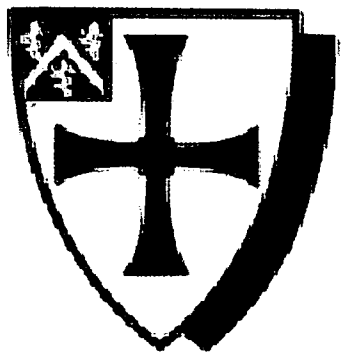
The clustering of galaxies on large scales

William James Frith

THE CLUSTERING OF GALAXIES ON LARGE SCALES

WILLIAM JAMES FRITH



A thesis submitted to the University of Durham
in accordance with the regulations for admission to the
Degree of Doctor of Philosophy.

UNIVERSITY OF DURHAM

2005

ABSTRACT: THE CLUSTERING OF GALAXIES ON LARGE SCALES

WILLIAM JAMES FRITH

We investigate the local large-scale structure of the Universe, addressing various possible issues confronting the ΛCDM paradigm. Primarily, we investigate the clustering statistics of the newly-completed 2 Micron All Sky Survey (2MASS), the largest all sky galaxy survey to date.

The 2MASS galaxy number counts over the $\approx$4000 deg^2 APM survey area are found to be low compared to predictions but are in good agreement with previous optical results. Surprisingly, the number counts over almost the entire sky ($|b|$ >20°, $\approx$ 27 000 deg^2) are also deficient compared to our predictions. These results do not appear to be significantly affected by systematic errors. Assuming a ΛCDM cosmology, the observed deficiencies in the APM survey area and for $|b|$ >20° represent $\approx 2.5\sigma$ and $\approx 4.0\sigma$ fluctuations in the local galaxy distribution respectively. These results are therefore potentially at odds with the form of clustering expected on large scales.

We examine the form of galaxy clustering to $r \lesssim 1000\,h^{-1}$ Mpc scales using the 2MASS angular power spectrum. We find a 3σ excess over mock ΛCDM results; however this is not enough to account for the observed number counts mentioned above. We determine the implied cosmological constraints; the 2MASS galaxy angular power spectrum is, in fact, in strong support of ΛCDM, with a measured power spectrum shape of $\Gamma_{\rm eff} = 0.14 \pm 0.02$. In addition, we determine a K_s-band galaxy bias of $b_K = 1.39 \pm 0.12$.

We determine high-order correlation functions of the 2MASS galaxy sample to extremely large scales ($r \lesssim 100\,h^{-1}$ Mpc). The results are in strong support of Gaussian initial conditions and hierarchical clustering; we reject primordial strong non-Gaussianity at the $\approx 2.5\sigma$ confidence level. Unlike all previous such analyses, our results are relatively robust to the removal of large superclusters from the sample. We also measure a K_s-band quadratic galaxy bias of $c_2 = 0.57 \pm 0.33$. This result differs significantly from previous negative constraints; we discuss a possible explanation for this apparent discrepancy.

Finally, we examine the extent of possible Sunyaev-Zeldovich contamination in the first year Wilkinson Microwave Anisotropy Probe (WMAP) data using various foreground galaxy cluster catalogues. We find evidence suggesting that the associated temperature decrements extend to $\gtrsim 1°$ scales. Such a result would indicate a much higher baryon density than the concordance value; in addition, CMB power spectrum fits and the associated cosmological constraints would also be compromised.

PREFACE

THE CLUSTERING OF GALAXIES ON LARGE SCALES

The work described in this thesis was undertaken between October 2002 and September 2005 whilst the author was a research student under the supervision of Prof. Tom Shanks in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at this (or any other) university.

The main collaborators involved in this thesis are Prof. T. Shanks, Dr. P.J. Outram, Dr. A.D. Myers, Prof. Sir A.W. Wolfendale, Dr. N. Metcalfe, Dr. R. Fong and Dr. G.S. Busswell.

