Mon. Not. R. Astron. Soc. 000, 000–000 (0000) Printed 12 April 1999 (MN LATEX style file v1.4)

Spherical Harmonic Analysis of the PSCz Galaxy Catalogue: Redshift distortions and the real–space power spectrum

H. Tadros,[1,2] W. E. Ballinger,[3,2] A. N. Taylor,[3], A. F. Heavens,[3] G. Efstathiou[4], W. Saunders[3], C. S. Frenk[5], O. Keeble[6], R. McMahon[4], S. J. Maddox[4], , S. Oliver[6], M. Rowan-Robinson[6], W. J. Sutherland[2] and S. D. M. White[7]

[1] *Astronomy Centre, University of Sussex, Falmer, Brighton UK.*
[2] *Department of Physics, University of Oxford, Keble Road, Oxford, UK.*
[3] *Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh.*
[4] *Institute of Astronomy, Cambridge, UK.*
[5] *Department of Physics, University of Durham, UK.*
[6] *Imperial College, University of London, UK.*
[7] *MPI-Astrophysik, Garching, Germany.*

12 April 1999

ABSTRACT

We apply the formalism of spherical harmonic decomposition to the galaxy density field of the IRAS PSCz redshift survey. The PSCz redshift survey has almost all–sky coverage and includes IRAS galaxies to a flux limit of 0.6 Jy. Using maximum likelihood methods to examine (to first order) the distortion of the galaxy pattern due to redshift coordinates, we have measured the parameter $\beta \equiv \Omega^{0.6}/b$. We also simultaneously measure (a) the undistorted amplitude of perturbations in the galaxy distribution when a parameterised power spectrum is assumed, or (b) the shape and amplitude of the real–space power spectrum if the band–power in a set of passbands is measured in a step–wise fashion. These methods are extensively tested on a series of CDM, ΛCDM and MDM simulations and are found to be unbiased.

We obtain consistent results for the subset of the PSCz catalogue with flux above 0.75 Jy, but inclusion of galaxies to the formal flux limit of the catalogue gives variations which are larger than our internal errors. For the 0.75 Jy catalogue we find, in the case of a parameterised power spectrum, $\beta = 0.58 \pm 0.26$ and the amplitude of the real space power measured at wavenumber $k = 0.1h\mathrm{Mpc}^{-1}$ is $\Delta_{0.1} = 0.42 \pm 0.03$. Freeing the shape of the power spectrum we find that $\beta = 0.47 \pm 0.16$ (conditional error), and $\Delta_{0.1} = 0.47 \pm 0.03$. The shape of the real–space power spectrum is consistent with a $\Gamma = 0.2$ CDM–like shape parameter, but does not strongly rule out a number of other models. Finally by combining our estimate of the amplitude of galaxy clustering and the distortion parameter we find the amplitude of mass fluctuations on a scale $k = 0.1\ h^{-1}$Mpc is $\Delta_\rho = 0.24\Omega_0^{-0.6}$, with an uncertainty of 50%.

Key words: cosmology: large-scale structure of Universe

1 INTRODUCTION

The pattern of galaxy clustering in redshift surveys is distorted because of peculiar (non-Hubble) velocities which contribute to the line of sight velocity measured by the galaxy redshift. This distortion makes the interpretation of the clustering more difficult, and must be corrected for when measuring statistical properties of the pattern, such as the 2-point correlation function or power spectrum. However the distortion itself contains important information about the motion of matter in the universe. Because the effect is only radial, the measured clustering pattern becomes anisotropic. The degree of anisotropy is determined by the distortion parameter⋆ $\beta \equiv \Omega^{0.6}/b$, where Ω is the cosmological mass–

⋆ A distinction must be made between the distortion parameters for different galaxy types, since their relative biases are known to density parameter and $b^2 \equiv P(k)_{\rm gal}/P(k)_{\rm mass}$ is the bias parameter, relating clustering in mass to clustering in galaxies.

By measuring the degree of distortion one can hope to constrain β, and hence the density parameter and bias. In addition, measurement of the real–space (i.e. undistorted) power spectrum can put strong constraints on theories of structure formation. A major aim of large–scale redshift surveys is to constrain the shape and position of the break in the power spectrum as it turns over into the primordial spectrum. This would demonstrate a strong link between the clustering responsible for the microwave background fluctuations and that found in redshift surveys, as well as putting strong constraints on the slope of the connecting spectrum. This in turn provides constraints on inflationary models for the generation of structure.

A comprehensive overview of the literature to date concerning the analysis of redshift space distortions is given by Hamilton (1997). Here we shall outline some of the main results leading up to this analysis.

Jackson (1972) first noticed the distortion when viewing a cluster of galaxies in redshift–space, where peculiar velocities are generated by virial motion. Kaiser (1987) derived the main features of the redshift distortion in the linear regime, simplifying the analysis by assuming plane density waves and a distant observer approximation. The major result was that power is boosted in redshift space by a factor $(1+\beta\mu_k^2)^2$, where μ_k is the cosine of the angle between wavevector and line of sight. Hamilton (1993) suggested expanding the distorted power in Legendre polynomials, since the main distortion is quadrupole in nature. Taking the ratio of quadrupole to monopole power yields a function in terms of β only, and one could hope to measure the distortion independently of power. Applying this to the IRAS 2Jy survey Hamilton (1993) found $\beta = 0.69 \pm 0.26$. Hamilton further applied this method to a merged QDOT + 1.2 Jy redshift survey and found $\beta = 0.69 \pm 0.20$. However Hamilton also suggested that the merged survey was systematically biased due to a difference in clustering between the near and far regions of the QDOT sample.

Cole, Fisher & Weinberg (1994) and Cole, Fisher & Weinberg (1995) tested the quadrupole to monopole estimator on simulations and applied it to the IRAS 1.2Jy and QDOT surveys and found $\beta = 0.52 \pm 0.15$. Peacock & Dodds (1994) used a radially projected Gaussian smoothing function to model the effects of nonlinearity in the small scale velocity field. They found that $\beta = 1.0 \pm 0.2$ for IRAS galaxies, using the APM survey to constrain the real-space power spectrum (Baugh & Efstathiou 1993), and allowing a relative bias between the galaxy samples.

Taylor & Hamilton (1996) and Fisher & Nusser (1996) suggested that the main contribution to nonlinearity in redshift space was not from an incoherent velocity dispersion, but rather from coherent infall in the translinear regime. Using the Zel'dovich approximation they calculated the scale dependence of the quadrupole to monopole ratio, finding that the amplitude was still only a function of β. Taylor & Hamilton (1996) also found a dependence on the initial power spectrum, and the scaling behaviour of the wavenumber at which the quadrupole is zero. They applied these results to the merged QDOT + 1.2Jy surveys and found $\beta = 0.6 \pm 0.2$, assuming a local spectral index of $n = -1.5$, while Fisher & Nusser (1996) found $\beta = 0.6 \pm 0.2$ from the 1.2Jy survey alone.

differ. Since we are concerned with IRAS selected galaxies in this paper, we shall reserve β and b to refer to IRAS galaxies.

Fisher, Scharf & Lahav (1994) and Heavens & Taylor (1995) both dropped the plane parallel approximation and used a spherical harmonic decomposition to match the spherical nature of the IRAS redshift surveys. Fisher, Scharf & Lahav (1994) used a Gaussian window function in the radial direction and, applying a maximum likelihood formalism, found $\beta = 0.96 \pm 0.19$ for the 1.2Jy survey. Heavens & Taylor (1995) used spherical Bessel functions to decompose the density field radially, since these combine the property of being orthonormal and eigenfunctions of the Laplacian, with the advantage that each mode only picks up power from a narrow range of wavenumbers. In addition they applied a radial smoothing to correct for the effects of nonlinearity. Applying a maximum likelihood analysis they measured simultaneously β and the amplitude of a fixed power spectrum (characterised by the fractional r.m.s. in counts in $8h^{-1}$Mpc spheres), finding $\beta = 1.1 \pm 0.5$ and $\sigma_8 = 0.68 \pm 0.05$, where we quote marginal errors. Ballinger, Heavens & Taylor (1995) extended the analysis by dropping the fixed power spectrum and measuring the real–space band–power in a series of passbands in a step-wise fashion. With a free power spectrum, they found $\beta = 1.04$ with a conditional error of 0.3.

In this paper we shall apply the methods of Heavens & Taylor (1995) (hereafter HT) and Ballinger, Heavens & Taylor (1995) (hereafter BHT) to measure the degree of distortion and the amplitude and shape of the undistorted power spectrum of the IRAS Point Source Catalogue redshift survey (PSCz; Saunders *et al.* 1998).

The contents of the paper are as follows. In Section 2 we discuss the methods used, including some technical improvements to the original analysis. In Section 3 we extensively test the methods on a suite of simulated PSCz redshift surveys. In Section 4 we give a brief outline of the PSCz catalogue itself and highlight the checks made on possible sources of contamination to the present analysis. In Section 5 we present the results of the analysis for the redshift distortion and amplitude of the real–space power spectrum, and the shape of the real–space power spectrum measured in a series of passbands. In Section 6 we discuss the implications of these results and present our conclusions. There are also two technical appendices outlining the spherical harmonic formalism, deriving the main equations used in the analysis and detailing the construction of the mixing matrices. We also explain, in more detail, the technical improvements made to the analysis in the present work. We begin by discussing the methods to be used.

2 METHODS

In this paper we present two complementary analyses of the redshift distortions in the PSCz survey. The first analysis, described in detail in Section 2.1, is a two parameter maximum likelihood analysis of the data to obtain the value of $\beta = \Omega^{0.6}/b$ and the amplitude of the real space power spectrum at a wavenumber $k = 0.1\ h\mathrm{Mpc}^{-1}$† (assuming a parametrised form for the shape of the power spectrum). The details of this analysis follow HT with some technical improvements to the method concerning how the sky mask and Local Group motion are dealt with. These modifications are explained in Section 2.1 and in Appendix A. The second analysis, described in Section 2.2 follows BHT and is an extension of that described above. Here we perform a likelihood fit for β and the full real space power spectrum, sampled in a series of passbands. This has the advantage of relaxing the assumption of a fixed spectral shape. We begin by discussing the galaxy density fluctuations in redshift–space.

2.1 Density fluctuations in redshift–space

Following HT, the density field of the galaxy distribution $\rho(\boldsymbol{s})$ is expanded in terms of spherical harmonics, $Y_{\ell m}$, and a discrete set of spherical Bessel functions, j_ℓ,

$$\hat{\rho}_{\ell mn} = c_{\ell n} \int d^3 s\, \rho(\boldsymbol{s}) w(s) j_\ell(k_{\ell n} s)\, Y^*_{\ell m}(\theta, \phi)\,, \tag{1}$$

where $w(s)$ is an adjustable weighting function and $\boldsymbol{s}$ is the redshift-space position variable. The inverse transform is

$$\rho(\boldsymbol{s}) = \sum_{\ell mn} c_{\ell n} \rho_{\ell mn} j_\ell(k_{\ell n} s) Y_{\ell m}(\theta, \phi), \tag{2}$$

where the $c_{\ell n}$ are normalization constants and k_{ln} are discrete wavenumbers (see Appendix C of HT and Appendix A in this work for a definition).

The domain of the integral in equation (1) can be transformed from redshift space $\boldsymbol{s}$ to real space $\boldsymbol{r}$ by continuity, $d^3 s\, \rho(\boldsymbol{s}) = d^3 r\, \rho(\boldsymbol{r})$, leaving only the argument of the spherical Bessel function as a function of the redshift-space distance s. Noting that redshift distortions only affect radial coordinates, the transformation can be written

$$s(\boldsymbol{r}) = r + \frac{u(\boldsymbol{r})}{H_0}, \tag{3}$$

where $u = \hat{\boldsymbol{r}}.(\boldsymbol{v}(\boldsymbol{r}) - \boldsymbol{v}_{\rm LG})$ is the radial component of the peculiar velocity field in the Local Group frame and H_0 is the Hubble parameter. Note that the expansion of the density field is done in this frame, since we need to assume that $|u| \ll H_0 r$, which will only be the case in the near-field limit if we choose a frame in which $\boldsymbol{v} - \boldsymbol{v}_{\rm frame} \to \boldsymbol{0}$ as $r \to 0$. In linear theory the velocity field is related to perturbations in the mass density by the continuity equation

$$\nabla.\boldsymbol{v} = -H_0 \Omega^{0.6} \delta_\rho, \tag{4}$$

where $\Omega^{0.6} \approx d\ln\delta_\rho / d\ln a$ is the linear growth index and δ_ρ is the fractional matter overdensity. Expanding the continuity equation in spherical harmonics we find that the radial velocity field can be expressed as

$$u(\boldsymbol{r}) = \Omega^{0.6} \sum_{\ell mn} c_{\ell n} \delta_{\ell mn} k^{-1}_{\ell n} j'_\ell(k_{\ell n} r) Y_{\ell m}(\theta, \phi), \tag{5}$$

where $j'_\ell(z) \equiv dj_\ell(z)/dz$.

Expanding equation (1) to first order in $u = r - s$ we find

$$\hat{\rho}_{\ell mn} = (\rho_0)_{\ell mn} + \sum_{\ell' m' n'} W^{mm'}_{\ell\ell'} \left(\Phi^{nn'}_{\ell\ell'} + \beta V^{nn'}_{\ell\ell'}\right) \delta_{\ell' m' n'}, \tag{6}$$

where the mean coefficients $(\rho_0)_{\ell mn}$ are given in Appendix A, and include a term for the Local Group velocity, although the effect of this for $\ell \neq 1$ is very small. The transition matrices $\boldsymbol{W}$, $\boldsymbol{\Phi}$, and $\boldsymbol{V}$ describe the effects of the sky mask, the radial selection function and the first order redshift space distortion, respectively. The transition matrices are derived and defined in Appendix A.

In HT it was assumed that the sky mask was azimuthally symmetric. This simplified the analysis since the mask mixing matrix, $\boldsymbol{W}$, was then real. Given the higher surface density of galaxies in the PSCz we drop this assumption and use the exact PSCz mask (see Figure 7). Also, we do not apply an artificial galactic cut, as was done conservatively for the 1.2Jy survey (HT). The mask used for this analysis is described briefly in Section 4 and in Appendix B we describe in detail the methods used for mask generation and application.

Throughout this paper we have also applied a radial correction for the effects of small–scale nonlinear peculiar velocity fields (see HT, Section 4.4 and Appendix B, therein, and Appendix A in this work). This assumes that the effects of the 'fingers of God' due to virialised clusters and redshift measurement errors can be modelled by adding an incoherent random displacement to each galaxy radial distance

$$s' = s + \varepsilon(\boldsymbol{r}), \tag{7}$$

where $\langle\varepsilon\rangle = 0$, $\langle\varepsilon(\boldsymbol{r})\varepsilon(\boldsymbol{r}')\rangle = (\sigma_v^2/3H_0^2)\delta_D(\boldsymbol{r} - \boldsymbol{r}')$ and σ_v is the three-dimensional velocity dispersion. If the distribution of random displacement is assumed to be drawn from a Maxwellian distribution we should ensemble average over this distribution when forming the correlators of the density field, $\langle\delta\delta'\rangle$. But since the random displacements are incoherent we need only calculate the effects on the mean value of modes, since $\langle XY\rangle = \langle X\rangle\langle Y\rangle$ if X and Y are uncorrelated. The averaging effect on a given mode can be expressed as a convolution of the density field by a scattering matrix, $\boldsymbol{S}$ (see Appendix A for a definition).

