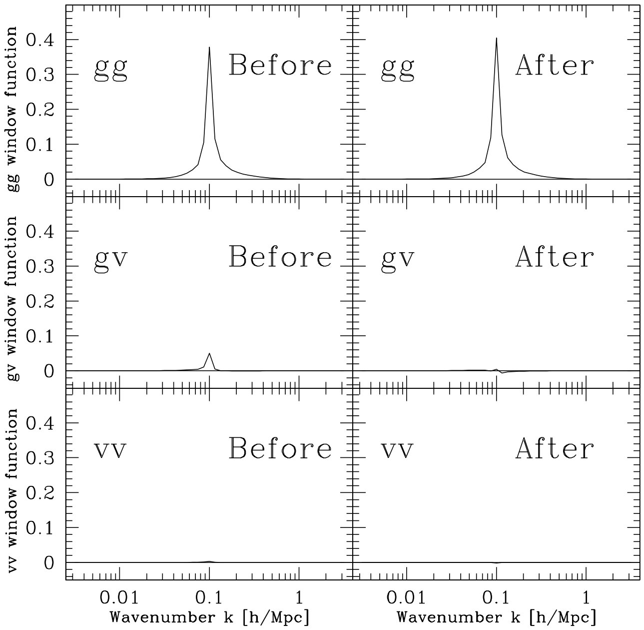

Figure 12. The window function for our measurement of the 25th band of the galaxy–galaxy power is shown before (left) and after (right) disentanglement. Whereas unwanted leakage of gv and vv power is present initially, these unwanted window functions both average to zero afterward. The success of this method hinges on the fact that since the three initial functions (left) have similar shape, it is possible to take linear combinations of them that almost vanish (right).

3.6.2 Disentangling the three power spectra

The fact that we are measuring three power spectra rather than one introduces an additional complication. As illustrated by Fig. 12, an estimate of the power in one of the three spectra generally picks up unwanted contributions (‘leakage’) from the other two, making it complicated to interpret. Although the above-mentioned $\mathbf{F}^{-1}$ method in principle eliminates leakage completely, the cost in terms of increased error bars is found to be prohibitive. We therefore follow HTP00 in adopting the following procedure for disentangling this three power spectra: For each of the 49 k-bands, we take linear combinations of the gg, gv and vv measurements such that the unwanted parts of the window functions average to zero. This procedure is mathematically identical to that used in Tegmark & de Oliveira-Costa (2001) for separating different types of CMB polarization, so the interested reader is referred there for the explicit equations. The procedure is illustrated in Fig. 12, and by construction has the property that leakage is completely eliminated if the true power has the same shape (not necessarily the same amplitude) as the prior. We find that this method works quite well (in the sense that the unwanted windows do not merely average to zero) for the most accurately measured band powers, in particular the central parts of the gg-spectrum, with slightly poorer leakage elimination for bands with larger error bars.

The window functions plotted in Fig. 11 are the gg-windows after disentanglement. Note that although our disentanglement introduces correlations between the gg, gv and vv measurements for a given k-band, different k-bands remain uncorrelated.

---

[1] The reader interested in mathematical challenges will be interested to know that it remains a mystery to the authors why this $\mathbf{F}^{1/2}$ method works so well. We have been unable to prove that $\mathbf{F}^{1/2}$ has no negative elements (indeed, counterexamples can be contrived), yet the method works like magic in practice in all LSS and CMB applications we have tried.

4 RESULTS

4.1 The three power spectra

Our basic results are shown in Fig. 13. The single most striking feature of this plot is clearly that the 2dFGRS is an amazing data set with unprecedented constraining power. The window functions in Fig. 11 are seen to be quite narrow despite the complicated survey geometry. The galaxy–galaxy power is constrained to 20 per cent or better over an order of magnitude in length scale, in about a dozen uncorrelated bands centred around $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$. Whereas the increase in error bars on large scales reflects the finite survey volume, the lack of information on small scales is caused by our analysis being limited to the first 4000 PKL modes. Comparing Fig. 13 with Fig. 9 serves as a sobering reminder of the importance of decorrelating and disentangling the measurements to avoid a misleadingly rosy picture of how well one can do.

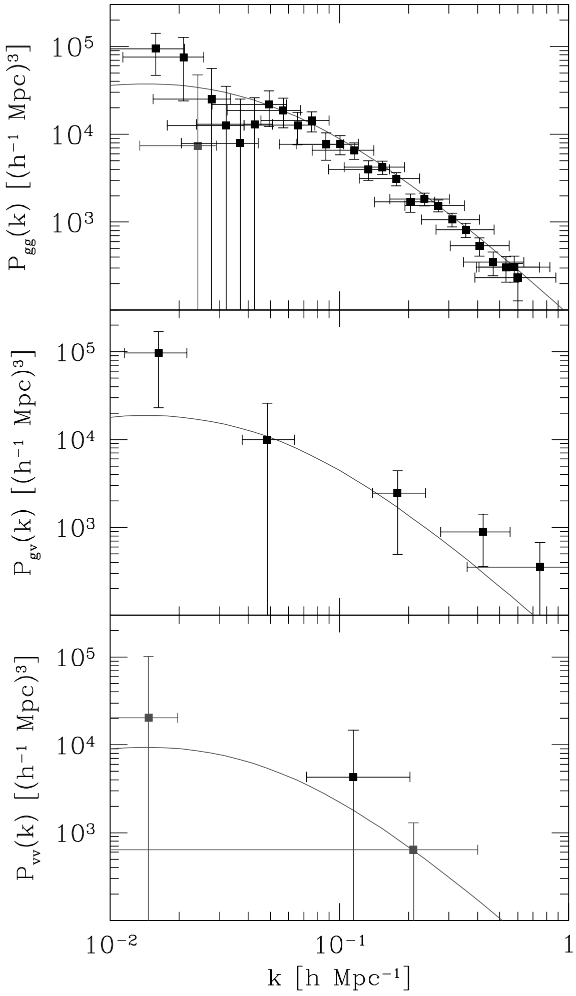

Figure 13. Decorrelated and disentangled measurements of the galaxy–galaxy power spectrum (top), the galaxy–velocity power spectrum (middle) and the velocity–velocity power spectrum (bottom) for the baseline galaxy sample. Grey points represent negative values – since the points are differences between two positive quantities (total power minus expected shot noise power), they can be negative when the signal-to-noise ratio is poor. Each points is plotted at the k-value that is the median of its window function, and 68 per cent of this function is contained within the range of the horizontal bars. The curves shows our prior power spectrum. Note that most of the information in the survey is on the galaxy–galaxy spectrum. Band-power measurements with very low information content have been binned into fewer (still uncorrelated) bands.

Whereas $P_{gg}(k)$ is well measured, Fig. 13 shows that the information about $P_{gv}(k)$ is quite limited and that on $P_{vv}(k)$ almost non-existent. To avoid excessive cluttering in Fig. 13, band-power measurements with very low information content have been binned into fewer (still uncorrelated) bands. The main cause of these large error bars is that the information on P_{vv} and P_{gg} comes from the quadrupole and hexadecapole moments of the clustering anisotropy, which are intrinsically small and hence poorly constrained quantities. However, the problem may be exacerbated by the lack of large contiguous angular regions in the current data, impeding accurate comparisons of angular and radial clustering (the situation is similar for the SDSS, Zehavi et al. 2002), and should improve as the survey nears completion and gets more filled in. This effect is evident from a comparison with the results from the much more contiguous PSCz survey (HTP00): the error bars on P_{gg} are appreciably larger for PSCz than 2dFGRS, but those on redshift distortions (say β) are comparable.

In the remainder of this paper, we will address two separate issues in turn: redshift-space distortions/biasing (β, r) and the detailed shape of the galaxy–galaxy power spectrum (model fits, evidence for baryonic wiggles, etc.).

4.2 Constraints on redshift space distortions

As seen from Fig. 13, the constraints on $P_{gv}(k)$ and $P_{vv}(k)$ from 2dFGRS are too weak to allow $\beta(k)$ and $r(k)$ to be measured reliably as a function of scale. As data on large-scale structure improve, it should become possible to accomplish such a measurement, and thereby to establish quantitatively the scale dependence of biasing at linear scales. In the meantime we limit ourselves to the less ambitious goal of measuring overall parameters β and r, simply treating them as scale-independent constants. This has not been previously done for the case of r. Such scale-independence of bias on linear scales is a feature of local bias models (Coles 1993; Fry & Gaztañaga 1993; Scherrer & Weinberg 1998; Coles, Melott & Munshi 1999; Heavens, Matarrese & Verde 1999).

For our redshift-distortion analysis, we employ a simple scale-invariant power spectrum $P_{gg}(k)$ of the BBKS form (Bardeen et al. 1986), parametrized by an amplitude σ_8 and a 'shape parameter' Γ that on a log plot shifts the curve vertically and horizontally, respectively. We will use more physically motivated power spectra with baryon wiggles etc. in Section 6.2 – we tried various alternative parametrizations, and found that the detailed form had essentially no effect on the (r, β)-constraints, since they come from the ratios of the three spectra, not from their shapes. Our model for the underlying band-power vector p thus depends on four parameters $(\Gamma, \sigma_8, \beta, r)$. We map out the likelihood function $L = \mathrm{e}^{-\chi^2/2}$ using equation (31) on a fine grid in this parameter space, and compute constraints on individual parameters by marginalizing over the other parameters as described in Tegmark et al. (2001), maximizing rather than integrating over them. The results are plotted in Figs 14, 15 and 16.