Results from this thesis have appeared in the following papers:

- Frith, W.J., Busswell, G.S., Fong, R., Metcalfe, N. & Shanks, T. 2003, MNRAS, 345, 1049
- Busswell, G.S., Shanks, T., Outram, P.J., Frith, W.J., Metcalfe, N. & Fong, R. 2004, MNRAS, 354, 991
- Frith, W.J. 2004, MNRAS, 348, 916
- Myers, A.D., Shanks, T., Outram, P.J., Frith, W.J. & Wolfendale, A.W. 2004, MNRAS, 347, L67 (Chap. 6)
- Frith, W.J., Shanks, T. & Outram, P.J. 2005a, MNRAS, 361, 701 (Chap. 2)
- Frith, W.J., Outram, P.J. & Shanks, T. 2005b, MNRAS, 364, 593 (Chap. 4)
- Frith, W.J., Outram, P.J. & Shanks, T. 2005c, submitted to MNRAS, astro-ph/0507704 (Chap. 5)
- Frith, W.J., Metcalfe, N. & Shanks, T. 2005d, submitted to MNRAS, astro-ph/0509875 (Chap. 3)

and in the conference proceedings:

- Frith, W.J., Outram, P.J. & Shanks, T. 2004, ASP Conf. Proc., Vol. 329, 49

ACKNOWLEDGEMENTS

I will begin by thanking my supervisor, Tom Shanks, without whose enthusiasm and encouragement none of the work presented here would have been possible. I have to thank him in particular for providing such a stimulating and enjoyable environment to work in and I hope that I have, to some extent, fulfilled the faith he invested in me at the beginning of this PhD.

I must also thank his extremely able deputy, Phil Outram, who has helped me through many aspects of this work. It is he I have to thank for answering some of the more basic questions I was too embarrassed to ask Tom, and in assisting with many of the more detailed problems which arose; I refer to work related issues here rather than the many hours devoted to discussing football, for which I thank him also.

On this point I also thank Nic Ross and Adam Myers, my two office partners, who helped to create a lively and intellectual atmosphere. Their taste in football teams is not great, but they both made working in the department very enjoyable. Along with them I should also mention many of my contemporaries who have helped to make the day-to-day aspect of research and Durham more agreeable and have in most cases either been with me or against me on the Cosmic five-a-side football team or taken large quantities of money from me through my terrible poker playing; Jose da Angela, Kris Beckwith, Craig Booth, Anthony Brown, Rob Crain, Greg Davies, Geraint Harker, John Helly, Noam Libeskind, Matt Middleton, Mark Norris, Georgis Mountrichas, Dave Radburn-Smith, Mark Swinbank, Marc Vallbe, Claudio della Vecchia and Rich Whitaker.

The department is also blessed with many extremely clever people with whom I've had the privilege either to collaborate or receive advice and discuss ideas with more generally. First, I should mention Dick Fong who, as my fourth year tutor first raised the idea of me continuing in research and encouraged me to come back over the summer after my graduation to publish some of our results. Dating from that time, I also frequently sought advice from Shaun Cole and Carlton Baugh, to whom I both thank and apologise to for bothering them so much. Both were patient in the face of most often banal questioning,

and very kind in allowing me access to their results. Also, I thank Sir Arnold Wolfendale for many fascinating discussions and helpful advice, and Nigel Metcalfe for his help in advancing my computing and observing skills dating back to the third year of my degree. I would also like to thank Geoff Busswell, Peter Draper, Alastair Edge, Pete Edwards, Vincent Eke, Adrian Jenkins, Alan Lotts, John Lucey, Peder Norberg and Chris Simpson who have all helped me along the way at some level.

Realising that I'm now beginning to over-run, I'll turn to thanking those who have not directly assisted in this work. First, all those I've lived with over the last three years; the New Belvedere lot with whom I enjoyed one of the best years I've had in seven years here, particularly Mital Patel, Matt Hill and Dan Wilkinson. And of course no acknowledgement would be complete without thanking Ruth Emerson, my girlfriend of two years who I met in New Belvedere on the first day of my PhD, and who has supported me throughout my time here. She had to endure numerous discussions on the Cardassian model and also astronomy in general; for the former in particular, I thank her. Also, I thank Greig Coull, Dave Haigh, Dan Kerry, Mark Morley-Fletcher, Zoe Waterhouse and Rich Whitaker who have all had to put up with me for several years, and have helped to keep me going, usually with a beer of some kind.

Finally, I thank all my other friends and of course my family. The greatest thanks I reserve for my parents without whom nothing I have achieved would have been possible.