The statistical properties of the redshift space density field can be encapsulated by the mode–mode correlation function

$$\langle \boldsymbol{D}\boldsymbol{D}^\dagger\rangle = \frac{1}{2}\sum \boldsymbol{S}\boldsymbol{W}(\Phi + \beta\boldsymbol{V})\boldsymbol{P}(\Phi + \beta\boldsymbol{V})^t \boldsymbol{W}^\dagger \boldsymbol{S}^t + \boldsymbol{N} \tag{8}$$

where the observable data vector, $\boldsymbol{D}$, is given by $D_{\ell mn} \equiv [\hat{\rho}_{\ell mn} - (\rho_0)_{\ell mn}]/\bar{\rho}$ and $\bar{\rho}$ is the mean number density, and $\boldsymbol{N}$ is the shot–noise contribution to the covariance matrix. We have assumed that the underlying Fourier modes of the density field are statistically isotropic and homogeneous:

$$\langle\delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle = (2\pi)^3 P(k)\delta_D(\boldsymbol{k} - \boldsymbol{k}'), \tag{9}$$

where δ_D is the Dirac delta function. Since the shot–noise term contains a part linear in the matrix $\boldsymbol{W}$, the mode–mode correlator is complex for a non-azimuthally-symmetric mask. This can be easily overcome by decomposing the transformed density field into real and imaginary parts:

$$\boldsymbol{D} = \Re e\boldsymbol{D} + i\Im m\boldsymbol{D}. \tag{10}$$

† Throughout this paper we shall use $h = H_0/100\,\mathrm{kms}^{-1}\ \mathrm{Mpc}^{-1}$, where H_0 is the Hubble parameter.

This is convenient since our hypothesis that $\boldsymbol{D}$ is a Gaussian random variable requires that the terms $\Re e\boldsymbol{D}$ and $\Im m\boldsymbol{D}$ are also Gaussian random variables. In the limit of full sky coverage these terms would also be uncorrelated, but the presence of a non-azimuthally symmetric mask mixes real and imaginary parts. Even so the likelihood function for the observed redshift density field can be written in Gaussian form:

$$\mathcal{L}[D|\beta, P(k)] = \frac{1}{(2\pi)^{N/2}|\boldsymbol{C}|^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{D}^t\boldsymbol{C}^{-1}\boldsymbol{D}\right], \qquad (11)$$

where $\boldsymbol{D} = (\Re e\boldsymbol{D}, \Im m\boldsymbol{D})$ is a data vector of real and imaginary parts of the density modes, with dimension its being set by the upper limit to the wavenumber. The data covariance matrix is

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{\Re e\Re e} & \boldsymbol{C}_{\Re e\Im m} \\ \boldsymbol{C}_{\Im m\Re e} & \boldsymbol{C}_{\Im m\Im m} \end{pmatrix} \qquad (12)$$

where $\boldsymbol{C}_{\Re e\Re e} = \langle \Re e\boldsymbol{D}\, \Re e\boldsymbol{D}^t\rangle$, $\boldsymbol{C}_{\Re e\Im m} = \langle \Re e\boldsymbol{D}\, \Im m\boldsymbol{D}^t\rangle$ and so on. From the symmetry relations given in Appendix B we see that this is a real, symmetric matrix.

Another refinement to the analysis of HT is the inclusion of $m = 0$ modes. Since these are real modes, which have twice the power as the $m \neq 0$ modes, we treat these separately.

Finally, we have also included a correction for the effect of the Local Group motion. In HT and BHT, we simply removed the dipole contribution to the harmonics. While this is exact for full sky coverage, for incomplete sky coverage there is a mixing of modes and the dipole leaks into higher modes, although at a level which was negligible for the PSCz mask used. Here we account for that leakage. Again technical details can be found in Appendix A.

The free parameters for the likelihood are then the redshift distortion parameter β and the real–space power spectrum $P(k)$. In this paper we shall consider simultaneous measurement of

(a) β and the normalisation of $P(k)$ for a fixed spectral shape,
(b) β and the shape and amplitude of $P(k)$.

The first method maximises the likelihood over a 2-dimensional parameter space, while the latter requires a maximisation over a 7-dimensional parameter space – one dimension for β and 6 dimensions for the step-wise power spectrum sampled uniformly in logarithmic pass–bands. In Section 2.2 we shall discuss the step–wise maximum likelihood method for measuring the real power spectrum. Before doing so there are two issues we shall deal with here – that of the correct weighting function to use in the transformation equation (1) and the choice of how to measure the normalisation of the power spectrum.

The minimum variance weighting scheme to use in the transformation equation (1) has been discussed by Feldman, Kaiser & Peacock (1994), HT, and Hamilton (1997). All three have found that for an undistorted, flux limited redshift survey the optimal weighting function for measuring power spectra is

$$w(r) = \frac{1}{1 + \phi(r)P_w(k)}, \qquad (13)$$

where $\phi(r)$ is the survey selection function in real space and $P_w(k)$ is the expected power spectrum. As can be seen one has to make an initial guess for the power spectrum to insert into the weighting function. This is not important: a poor choice would lead to larger than necessary error bars, but should not bias the solution. In any case, if one was concerned, one could iterate. We have chosen to use a $\Gamma = 0.2$ power spectrum for the analysis of the PSCz, where Γ is the CDM-like spectral shape parameter. In the tests on mock PSCz catalogues (Section 3) we used the $P_w(k)$ appropriate to the simulations. In both cases we have used the full k-dependence in the weighting scheme (cf HT, who used a k-independent P(k) in the analysis of the 1.2Jy survey). Note that this weighting function is strictly only optimal for estimating power spectra, and when $\beta = 0$. HT derived an approximate weighting scheme optimised for measuring β, with the main difference being to correct for the misplacement of weights in redshift space and to boost the power by a factor $(1 + \beta)$. Since these seem to have little effect on the maximum likelihood solution, we used the weighting function optimised for power.

In the two-parameter fits, we have chosen to express the power normalisation in terms of $\Delta_{0.1}$, where

$$\Delta_k^2 \equiv P(k)k^3/(2\pi^2) \qquad (14)$$

is the contribution to the overdensity variance per $\ln(k)$. This choice is motivated by the fact that we choose an upper wavenumber limit $k \simeq 0.1h\mathrm{Mpc}^{-1}$ to avoid nonlinear scales, and most of the modes are concentrated near the upper wavenumber limit. Hence it is the power at this wavenumber which is most strongly constrained and is the most insensitive to the shape of the power spectrum assumed. Since it is common to normalise spectra by the variance in spheres of radius 8 h^{-1}Mpc,

$$\sigma_8^2 = \int \frac{dk}{k}\Delta^2(k)\left[\frac{3j_1(kr_8)}{kr_8}\right]^2, \qquad (15)$$

we have fitted the relationship between σ_8 and $\Delta_{0.1}$ as a function of the CDM shape parameter, Γ. A fit accurate to a few percent is given by

$$\sigma_8^2 = 0.616\left[1 - 0.84\left(1 - \frac{\Gamma}{0.2}\right)\right]\left(\frac{\Delta_{0.1}}{0.5}\right)^2. \qquad (16)$$

A further consideration is the validity of the linear bias model used to relate the galaxy distribution to the peculiar velocity field, which appears in β. While a wavenumber of $k = 0.1h\mathrm{Mpc}^{-1}$ corresponds to a scale of about 60 h^{-1}Mpc, and one would hope biasing could be described by a local and linear model, we have no concrete guarantee that biasing is not still scale–dependent or non-local, as in the case of cooperative biasing (Bower *et al.* 1993). However realistic models of biasing mechanisms do seem to give some support to local biasing that is approximately linear on these scales (Kauffman, Nusser & Steinmetz 1997, Mann, Peacock & Heavens 1998, Heavens, Matarrese & Verde 1998).

Having described the main formalism for the measurement of distortions and real space power in this section, we now turn to the methods used to measure the shape and amplitude of the power spectrum in a stepwise fashion.

2.2 Stepwise maximum likelihood

This method fits the shape of a function in an essentially model-independent way – see Efstathiou, Ellis & Peterson

(1988). The function is divided into bins, and the amplitude of these bins is allowed to vary.

We exploit the fact that we have a good estimate for the power spectrum from previous studies, $P_0(k)$, so we write

$$P(k) = P_0(k) \times 10^{f_i}, \tag{17}$$

where the original power spectrum (CDM with $\Gamma = 0.2$ and $\Delta^2_{0.1} = 0.18$) is multiplied by the free function 10^{f_i}. The subscript $i = 1, 2, 3, \ldots$ labels each wavenumber bin corresponding to the range $k_i^{\rm min} < k < k_i^{\rm max}$. The parameters are introduced in this exponential form to ensure that the measured power spectrum is positive.

The parameters for the maximum likelihood fit are the distortion parameter β and the coefficients f_i; $i = 1, \ldots, n$. This construction essentially amounts to specifying the power spectrum at a series of wavenumbers, but the form chosen allows us to start the maximum likelihood search conveniently at $f_i = 0$, since we have a fairly good idea of what to expect. Note that, although the form suggests that the break scale is fixed, adjustment of the f_i can alter this in an essentially arbitrary way, which is the main reason for parametrising the power spectrum in bins. In HT, and in Fisher, Scharf & Lahav (1994), the power spectrum was parametrised to have a specific shape, whereas here it has a free form (apart from bin discreteness).

We begin by expanding the covariance matrix of equation (12) in powers of β;

$$\boldsymbol{C} = \sum_{n=0}^{2} \beta^n \boldsymbol{A}_n + \boldsymbol{N} \tag{18}$$

The covariance matrix can now be constructed with little computational cost as each sub-matrix can be written as a linear combination of contributions from each power-spectrum bin;

$$\begin{aligned} \boldsymbol{A}_0 &= 10^{f_1}\boldsymbol{A}_{01} + 10^{f_2}\boldsymbol{A}_{02} + \ldots, \\ \boldsymbol{A}_1 &= 10^{f_1}\boldsymbol{A}_{11} + 10^{f_2}\boldsymbol{A}_{12} + \ldots, \\ \boldsymbol{A}_2 &= 10^{f_1}\boldsymbol{A}_{21} + 10^{f_2}\boldsymbol{A}_{22} + \ldots, \end{aligned} \tag{19}$$

where $\boldsymbol{A}_{01}$ is the contribution to $\boldsymbol{A}_0$ from the first bin, and so on. The matrices $\boldsymbol{A}_{0i}$, $\boldsymbol{A}_{1i}$ and $\boldsymbol{A}_{2i}$ need only be constructed once and then the full A-matrices can be created each time from a simple linear sum. This requires far less computer time than fitting for a standard power spectrum parameter such as Γ. That would require the A-matrices to be constructed afresh from window function convolutions each time a parameter changed.

The maximizing procedure is straightforward, the $(n+1)$D parameter space is searched until a maximum of the likelihood function $\mathcal{L}[D|\beta, f_i; i = 1, \ldots, n]$ is found. This generalizes the method of HT, where, in effect, the f_i were all constrained to move up or down together.

If n is high, exploring the likelihood space with a grid can become computationally expensive. However, the maximum likelihood solution is easily found, since we have a good starting guess for the parameters, and can use a standard library routine to maximise the likelihood function with respect to the parameters. Ideally, we would like to construct complete hypersurfaces of constant likelihood, but this is impractical because of the dimensionality of the problem. Errors can be obtained by searching along the parameter axes until the likelihood drops by $\sqrt{e}$, but note that this ignores the correlations between parameters. This means that the errors are *conditional*: for each parameter they are calculated assuming that the other parameter values are known exactly – see Press *et al.* (1992; §15) for a discussion. Note in particular that there is a strong covariance between the amplitude of the power spectrum and β, and these errors do not properly reflect this (see HT, Fisher, Scharf & Lahav 1994).

By parametrising in bins, the fitted power spectrum is discontinuous. In principle there might be a case for adjusting the shape of the power spectrum so that it passes continuously through the maximum likelihood points; in the event there is little to be achieved by this, since the error bars are large enough so that a smooth curve is a perfectly acceptable fit.

3 TESTS ON SIMULATED SURVEYS

The methods described in Sections 2.1 and 2.2 have been applied to mock PSCz redshift surveys drawn from large N-body simulations. Before presenting the results from the PSCz redshift survey, we describe the result of applying these methods to the mock PSCz data sets.

3.1 Numerical simulations

The dissipationless N-body simulations that we have used to create the mock data-sets consist of 3 ensembles of 9 realizations each. The first two ensembles are of CDM-like universes and contain 160^3 particles within a periodic computational box of length $\ell_b = 600\ h^{-1}$Mpc. These simulations are similar to the ones used and discussed in detail by Croft & Efstathiou (1994a) but employ more particles within a larger computation box (the simulations of Croft and Efstathiou use 100^3 particles within a box of size $\ell_b = 300\ h^{-1}$Mpc). The simulations were run with the particle-particle-particle-mesh (P^3M) code described by Efstathiou *et al.* (1985) and model the gravitational clustering of dark matter in a CDM dominated universe with scale invariant initial density fluctuations. These two ensembles are as follows: the standard CDM model (Davis *et al.* 1985), *i.e.* a spatially flat universe with $\Omega_0 = 1$ and $h = 0.5$ (the SCDM ensemble); a spatially flat low density CDM universe with $\Omega_0 = 0.2$, $h = 1.0$, and a cosmological constant contribution $\Omega_\Lambda = \Lambda/(3H_0^2) = (1 - \Omega_0) = 0.8$ (the ΛCDM ensemble). The initial power spectra of the models are generated from the fitting function

$$P\,(k) \propto \frac{k}{\left[1 + \left[ak + (bk)^{\frac{3}{2}} + (ck)^2\right]^\nu\right]^{\frac{2}{\nu}}}, \tag{20}$$

where $\nu = 1.13$, $a = (6.4/\Gamma)\ h^{-1}$Mpc, $b = (3.0/\Gamma)\ h^{-1}$Mpc and $c = (1.7/\Gamma)\ h^{-1}$Mpc. Equation (18) is a good approximation to the linear power spectrum of scale-invariant CDM models with low baryon density, $\Omega_b \ll \Omega_0$ (Bond & Efstathiou 1984). The parameter Γ in equation (18) is equal to 0.5 for the SCDM ensemble and $\Gamma = 0.2$ for the ΛCDM ensemble. For pure CDM-type models the shape parameter is given by $\Gamma = \Omega h$. But if the model contains baryonic matter the shape parameter can be rescaled so $\Gamma = \Omega h e^{-\Omega_B/2}$, where Ω_B is the baryon density parameter. (Peacock 1992).

The third ensemble of 9 simulations is a spatially flat mixed dark matter model (the MDM ensemble) in which CDM contributes $\Omega_{\rm CDM} = 0.6$, baryons contribute $\Omega_b = 0.1$ and massive neutrinos contribute $\Omega_\nu = 0.3$. This set of simulations was run in a box of length 300 h^{-1}Mpc with 100^3 particles. The initial conditions for these simulations are generated from the power spectrum given in equation (1) of Klypin *et al.* (1993). In the MDM models, the thermal motions of the neutrinos are ignored and so they follow the evolution of a collisionless cold component with $\Omega_0 = 1$. The MDM simulations are exactly those used by Croft & Efstathiou (1994b).

In the analysis presented below, we use data and simulated data out to a comoving distance of at least 240 h^{-1}Mpc, this means that the MDM models are not large enough to properly simulate the survey without a significant amount of wrapping of the simulation box. The larger CDM-like models however are superior in this respect and no wrapping is required to simulate the data.

The final output times of the models are chosen to approximately match the microwave background anisotropies measured in the first year COBE maps (Smoot *et al.* 1992) ignoring any contribution from gravitational waves;

$$\Delta(k) = 180k^2T(k). \tag{21}$$

Thus the *rms* mass fluctuations in spheres of radius 8 h^{-1}Mpc are $\sigma_8 = 1$ for the CDM-like models and $\sigma_8 = 0.67$ for the MDM model. The amplitude of the temperature anisotropies measured from the four year COBE maps (Wright *et al.* 1996) is slightly higher than from the first year results. However, for the purposes of testing the techniques described above, this discrepancy is unimportant.

3.2 Construction of mock PSCz surveys

We draw mock PSCz redshift surveys from the N-body simulation distribution as follows:

(i) An observer, situated at an arbitrary position in the simulation box, carves out a sphere of radius $r_{\rm max} = 240$ h^{-1}Mpc.

(ii) Mass points are selected according the PSCz selection function:

$$\phi(\Delta) = \phi_\star \frac{10^{(1-\alpha)\Delta}}{(1+10^{\gamma\Delta})^{\frac{\eta}{\gamma}}}, \tag{22}$$

where $\Delta = \log_{10}(r/r_0)$ and the values of the parameters are $\phi_\star = 0.006794\,h^3{\rm Mpc}^{-3}$, $r_0 = 1.95113$ h^{-1}Mpc, $\alpha = 1.88654$, $\eta = 4.38595$ and $\gamma = 1.54554$. These parameters were derived using the methods described by Mann, Saunders & Taylor (1996).