Fig. 14 assumes $\Gamma = 0.14$, $r = 1$ (the best-fitting values) and explores how the results change as we include information from smaller and smaller scales. As will be discussed in more detail in Section 6, non-linear effects invalidate the Kaiser approximation for redshift space distortions on small scales. A smoking gun signature of such non-linearities is r and hence the best-fitting β dropping and ultimately going negative, as non-linear 'fingers of god' (FOGs) reverse the effect of linear redshift distortions. The fact that Fig. 14

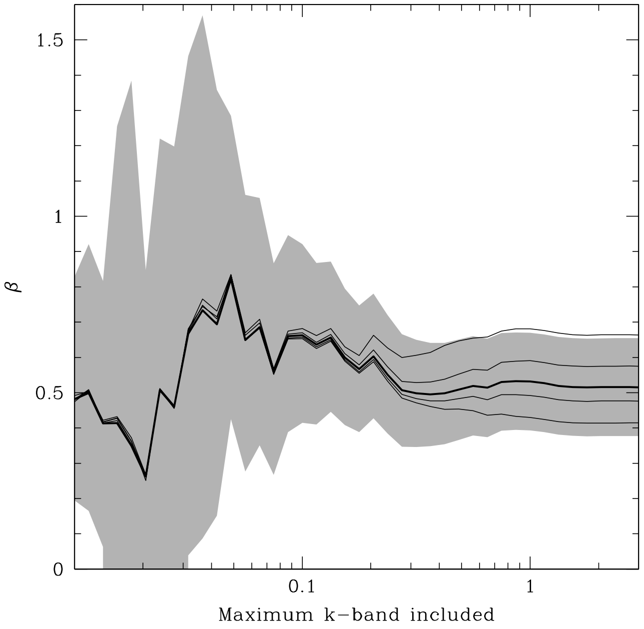

Figure 14. The grey band shows the 1σ allowed range for β, assuming $r = 1$ and the *shape* of the prior $P_{\rm gg}(k)$ but marginalizing over the power spectrum normalization, using FOG compression with density threshold $1+\delta_{\rm c} = 100$. These fits are performed cumulatively, using all measurements for all wavenumbers $\leqslant k$. From bottom to top, the five curves show the best-fitting β for FOG thresholds $1+\delta_{\rm c} = \infty$ (no FOG compression), 200, 100 (heavy), 50 and 25.

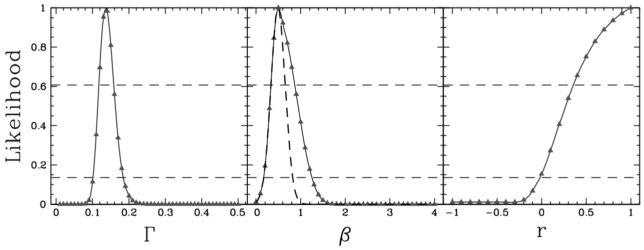

Figure 15. One-dimensional likelihood curves for Γ, β and r are shown after marginalizing over the power spectrum normalization and the other parameters using our baseline $(1+\delta_{\rm c} = 100)$ finger-of-God compression. The 68 per cent and 95 per cent constraints are where the curves intersect the dashed horizontal lines. The dashed curve in the middle panel shows how the β-constraints tighten up when assuming $r = 1$.

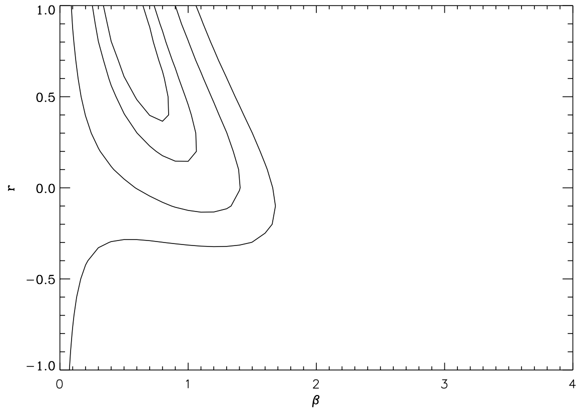

Figure 16. Constraints in the (β, r) plane are shown for our baseline $(1+\delta_{\rm c} = 100)$ finger-of-God compression, using all measurements with $k < 0.3$h/Mpc and marginalizing over the power spectrum normalization for fixed spectral shape. The four contours correspond to $\Delta\chi^2 = 12.29$, 6.18 and 11.83, and would enclose 39 per cent, 68 per cent, 95 per cent and 99.8 per cent of the probability, respectively, if the likelihood function were Gaussian.

does not show this effect is reassuring evidence that little small-scale information is present in our data. This is of course by design, since our PKL modes contain contributions only from $\ell \leqslant 40$, corresponding to a comoving distance around 20 h^{-1} Mpc at the characteristic survey depth of 400 h^{-1} Mpc. This lack of small-scale information in our PKL modes is also reflected in the error bars on β, which are seen to stop decreasing around $k \sim 0.2\,h\,{\rm Mpc}^{-1}$.

Fig. 14 also shows how the results depend on the FOG removal described in Section 3.1. The curves are seen to diverge markedly around $k \sim 0.2\ h\,{\rm Mpc}^{-1}$, with the FOG-related uncertainty becoming as large as the statistical error bars for $k \sim 1\ h\,{\rm Mpc}^{-1}$. We will return to these non-linearity issues in Section 5.4 below.

Fig. 15 shows the constraints on Γ, β and r after marginalizing over the other parameters. The best-fitting model is $\Gamma = 0.14$, $\beta = 0.50, r = 1, \sigma_8 = 0.99$. The reason that the constraints on β are so weak is illustrated in Fig. 16: there is a degeneracy with r. Fig. 13 shows that our information about redshift distortions is coming predominantly from $P_{\rm gv}(k)$, not from the poorly constrained $P_{\rm vv}(k)$, so we are to first order measuring the combination βr rather than β and r individually. Imposing the prior $r = 1$, as was implicitly done in Peacock et al. (2001) and almost all prior work, therefore tightens the upper limit on β substantially, as shown by the dashed curve in Fig. 15.

4.3 The galaxy–galaxy power spectrum alone

The previous subsection discussed the 2dFGRS constraints on redshift-space distortions, essentially the ratios of the power spectra $P_{\rm gg}(k)$, $P_{\rm gv}(k)$ and $P_{\rm vv}(k)$, without regard to their shape. Let us now do the opposite, and focus on the shape of the galaxy power spectrum $P_{\rm gg}(k)$. The success of the disentanglement scheme illustrated in Fig. 12 implies that the galaxy power spectrum plotted in Fig. 13 is robust, essentially independent of what the power spectra $P_{\rm gv}(k)$ and $P_{\rm vv}(k)$ are doing. However, this robustness came at a price in terms of increased error bars. Assuming that all three power spectra have essentially the same shape, but not the same amplitudes, we compute a more accurate estimate of $P_{\rm gg}(k)$ as follows.

We first assume some fixed values for β and r. This allows us to eliminate $P_{\rm gv}(k)$ and $P_{\rm vv}(k)$ using equation (22), reducing the size of our parameter vector $\boldsymbol{p}$ from $3 \times 49 = 147$ to 49 and our Fisher matrix to size 49×49, and gives 49 decorrelated estimators of $P_{\rm gg}(k)$. The result assuming $\beta = 0.5$, $r = 1$ (our best-fitting values) is shown in Fig. 17. We perform no binning here except averaging the noisy bands with $k < 0.02$ and $k > 0.8$ into single bins to reduce clutter. We then repeat this exercise for a range of values of β and r consistent with our analysis in the previous subsection to quantify the uncertainty these parameters introduce. We find these uncertainties to be quite small, as expected considering the small initial leakage of gv and vv power (see Fig. 12), and can therefore quantify the added uncertainty $\delta P_{\rm gg}$ to first order as

$$\delta \ln P_{\rm gg}(k) = \left|\frac{\partial \ln P_{\rm gg}(k)}{\partial(\beta r)}\right| \delta(\beta r) + \left|\frac{\partial \ln P_{\rm gg}(k)}{\partial(\beta^2)}\right| \delta(\beta^2). \qquad (32)$$

Numerically, we find these two derivatives to be approximately -0.2 and -0.04, respectively, essentially independent of k. This