(iii) Mass points are moved to their redshift space positions, $\boldsymbol{s}$, via the equation

$$\boldsymbol{s} = \boldsymbol{r}\left[1 + \frac{(\boldsymbol{v}-\boldsymbol{v}_0)\cdot\hat{\boldsymbol{r}}}{H_0 r}\right] \tag{23}$$

where we have full information on the velocity $\boldsymbol{v}$ of each particle. The observer velocity $\boldsymbol{v}_0$ is taken to be zero for the simulations, but is more correctly the Local Group velocity in the survey (so the second term in brackets does not formally diverge as $r \to 0$).

(iv) Cartesian x, y, z coordinates are converted to galactic latitude and longitude and a PSCz sky mask (see Section 4.1) is applied to the distribution.

(v) Mass points are drawn at random to have the correct selection function (22).

(vi) Mass points in the masked region of sky are discarded. The mask used here is in fact slightly less conservative than that finally used for the real data.

Using this procedure we obtain mock PSCz surveys with approximately the same density (and therefore the same number) of mass points as there are galaxies in the PSCz survey to a proper distance of $r_{\rm max}$ h^{-1}Mpc. Mass points are taken to represent galaxies throughout the analysis, so the linear bias parameter is unity. We have made no attempt to choose 'galaxies' from the simulation density distribution as this introduces extra complications and is unnecessary for the statistical testing purposes for which we employ the mock catalogues.

3.3 Results of analysing mock PSCz catalogues

Figures 1, 2 and 3 show the result of applying the analysis described in Section 2.1 to the mock PSCz catalogues. In all cases, we have used $r_{\rm max} = 240$ h^{-1}Mpc and have analysed modes up to $k = 0.1$ hMpc^{-1}. The maximum ℓ and n values are 21 and 8. For the convolutions we have included modes up to $\ell = 30$, $n = 20$.

For the calculation of the scattering matrix (equation 52, Appendix A), the three-dimensional velocity dispersion σ_v was measured from the simulations for each model. We have used the values: $\sigma_v = 1160\,{\rm km s^{-1}}$, $587\,{\rm km s^{-1}}$ and $859\,{\rm km s^{-1}}$ for the SCDM, ΛCDM and MDM models respectively.

In the analysis of the CDM-like mock catalogues we have assumed a real-space power spectrum of the form given in equation (20). For the MDM model, a power spectrum of the form of equation (1) of Klypin *et al.* (1993) was used. These power spectra are needed for the calculation of the covariance matrix (equations 52 – 54, Appendix A) as well as for the weighting function (equation 13).

Figures 1, 2 and 3 show contours of likelihood in the $\beta - \Delta_{0.1}$ plane for the SCDM, ΛCDM and MDM mock catalogues respectively. The contours are plotted at intervals of 0.5 in log–likelihood and the x-axis is labelled at intervals of 0.1 in $\Delta_{0.1}$. The true values of β and $\Delta_{0.1}$ are indicated by points. Figure 4 shows the ensemble average of the individual likelihoods and demonstrates the correct recovery of the parameters

The average recovered value of β and $\Delta_{0.1}$ over the nine realizations of the PSCz survey for each of the cosmological models, together with the true values of these parameters are presented in Table 1. The error on the recovered value of β and $\Delta_{0.1}$ is the error on the mean value (obtained from averaging the estimates, rather than from the ensemble averaged likelihood plots of Fig 4) multiplied by $\sqrt{9}$ i.e. this is the error appropriate to the recovered values of β and $\Delta_{0.1}$ from a single realization of the PSCz survey. This error bar provides us with an 'external' estimate of the error. The formal 'internal' errors on the values of β and $\Delta_{0.1}$ are given by the projection of $\delta \ln \mathcal{L} = -0.5$ onto the parameter axes, i.e. the projection of the first plotted contour. Evalu-

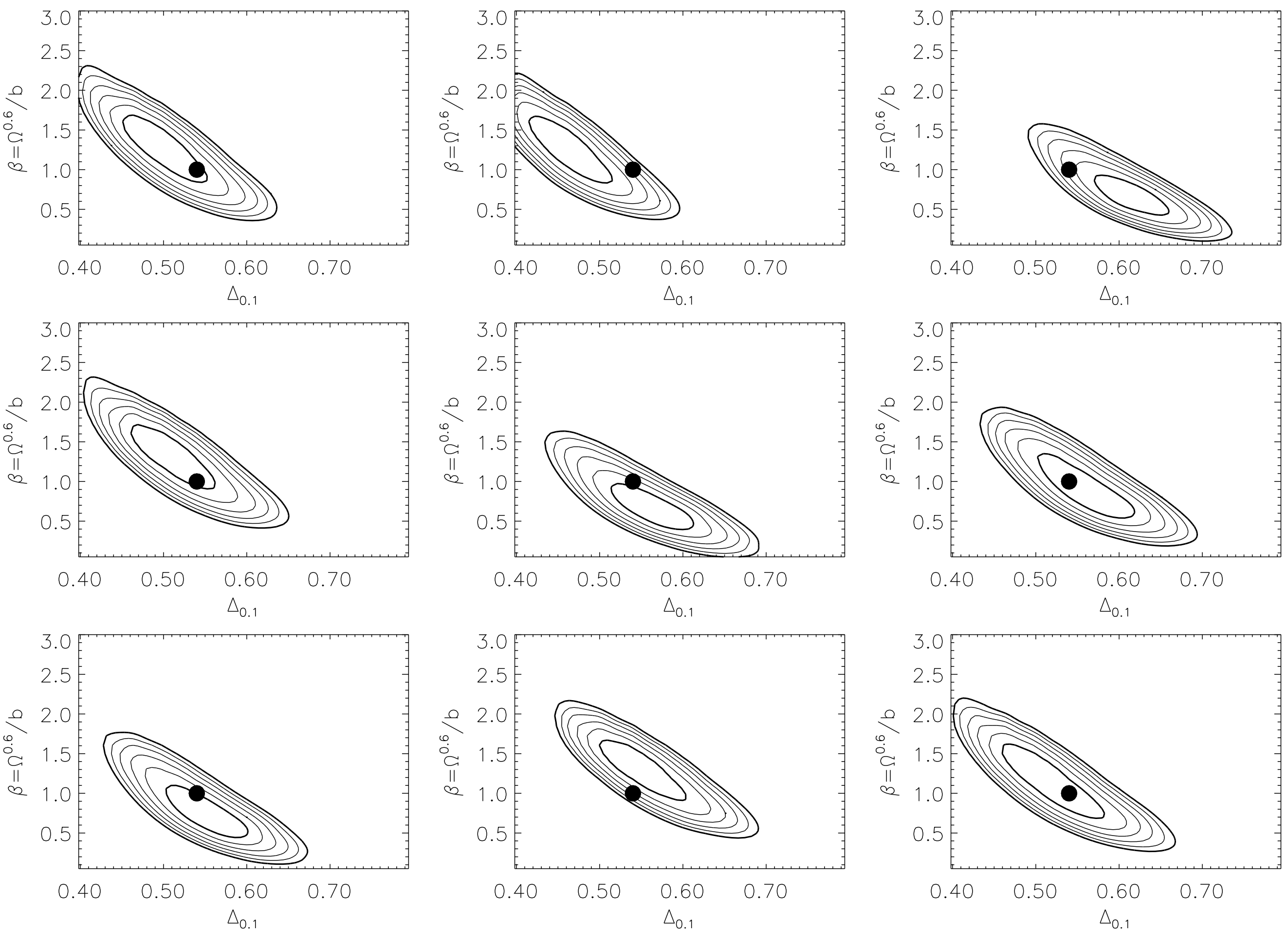

Figure 1. Figure shows contours of likelihood in the $\beta - \Delta_{0.1}$ plane for 9 realizations of the PSCz redshift survey drawn from SCDM N-body simulations. Points show the true parameter values. The contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$. The joint distribution has (β, $\Delta_{0.1}$) within 2.3 contours with 68% probability, and within 6.17 contours with 95% probability. The 1-σ error bars on the parameters individually are obtained by projecting the inner contour onto the axes (Press *et al.* 1992).

ating these points for each of the plots shown in Figures 1, 2 and 3 gives the internal error estimates for β and $\Delta_{0.1}$, together with the one sigma range on this error. These errors are shown in columns 3 and 6 of Table 1 for β and $\Delta_{0.1}$, respectively. Note that the parameter estimates for the external error determination were made from likelihood surfaces which were rather crudely sampled, so the mean of these estimates is artificially close to the true solution.

The recovered values of β and $\Delta_{0.1}$ are in very good agreement with the true values. In each model, the correct values of β and $\Delta_{0.1}$ are recovered to within the 1σ error. The internal error measured from the projection of the first contour level provides a reasonable estimate of the true error, as can be seen from the fact that the external and internal error estimates are in good agreement. The only case where this is not true is for $\Delta_{0.1}$ in the MDM model where the external error appears to be too large. This is possibly due to the fact that these simulations are not large enough to model the PSCz redshift survey out to a proper distance of $r_{\max} = 240\ h^{-1}$Mpc.

It is remarkable that the correction applied here for the non-linear effects introduced by the small-scale velocity field is adequate even for the SCDM model, which has a very high three-dimensional velocity dispersion ($\sigma = 1160\,\mathrm{kms}^{-1}$). Without this correction we would expect the recovered value of β to be significantly lower than the true value as non-linear ("finger of god") effects become important even for relatively small wavenumbers ($k \sim 0.06\ h\mathrm{Mpc}^{-1}$ – see e.g. Tadros & Efstathiou 1996).

We have also tested the stepwise maximum likelihood technique using the mock PSCz surveys described above. Figure 5 shows the results for individual SCDM mock catalogues and Figure 6 shows the ensemble averaged results for all of the SCDM, ΛCDM and MDM models. As expected in individual simulations the uncertainty grows with decreasing wavenumber, but the ensemble averaged results show that the stepwise maximum likelihood method is unbiased.

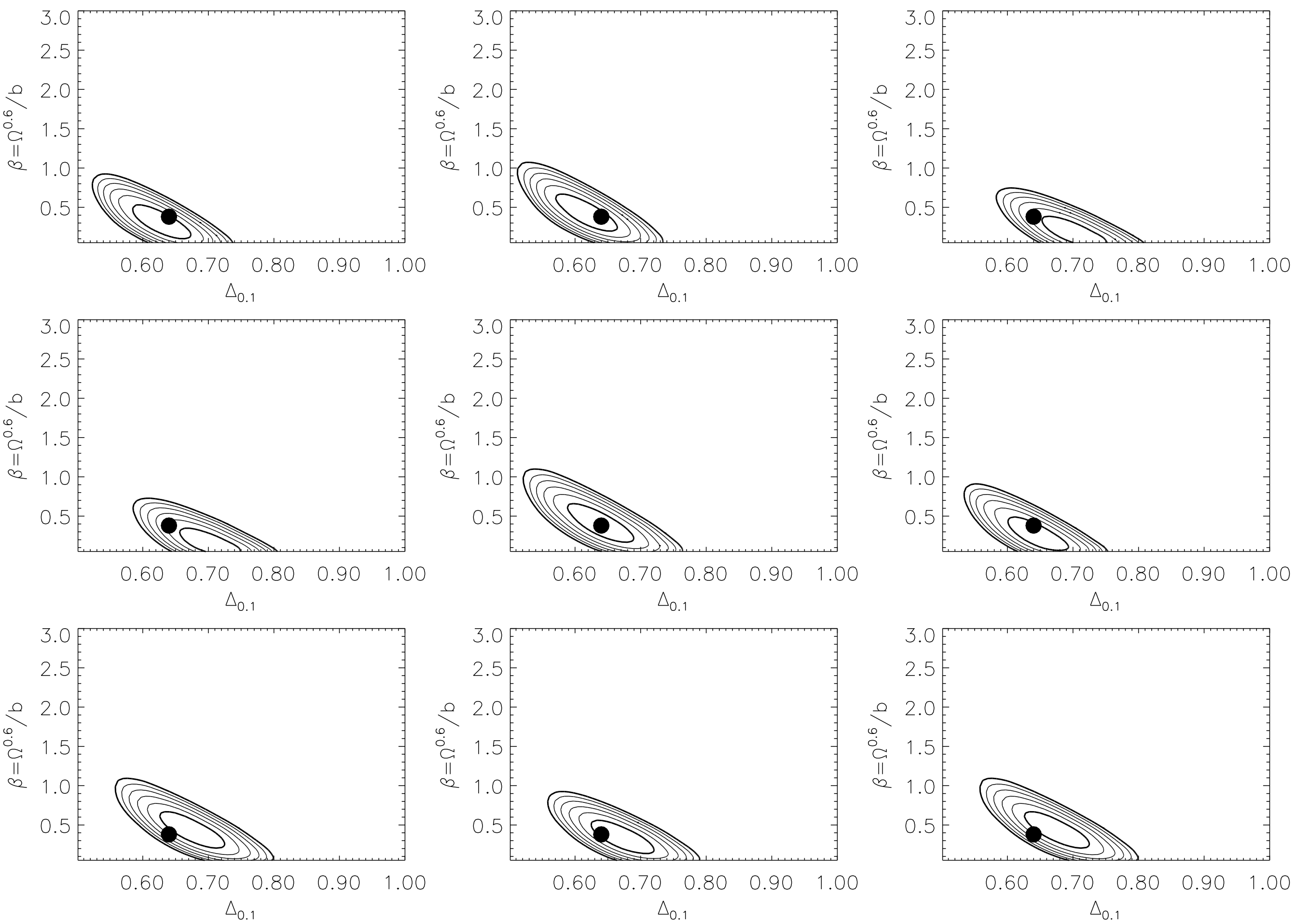

Figure 2. Contours of likelihood in the β – $\Delta_{0.1}$ plane for 9 realizations of the PSCz redshift survey drawn from ΛCDM N-body simulations. Again, as in all the following figures, contours are plotted at intervals of $\delta \ln \mathcal{L} = 0.5$. Points show the true parameter values.

Note in particular that the longest measured wavelength bin shows no significant bias, since we do not include the $\ell = 0$ mode. This has caused problems in the past for power spectrum analysis where the mean of the survey is estimated from the survey itself, suppressing the longest wavelength modes (Tadros & Efstathiou 1996).

From these tests on simulated PSCz redshift surveys we can be confident that the analysis described in this paper provides an unbiased estimate of the parameters β and $\Delta_{0.1}$. Furthermore, the errors obtained by examining the likelihood contours provide a reasonable estimate of the true error on β and $\Delta_{0.1}$.

4 THE IRAS PSCZ REDSHIFT SURVEY

We are now in a position to apply the analysis described above to the newly completed IRAS PSCz redshift survey. In Section 4.1 we briefly describe the survey itself, and in Sections 4.2 & 4.3 we test for various systematic effects in the catalogue that could bias the recovered value of β. The results of our analysis are presented in Section 5.

4.1 The Dataset

The IRAS PSCz survey is the largest and deepest all sky redshift survey of galaxies to date. It was conducted as a follow up to the 1 in 6 sparsely sampled QDOT redshift survey (Lawrence *et al.* 1998). The aim was to determine redshifts for all galaxies in the IRAS Point Source Catalogue (PSC – Joint IRAS Science Working Group 1988) to a $60\mu m$ flux limit of 0.6Jy in the area of sky where reliable optical identifications and redshift acquisition were feasible. An outline of survey construction is given in Saunders *et al.* (1995). About 4600 new redshifts were measured in the course of the project, the rest being taken from the literature or obtained via private communications. Details of the data reduction process will be published elsewhere – (Keeble *et al.* in preparation). Figure 8 shows the $N(z)$ distribution for galaxies in the PSCz redshift survey. The median redshift of galaxies in the survey is ~ 0.03, but there is a tail of galax-

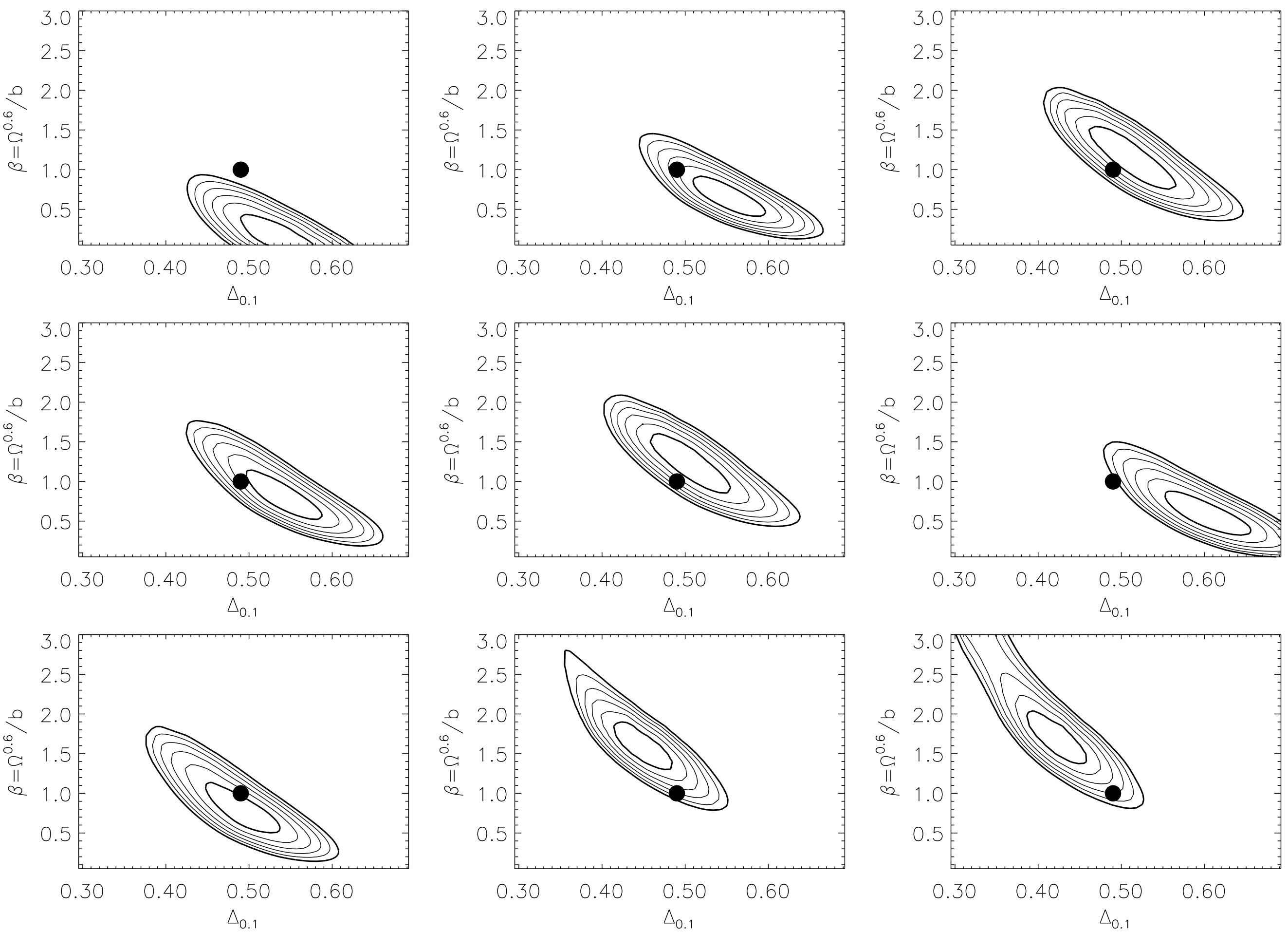

Figure 3. Figure shows contours of likelihood in the β – $\Delta_{0.1}$ plane for 9 realizations of the PSCz redshift survey drawn from MDM N-body simulations.

Table I

Model	1.β_{rec}	2.$Error_I$	3.β_{true}	4.Δ_{rec}	5.$Error_I$	6.Δ_{true}
SCDM	1.01 $\pm$ 0.24	0.17 $\pm$ 0.05	1.00	0.53 $\pm$ 0.03	0.04 $\pm$ 0.01	0.54
ΛCDM	0.36 $\pm$ 0.12	0.15 $\pm$ 0.04	0.38	0.66 $\pm$ 0.03	0.05 $\pm$ 0.01	0.64
MDM	1.00 $\pm$ 0.60	0.20 $\pm$ 0.05	1.00	0.51 $\pm$ 0.06	0.03 $\pm$ 0.01	0.49

Table 1. The results of applying the analysis described in Section 2.1 to mock PSCz redshift surveys drawn from three cosmological models. Columns 1 and 4 give the recovered values of β and $\Delta_{0.1}$, averaged over nine realizations, with errors derived from the scatter in the recovered values (external errors). These are to be compared with the true values of the parameters, given in columns 3 and 6. Columns 2 & 5 show the mean and scatter of the 1-σ projected error bars (internal errors).

ies out to redshifts as large as ~ 0.2. The solid curve shows the predicted $N(z)$ using the selection function of equation (22). The total error on the measured velocities is $\sim 100-150\,\mathrm{kms}^{-1}$, which is negligible for this analysis.

The PSCz was undertaken for various analyses, some of which demand the largest sky coverage possible. The default PSCz mask defining the survey consists of the IRAS coverage gaps, the LMC and SMC, and areas estimated on the basis of their $100\mu m$ sky brightness (Rowan-Robinson *et al.* 1986) to have $A_B > 2^m$, altogether 16% of the sky (Saunders *et al.* 1998). At low latitudes there is inevitable degradation in the survey due to incomplete identifications and the greater difficulty of getting redshifts. For an analysis such as this, where we are looking for weak clustering over large scales, uniformity is paramount. We therefore created more conservative masks, corresponding to optical extinctions of $A_B = 0.5^m, 0.75^m, 1^m$, for use in this analysis. The results presented are all for $A_B = 0.75^m$ (excluding 35% of

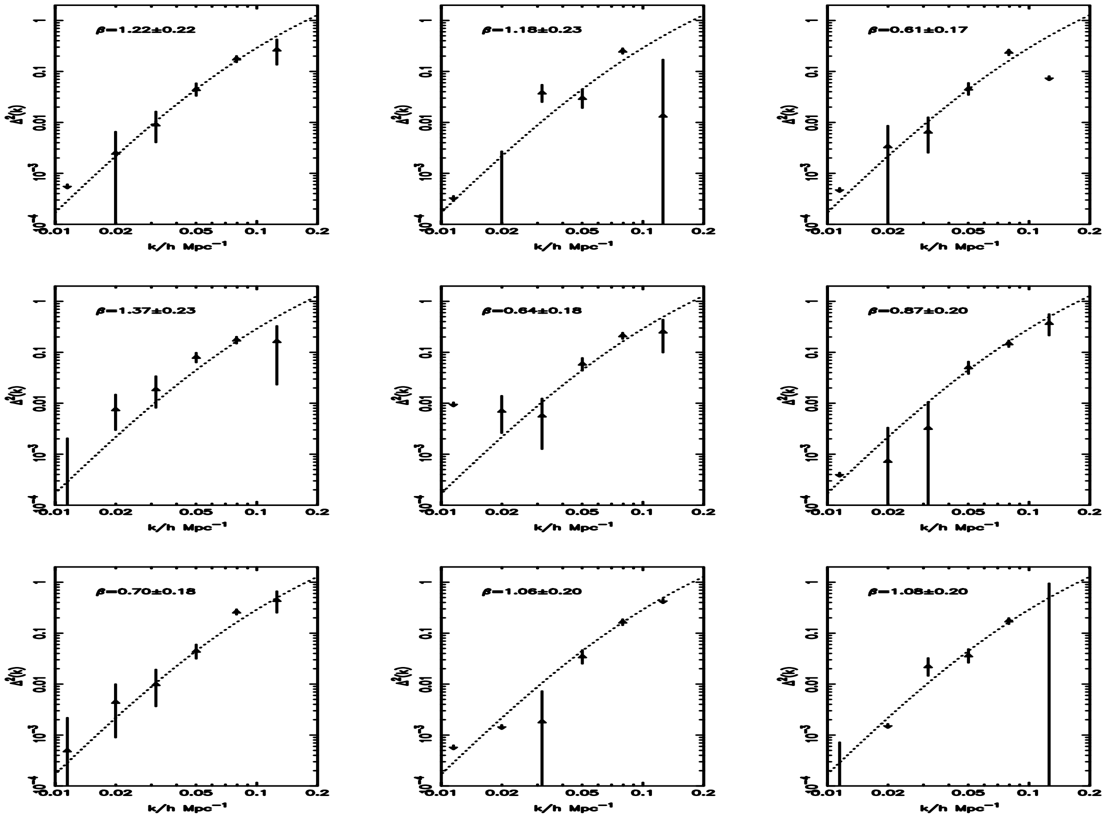

Figure 5. The real-space power spectrum obtained from an ensemble of SCDM mock PSCz simulations (data points), with the true power spectrum dotted. For this analysis, $r_{\rm max} = 300h^{-1}$ Mpc and $k_{\rm max} = 0.87h$ Mpc^{-1}.

the sky), but are robust to which of these three latter masks is used.

There is also incompleteness at high redshift, in that galaxies with than $A_{bJ} > 19^m$ were not systematically pursued spectroscopically. Based on the joint optical/60μm luminosity functions of Saunders *et al.* (1990) and the ULIRG survey of Clements *et al.* (1996), we believe that a negligible number of these faint galaxies are closer than $z = 0.1$, for areas with $A_B < 1^m$.

Within the volume defined by these cuts in area and distance, there are potentially variations in the IRAS data quality due to: the number of scans made by the satellite (HCONs), the South Atlantic Anomaly, Malmquist effects, hysteresis, confusion and noise lagging (Joint IRAS Science Working Group 1988). The most serious of these is the first, causing known incompleteness at faint fluxes. We restored most of this by reclaiming galaxies from the PSC reject file (Saunders *et al.* 1998). We investigated the residual quality variations both theoretically and empirically (through source counts), and were unable to find evidence for any non-uniformities above a level of 2% in flux or 3% in density. Several effects are expected to come in below this level, and so these numbers represent our overall best-guess estimates of the PSCz uniformity. Most of the concerns about the uniformity of the PSC affect sources near our flux limit of 0.6Jy, and so our final quoted results refer to a sample cut conservatively at 0.75Jy.

The distribution of the PSCz galaxies used in this analysis on the sky is shown in Figure 7.

4.2 Internal consistency checks

The ability to make high-precision estimates of the power spectrum with this method allows us to look for systematic effects in the catalogue. We do this by making a number of changes in the sample and the modes analysed and looking for inconsistencies in the recovered power spectrum and β. In this section, we give details of the various tests for systematics we have applied to the PSCz, and in Section 4.3 we describe the results.

We have tested the catalogue (and the method) for sensitivity to the following:

- Changes in the upper wavenumber limit, $k_{\rm max}$
- Changes to the outer boundary, $r_{\rm max}$
- Evidence for excess angular variance (explained below)
- Changes to the flux limit
- Changes from point source fluxes to addscan fluxes
- Changes between a high flux sample and low flux sample.

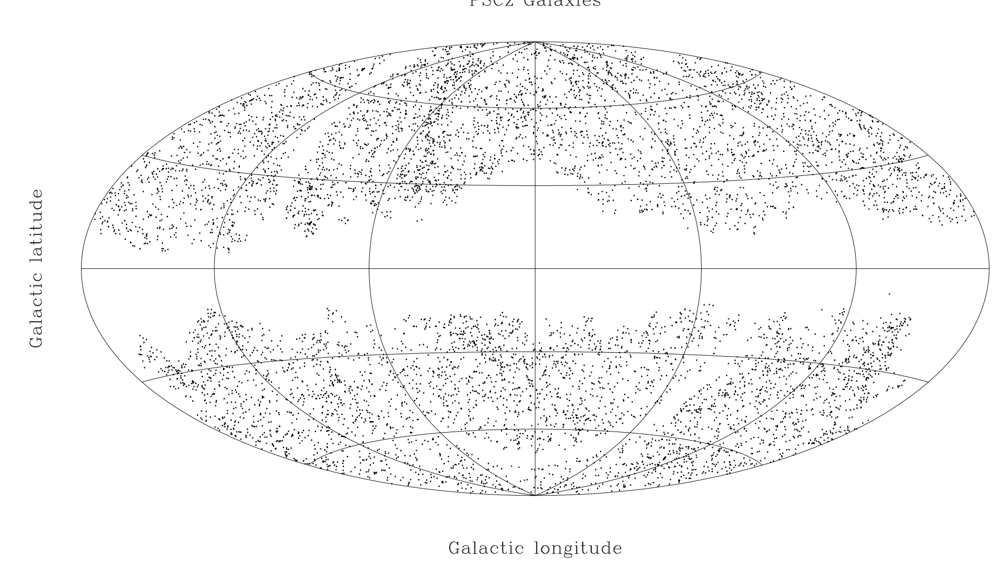

Figure 7. The sky-distribution of galaxies in the PSCz survey with $60\mu m$ flux above 0.75Jy. Here we have applied the conservative sky-mask used in our analysis. No cut in redshift has been imposed.

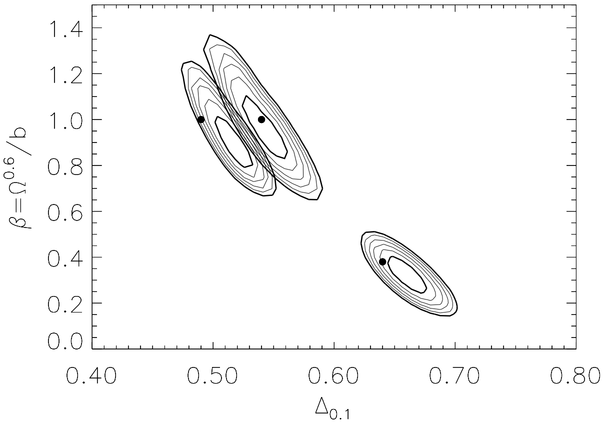

Figure 4. Parameter likelihoods for ensemble-averaged mock PSCz redshift surveys drawn from SCDM simulations (top right), ΛCDM simulations (bottom right) and MDM simulations (top left). True parameter values are marked with a point.

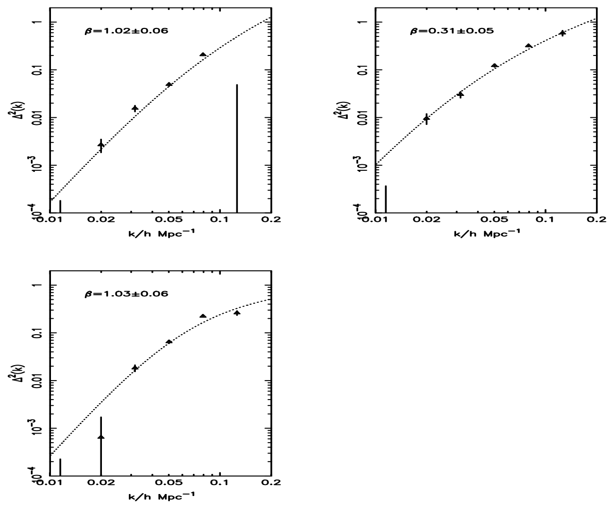

Figure 6. The ensemble average of real-space power spectrum estimates from the masked SCDM, ΛCDM and MDM mock surveys, along with the true power spectrum in each case. The recovered value of β, whose true values are 1.0, 0.38, and 1.0 are also shown.

By 'addscan' fluxes, we mean fluxes determined from the coadded raw IRAS data (kindly provided by IPAC), using software provide by Amos Yahil. The last test mentioned above is an important one, and needs further explanation. We first divide the PSCz catalogue into thin radial shells (width 100 km s^{-1}). The median flux of galaxies in a shell is calculated and galaxies with flux above this value are put into a 'bright' catalogue (the high-split catalogue) and the galaxies with flux below the median go into the low-split catalogue. This effectively divides the PSCz by luminosity while ensuring that the two sub-samples have the same radial distribution. Consequently, any inconsistencies between the results from the high and low-split catalogues cannot be due simply to changes in selection function or, equivalently, the radial distribution of points. We should point out that demonstrating consistency is not always straightforward: even disjoint samples of galaxies are not independent, as they are (by assumption) drawn from the same underlying density field. They do, however, have independent shot noise.