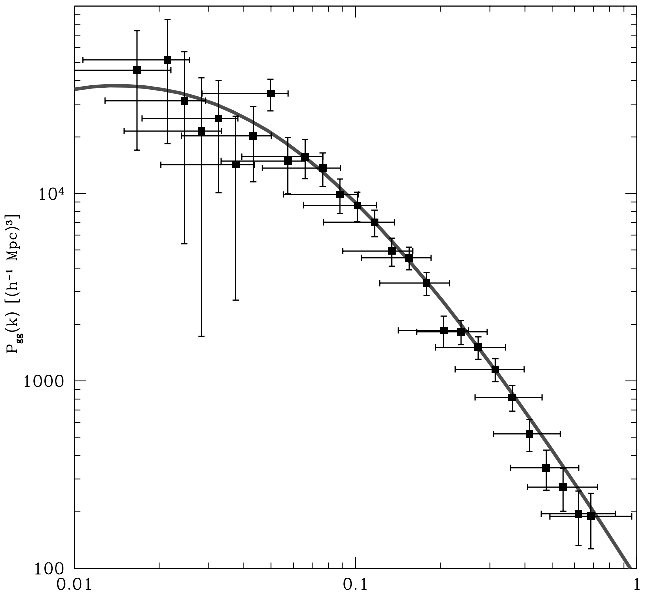

Figure 17. The decorrelated galaxy–galaxy power spectrum is shown for the baseline galaxy sample assuming $\beta = 0.5$ and $r = 1$. As discussed in the text, uncertainty in β and r contributes to an overall calibrations uncertainty of order 12 per cent which is not included in these error bars.

scale-independence is not surprising in the small-angle limit, where these derivatives would involve simply various average moments of μ, the angle between the $\boldsymbol{k}$-vector and the line of sight. Assuming uncertainties $\delta\beta = 0.15$ and and $\delta r = 0.5$, equation (32) thus gives $\delta \ln P_{\mathrm{gg}}(k) \approx 0.12$, the second term being negligible relative to the first. In conclusion, the uncertainties in Fig. 17 induced by uncertainties about β and r can be summarized as simply an overall multiplicative calibration error of order 12 per cent for the measured power spectrum.

5 HOW RELIABLE ARE OUR RESULTS?

How reliable are the results presented in the previous section? In this section, we perform a series of tests, both of our software and algorithms and of potential systematic errors. We also discuss the underlying assumptions that are likely to be most important for interpreting the results.

5.1 Validation of method and software

Since our analysis consists of a number of numerically non-trivial steps, it is important to test both the software and the underlying methods. We do this by generating $N_{\mathrm{monte}} = 100$ Monte Carlo simulations of the 2dFGRS catalogue with a known power spectrum, processing them through our analysis pipeline and checking whether they give the correct answer on average and with a scatter corresponding to the predicted error bars. We found this end-to-end testing to be quite useful in all phases of this project – indeed, we had to run the pipeline 43 times until everything finally worked.

5.1.1 The mock survey generator

Standard *N*-body simulations would not suffice for our precision test, because of a slight catch-22 situation: the true non-linear power spectrum of which an *N*-body simulation is a realization (with shot noise added) is not known analytically, and is usually estimated by measuring it from the simulation – but this is precisely the step that we wish to test. We therefore generate realizations that are firmly in the linear regime, returning to non-linearity issues below. We do this as described in PTH01, with a test power spectrum of the simple Gaussian form $P(k) \propto \mathrm{e}^{-(Rk)^2/2}$ with $R = 32\ h^{-1}$ Mpc, normalized so that the rms fluctuations $\langle \delta^2 \rangle^{1/2} = 0.2$.

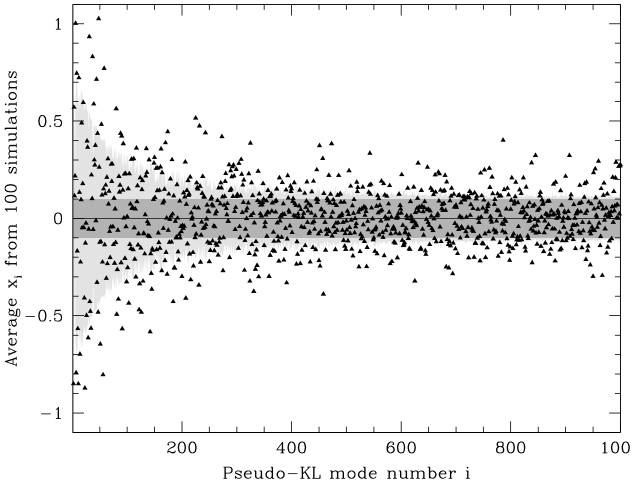

Figure 18. The triangles show the elements x_i of the data vector $\boldsymbol{x}$ (the pseudo-KL expansion coefficients) averaged over 100 Monte-Carlo simulations of the baseline galaxy sample. If the algorithms and software are correct, then their mean should be zero and about 68 per cent of them should lie within the shaded grey region giving their standard deviation.

5.1.2 Testing the PKL pixelization

Fig. 18. shows the result of processing the Monte Carlo simulations through the first step of the analysis pipeline, i.e. computing the corresponding Pseudo-KL expansion coefficients x_i. This is a sensitive test of the mean correction given by equation (9), which can be a couple of orders of magnitude larger than the scatter in Fig. 18 for some modes. A number of problems with the radial selection function integration and the spherical harmonic expansion of the angular mask in our code were discovered in this way. After fixing these problems, the coefficients x_i became consistent with having zero mean as seen in the figure.

Fig. 18 also shows that the scatter in the modes is consistent with the predicted standard deviation $\sigma_i = (\mathbf{C}_{ii}/N_{\mathrm{monte}})^{1/2}$ (shaded region), with most of the the fluctuations being localized to modes probing large scales (with i being small). A more sensitive test of this scatter is shown in Fig. 19, which shows that the theoretically predicted variance for each mode agrees with what is observed in the 100 Monte Carlo realizations. Since crowding makes it hard to verify all modes in this plot, an alternative representation of this test is shown in Fig. 20.

Although these tests verify that the mean and variance of each mode come out as they should, they are not sensitive to errors in the off-diagonal elements of the covariance matrix $\mathbf{C}$, i.e. to incorrect correlations between the mode coefficients. To close this loophole, Fig. 21. shows the scatter in the true KL modes ($\boldsymbol{y} = \mathbf{B}\boldsymbol{x}$), illustrating agreement with the theoretical variance prediction even in this alternative basis where all coefficients y_i should be uncorrelated. Note that the expected variance decreases monotonically here, as

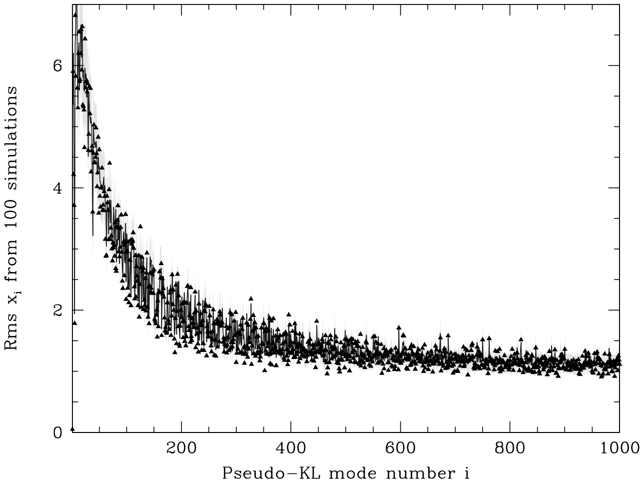

Figure 19. The triangles show the rms fluctuations of the elements x_i from 100 Monte-Carlo simulations. If the algorithms and software are correct, then the expectation value of this rms is given by the thin dark grey curve, and most of them should scatter in the grey region.

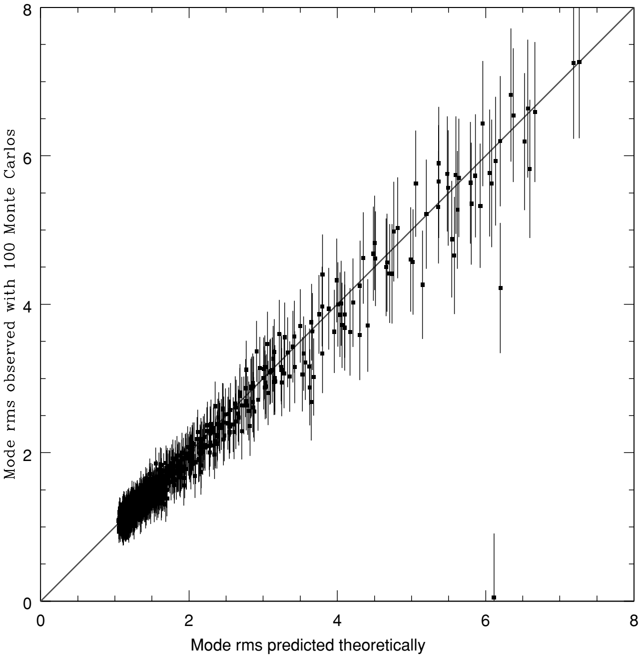

Figure 20. In this alternative representation of the test from Fig. 19, most of the vertical lines should intersect the 45° line if the algorithms and software are correct.

opposed to in Fig. 19, since the true KL modes are strictly sorted by decreasing variance.