4.3 Results of consistency tests

For speed, we generally chose a relatively small wavenumber cutoff, $k_{\rm max} = 0.08h$ Mpc^{-1}, when performing these consistency tests. We find no evidence for any systematic changes in the results for changes in $r_{\rm max}$, or $k_{\rm max}$, or from chang-

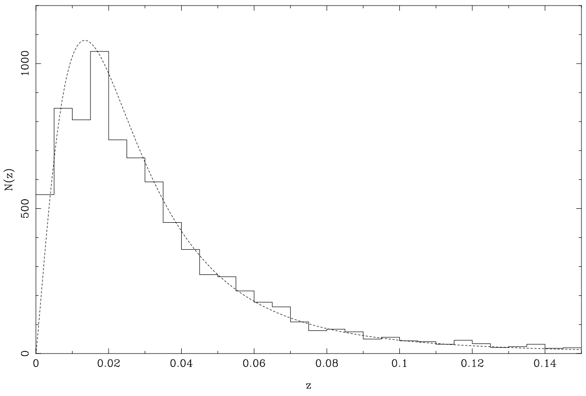

Figure 8. The redshift distribution for the galaxies in the >0.75Jy PSCz sample, in the unmasked regions of sky. For the analysis, the data are also cut at various values of z_{max} around 0.08. The line shows the selection function used.

ing point source fluxes to addscan fluxes. However when we changed the flux limit of the PSCz sample from 0.6Jy to 0.75Jy the results for $\Delta_{0.1}$ from these catalogues (hereafter PSCz0.6 and PSCz0.75) differed significantly. The high/low-split catalogue results show this inconsistency more clearly as explained in the next section.

We can make high and low-split catalogues from both PSCz0.6 and PSCz0.75. Figures 9 and 10 show parameter estimates for the high-split and low-split catalogues for these two flux limits. Figure 9 shows clear problems for the PSCz0.6 sample, as evidenced by the inconsistent parameter estimation from the high and low split subsamples. PSCz0.75 on the other hand, gives consistent results. We have performed a counts-in-cells analysis (c.f Efstathiou 1995), and similarly find an excess variance in the low-flux data.

The PSC has known imperfections at our flux limit of 0.6Jy. These include incompleteness in 2HCON areas, a large and crude 'faint flux overestimation' correction factor, and low and variable signal-to-noise (all Beichman *et al.* 1988). These problems have been investigated and, where possible corrected for in the galaxy catalogue used for this survey (Saunders *et al.* 1998). Nevertheless, the completeness and uniformity, which are paramount to this analysis, are certain to deteriorate towards the flux limit.

Because we find that our results are sensitive to the flux limit we use we have conservatively chosen to limit the catalogue to 0.75Jy in our quoted results.

However, despite extensive tests, we have not been able to find any 'smoking gun' in the sense of identifying where the excess clustering is coming from at lower fluxes. In fact, PSCz0.6 gives a result for β which is consistent with the PSCz0.75 sample; it is the amplitude which is the problem, being enhanced over PSCz0.75 by 15%. Given the smallness of the errors in the amplitude such a discrepancy is significant.

We have also searched for, and failed to find, a signature of direction-dependent flux errors, as proposed by Hamilton (1995). This arose from his analysis of the redshift-space correlation function of various IRAS surveys. He found an elongation along the line-of-sight on larger scales than plausibly accounted for by fingers-of-god. This was particularly marked for the QDOT survey in the region beyond 80 h^{-1}Mpc, and one explanation proposed was a direction-dependent flux error of $\sim 0.1Jy$, varying on angular scales of $7°$.

We can check partly for this effect by inspecting the $n = 1$ modes for signs of excess angular variance. These modes, which oscillate once radially, should be enhanced, as flux errors would affect the distant part of the catalogue more than the nearby region. We see no such signature: the power in these modes is consistent with the model solution and shot noise. It should be stressed however that our large-scale analysis does not probe angular scales as small as $7°$ ($\ell = 50$); it is only sensitive to variations on rather larger angular scales.

We also rule out a straightforward radially-dependent effect as the cause of the discrepancy between the high and low luminosity sub-samples of PSCz0.6, since these sub-samples have the same mean radial distributions by construction. Sutherland et al. (1998) have analysed the redshift-space power spectrum and find the same behaviour as seen here, i.e. a higher amplitude when all galaxies down to the 0.6 Jy flux limit are included. The effect is smaller in redshift space, but this would be expected if the lower flux galaxies had an excess variance, for whatever reason.

If it is a real luminosity effect (i.e. faint galaxies cluster more strongly than bright galaxies, introducing a differential bias), it must be very large, since in the analysis, the samples with different flux limits differ little in the weight given to galaxies of different luminosity. Moreover, if it is a true luminosity effect we would also expect to see it in PSCz0.75.

In our desire to leave no stone unturned, we also investigated the selection function itself, as the method of determining it (Mann, Saunders & Taylor 1996) provides the selection function in redshift space, not real space. The effects of this can be accounted for through a Taylor expansion and the addition of an extra term in the equations, but inclusion of this extra term makes an infinitesimal difference to the parameter estimation, so we neglect it.

5 THE REDSHIFT DISTORTION OF THE PSCZ AND THE AMPLITUDE OF THE REAL SPACE POWER SPECTRUM

We have analysed the PSCz0.75 survey using the method described in Section 2.1. Here we present results for a boundary at $r_{\max} = 200\ h^{-1}$Mpc. With the sky mask as defined in Section 4.1 we have 7042 galaxies in the data-set to the flux limit of 0.75Jy. We use a wavenumber limit of $k_{\max} = 0.13h$ Mpc^{-1}. These modes have ℓ up to 25, and n up to 9, but not all of the modes with these limits are used, as many fall outside the wavenumber range. We include modes up to $\ell = 30$ and $n = 20$ for the convolutions.

For the calculation of the covariance matrices we have assumed a CDM-like real space power spectrum with $\Gamma = 0.2$ (equation 20). As explained in Section 2.1, although the shape of the real space power spectrum in this analysis is

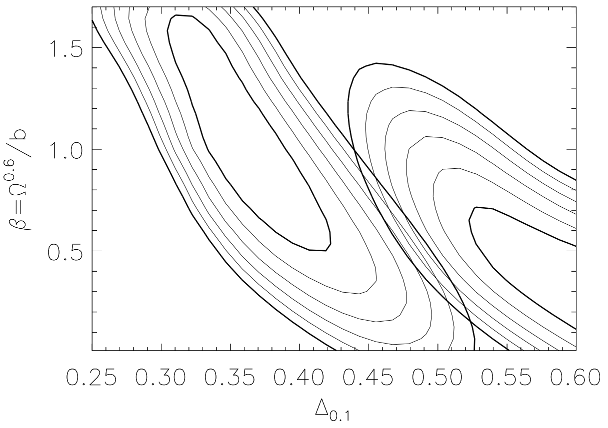

Figure 9. Likelihood contours for the two subsamples of PSCz0.6 (galaxies with 60μ flux about 0.6Jy) split by luminosity. Here we have used a restricted set of modes, $k_{max} = 0.08$, and a boundary at 200 h^{-1} Mpc. The high and low luminosity subsamples are created as described in the text. Contours for the low luminosity sub-sample are to the bottom right of the plot. The amplitude, $\Delta_{0.1}$, is seen to be significantly higher for the low luminosity subsample.

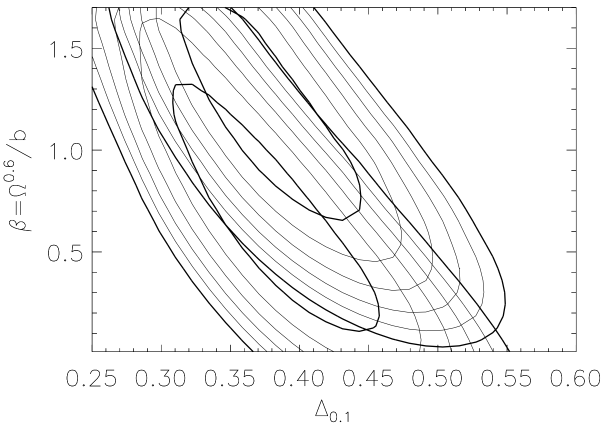

Figure 10. Likelihood contours for the two subsamples of PSCz0.75 (galaxies with 60μ flux about 0.75Jy) split by luminosity. Again, we have used a restricted set of modes, $k_{max} = 0.08$, and a boundary at 200 h^{-1} Mpc. It can be seen that by excluding the galaxies in PSCz with 60μm flux between 0.6Jy and 0.75Jy before splitting the sample by luminosity we obtain parameter estimates for the high/low luminosity subsamples which are consistent.

fixed, the amplitude at a wavenumber $k = 0.1\,h\text{Mpc}^{-1}$ is the second free parameter in the likelihood fit. A wavenumber-dependent weighting function (equation 13) has been employed. Again, we have used a CDM-like power spectrum with $\Gamma = 0.2$ for $P_w(k)$. The normalization of $P_w(k)$ is such that the amplitude $\Delta^2_{0.1}$ is 0.18. In this minimum variance

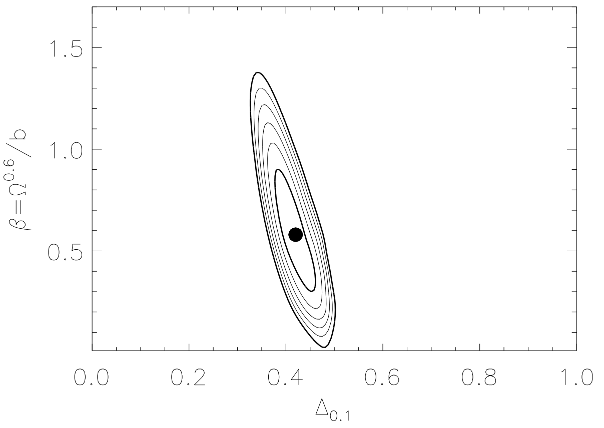

Figure 11. Likelihood contours for the two parameters $\beta \equiv \Omega_0^{0.6}/b$ (where b is the IRAS bias parameter) and the amplitude of the power spectrum (parametrised by $\Delta^2(k) \equiv k^3P(k)/(2\pi^2)$ at a wavenumber $k = 0.1\,h$ Mpc^{-1}) for the PSCz survey flux-limited to 0.75Jy. The outer boundary is 200 h^{-1} Mpc, and all wavenumbers up to $k = 0.13h$ Mpc^{-1} have been analysed. A conservative mask with extinction $A_B < 0.75$ has been applied, leaving a total of 7042 galaxies in the sample.

weighting scheme, the exact form and amplitude of the function $P_w(k)$ will not be critical.

We have employed a value of $\sigma_v = 420\,\text{kms}^{-1}$ for the three-dimensional velocity dispersion of galaxies in order to calculate the scattering matrices and correct for non-linearities (equation 52). This is the value of σ_v obtained by adopting a pairwise velocity dispersion of $340\,\text{kms}^{-1}$ (Huchra *et al.* 1983, Fisher *et al.* 1994) and assuming that the small scale peculiar velocity field is incoherent.

Figure 11 shows the contours of likelihood in the $\beta - \Delta_{0.1}$ plane, plotted at intervals $\delta \ln \mathcal{L} = 0.5$ for the PSCz survey. The recovered values of β and $\Delta_{0.1}$ are:

$$\beta = 0.58 \pm 0.26$$
$$\Delta_{0.1} = 0.42 \pm 0.03.$$

We have analysed PSCz0.6 to a wavenumber limit of 0.12 h Mpc^{-1}; this gives $\beta = 0.57 \pm 0.25$ but $\Delta_{0.1} = 0.48 \pm 0.04$. However, as mentioned earlier, there are inconsistencies in detail which lead us not to trust the results from the full 0.6 Jy sample.

5.1 The shape and amplitude of the real–space power spectrum of IRAS galaxies

The stepwise power spectrum was fitted as described in Section 2.2 and has been applied to the 1.2Jy survey (BHT). Unlike the BHT analysis, we have chosen that the right hand bin should begin at $k_{\rm max}$, to show how information beyond the observed modes can be analysed. There is a constraint on this bin due to the effect of mixing, but this is quite weak as can be seen by the large error bar on the final point in

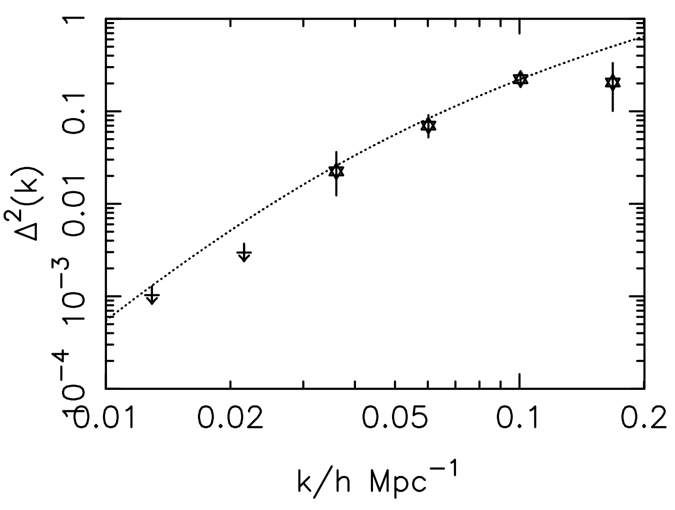

Figure 12. The real-space power spectrum of the PSCz redshift survey in dimensionless units. The curve is a CDM model with $\Gamma = 0.2$. The estimated redshift distortion is $\beta = 0.47 \pm 0.16$ (conditional error).

Figure 12. Again, reassuringly, the redshift distortion result is similar to that of the two-parameter fit with

$$\beta = 0.47 \pm 0.16, \tag{24}$$

although note that this is a conditional error, as are the error bars on the power spectrum. The analysis has $r_{\max} = 200\ h^{-1}$Mpc and $k_{\max} = 0.13h$ Mpc^{-1}. Reducing the wavenumber limit to 0.12 gives a consistent result of $\beta = 0.60 \pm 0.18$.

The power spectrum again fits low Γ CDM models very well, although there is some evidence that the break is slightly sharper as suggested by Peacock (1997). It is perhaps possible that the final, high-k point is suppressed by nonlinearities, although the random smoothing model was used. A particular point to note is the very tight lower limit on the first two bins, close to the value expected in current models. This demonstrates that the sample we finally use after the cuts discussed in Section 4.3 is uniform (in the sense of rms variations) to better than 3% on scales of 500 h^{-1}Mpc.

6 DISCUSSION AND CONCLUSIONS

We have presented the results of a likelihood analysis of a spherical harmonic expansion of the PSCz galaxy survey. The aim has been to establish both the real-space power spectrum and the degree of redshift distortion from the distribution of galaxies in redshift space. The main assumptions in the analysis are that structure grows via gravitational instability, and that the galaxies are linearly biased tracers of the underlying mass distribution, at least on the scales which we analyse (wavelengths $> 48h^{-1}$ Mpc). There are many attractive features about this analysis, and it is worth reiterating them here. Firstly, the expansion in spherical polar coordinates is ideal for many redshift surveys, but particularly for the IRAS galaxy surveys, which cover a large fraction of the sky. The effects of redshift distortion are radial, so are naturally accounted for in this system; in addition the masked regions of sky, due to the galactic plane, unobserved regions of sky, and patches badly affected by cirrus, are specified in terms of the angular coordinates alone. These facts allow a relatively simple translation from real- to redshift-space expansion coefficients in linear theory. This is a major advantage for relatively shallow IRAS surveys, since the linear modes typically subtend large angular scales. A general feature of redshift distortion studies to obtain β is that all the information is obtained from the galaxy positions, rather than relying on peculiar velocity measurements. The particular expansion we have used here is effectively a power spectrum analysis in spherical coordinates. Because of the survey geometry and redshift distortions we do not measure the power spectrum directly, but the window functions converting real-space power to redshift-space expansion coefficients are narrow, and we are able to measure with some accuracy the power spectrum on large scales between 40 and 270 h^{-1} Mpc. In addition constraints on the power can be put on scales up to $\sim 500\ h^{-1}$Mpc.

The analysis we have done follows closely the work of Heavens & Taylor (1995) and Ballinger, Heavens & Taylor (1995), although we have made some technical improvements in the treatment of the mask and the local group motion, but the effects of the approximations in the previous papers are in fact negligibly small.