5.1.3 Testing the quadratic compression, Fisher decorrelation and disentanglement

Figs 22 and 23 show the result of processing the Monte Carlo simulations through the remaining steps of the analysis pipeline, i.e. computing the raw quadratic estimator vector $\boldsymbol{q}$ and, from it, the decorrelated and disentangled band-power vector $\hat{\boldsymbol{p}}$. The mean recovered power spectra are seen to be in excellent agreement with the Gaussian prior used in the simulations (Fig. 22) convolved with the

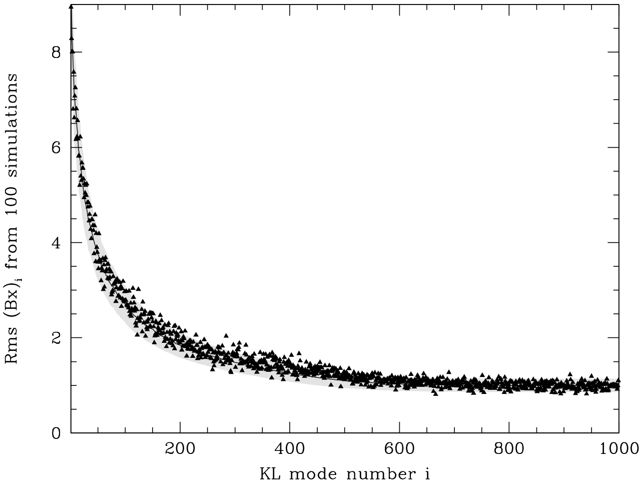

Figure 21. The triangles show the rms fluctuations of the elements $(\mathbf{B}\boldsymbol{x})_i$ from 100 Monte-Carlo simulations. If the algorithms and software are correct, then the expectation value of this rms is given by the thin dark grey curve, and most of them should scatter in the light grey banana-shaped region.

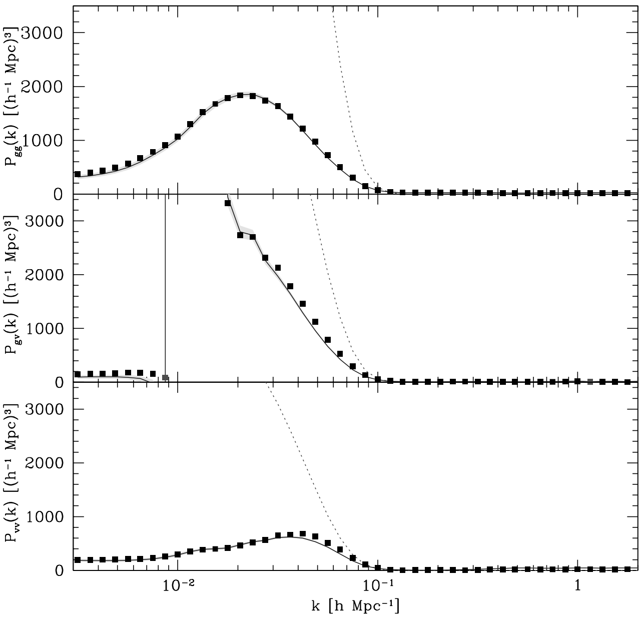

Figure 22. The triangles show the decorrelated and disentangled band-power estimates $\hat{p}_i$, averaged over 100 Monte-Carlo simulations of the baseline galaxy sample. If the algorithms and software are correct, then this should recover the window-convolved input power spectrum $\mathbf{W}\boldsymbol{p}$, plotted as a thin blue line. The thin shaded light grey band indicates the expected scatter. The harmless discontinuity in the middle panel is an artefact of the disentangled galaxy–velocity windows having negative area on the largest scales where there is essentially no information available.

window functions, and the observed scatter is seen to be consistent with the predicted error bars (Fig. 23). These two figures therefore constitute an end-to-end test of our data analysis pipeline, since errors in any of the many intermediate steps would have shown up here at some level. Since information from large numbers of modes contributes to each p_i, the scatter is seen to be small. Therefore, even quite subtle bugs and inaccuracies can be (and were!) discovered and remedied as a result of this test.

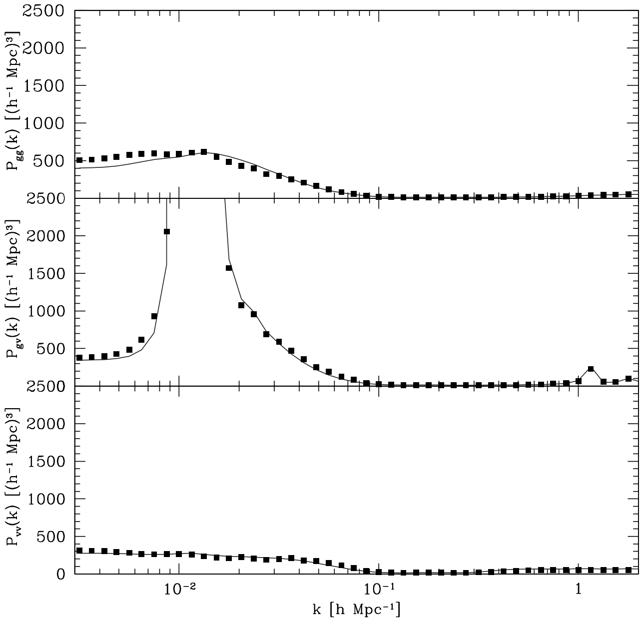

Figure 23. Same as the previous figure, but testing the error bars Δp_i rather than the power itself. The triangles show the observed rms of the power spectrum estimates from 100 simulations and the solid dark grey curve shows the predicted curve around which they should scatter.

5.2 Robustness to method details

Our analysis pipeline has a few 'knobs' that can be set in more than one way. This section discusses the sensitivity to such settings.

5.2.1 *Effect of changing the prior*

The analysis method employed assumes a 'prior' power spectrum via equation (23), both to compute band power error bars and to find the galaxy pair weighting that minimizes them. As mentioned, an iterative approach was adopted starting with a simple BBKS model, then shifting it vertically and horizontally to better fit the resulting measurements and recomputing the measurements a second time. To what extent does this choice of prior affect the results? On purely theoretical grounds (e.g. Tegmark et al. 1997), one expects a grossly incorrect prior to give unbiased results but with unnecessarily large variance. If the prior is too high, the sample-variance contribution to error bars will be overestimated and vice versa. This hypothesis has been extensively tested and confirmed in the context of power spectrum measurements from both the Cosmic Microwave Background (e.g. Bunn 1995) and galaxy redshift surveys (PTH01), confirming that the correct result is recovered on average even when using a grossly incorrect prior. In our case, the prior by construction agrees quite well with the actual measurements (see Fig. 13), so the quoted error bars should be reliable as well.

5.2.2 *Effect of changing the number of PKL modes*

We have limited our analysis to the first $N = 4000$ PKL modes whose angular part is spanned by spherical harmonics with $\ell \leqslant 40$. This choice was a tradeoff between the desire to capture as much information as possible about the galaxy survey and the need to stay away from small scales where non-linear effects invalidate the Kaiser approximation to redshift distortions. To quantify our sensitivity to these choices, we repeated the entire analysis using 500, 1000, 2000 and 4000 modes. Our power spectrum measurements on the very largest scales were recovered even with merely 500 modes. As we added more and more modes (more and mode small-scale information), the power measurements converged to those in Fig. 13 for larger and larger k. The rising part of the envelope in Fig. 10 remained essentially unchanged, merely continuing to grow further as more modes were added, so the turnover of this envelope directly shows the k-scale beyond which we start running out of information. The version of Fig. 10 shown in this paper indicates that our 4000 PKL modes have captured essentially all cosmological information from the 2dFGRS for $k \lesssim 0.1$.

5.2.3 *Numerical issues*

The computation of the matrices $\mathbf{P}_i$ involves a summation over multipoles ℓ that should, strictly speaking, run from $\ell = 0$ to $\ell = \infty$, since the angular mask itself has sharp edges involving harmonics to $\ell = \infty$. In practice, this summation must of course be truncated at some finite multipole $\ell_{\rm cut}$. To quantity the effect of this truncation, we plot the diagonal elements of the $\mathbf{P}$ matrices as a function of $\ell_{\rm cut}$ and study how they converge as $\ell_{\rm cut}$ increases. We define a given PKL mode as having converged by some multipole if subsequent ℓ values contribute less than 1 per cent of its variance. Fig. 24 plots the number of usable PKL modes as a function of wavenumber k, defining a mode to be usable for our analysis only if it is converged for all smaller wavenumbers $k' < k$ for all three power flavours ($P_{\rm gg}$, $P_{\rm gv}$ and $P_{\rm vv}$). We use $\ell_{\rm cut} = 260$ in our final analysis, since this guarantees that all 4000 modes are usable for wavenumbers k in the range $0 - 0.5\,h\,{\rm Mpc}^{-1}$, i.e. comfortably beyond the large scales $0 - 0.3\,h\,{\rm Mpc}^{-1}$ that are the focus of this paper. With this cut-off, the computation of the P-matrices (which scales as $\ell_{\rm cut}^2$ asymptotically), took about a week on a SunBlade1000 workstation. Our power spectrum estimates are likely to remain fairly accurate as far out as we plot them, i.e. to $k \sim 1\,h\,{\rm Mpc}^{-1}$, since Fig. 24 shows most modes remaining usable out to this scale, and since we find that even the ones that do not meet our strict 1 per cent convergence criterion at every single band are generally fairly accurately treated. Indeed, we repeated our entire analysis with $\ell_{\rm cut} = 120$ and obtained almost indistinguishable power spectra.

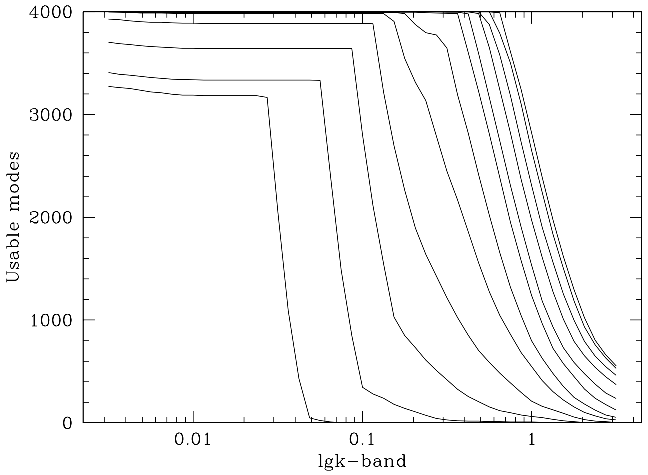

Figure 24. Numerical convergence. The figure shows for how many of our 4000 PKL modes the numerical calculations are converged to accurately measure the power up to a given wavenumber k. From left to right, the 12 curves correspond to truncation at $\ell_{\rm cut} = 20$, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220 and 240.

5.3 Tests for problems with data modelling

In Section 2, we performed detailed modelling of the way in which the 2dFGRS data was selected, and produced a uniform galaxy sample fully characterized by a selection function $\bar{n}(\boldsymbol{r})$ of the separable form of equation (1). Let us now assess how sensitive our results are to potential misestimates of $\bar{n}$, both angularly and radially, by discarding purely angular and radial modes from our analysis.

5.3.1 Robustness to angular problems

Angular modulations caused by dust extinction tend to have a power spectrum rising sharply toward the largest scales (Vogeley 1998), and is therefore of particular concern for the interpretation of our leftmost bandpower estimates. The galaxy magnitudes are extinction corrected by the 2dFGRS team, using extinction map produced by Schlegel et al. (1998), so any inaccuracies in this extinction model would masquerade as excess large-scale power. Inaccuracies in zero-point offsets or in the magnitude dependent completeness correction that we applied in Section 2.2.2 could also introduce spurious angular power.

Of our 4000 modes, 147 are purely angular (see Fig. 6 for an example), and as described in Section 3.2, the remaining 3853 are orthogonal to them. This means that to first order, angular problems affect only these 147 PKL coefficients x_i. We repeated our entire analysis with these coefficients discarded, and found that the error bars became so large for $k \lesssim 0.03\,h\,\mathrm{Mpc}^{-1}$ that no signal could be detected there. In other words, the information on the power spectrum on the very largest scales comes mainly from the purely angular modes. On smaller scales, the measured power spectrum remained essentially unchanged. Although we have no indication that angular problems are actually present, it may be prudent to follow the 2dFGRS team and discard the information on the very largest scales – to be conservative, we therefore use only the measurements for $k \geqslant 0.01\,h\,\mathrm{Mpc}^{-1}$ in our likelihood analyses (for β, r and cosmological parameters).

5.3.2 Robustness to problems with the radial selection function

45 of our 4000 modes are purely radial (see Fig. 6 for an example), and are to first order the only ones affected by mis-estimates of the radial selection function $\bar{n}(r)$. Since accurate k-corrections and evolution modelling are notoriously challenging to perform, we repeated our entire analysis with these 45 modes omitted as a precaution. This resulted in a slight increase in error bars on the largest scales, but much less noticeable than when we removed the angular modes as described above. This can be readily understood geometrically. If we count the number of modes that probe mainly scales $k < k_*$, then the number of purely radial, purely angular and arbitrary modes will grow as k_*, k_*^2 and k_*^3, respectively. This means that 'special' modes (radial and angular) will make up a larger fraction of the total pool on large scales (at small k), and that the purely radial ones will be outnumbered by the purely angular ones.

Percival et al. (2001) report that slight changes in $\bar{n}(r)$ did not have a strong effect on the recovered 2dFGRS power spectrum, and we confirm this. We repeated our analysis with a number of different radial selection functions $\bar{n}(r)$, including the one from Colless et al. (2001) (the dashed curve in Fig. 3), finding only changes smaller than the error bars for $P(k)$ on the largest scales and no noticeable changes for larger k.

A final end-to-end test for problems with any special (angular, radial, or local group) modes is shown in Fig. 25. Here we have repeated the entire analysis twice, once excluding all the special modes and once using *only* the special modes (except the monopole). The latter is seen to give quite large error bars since only 196 modes are used (4 local group, 147 angular and 45 radial), but all three are seen to be reassuringly consistent. In contrast, systematic problems with any special modes would tend to add power to the special modes. This shows that any misestimates of special modes is having a negligible impact on our final results.

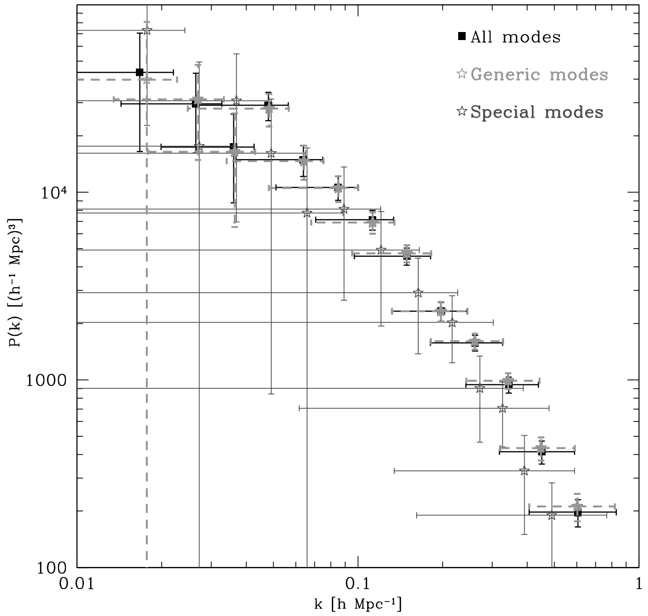

Figure 25. Constraints on excess power in special modes. Our 2dF power spectrum measurements from Fig. 17 are averaged into fewer bands and compared with measurements using only special (radial, angular and local group) modes and only generic (the remaining) modes (dashed).

5.4 Non-linearity issues

A key assumption (essentially the only one) underlying our analysis is that the Kaiser (1987) linear perturbation theory approach to redshift space distortions is valid. This approximation is known to break down on small scales where non-linear effects become important, which is why we have limited our analysis to large scales.

To be more precise, our basic measurement of $P_{\mathrm{gg}}(k)$, $P_{\mathrm{gv}}(k)$ and $P_{\mathrm{vv}}(k)$ assumes nothing at all, and measures the quantities that reduce to the monopole, quadrupole and hexadecapole of power in the small-angle approximation (Hamilton 1997c). However, relating these three measured functions to $\beta(k)$ and $r(k)$ via equation (22) does require the Kaiser approximation to be valid.

Substantial progress has recently been made in quantifying non-linear effects on redshift distortions, using both perturbation theory, gravitational N-body simulations and semianalytic galaxy formation theory (Hatton & Cole 1997, 1999; Scoccimarro, Couchman & Frieman 1999a; Heavens et al. 1999; Scoccimarro, Zaldarriaga & Hui 1999b; Hamilton 2000; Seljak 2001; Scoccimarro & Sheth, in preparation). The consensus is that non-linear effects may be important even on scales as large as $k \sim 0.1 - 0.3\,h\,\mathrm{Mpc}^{-1}$, although the critical scale is sensitive to the type of galaxies involved via their bias properties (Seljak 2001). Moreover, a generic smoking-gun signature of non-linear effects is found to be that the ratio $P_{\mathrm{gv}}(k)/P_{\mathrm{gg}}(k)$ starts dropping and eventually becomes negative, as non-linear fingers-of-god reverse the signature of linear infall. The ratio $P_{\mathrm{vv}}(k)/P_{\mathrm{gg}}(k)$ increases sharply in this regime.