The results we present are conservative in a number of ways. We have firstly applied a rather conservative mask, and consider only regions of sky where the extinction $A_B < 0.75$. Secondly, we found that we were able to obtain robust results only for a subset of the catalogue cut at a higher flux than the formal 0.6 Jy limit. The most striking example of the inconsistency near the flux limit is illustrated in Figure 9, where we have taken care to analyse two subsamples with the same radial and angular distribution to minimise other sources of error. This graph shows why we do not trust the catalogue at the lowest flux limits, although we have not been able to pin down the source of the discrepancies. What we do find is that the results are quite stable if we include galaxies down to a flux limit of 0.75 Jy. The price we pay for these conservative choices is that the galaxy sample is cut back to about 7000 galaxies, and the error bars are larger than we would have liked. However, we believe that the resulting estimates and errors can be trusted. In the first analysis, we assume that the power spectrum has the shape of a $\Gamma = 0.2$ CDM spectrum, at least on scales larger than 40 h^{-1}Mpc, and estimate simultaneously β and the amplitude (characterised by the dimensionless power at 0.1 h^{-1}Mpc). This analysis yields

$$\begin{aligned} \beta &= 0.58 \pm 0.26 \\ \Delta_{0.1} &= 0.42 \pm 0.03. \end{aligned} \tag{25}$$

If we allow the power spectrum to be free (or at least characterized by a large number of parameters), the results change only slightly:

$$\beta = 0.47 \pm 0.16$$

$$\Delta_{0.1} = 0.47 \pm 0.03 \tag{26}$$

but note that in this case the quoted errors are conditional, and are not reflective of the true marginal errors.

The value of β which we derive is lower than the values of 1.1 and 1.0 obtained by similar analysis of the IRAS 1.2 Jy catalogue (Heavens & Taylor 1995, Ballinger, Heavens & Taylor 1995), and by the analysis, also using spherical harmonics, of Fisher, Scharf & Lahav (1994). However, the errors in our previous analyses are large enough to allow β to be as low as 0.5. Our result is consistent with determinations based on the plane-parallel approximation (e.g. Cole, Fisher & Weinberg 1994,Cole, Fisher & Weinberg 1995) and correlation function studies (Hamilton 1992). A full discussion of results in this area can be found in Hamilton (1997). Methods based on comparison between velocity fields and IRAS density fields also show a (formally inconsistent) spread of results, from 0.49 ± 0.07 (Willick *et al.* 1997) to 0.89 ± 0.12 (Sigad *et al.* 1998). One possible way to reconcile discrepant results in general is if bias is scale-dependent (e.g. Coles 1993, Mann, Peacock & Heavens 1998), and the measurements are effectively probing different scales, but it is not clear whether this is appropriate here.

Since we compute the power spectrum and the redshift distortion simultaneously, and assuming the linear bias model, we are able to express the answer in terms of the strength of clustering of the mass, as has been done from the abundance of clusters (White, Efstathiou & Frenk 1993). We express the amplitude of matter clustering in terms of the dimensionless matter power spectrum at a wavenumber of 0.1 h Mpc^{-1}:

$$\Delta_{0.1,\rho} = 0.24\Omega_0^{-0.6} \tag{27}$$

with an error of about 50%. If we convert this to the r.m.s. fluctuations in an 8 h^{-1}Mpc sphere, assuming a $\Gamma = 0.2$ shape parameter (using the approximation equation 16), we get $\sigma_8 = 0.38\Omega_0^{-0.6}$, again with an error of about 50%. This rises to 0.57 if $\Gamma = 0.5$.

What should we conclude from this? The simplest, although not the only, interpretation is that IRAS galaxies are unbiased, and $\Omega_0 \simeq 0.4$. The strongest indications that the Universe might be Einstein-de Sitter have in the past come from the high values of β from peculiar velocity and redshift distortion studies. The picture is still not clear, but the evidence is shifting away from a high β towards low. In any case, it is impossible to make firm statements about Ω_0 based on these studies because of our ignorance of the bias parameter, and we shall almost certainly have to wait until large redshift surveys of $10^5 - 10^6$ galaxies are completed, when higher-order statistics will lift the degeneracy between Ω_0 and b (e.g. Matarrese, Verde & Heavens 1997, Verde *et al.* 1998). Indications of a low value of Ω_0 come from the supernova Hubble diagram (Perlmutter *et al.* 1998, Riess *et al.* 1998) combined with rather uncertain flatness constraints from the microwave background, although the prospects are good (Tegmark, Eisenstein & Hu 1998).

ACKNOWLEDGMENTS

We are grateful to the many people who provided extensive and invaluable assistance in terms of telescope observations, provision of redshifts, optical and IRAS information and reduction software. The PSCz used observations from the INT, AAT and CTIO 1.5m telescopes, had extensive assistance from the ROE and IPAC, and made great use of the NED, SIMBAD and ZCAT databases. Computations were carried out on STARLINK facilities.

BIBLIOGRAPHY

Ballinger W. E., Heavens A. F., Taylor A. N., 1995. *Mon. Not. R. astr. Soc.*, **276**, 59P.

Baugh C. M., Efstathiou G., 1993. *Mon. Not. R. astr. Soc.*, **265**, 145.

Beichman C., Habing G. N. H. J., Clegg P. E., Chester T. J., 1988. *IRAS Explanatory Supplement, NASA RP-1190, Vol 1*, U. S. Government Printing Office, Washington D. C.

Bond J. R., Efstathiou G., 1984. *Ap. J. Lett.*, **285**, L45.

Bower R. G., Coles P., Frenk C., White S., 1993. *Ap. J.*, **405**, 403.

Clements D., Sutherland W., Saunders W., Efstathiou G., McMahon R., Maddox S., Lawrence A., Rowan-Robinson M., 1996. *Mon. Not. R. astr. Soc.*, **279**, 459.

Cole S., Fisher K. B., Weinberg D. H., 1994. *Mon. Not. R. astr. Soc.*, **267**, 785.

Cole S., Fisher K. B., Weinberg D. H., 1995. *Mon. Not. R. astr. Soc.*, **275**, 515.

Coles P., 1993. *Mon. Not. R. astr. Soc.*, **262**, 1065.

Croft R. A. C., Efstathiou G., 1994a. *Mon. Not. R. astr. Soc.*, **267**, 390.

Croft R. A. C., Efstathiou G., 1994b. *Mon. Not. R. astr. Soc.*, **268**, 23P.

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985. *Ap. J.*, **292**, 371.

Edmonds A., 1957. *Angular Momentum in Quantum Mechanics*, Princeton University Press; Princeton.

Efstathiou G., Ellis R. S., Peterson B. A., 1988. *Mon. Not. R. astr. Soc.*, **232**, 431.

Efstathiou G., Davis M., Frenk C. S., White S. D. M., 1985. *Ap. J. Suppl.*, **57**, 241.

Efstathiou G., 1995. *Mon. Not. R. astr. Soc.*, **276**, 1425.

Feldman H. A., Kaiser N., Peacock J. A., 1994. *Ap. J.*, **426**, 23.

Fisher K. B., Nusser A., 1996. *Mon. Not. R. astr. Soc.*, **279**, 1P.

Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J. P., 1994. *Mon. Not. R. astr. Soc.*, **267**, 927.

Fisher K. B., Scharf C., Lahav O., 1994. *Mon. Not. R. astr. Soc.*, **266**, 219.

Hamilton A. J. S., 1992. *Ap. J. Lett.*, **385**, L5.

Hamilton A. J. S., 1993. *Ap. J. Lett.*, **406**, L47.

Hamilton A. J. S., 1995. In: *Clustering in the Universe*, 143, eds Maurogordato S., Balkowski C., Tao C., Tran Thanh Van J., Editions Frontieres.

Hamilton A. J. S., 1997. In: *Ringberg Workshop on Large Scale Structure 1996*, ed. Hamilton D., Kluwer Academic, Dordrecht.

Heavens A. F., Taylor A. N., 1995. *Mon. Not. R. astr. Soc.*, **275**, 483.

Heavens A. F., Matarrese S., Verde L., 1998. *Mon. Not. R. astr. Soc.*, **in press (astro-ph/9808016)**.

Huchra J., Davis M., Latham D., Tonry J., 1983. *Ap. J. Suppl.*, **52**, 89.

Jackson J., 1972. *Mon. Not. R. astr. Soc.*, **156**, 1P.

Joint IRAS Science Working Group, 1988. *IRAS Point Source Catalogue, Version 2*, U. S. Government Printing Office, Washington D. C.

Kaiser N., 1987. *Mon. Not. R. astr. Soc.*, **227**, 1.

Kauffman G., Nusser A., Steinmetz M., 1997. *Mon. Not. R. astr. Soc.*, **286**, 795.

Klypin A., Holtzman J., Primack J., Regos E., 1993. *Ap. J.*, **416**, 1.

Landau L., Lifshitz E., 1977. *Quantum Mechanics*, Pergamon Press; Oxford.

Lawrence A., Rowan-Robinson M., Saunders W., Parry I. R., Xiaoyang X., Ellis R. S., Frenk C. S., Efstathiou G., Kaiser N., Crawford J., 1998. *Mon. Not. R. astr. Soc., submitted.*

Mann R., Peacock J., Heavens A., 1998. *Mon. Not. R. astr. Soc.*, **293**, 209.

Mann R. G., Saunders W., Taylor A. N., 1996. *Mon. Not. R. astr. Soc.*, **279**, 636.

Matarrese S., Verde L., Heavens A., 1997. *Mon. Not. R. astr. Soc.*, **290**, 651.

Peacock J. A., Dodds S. J., 1994. *Mon. Not. R. astr. Soc.*, **267**, 1020.

Peacock J. A., 1992. *Mon. Not. R. astr. Soc.*, **258**, 581.

Perlmutter S., Aldering G., Della Valle M., Deustua S., Ellis R., Fabbro S., Fruchter A., Goldhaber G., Groom D. E., Hook I., Kim A. G., Kim M., Knop R. A., Lidman C., Mcmahon R. G., Nugent P., Pain R., Panagia N., Pennypacker C. R., Ruiz-Lapuente P., Schaefer B., Walton N., 1998. *Nature*, **391**, 51.

Press W. H., Teukolsky S., Vetterling W., Flannery B., 1992. *Numerical Recipes in Fortran*, Cambridge University Press.

Riess A. G., Filippenko A. V., Challis P., Clocchiattia A., Diercks A., Garnavich P. M., Gilliland R. L., Hogan C. J., Jha S., Kirshner R. P., Leibundgut B., Phillips M. M., Reiss D., Schmidt B. P., Schommer R. A., Smith R. C., Spyromilio J., Stubbs C., Suntzeff N. B., Tonry J., 1998. *Astron. J.*, **(in press)**, astroph/9805201.

Saunders W., Rowan-Robinson M., Lawrence A., Efstathiou G., Kaiser N., Ellis R., Frenk C., 1990. *Mon. Not. R. astr. Soc.*, **242**, 318.

Saunders W., Sutherland W., Efstathiou G., Tadros H., Maddox S., McMahon R., White S., Oliver S., Keeble O., Rowan-Robinson M., Frenk C., 1995. In: *Wide Field Spectroscopy and the Distant Universe, The* 35^{th} *Herstmonceux Conference*, 88, eds Maddox S. J., Aragon-Salamanca A.

Saunders W., Sutherland W., Efstathiou G., Tadros H., Maddox S., McMahon R., White S., Oliver S., Keeble O., Rowan-Robinson M., Frenk C., 1998. *Mon. Not. R. astr. Soc., in preparation.*

Scharf C., Hoffman Y., Lahav O., Lynden-Bell D., 1992. *Mon. Not. R. astr. Soc.*, **256**, 229.

Sigad Y., Eldar A., Dekel A., Strauss M., Yahil A., 1998. *Ap. J.*, **495**, 516.

Smoot G. F., Bennett C. L., Kogut A., Wright E. L., Aymon J., Boggess N. W., Cheng E. S., De Amici G., Gulkis S., Hauser M. G., Hinshaw G., Jackson P. D., Janssen M., Kaita E., Kelsall T., Keegstra P., Lineweaver C., Loewenstein K., Lubin P., Mather J., Meyer S. S., Moseley S. H., Murdock T., Rokke L., Silverberg R. F., Tenorio L., Weiss R., Wilkinson D. T., 1992. *Ap. J. Lett.*, **396**, L1.

Sutherland et al., 1998. *Mon. Not. R. astr. Soc., in press.*

Tadros H., Efstathiou G., 1996. *Mon. Not. R. astr. Soc.*, **282**, 1381.

Taylor A. N., Hamilton A. J. S., 1996. *Mon. Not. R. astr. Soc.*, **282**, 767.

Tegmark M., Eisenstein D., Hu W., 1998. In: *Fundamental parameters in Cosmology (Rencontres de Moriond).*

Verde L., Heavens A., Matarrese S., Moscardini L., 1998. *Mon. Not. R. astr. Soc.*, **in press (astro-ph/9806028)**.

White S. D. M., Efstathiou G., Frenk C., 1993. *Mon. Not. R. astr. Soc.*, **262**, 1023.

Willick J., Strauss M., Dekel A., Kolatt T., 1997. *Ap. J.*, **486**, 629.

Wright E. L., Bennett C. L., Gorski K., Hinshaw G., Smoot G. F., 1996. *Ap. J. Lett.*, **464**, L21.

APPENDIX A: OUTLINE OF THE SPHERICAL HARMONIC FORMALISM

In this Appendix we outline the spherical harmonic formalism used to estimate the power spectrum and redshift space distortion in the PSCz. The general derivation is similar to that given in HT, but is somewhat more comprehensive.

A.1 Definitions

Following HT we expand the density field of the galaxy distribution in spherical harmonics, $Y_{\ell m}$, and spherical Bessel functions, j_ℓ,

$$\hat{\rho}_{\ell mn} = c_{\ell n} \int d^3 s\, \rho(\boldsymbol{s}) w(s) j_\ell\left(k_{\ell n} s\right) Y^*_{\ell m}\left(\theta, \phi\right), \tag{28}$$

where $w(s)$ is an adjustable weighting function and $\boldsymbol{s}$ is the redshift-space position variable. $c_{\ell n}$ are normalisation constants, and $k_{\ell n}$ are discrete wavenumbers (see A2). The inverse transformation is

$$\rho(\boldsymbol{s}) = \sum_{\ell mn} c_{\ell n} \rho_{\ell mn} j_\ell(k_{\ell n} s) Y_{\ell m}(\theta, \phi). \tag{29}$$

We use the Binney and Tremaine (1987) definition of spherical harmonics which uses the Condon–Shortley phase condition;

$$Y_{\ell m}(\theta,\phi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-|m|)!}{(\ell+|m|)!}} P_\ell^{|m|}(\cos\theta) e^{im\phi} \times \begin{cases} (-1)^m & m \geq 0 \\ 1 & m < 0 \end{cases} \tag{30}$$

where $P_\ell^{|m|}(x)$ is the Legendre polynomial defined (for positive m only) by

$$P_\ell^m(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_\ell(x). \tag{31}$$

Note that some definitions add an extra $(-1)^m$ term (e.g. Press et al. 1980). With the Binney and Tremaine definition the Hermitian relation,

$$Y_\ell^{-m} = (-1)^m Y_\ell^{m*}, \tag{32}$$

is automatically satisfied.