Ideally, to do full justice to the 2dFGRS data set, one would like to perform a suite of non-linear simulations until a realistic biasing scheme is found that reproduces all observed characteristics of the data. The fast PTHalos approach (Scoccimarro & Sheth 2002) suggests that such an ambitious approach may ultimately be feasible. In the interim, the results obtained with analytic approximations must be interpreted with great caution. Peacock et al. (2001) use the widespread approach of adding a nuisance parameter to the Kaiser formula, interpreted as a small-scale velocity dispersion, and marginalizing over it. This gives $\beta = 0.43 \pm 0.07$ from 141 000 2dFGRS galaxies. Hatton & Cole (1999) and Scoccimarro & Sheth (in preparation) argue that this is approximation is inaccurate, underestimating the non-linear corrections (hence underestimating β) on large scales, and that the approximation of Hatton & Cole (1999) is preferable.

Given these important uncertainties, we adopt a more empirical approach, using the above-mentioned P_{gv}-drop in the data as an indicator of where to stop trusting the results. This was also done in the PSCz analysis of HTP00, where β was found to start dropping for $k \gtrsim 0.3\,h\,\mathrm{Mpc}^{-1}$. Fig. 13 shows no indication of $P_{\mathrm{gv}}(k)/P_{\mathrm{gg}}(k)$ (basically the quadrupole-to-monopole ratio) dropping, suggesting that our linear approximation is not seriously biasing our results on the large scales probed by our PKL modes (which recover information fully down to $k \sim 0.1$ as described above).

To quantify further the effect of non-linearities empirically, we performed our entire analysis five times with different levels of finger-of-god (FOG) compression as described in Section 3.1. The five curves in Fig. 14 correspond to progressively more aggressive compression with overdensity cut-offs $1 + \delta_{\mathrm{c}} = \infty$, 200, 100, 50 and 25, respectively. This corresponds to 0,6677, 7820, 8643 and 9124 FOGs compressed, involving 0,18 544, 24 031, 29 807 and 36 098 galaxies, respectively. Fig. 14 shows that more aggressive FOG-compression has an effect with the expected sign, increasing the best-fitting β-value for $k \gtrsim 0.1$, and that the effect is reassuringly small compared with the statistical error bars. Since a cluster is expected to have an overdensity around 200 when it virializes, higher later because the background density drops, thresholds $1 + \delta_{\mathrm{c}} < 100$ are likely to be overkill – we included the cases $1 + \delta_{\mathrm{c}} = 50$ and 25 in the figure merely to explore an extreme range of remedies. By removing essentially all structures that are elongated along the line of sight, one of course creates an artificial excess of flattened structures, leading to an overestimate of β. In conclusion, we believe that our estimate $\beta = 0.49 \pm 0.16$ is not severely affected by non-linearities. A conservative approach would be to take our measurement without FOG compression and use it merely as a lower limit, giving $\beta > 0.26$ at 90 per cent confidence.

Non-linearities affect our analysis in a different way as well, leading to slight underestimates of error bars. Our power spectrum measurements are simply certain second moments of the data, and remain valid regardless of whether the underlying density field is Gaussian or not. The power spectrum variance, however, involves fourth moments, and we have computed our error bars by making the Gaussian approximation to calculate these moments. The standard rule of thumb is that this approximation underestimates the error bars on the correlation function $\xi(r)$ by a factor $[1 + \xi(r)]^{1/2}$. Norberg et al. (2001a) fit the 2dFGRS correlation function to a power law $\xi(r) = (r/r_*)^{-\gamma}$ with correlation length $r_* = 4.9\,h^{-1}$ Mpc and slope $\gamma = 1.71$. Taking $k \sim \pi/r$, this gives error bar correction factors $[1 + (r_* k/\pi)^{\gamma})]^{1/2} \approx 2$ per cent, 7 and 13 per cent at 0.1, 0.2 and $0.3\,h\,\mathrm{Mpc}^{-1}$, respectively. Here $\xi(r)$ should refer to the correlation function of the matter, not of the galaxies, so if the 2dFGRS galaxies are biased with $b > 1$, the correction factors will be smaller. In conclusion, although non-linear error bar corrections certainly become important on very small scales, they are likely to be of only minor importance on the large scales $k < 0.3\,h\,\mathrm{Mpc}^{-1}$ that are the focus of this paper.

5.5 Bias issues

Although our basic measurement of $P_{\mathrm{gg}}(k)$, $P_{\mathrm{gv}}(k)$ and $P_{\mathrm{vv}}(k)$ assumes nothing about biasing, a bias model is obviously necessary before the results can be used to constrain cosmological models. We therefore comment briefly on the bias issue here.

Substantially larger data sets such as the complete SDSS catalogue hold the promise of measuring $\beta(k)$ and $r(k)$ with sufficient accuracy to quantity their scale-dependence, if any. Fig. 13 shows that our present sample is still not quite large enough to place strong constraints of this type.

An alternate route to constraining $b(k)$ involves comparing the clustering amplitudes of various subsamples, selected by, say, luminosity or spectral type. Such comparisons can also constrain r directly (Tegmark & Bromley 1999; Blanton 2000). It has been long known that bright elliptical galaxies are more clustered than spirals, presumably because the former are more likely to reside in clusters. Recent subsample analysis of the 2dFGRS (Norberg et al. 2001a) and SDSS (Zehavi et al. 2002) have confirmed and further quantified this effect.

Since recent cosmological parameter analyses using $P(k)$-measurements (most recently Wang, Tegmark & Zaldarriaga 2002; Efstathiou et al. 2002) have assumed that the bias factor b is scale-independent on linear scales, it is important to note that slight scale-dependence of bias is likely to be present in $P_{\mathrm{gg}}(k)$-measurements from a heterogeneous galaxy sample such as the 2dFGRS. Most of the information about $P_{\mathrm{gg}}(k)$ on large scales comes from distant parts of the survey, where bright ellipticals are over-represented since dimmer galaxies get excluded by the faint magnitude limit. This could cause $b(k)$ to rise as $k \to 0$. If uncorrected, this effect could masquerade as evidence for a redder power spectrum, i.e. one with a smaller spectral index n.

Fig. 17 indeed suggests slightly more 2dFGRS power on the largest scales than currently favoured cosmological models with constant bias would suggest, although this excess may also be caused by the angular or radial issues mentioned above. Detailed power spectrum analysis of subsamples should settle this issue.

6 DISCUSSION AND CONCLUSIONS

To place our results in context, we will now briefly discuss how they compare with other recent power spectrum measurements and with cosmological models.

6.1 Comparison with other surveys

Fig. 26 compares our 2dFGRS power spectrum measurements from Fig. 17 (averaged into fewer bands to reduce clutter) with measurements from other recent surveys. The PSCz and UCZ redshift surveys were analysed with the same basic method that we have

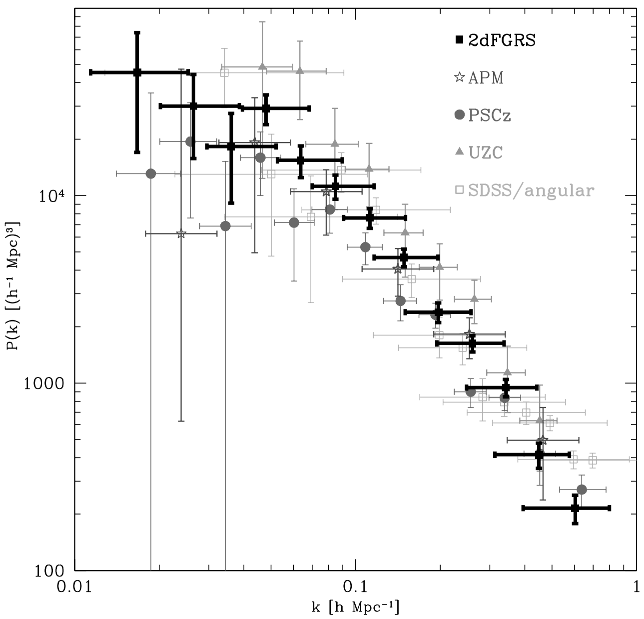

Figure 26. Comparison with other power spectrum measurements. Our 2dF power spectrum measurements from Fig. 17 are averaged into fewer bands and compared with measurements from the PSCz (HTP00) and UZC (this work) redshift surveys as well as angular clustering in the APM survey (Efstathiou & Moody 2001) and the SDSS (the points are from Tegmark et al. 2002) for galaxies in the magnitude range $21 < r' < 22$ – see also Dodelson et al. (2002).

employed here,[2] so a direct comparison involves no method-related interpretational issues. The 2dFGRS sample is seen to be slightly more biased than PSCz, but slightly less biased than UZC. Fig. 26 also suggests that 2dFGRS may have a slightly redder power spectrum than PSCz. This would also be consistent with the scale-dependent bias scenario mentioned above – the PSCz survey would probably be less afflicted than 2dFGRS, since the *IRAS*-selected galaxies in PSCz tend to avoid clusters.

Although the 2dFGRS error bars are seen to be small compared the PSCz and UZC ones, due to the larger sample size and survey volume, the horizontal bars show that the 2dFGRS window functions are somewhat broader. This is easy to understand: whereas PSCz and UZC cover large contiguous sky regions, the 2dFGRS sky coverage is currently fragmented into a multitude of regions of small angular extent, exacerbating aliasing problems. Indeed, since the characteristic width of 2dFGRS patches in the narrowest direction is more than an order of magnitude smaller than for PSCz or UZC (of order 2° rather than ∼60°), the fact that the windows are only 2–3 times wider reflects the quality of the 2dFGRS survey design and the power of the quadratic estimator method.

The remaining two power spectra are interesting since they were measured without use of redshift information and thus without the additional complications introduced by redshift space distortions. The APM points are from the likelihood analysis of Efstathiou & Moody (2001), using a few million galaxies, and reflect the full uncertainty even on the largest scales. Here the horizontal bands have a different interpretation, indicating the bands used in the likelihood analysis. Note that although the 2dFGRS galaxies are a subset of the APM galaxies, they need not have the exact same bias. Since the 2dFGRS subset involves on average brighter and more luminous galaxies, one might expect them to be slightly more clustered. The SDSS points (from Tegmark et al. 2002) are for about a million galaxies in the magnitude range $21 < r' < 22$, and the horizontal bars have the same interpretation as for the 2dFGRS points (redshift information obviously helps tighten up the windows). In contrast, the parametrized SDSS power spectrum in Dodelson et al. (2002) can be interpreted like the APM one.

A direct comparison of our power spectrum results with those reported by the 2dFGRS team (Percival et al. 2001) is unfortunately not possible at this time, since their window functions are of crucial importance and have not yet been made publicly available. However, an indirect comparison is possible as described in the next section, indicating good agreement. Our β-constraints are consistent with those reported in Peacock et al. (2001).

6.2 Cosmological constraints

Fig. 27 compares our 2dFGRS measurements with theoretical predictions from a series of models. No corrections have been made for non-linear evolution or scale-dependent bias. The measurements are seen to be in good agreement with both our simple BBKS prior and the recent concordance model from Efstathiou et al. (2002) – specifically, this is fit B from their paper, a flat scale-invariant scalar model with $\Omega_\Lambda = 0.71, h = 0.69$, baryon density $\omega_b = 0.021$ and dark matter density $\omega_d = 0.12$. ($\omega_b \equiv h^2\Omega_b$, $\omega_d \equiv h^2\Omega_d$.) Both of these are of course good fits by construction: we iterated our analysis until we found a prior that was consistent with the data, and Efstathiou et al. (2002) searched for models fitting both the 2dFGRS power

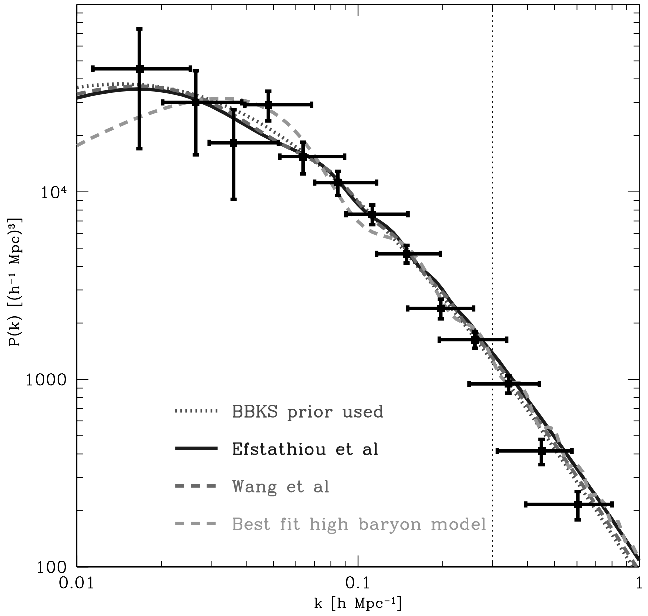

Figure 27. Our 2dF power spectrum measurements from Fig. 17 are averaged into fewer bands and compared with theoretical models. The BBKS model is the wiggle-free prior used for our calculation. The flat ΛCDM ‘concordance’ models from Wang et al. (2002) and Efstathiou et al. (2002), both renormalized to our 2dF measurements, are seen to be quite similar. The wigglier curve corresponds to the best-fit high baryon model in the upper right corner of Fig. 28. Only data to the left of the dashed vertical line are included in our fits.

---

[2] Since the UZC analysis in PTH01 did not include redshift space distortions, we performed a complete reanalysis of that data set for this figure, expanding the 13 342 galaxies surviving the cuts described in PTH01 in 1000 PKL modes.

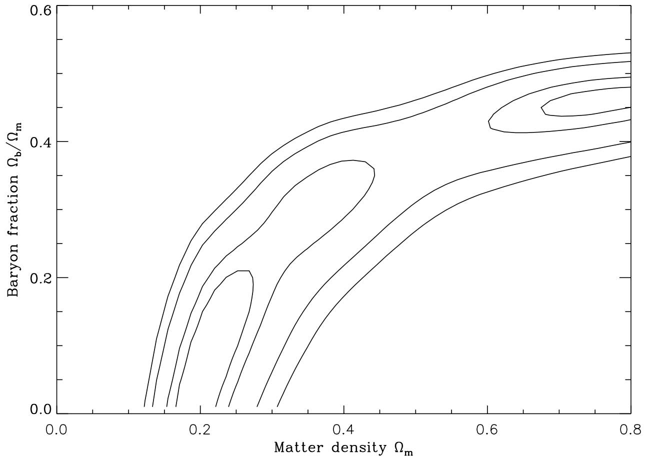

Figure 28. Constraints in on the matter density Ω_m and the baryon fraction Ω_b/Ω_m from the linear power spectrum over the range $0.01\,h\,\mathrm{Mpc}^{-1} < k < 0.3\,h\,\mathrm{Mpc}^{-1}$, after marginalizing over the power spectrum amplitude. These constraints assumes a flat, scale-invariant cosmological model with $h = 0.72$. For comparison with Percival et al. (2001), contours have been plotted at the level for one-parameter confidence of 68 per cent and two-parameter confidence of 68 per cent, 95 per cent and 99 per cent (i.e., $\chi^2 - \chi^2_{\rm min} = 1, 2.3, 6.0, 9.2$. Marginalizing over the Hubble parameter h and limiting the analysis to scales $k < 0.15h$/Mpc as in Percival et al. (2001) further weakens the constraints.

spectrum and CMB data. However, the fact that the Efstathiou et al. (2002) model fits our data so well provides an important cross-check between the 2dFGRS team power spectrum measurement (Percival et al. 2001) and ours, indicating good agreement.

Fig. 27 also shows the concordance model from Wang et al. (2002), resulting from a fit to all CMB data and the PSCz galaxy power spectrum. It is a flat scalar model with $\Omega_\Lambda = 0.66$, $h = 0.64$, baryon density $\omega_b = 0.020$, dark matter density $\omega_d = 0.12$ and a slight red-tilt, $n_s = 0.91$, here renormalized to the PSCz data. The fact that these pre-2dF and post-2dF concordance models agree so well is a reassuring indication that such multiparameter analyses are converging to the correct answer, and that the final numbers are not overly sensitive to bias issues or methodological technicalities.

A full multiparameter analysis of our results along the lines of Wang et al. (2002) and Efstathiou et al. (2002) is clearly beyond the scope of the present paper. However, since evidence for baryonic wiggles in the galaxy power spectrum has generated strong recent interest, first from the PSCz data (HTP00) and then more strikingly from the 2dF data (Percival et al. 2001; Miller et al. 2001), we perform a limited analysis to address the baryon issue.

We consider flat scale-invariant scalar models parametrized by the total matter content Ω_m, the baryon fraction Ω_b/Ω_m, the hubble parameter h and the spectral index n_s. We map out the likelihood function $L = e^{-\chi^2/2}$ using equation (31) on a fine grid in this parameter space, and compute constraints on individual parameters by marginalizing over the other parameters. Fig. 28 shows the result of fixing $n_s = 1$ and $h = 0.72$, the best-fitting value from Freedman et al. (2001). Here the axes have been chosen to facilitate comparison with Fig. 5 from Percival et al. (2001).[3] The general agreement between the two figures is seen to be good, both in terms of the shape and location of the banana-shaped degeneracy track, and in that there are two distinct favoured regions – a low-baryon solution like the concordance models in Fig. 27 and a high-baryon solution that is inconsistent with both Big Bang Nucleosynthesis (Burles, Nollett & Turner 2001) and CMB constraints. To illustrate the nature of the banana degeneracy in Fig. 28, we have plotted the best-fitting high-baryon model in Fig. 27. It has $\Omega_m = 0.75$ and $\omega_b = 0.18$, and is seen to provide a slightly better fit to the data around $k = 0.04\,h\,\mathrm{Mpc}^{-1}$ at the expense of slight difficulties on smaller scales.

There is, however, one notable difference between Fig. 28 and its twin in Percival et al. (2001). Whereas the latter excluded $\Omega_b/\Omega_m = 0$, we find no significant detection of baryons. This is of course not an indication of problems with either analysis, since the Percival et al. figure excludes zero baryons only at modest significance. Most importantly, as emphasized by Efstathiou et al. (2002), the constraints get much weaker when allowing small variations in other parameters, most strikingly the spectral index n_s. We confirm this effect by marginalizing over n_s and h with various priors. This means that the full statistical power of the complete 2dF and SDSS data sets will be needed to provide unequivocal evidence for baryonic signatures in the galaxy distribution.