A.2 Normalisation and boundary conditions

The $c_{\ell n}$ are normalization constants introduced to preserve the normalisation of the spherical Bessel function in a finite sphere. They are derived via the application of the orthogonality relations for spherical Bessel functions, for given boundary conditions (see HT). The boundary conditions used for this analysis are Neumann boundary conditions (i.e. that the first derivative of the potential field normal to the boundary is zero). This implies that the velocity field is zero on the boundary i.e. the boundary is undistorted. We therefore need not include surface effects. The boundary condition gives us the wavenumbers $k_{\ell n}$:

$$\frac{d}{dr} j_\ell\left(k_{\ell n} r\right)|_{r=r_{\max}} = 0, \tag{33}$$

and the normalization constants

$$c_{\ell n} = \frac{k_{\ell n}}{|j_\ell(k_{\ell n} r_{\max})| \sqrt{\frac{r}{2}\left[\frac{1}{4} + k_{\ell n}^2 r_{\max}^2 - \left(\ell + \frac{1}{2}\right)^2\right]}}. \tag{34}$$

A.3 Redshift Distortions

As described in the main text, the transformation to redshift space can be carried out by the identity $d^3 s\, \rho(\boldsymbol{s}) = d^3 r\, \rho(\boldsymbol{r})$, leaving redshifted coordinates in the weighting and spherical Bessel functions. HT then expand $w(s) j_\ell(k_{\ell n} s)$ in equation (28) to first order in $s - r$,

$$w(s) j_\ell(k_{\ell n} s) \simeq w(r) j_\ell(k_{\ell n} r) + u(\mathbf{r}) \frac{d}{dr}\left[w(r) j_\ell(k_{\ell n} r)\right], \tag{35}$$

where $u(\boldsymbol{r})$ is the radial component of the peculiar velocity field, $\boldsymbol{v}(\boldsymbol{r}) - \boldsymbol{v}_{\rm LG}$. Linear theory provides a relation between the peculiar velocity field $\boldsymbol{v}(\boldsymbol{r})$ and the density field via the continuity equation

$$\nabla . \boldsymbol{v}(\boldsymbol{r}) = -H_0 \Omega_0^{0.6} \delta_\rho(\boldsymbol{r}) \tag{36}$$

to first order. In Appendix A of HT it is shown that the radial component of this field is

$$u(\boldsymbol{r}) = \Omega_0^{0.6} \sum_{\ell mn} c_{\ell n} k_{\ell n}^{-1} \delta_{\ell mn} j'_\ell(k_{\ell n} r) Y_{\ell m}(\theta,\phi) - \boldsymbol{v}_{\rm LG} \cdot \hat{\boldsymbol{r}}, \tag{37}$$

where $j'(z) = dj_\ell(z)/dz$ and $\delta_{\ell mn}$ is the transform of the galaxy number over-density field $\delta(\boldsymbol{r})$. Note that in this expression we have added the contribution from the Local Group dipole. This was neglected in the analysis of HT since it only contributed significantly to the $\ell = 1$ mode, which was not included in the analysis. We include it here for completeness.

Substituting equation (37) into equation (35) gives

$$w(s) j_\ell(k_{\ell n} s) \simeq w(r) j_\ell(k_{\ell n} r) + \beta \sum_{\ell' m' n'} c_{\ell' n'} k_{\ell' n'}^{-1} \delta_{\ell' m' n'} j'_{\ell'}(k_{\ell' n'} r) \frac{d}{dr}\left[w(r) j_\ell(k_{\ell n} r)\right] Y_{\ell' m'} - (\boldsymbol{v}_{\rm LG} \cdot \hat{\boldsymbol{r}}) \frac{d}{dr}\left[w(r) j_\ell(k_{\ell n} r)\right]. \tag{38}$$

A.4 Transforming real space harmonic modes to redshift space

From the definition of perturbations in the galaxy density field we have

$$\rho(\boldsymbol{r}) = \rho_0(\boldsymbol{r})\left[1 + \delta(\boldsymbol{r})\right] = \rho_0(\boldsymbol{r}) \left(1 + \sum_{\ell mn} c_{\ell n} \delta_{\ell mn} j_\ell(k_{\ell n} r) Y_{\ell m}(\theta,\phi)\right), \tag{39}$$

where $\rho_0(\boldsymbol{r})$ is the observed mean density of galaxies in the survey.

Substituting equations (38) and (39) into equation (28) we find

$$\hat{\rho}_{\ell mn} = (\rho_0)_{\ell mn} + \sum_{\ell' m' n'} \left(\Phi_{\ell\ell' nn'}^{mm'} + \beta V_{\ell\ell' nn'}^{mm'}\right) \delta_{\ell' n'}^{m'}. \tag{40}$$

The mean–field harmonics are defined as

$$(\rho_0)_{\ell mn} = c_{ln} \int d^3 r\, \rho_0(\boldsymbol{r}) w(r) j_\ell(k_{\ell n} r) Y_{\ell m}^*(\theta,\phi) - c_{ln} \int d^3 r\, (\boldsymbol{v}_{\rm LG} \cdot \hat{\boldsymbol{r}}) \rho_0(\boldsymbol{r}) \frac{\mathrm{d}}{\mathrm{d}r}\left[w(r) j_\ell(k_{\ell n} r)\right] Y_{\ell m}^*(\theta,\phi), \tag{41}$$

where we have incorporated the the contribution of the dipole term, since this is independent of $\delta(\boldsymbol{r})$. In the rest of the analysis we assume a Local Group velocity of 622 km s^{-1} towards $\ell = 277°$, $b = 30°$.

The remaining matrices in equation (40) account for the mixing of modes due to the angular and radial geometry of the survey:

$$\Phi_{\ell\ell' nn'}^{mm'} = c_{\ell n} c_{\ell' n'} \int d^3 r\, \rho_0(\boldsymbol{r}) w(r) j_\ell(k_{\ell n} r) Y_{\ell m}^*(\theta,\phi) j_{\ell'}(k_{\ell' n'} r) Y_{\ell' m'}(\theta,\phi) \tag{42}$$

and the angular geometry and radial distortion:

$$V_{\ell\ell' nn'}^{mm'} = \frac{c_{\ell n} c_{\ell' n'}}{k_{\ell' n'}^2} \int d^3 r\, \rho_0(\boldsymbol{r}) \frac{d}{dr}\left[w(r) j_\ell(k_{\ell n} r)\right] Y_{\ell m}^*(\theta,\phi) j'_{\ell'}(k_{\ell' n'} r) Y_{\ell' m'}(\theta,\phi). \tag{43}$$

Defining the observable data vector $D_{\ell mn} = (\hat{\rho}_{\ell mn} - (\hat{\rho}_0)_{\ell mn})/\overline{\rho}$, and $\boldsymbol{\delta}$, the vector of underlying density field values, equation (40) can be written in a compact form as

$$\boldsymbol{D} = (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{\delta}. \tag{44}$$

where we have adopted a matrix notation.

A.5 Separation into angular and radial terms

The mean observed density field, $\rho_0(\boldsymbol{r})$, may be separated into angular and radial components

$$\rho_0(\boldsymbol{r}) = M(\theta,\phi)\, \rho_0(r), \tag{45}$$

with $M(\theta,\phi)$ being the sky mask of the survey and $\rho_0(r)$ containing the radial selection function. The 'transition matrices' $\boldsymbol{\Phi}$ and $\boldsymbol{V}$ now separate to become

$$\Phi_{\ell\ell' nn'}^{mm'} = W_{\ell\ell'}^{mm'} \Phi_{\ell\ell'}^{nn'} \tag{46}$$

$$V_{\ell\ell' nn'}^{mm'} = W_{\ell\ell'}^{mm'} V_{\ell\ell'}^{nn'} \tag{47}$$

where

$$\Phi_{\ell\ell'}^{nn'} = c_{\ell n} c_{\ell' n'} \int_0^{r_{\rm max}} r^2 dr\, \rho_0(r) w(r) j_{\ell'}(k_{\ell' n'} r) j_\ell(k_{\ell n} r), \tag{48}$$

$$V_{nn'}^{\ell\ell'} = \frac{c_{\ell n} c_{\ell' n'}}{k_{\ell' n'}^2} \int_0^{r_{\rm max}} r^2 dr\, \rho_0(r) \frac{d}{dr}\left[w(r) j_l(k_{ln} r)\right] \frac{d}{dr} j_{\ell'}(k_{\ell' n'} r) \tag{49}$$

and

$$W_{\ell\ell'}^{mm'} = \int_{4\pi} Y_{\ell'}^{m'}(\Omega) M(\Omega) Y_{\ell}^{*m}(\Omega) d\Omega. \quad (50)$$

A.6 "Fingers of god" and nonlinearities

Throughout the paper we have also applied the correction (described by HT) for the effect of the small scale non-linear peculiar velocity field (see their Appendix B). In this procedure, it is assumed that the real-space position of a galaxy is perturbed slightly by a peculiar velocity such that the radial redshift-space position becomes

$$s(\boldsymbol{r}) \rightarrow s'(\boldsymbol{r}) = s(\boldsymbol{r}) + \epsilon(\boldsymbol{r}). \quad (51)$$

The quantity $\epsilon(\boldsymbol{r})$ is a random variable drawn from a Maxwellian distribution with mean zero and $\langle \epsilon(\boldsymbol{r})\epsilon(\boldsymbol{r}')\rangle = (\sigma_v^2/3H_0^2)\delta_D(\boldsymbol{r}-\boldsymbol{r}')$ and where σ_v is the 3-d velocity dispersion. Errors in the measurement of redshift, σ_z^2 can be incorporated into the formalism by adding the variances in quadrature; $\sigma_{total}^2 = (\sigma_v^2/3) + \sigma_z^2$. Since this distortion is incoherent, the effect is to convolve each mode with a radial Gaussian distribution. From this HT derived a scattering matrix $\boldsymbol{S}$ defined by

$$S_{\ell\ell' nn'}^{mm'} = \frac{c_{\ell n} c_{\ell' n'} \delta_{\ell\ell'}^K \delta_{mm'}^K}{V\pi} \int\int \frac{e^{-\frac{(r-y)^2}{(2\sigma)^2}}}{\sqrt{2\pi}\sigma} j_\ell(k_{\ell n} r)\, j_{\ell'}(k_{\ell' n'} y)\; r dr\, y dy \quad (52)$$

which is to be convolved with the density field. δ^K is the Kronecker delta function. The amplitudes, $D_{\ell mn}$, corrected for a small–scale incoherent velocity field are now given by

$$D_{\ell mn} = S_{\ell\ell'' nn''}^{mm''} W_{\ell''\ell'}^{mm'} \left(\Phi_{\ell''\ell'}^{n''n'} + \beta V_{\ell''\ell'}^{n''n'}\right) \delta_{\ell' m' n'}, \quad (53)$$

or in matrix notation,

$$\boldsymbol{D} = \boldsymbol{SW}(\boldsymbol{\Phi} + \beta \boldsymbol{V})\delta. \quad (54)$$

A.7 Correlations

HT assumed that the sky mask was azimuthally symmetric, making the mask mixing matrices $\boldsymbol{W}$ real. In the current analysis, we drop this assumption as the sky mask for the PSCz survey is not symmetric (see Figure 7). This complicates the analysis, since the noise matrix is no longer real. We can avoid this problem by constructing the covariance matrix for the real and imaginary parts of $\boldsymbol{D}$. Since these are also Gaussian random, there are now four terms to the covariance matrix $\boldsymbol{C}$,

$$\begin{aligned} \boldsymbol{C}_{\Re e \Re e} &= \langle \Re e \boldsymbol{D}\, \Re e \boldsymbol{D}^t \rangle, & (55)\\ \boldsymbol{C}_{\Re e \Im m} &= \langle \Re e \boldsymbol{D}\, \Im m \boldsymbol{D}^t \rangle, & (56)\\ \boldsymbol{C}_{\Im m \Re e} &= \langle \Im m \boldsymbol{D}\, \Re e \boldsymbol{D}^t \rangle, & (57)\\ \boldsymbol{C}_{\Im m \Im m} &= \langle \Im m \boldsymbol{D}\, \Im m \boldsymbol{D}^t \rangle, & (58) \end{aligned}$$

where $\Re e$ signifies the real part and $\Im m$ the imaginary part. These terms are given by the following summations:

$$C_{\Re e \Re e} = \frac{1}{2}\sum_\alpha \left[\Re e(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Re e(\Phi_\nu^\alpha + \beta V_\nu^\alpha) + \Im m(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Im m(\Phi_\nu^\alpha + \beta V_\nu^\alpha)\right] P(k_\alpha) + N_{\mu\nu}, \quad (59)$$

$$C_{\Im m \Im m} = C_{\Re e \Re e}, \quad (60)$$

$$C_{\Re e \Im m} = \frac{1}{2}\sum_\alpha \left[\Re e(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Im m(\Phi_\nu^\alpha + \beta V_\nu^\alpha) - \Im m(\Phi_\mu^\alpha + \beta V_\mu^\alpha)\Re e(\Phi_\nu^\alpha + \beta V_\nu^\alpha)\right] P(k_\alpha) + N_{\mu\nu}, \quad (61)$$

$$C_{\Im m \Re e} = -C_{\Re e \Im m}, \quad (62)$$

where $\mu = (\ell mn)$, $\nu = (\ell' m' n')$ and $\alpha = (\ell'' m'' n'')$. The noise matrix, $\boldsymbol{N}$, is given by the equation

$$N_{\mu\nu} = c_{\ell n} c_{\ell' n'} \int \rho_0(r) w^2(r) j_\ell(k_{\ell n} r)\, j_{\ell'}(k_{\ell' n'} r)\, r^2 dr \int \mathcal{P}_\mu\left[Y_{\ell m}(\Omega)\right] M(\Omega) \mathcal{P}_\nu\left[Y_{\ell' m'}^{\star}(\Omega)\right] d\Omega \quad (63)$$

as in HT, where the $\mathcal{P}_{\mu,\nu}$ represent the real or imaginary parts of the spherical harmonics according to whether we are using the real or imaginary parts of $D_{\mu,\nu}$. In Appendix B we discuss in more detail the construction of the mask and its decomposition into correlations between real and imaginary parts of the field $\boldsymbol{D}$.

Now we have the covariance matrix, a likelihood functional is constructed and we use an assumed real-space power spectrum $P(k)$ in equations (59) to (62). We maximize the likelihood for the value of β and the amplitude of the power spectrum, as described in the main text. The amplitude of the power spectrum is defined in terms of $\Delta_{0.1}$ where $\Delta^2(k) =$

$k^3 P(k)/2\pi^2$. The likelihood function (unnormalised) we use is

$$\begin{aligned}\mathcal{L}[\boldsymbol{D}|\beta, P(k_{ln})] &= [\det(\boldsymbol{C}_{\Re e \Re e}) \det(\boldsymbol{C}_{\Re e \Im m}) \det(\boldsymbol{C}_{\Im m \Re e}) \det(\boldsymbol{C}_{\Im m \Im m})]^{-\frac{1}{2}} \\ &\times \ \exp\left(-\frac{1}{2}\Re e \boldsymbol{D}^t \boldsymbol{C}^{-1}_{\Re e \Re e} \Re e \boldsymbol{D} - \frac{1}{2}\Im m \boldsymbol{D}^t \boldsymbol{C}^{-1}_{\Im m \Re e} \Re e \boldsymbol{D}\right) \\ &\times \ \exp\left(-\frac{1}{2}\Re e \boldsymbol{D}^t \boldsymbol{C}^{-1}_{\Re e \Im m} \Im m \boldsymbol{D} - \frac{1}{2}\Im m \boldsymbol{D}^t \boldsymbol{C}^{-1}_{\Im m \Im m} \Im m \boldsymbol{D}\right). \end{aligned} \tag{64}$$

APPENDIX B: CONSTRUCTION OF THE ANGULAR WINDOW MATRIX

In this appendix we describe in detail the formulation and calculation of the angular window matrix. Since we drop the assumption that the angular mask is axisymmetric, calculation of the data covariance matrices is somewhat more complicated than that found in HT (see Appendix A.7). In this Appendix we define the window transition matrix, $\boldsymbol{W}$, and describe how we can calculate it efficiently with the use of Clebsch–Gordan, or Wigner 3j coefficients. We then show that formally the data covariance matrix is imaginary, due to the non-axisymmetric mask. We expand the window matrix and show how the data covariance matrix can be made positive definite by taking the covariances of the real and imaginary parts of the data vector, $\boldsymbol{D}$. Finally, we present the explicit expressions used to calculate the full covariance matrix, including the noise terms.