6.3 Outlook

We have computed the real-space power spectrum and the redshift-space distortions of the first 10^5 galaxies in the 2dFGRS using pseudo-Karhunen-Loève eigenmodes and the stochastic bias formalism, providing easy-to-interpret uncorrelated power measurements with narrow and well-behaved window functions in the range $0.01\,h\,\mathrm{Mpc}^{-1} < k < 1\,h\,\mathrm{Mpc}^{-1}$. A battery of systematic error tests indicate that the survey is not only impressive in size, but also unusually clean.

Galaxy redshift surveys are living up to expectations. The striking early successes of the 2dFGRS and SDSS projects have firmly established galaxy redshift surveys as a precision tool for constraining cosmological models. However, it is important to bear in mind that this is only the beginning, and that many of the most exciting cosmological applications of these surveys still lie ahead. As discussed above, detailed comparisons with grids of fast simulations are likely to place information extracted from redshift distortions on a firmer footing and allow substantially more velocity information to be extracted from translinear scales. A bivariate analysis of how clustering depends jointly on both spectral type and luminosity should improve our quantitative understanding of biasing and allow possibilities such as the above-mentioned artificial red-tilt to be quantified and eliminated. With such progress combined with an order-of-magnitude increase in sample size, to more than 10^6 galaxies from 2dFGRS and SDSS combined, exciting opportunities will abound over the next few years, from definitive constraints on baryons and neutrinos to things that have not even been thought of yet.

ACKNOWLEDGMENTS

The authors wish to thank the 2dFGRS team for kindly making the data from this superb survey public and Shaun Cole, Mathew Colless and Karl Glazebrook in particular for helpful information about technical survey details. Thanks to Martin Kunz and Michael Vogeley for helpful comments. Support for this work was provided by NSF grants AST-0071213 & AST-0134999, NASA grants NAG5-9194, NAG5-11099 & NAG5-10763, the University of

---

[3] As a technical point, Percival et al. included band powers up to a nominal wavenumber $k = 0.15$ in their figure. Since our window functions are narrower, we have included band powers up to $k = 0.3$ in Fig. 28 to ensure that we do not use less small-scale information.

Pennsylvania Research Foundation, the Zaccheus Daniel Foundation and the David and Lucile Packard Foundation. MT is a Cottrell Scholar of Research Corporation.

REFERENCES

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15 (BBKS)
Blanton M., 2000, ApJ, 544, 63
Blanton M., Cen R., Ostriker J. P., Strauss M. A., Tegmark M., 2000, ApJ, 531, 1
Bond J. R., 1995, Phys. Rev. Lett., 74, 4369
Bond J. R., Jaffe A. H., Knox L. E., 2000, ApJ, 533, 19
Bunn E. F., 1995, PhD Thesis, U.C. Berkeley
Burles S., Nollett K. M., Turner M. S., 2001, ApJ, 552, L1
Cole S. et al., 2001, MNRAS, 326, 255
Coles P., 1993, MNRAS, 262, 1065
Coles P., Melott A., Munshi D., 1999, ApJ, 521, 5
Colless M. et al., 2001, MNRAS, 328, 1039, C01
Connolly A. et al., 2001, ApJ, in press (astro-ph/0107417)
Courteau S., van den Bergh S., 1999, AJ, 118, 337
Cross N. et al., 2001, MNRAS, 324, 825
Davis M., Geller M. J., 1976, ApJ, 208, 13
Dekel A., Lahav O., 1999, ApJ, 520, 24
Dodelson S. et al., 2002, ApJ, 572, 140
Dressler A., 1980, ApJ, 236, 351
Efstathiou G., Moody S. J., 2001, MNRAS, 325, 1603
Efstathiou G. et al., 2002, MNRAS, 330, L29
Falco E. E. et al., 1999, PASP, 111, 438
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1993, ApJ, 402, 42
Fisher K. B., Scharf C. A., Lahav O., 1994, MNRAS, 266 219
Freedman W. L., 2001, ApJ, 553, 47
Fry J. N., 1996, ApJ, 461, L65
Fry J. N., Gaztañaga E., 1993, ApJ, 413, 447
Giavalisco M., Steidel C. C., Adelberger K. L., Dickinson M. E., Pettini M., Kellogg M., 1998, ApJ, 503, 543
Hamilton A. J. S., 1993, ApJ, 417, 19
Hamilton A. J. S., 1997a, MNRAS, 289, 285
Hamilton A. J. S., 1997b, MNRAS, 289, 295
Hamilton A. J. S., 1997c, astro-ph/9708102
Hamilton A. J. S., 2000, MNRAS, 312, 257
Hamilton A. J. S., 2001, MNRAS, 322, 419
Hamilton A. J. S., Culhane M., 1996, MNRAS, 278, 73
Hamilton A. J. S., Tegmark M., 2000, MNRAS, 312, 285
Hamilton A. J. S., Tegmark M., 2002, MNRAS, 330, 506
Hamilton A. J. S., Tegmark M., Padmanabhan N., 2000, MNRAS, 317, L23, HTP00
Hatton S. J., Cole S., 1998, MNRAS, 296, 10
Hatton S. J., Cole S., 1999, MNRAS, 310, 1137
Heavens A. F., Taylor A. N., 1995, MNRAS, 483, 497
Heavens A. F, Matarrese S., Verde L., 1999, MNRAS, 301, 797
Huchra J. P., Geller M. J., de Lapparent V., Corwin G., Jr., 1990, ApJS, 72, 433
Kaiser N., 1987, MNRAS, 227, 1
Karhunen K., 1947, Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Kirjapaino oy. sana, Helsinki
Lineweaver C. H., Tenorio L., Smoot G. F., Keegstra P., Banday A. J., Lubin P., 1996, ApJ, 470, 38
Lynden-Bell D., 1971, MNRAS, 155, 95
Madgewick D. et al., 2002, MNRAS, 333, 133
Mann R. G., Peacock J. A., Heavens A. F., 1998, MNRAS, 293, 209
Matsubara T., Szalay A. S., Landy S. D., 2000, ApJ, 535, 1
Miller C. J., Nichol R. C., Batuski D. J., 2001, Science, 292, 2302
Norberg P. et al., 2001a, MNRAS, 328, 64
Norberg P. et al., 2001b, astro-ph/0111011
Padmanabhan N., Tegmark M., Hamilton A. J. S., 2001, ApJ, 550 52
Peacock et al., 2001, Nat, 410, 169
Pen U., 1998, ApJ, 504, 601
Percival W. J. et al., 2001, MNRAS, 327, 1297
Saunders W. et al., 2000, MNRAS, 317, 55
Schectman S. A., Landy S. D., Tucker A. D. L., Lin H., Kirschner R. P., Schechter P. L., 1996, ApJ, 470, 172
Scherrer R. J., Weinberg D. H., 1998, ApJ, 504, 607
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629
Scoccimarro R., Couchman H. M. P., Frieman J. A., 1999a, ApJ, 517, 531
Scoccimaro R., Zaldarriaga M., Hui L., 1999b, ApJ, 527, 1
Scranton R. et al., 2002, ApJ, in press (astro-ph/0107416)
Seljak U., 2001, MNRAS, 325, 1359
Szalay A. et al., 2002, ApJ, in press (astro-ph/0107419)
Taylor A. N., Ballinger W. E., Heavens A. F., Tadros H., 2001, MNRAS, 327, 689
Tegmark M., 1997, Phys. Rev. D, 55, 5895
Tegmark M., 1998, in Colombi S., Mellier Y., eds, Wide Field Surveys in Cosmology. Editions Frontières, Paris, p. 43
Tegmark M., Bromley B. C., 1999, ApJ, 518, L69
Tegmark M., de Oliveira-Costa A., 2001, Phys. Rev. D, 64, 063001
Tegmark M., Hamilton A. J. S., 1998, in Olinto A. V., Frieman J. A., Schramm D., eds, Relativistic Astrophysics Cosmology. World Scientific, Singapore, p. 270
Tegmark M., Peebles P. J. E., 1998, ApJ, 500, 79
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Tegmark M., Hamilton A. J. S., Strauss M. A., Vogeley M. S., Szalay A. S., 1998, ApJ, 499, 555 (THSVS98)
Tegmark M., Zaldarriaga M., Hamilton A. J. S., 2001, Phys. Rev. D, 63, 43007
Tegmark M. et al., 2002, ApJ, 571, 191
Turner E. L., 1979, ApJL, 228, L51
Vogeley M. S., 1998, in Hamilton D., ed., Ringberg Workshop on Large-Scale Structure. Kluwer, Amsterdam, astro-ph/9805160
Vogeley M. S., Szalay A. S., 1996, ApJ, 465, 34
Wang X., Tegmark M., Zaldarriaga M., 2002, Phys. Rev. D, 65, 123001
York D. et al., 2000, AJ, 120, 1579
Zehavi I. et al., 2002, ApJ, 571, 172

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