B.1 Definitions and conventions

An angular mask function is defined as

$$M(\Omega) = \begin{cases} 1 & \text{if observable area of sky,} \\ 0 & \text{otherwise} \end{cases} \tag{65}$$

From this we can define the transition matrix

$$W_{\ell\ell'}^{mm'} = \int_{4\pi} Y_{\ell'}^{m'}(\Omega) M(\Omega) Y_{\ell}^{m*}(\Omega)\, d\Omega, \tag{66}$$

In this Section we shall abbreviate our notation by dropping ℓ, m subscripts and $\int d\Omega$ except where needed for clarity. Note our convention for dashes and complex conjugates in the definition of $\boldsymbol{W}$. The $\boldsymbol{W}$ matrix operates on the density field by mixing modes;

$$D'_{\ell mn} = \sum_{\ell' m'} W_{ll'}^{mm'} D_{\ell' m' n}. \tag{67}$$

Rather than directly performing the integration in equation (66), which is a prohibitively slow process for large ℓ, it is faster to decompose the calculation into a two-step process, as outlined by Scharf *et al.* (1992). First one calculates the single transform of the mask function

$$M_{\ell m} = \int_{4\pi} M(\Omega) Y_\ell^{m*}(\Omega)\, d\Omega, \tag{68}$$

which can then be use to calculate the full transition matrix by the transformation ($Y' \equiv Y_{\ell'}^{m'}$ etc.)

$$W_{\ell\ell'}^{mm'} = \sum_{\ell'' m''} M_{\ell''}^{m''} \int Y' Y'' Y^*. \tag{69}$$

B.2 Clebsch–Gordan & Wigner 3j coefficients

The angular integral, $\int Y' Y'' Y^*$ can be calculated analytically by the Clebsch–Gordan or Wigner $3j$ matrices (Edmonds 1957). The Wigner $3j$ symbols are related to the Clebsch–Gordan coefficients by

$$\begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix} = (-1)^{\ell-\ell'-m''}(2\ell''+1)^{-1/2}\langle \ell m \ell' m' | \ell \ell' \ell'' - m''\rangle \tag{70}$$

Here we choose to express relations in terms of the $3j$ matrices. Hence

$$\int Y Y' Y'' = \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ m & m' & m'' \end{pmatrix}$$

with the triangle conditions $m + m' + m'' = 0$ and $|\ell - \ell'| \le \ell'' \le (\ell + \ell')$, plus permutations. Integrals containing complex conjugates of Y's can be derived using the hermicity of the spherical harmonics (equation (32)).

The Wigner $3j$ symbols are defined by Landau & Lifshitz (1977)

$$\begin{aligned}\left(\begin{array}{ccc}\ell & \ell' & \ell''\\ m & m' & m''\end{array}\right) = \delta^K_{m_1+m_2,m} \quad & (-1)^{\ell-\ell'-m''}\sqrt{\frac{(\ell+\ell'-\ell'')!(\ell-\ell'+\ell'')!(-\ell+\ell'+\ell'')!}{(\ell+\ell'+\ell''+1)!}}\\ & [(\ell+m)!(\ell-m)!(\ell'+m')!(\ell'-m')!(\ell''+m'')!(\ell''+m'')!]^{1/2}\\ & \sum_z (-1)^z [z!(\ell+\ell'-\ell''-z)!(\ell-m-z)!(\ell'+m'-z)!\\ & (\ell''-\ell'+m+z)!(\ell''-\ell'-m'+z)!]^{-1}\end{aligned} \tag{71}$$

where the summation over z is zero when any factorial terms in the summation are negative. When $m=m'=m''=0$ then

$$\begin{aligned}\left(\begin{array}{ccc}\ell & \ell' & \ell''\\ 0 & 0 & 0\end{array}\right) &= (-1)^{L/2}\sqrt{\frac{(L-2\ell)!(L-2\ell')!(L-2\ell'')!}{(L+1)!}}\frac{(L/2)!}{(L/2-\ell)!(L/2-\ell')!(L/2-\ell'')!}, \qquad \ell \quad \text{even},\\ &= 0, \qquad \ell \quad \text{odd}.\end{aligned} \tag{72}$$

where $L=\ell+\ell'+\ell''$. Practically it is easier to calculate these terms from $\left(\begin{array}{ccc}\ell & \ell & 0\\ 0 & 0 & 0\end{array}\right)$ using the recursive relations

$$\left(\begin{array}{ccc}\ell & \ell' & \ell''\\ 0 & 0 & 0\end{array}\right) = \sqrt{\frac{(L-2\ell'-1)(L-2\ell''+2)}{(L-2\ell')(L-2\ell''+1)}}\left(\begin{array}{ccc}\ell & \ell'+1 & \ell''-1\\ 0 & 0 & 0\end{array}\right). \tag{73}$$

With these relations, the computation time for the mask mixing matrix is reduced from days to a few minutes for ℓ up to 100.

B.3 Covariance Matrix Terms

Armed with the basic definitions of the window function transformation, and methods to generate it efficiently, we now begin to enumerate the various terms that will appear in the data covariance matrix. As stated above the problem of constructing a data covariance matrix is complicated by the combined presence of noise and a non–axisymmetric mask, leading to an imaginary covariance matrix. This can be dealt with in linear theory by decomposing the harmonic modes into real and imaginary parts and calculating their covariances. We begin with a formal derivation of the data covariance matrix.

The covariance matrix of the observed modes can be written in matrix notation as (c.f. equation (8))

$$\boldsymbol{C} = \langle \boldsymbol{D}\boldsymbol{D}^\dagger\rangle = \boldsymbol{S}\boldsymbol{W}(\boldsymbol{\Phi}+\beta\boldsymbol{V})\boldsymbol{P}(\boldsymbol{\Phi}+\beta\boldsymbol{V})^t\boldsymbol{W}^\dagger\boldsymbol{S}^t + \boldsymbol{R}\boldsymbol{W}, \tag{74}$$

where $\boldsymbol{P} = \mathrm{diag}[P(k_{\ell n})]$ and the shot noise term is given by $\boldsymbol{N} = \boldsymbol{R}\boldsymbol{W}$, which we have decomposed into radial and angular parts (c.f. equation (63)),

$$\boldsymbol{R} = c_{\ell n}c_{\ell' n'}\int \rho_0(r)w^2(r)j_\ell(k_{\ell n}r)\,j_{\ell'}(k_{\ell' n'}r)\,r^2dr \tag{75}$$

and

$$\boldsymbol{W} = \int Y' M Y^*. \tag{76}$$

Note that here we use the full window transition matrix.

Expanding $\boldsymbol{W}$ into real and imaginary parts, $\boldsymbol{W} = \Re e\boldsymbol{W} + i\Im m\boldsymbol{W}$, yields

$$\begin{aligned}\boldsymbol{C} \;=\;& \boldsymbol{S}\,\Re e\boldsymbol{W}(\boldsymbol{\Phi}+\beta\boldsymbol{V})\boldsymbol{P}(\boldsymbol{\Phi}+\beta\boldsymbol{V})^t\,\Re e\boldsymbol{W}^t\boldsymbol{S}^t + \boldsymbol{S}\,\Im m\boldsymbol{W}(\boldsymbol{\Phi}+\beta\boldsymbol{V})\boldsymbol{P}(\boldsymbol{\Phi}+\beta\boldsymbol{V})^t\,\Im m\boldsymbol{W}^t\boldsymbol{S}^t + \boldsymbol{R}\,\Re e\boldsymbol{W}\\ +\;& i\boldsymbol{S}\,\Im m\boldsymbol{W}(\boldsymbol{\Phi}+\beta\boldsymbol{V})\boldsymbol{P}(\boldsymbol{\Phi}+\beta\boldsymbol{V})^t\,\Re e\boldsymbol{W}^t\boldsymbol{S}^t - i\boldsymbol{S}\,\Re e\boldsymbol{W}(\boldsymbol{\Phi}+\beta\boldsymbol{V})\boldsymbol{P}(\boldsymbol{\Phi}+\beta\boldsymbol{V})^t\,\Im m\boldsymbol{W}^t\boldsymbol{S}^t + i\boldsymbol{R}\,\Im m\boldsymbol{W}\end{aligned} \tag{77}$$

where the real and imaginary parts of $\boldsymbol{W}$ can also be written

$$\Re e\boldsymbol{W} = \mathrm{Sym}\,\boldsymbol{W}, \tag{78}$$

and

$$\Im m\boldsymbol{W} = \frac{1}{i}\mathrm{Asym}\boldsymbol{W}, \tag{79}$$

where the operators Sym and Asym symmetrize and anti-symmetrize the matrices.

B.4 Expansion of the window matrix

Although imaginary terms exist in the data covariance matrix, we can use the property of a Gaussian field that the real and imaginary parts will be independent and each is independently Gaussian distributed to construct total the probability distribution function. We later show that the covariance matrix can be further expanded into independent terms that will allow us to define a total covariance matrix which is both real and symmetric.

Expanding the Y's into real and imaginary parts we find that

$$\boldsymbol{W} = (\Re e\Re e\boldsymbol{W} + \Im m\Im m\boldsymbol{W}) + i(\Re e\Im m\boldsymbol{W} - \Im m\Re e\boldsymbol{W}), \quad (80)$$

where we introduce the notation $\Re e\Re e\boldsymbol{W} = \int \Re eYM\Re eY'$, $\Im m\Im m\boldsymbol{W} = \int \Im mYM\Im mY'$, $\Re e\Im m\boldsymbol{W} = \int \Re eYM\Im mY'$ and $\Im m\Re e\boldsymbol{W} = \int \Im mYM\Re eY'$. The real and imaginary parts of the spherical harmonics can be found from

$$\Re eY = \frac{1}{2}(Y + Y^*) \quad (81)$$

and

$$\Im mY = \frac{1}{2i}(Y - Y^*) \quad (82)$$

The terms $\Re e\Re e\boldsymbol{W}$ can then be calculated from the full window matrix by

$$\begin{aligned}
\Re e\Re e\boldsymbol{W} &= \frac{1}{4}\left(\int YMY' + \int Y^*MY' + \int YMY'^* + \int Y^*MY'^*\right),\\
\Im m\Im m\boldsymbol{W} &= \frac{1}{4}\left(\int Y^*MY' + \int YMY'^* - \int YMY' - \int Y^*MY'^*\right),\\
\Re e\Im m\boldsymbol{W} &= \frac{1}{4i}\left(\int YMY' - \int YMY'^* + \int Y^*MY' - \int Y^*MY'^*\right),
\end{aligned} \quad (83)$$

where the various permutations can be calculated from the window function matrix using the Hermicity relations. These are the terms that will appear in the noise matrix, equation (63). Note the ordering of dashed terms as these terms are not in general symmetric in their indices. Clearly then $\Re e\boldsymbol{W} = \Re e\Re e\boldsymbol{W} + \Im m\Im m\boldsymbol{W}$ and $\Im m\boldsymbol{W} = \Re e\Im m\boldsymbol{W} - \Im m\Re e\boldsymbol{W}$.

B.3 Transpose relations

Here we list without proof the transpose relations for the three window matrices.

$$\begin{aligned}
\Re e\Re e\boldsymbol{W} &= (\Re e\Re e\boldsymbol{W})^T,\\
\Im m\Im m\boldsymbol{W} &= (\Im m\Im m\boldsymbol{W})^T,\\
\Im m\Re e\boldsymbol{W} &= (\Re e\Im m\boldsymbol{W})^T.
\end{aligned} \quad (84)$$

Given the symmetry relation $\Im m\Re e\boldsymbol{W} = (\Re e\Im m\boldsymbol{W})^T$, we only need explicitly to calculate the three window matrices $\Re e\Re e\boldsymbol{W}$, $\Im m\Im m\boldsymbol{W}$ and $\Im m\Re e\boldsymbol{W}$.

B.4 Parity Relations

Here we list without proof the parity relations ($m \to -m$) for the three basic window matrices.

$$\begin{aligned}
\Re e\Re eW_{m,-m'} &= (-1)^{m'}\Re e\Re eW_{mm'},\\
\Re e\Re eW_{-m,m'} &= (-1)^{m}\Re e\Re eW_{mm'},\\
\Re e\Re eW_{-m,-m'} &= (-1)^{m'+m}\Re e\Re eW_{mm'},
\end{aligned} \quad (85)$$

$$\begin{aligned}
\Im m\Im mW_{m,-m'} &= (-1)^{m'+1}\Im m\Im mW_{mm'},\\
\Im m\Im mW_{-m,m'} &= (-1)^{m+1}\Im m\Im mW_{mm'},\\
\Im m\Im mW_{-m,-m'} &= (-1)^{m+m'}\Im m\Im mW_{mm'},
\end{aligned} \quad (86)$$

$$\begin{aligned}
\Im m\Re eW_{m,-m'} &= (-1)^{m'}\Im m\Re eW_{mm'},\\
\Im m\Re eW_{-m,m'} &= (-1)^{m+1}\Im m\Re eW_{mm'},\\
\Im m\Re eW_{-m,-m'} &= (-1)^{m+m'+1}\Im m\Re eW_{mm'}.
\end{aligned} \quad (87)$$

B.5 Expansion of covariance matrix

We can re-arrange the covariance function by reducing the data vector into real and imaginary parts; $\boldsymbol{D} = \Re e\boldsymbol{D} + i\Im m\boldsymbol{D}$. This is convenient because our hypothesis is that $\boldsymbol{D}$ is a Gaussian random variable and so the terms $\Re e\boldsymbol{D}$ and $\Im m\boldsymbol{D}$ will also be Gaussian random variables. In this case we need terms like $\langle \Re e\boldsymbol{D}\,\Re e\boldsymbol{D}'\rangle$. This leads us to

$$\boldsymbol{C} = (\langle \Re e\boldsymbol{D}\,\Re e\boldsymbol{D}'^t\rangle + \langle \Im m\boldsymbol{D}\,\Im m\boldsymbol{D}'^t\rangle) + i(\langle \Im m\boldsymbol{D}\,\Re e\boldsymbol{D}'^t\rangle - \langle \Re e\boldsymbol{D}\,\Im m\boldsymbol{D}'^t\rangle). \quad (88)$$

We can further expand terms by expanding the real and imaginary terms of the $\boldsymbol{W}$ matrix and the underlying field, δ. Hence $\Re e D = \Re e \boldsymbol{W} \Re e \delta - \Im m \boldsymbol{W} \Im m \delta$ and $\Im m D = \Im m \boldsymbol{W} \Re e \delta + \Re e \boldsymbol{W} \Im m \delta$. Under the assumption of statistical independence between the real and imaginary δ's we find

$$\begin{aligned}
\langle \Re e \boldsymbol{D} \Re e \boldsymbol{D}^t \rangle &= \frac{1}{2} \boldsymbol{S} \Re e \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Re e \boldsymbol{W}^t \boldsymbol{S}^t + \boldsymbol{S} \Im m \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Im m \boldsymbol{W}^t \boldsymbol{S}^t + \boldsymbol{R} \Re e \Re e \boldsymbol{W}, \\
\langle \Im m \boldsymbol{D} \Im m \boldsymbol{D}^t \rangle &= \frac{1}{2} \boldsymbol{S} \Re e \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Re e \boldsymbol{W}^t \boldsymbol{S}^t + \boldsymbol{S} \Im m \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Im m \boldsymbol{W}^t \boldsymbol{S}^t + \boldsymbol{R} \Im m \Im m \boldsymbol{W}, \\
\langle \Re e \boldsymbol{D} \Im m \boldsymbol{D}^t \rangle &= \frac{1}{2} \boldsymbol{S} \Re e \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Im m \boldsymbol{W}^t \boldsymbol{S}^t - \boldsymbol{S} \Im m \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Re e \boldsymbol{W}^t \boldsymbol{S}^t - \boldsymbol{R} \Re e \Im m \boldsymbol{W}, \\
\langle \Im m \boldsymbol{D} \Re e \boldsymbol{D}^t \rangle &= \frac{1}{2} \boldsymbol{S} \Im m \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Re e \boldsymbol{W}^t \boldsymbol{S}^t - \boldsymbol{S} \Re e \boldsymbol{W} (\boldsymbol{\Phi} + \beta \boldsymbol{V}) \boldsymbol{P} (\boldsymbol{\Phi} + \beta \boldsymbol{V})^t \Im m \boldsymbol{W}^t \boldsymbol{S}^t - \boldsymbol{R} \Im m \Re e \boldsymbol{W}.
\end{aligned} \tag{89}$$

Substituting these expressions into equation (88) we reproduce equation (77). In addition we have shown that each of these terms are independent and Gaussian distributed. Therefore we can construct a real covariance matrix;

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{\Re e \Re e} & \boldsymbol{C}_{\Re e \Im m} \\ \boldsymbol{C}_{\Im m \Re e} & \boldsymbol{C}_{\Im m \Im m} \end{pmatrix} \tag{90}$$

From the symmetry relations given earlier we see that this is a real, symmetric matrix.