This thesis is dedicated to my parents

Contents

List of Figures

List of Tables

PROLOGUE

As the results and analysis in this thesis have already been written as separate papers, chapters 2 to 5 of this thesis are presented largely in the form that they are published (Frith et al. 2005a,b,c,d); chapter 6 has been rewritten as I am not first author on this paper (Myers et al. 2004). Each contains a separate introduction which explores the background and relevance of the work presented in that chapter. As a result, the first chapter is not a self-contained introduction to the whole thesis, but rather a brief overview of the current status of cosmology. Within the first three chapters, work carried out in Frith et al. (2003) and Busswell et al. (2004) is also referred to as part of the background on bright number counts. Similarly, some remarks on an alternative cosmological model, which are presented in Frith (2004), are also incorporated in the concluding chapter.

CHAPTER 1 | INTRODUCTION

1.1 THE STANDARD COSMOLOGICAL MODEL

The standard cosmological model has emerged over the last century through tandem theoretical and observational developments. As data has accumulated in support of this world picture, so has the belief among the astronomical community that it provides a useful working approximation of the large-scale nature of the Universe. The model may be summarised in four key elements:

1.1.1 THE COSMOLOGICAL PRINCIPLE

The most fundamental supposition of the standard cosmological model is that the Universe appears the same on the large-scale average, in a statistical sense, when undergoing a rotational or translational transform. It is certain that prominent departures from the mean occur at relatively small scales ($r \lesssim 10\,h^{-1}\,\mathrm{Mpc}$); much of this thesis is concerned with determining to what scales this 'cosmological principle' may be applied.

The simplest expression of large-scale isotropy and homogeneity is the similar nature of galaxy number counts observed in different directions. Since this is a topic to which we return in chapters 2 and 3, some information about this statistic is detailed here. We consider the relation between the observed flux of a galaxy and its luminosity in a Euclidean Universe, $f = L/4\pi r^2$. The number of galaxies per steradian above some flux limit is then:

$$N(>f) = nV = \frac{nr^3}{3} = \frac{n}{3}\left(\frac{L}{4\pi f}\right)^{3/2} \tag{1.1}$$

where n represents the mean number of galaxies per unit volume. The observed brightness of a galaxy is conventionally expressed by its apparent magnitude, m, where $f \propto 10^{-0.4m}$. The expected number of galaxies as a function of apparent magnitude for a spatially homogeneous distribution is therefore:

$$N(<m) \propto 10^{0.6m} \tag{1.2}$$

Galaxy number counts are observed to have a similar slope in all directions at magnitudes where the small-scale departures from the cosmological principle are expected to be averaged over the large surveyed volumes (see Fig. 3.1), suggesting that the galaxy distribution, at least, is homogeneous and isotropic on large scales.

1.1.2 The Expansion of the Universe

Soon after the approximately Euclidean nature of galaxy number counts at faint magnitudes was determined, Hubble made the surprising discovery that the distances to local galaxies, r, and their apparent velocities, v, were linearly related (Hubble 1929), such that:

$$r = \frac{v}{H_0} = \frac{cz}{H_0} \tag{1.3}$$

where the present value of the Hubble parameter, H_0, represents the slope of the correlation, and the redshift, z, defines the factor by which the wavelength of the observed light is distorted by the apparent velocity of the galaxy, $(1+z)$, such that if the wavelength of the light emitted by a galaxy and the observed wavelength are denoted by λ_1 and λ_0 respectively, $z = (\lambda_0 - \lambda_1)/\lambda_1$.

Ignoring peculiar motions caused by local irregularities, galaxies are observed to be receding away from us in this manner; considering the cosmological principle this implies that the Universe is uniformly expanding. Hubble's original constraint on the local rate of this expansion of $H_0 \approx 550\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ has since been shown to be an over-estimate; recent constraints suggest a value for the Hubble constant of $h = H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}} = 0.5 - 1.0$.

1.1.3 The Dynamics of the Universe

The standard cosmological model assumes that the dynamics of the Universe are described by Einstein's theory of general relativity. An approximate argument follows: Assuming that the radiation and material pressure are small, we consider the gravitational mass inside a sphere of radius a predicted by Newtonian mechanics:

$$M = \rho V = \frac{4}{3}\pi\rho a^3 \tag{1.4}$$

Employing the usual inverse square law, the acceleration due to gravity at the surface of this sphere is:

$$\ddot{a} = -\frac{GM}{a^2} = -\frac{4}{3}\pi G\rho a \tag{1.5}$$

The Hubble parameter, H, may be expressed more generically as the rate of the expansion at any particular moment in the history of the Universe, rather than the current (or local) value commonly referred to as the Hubble constant, H_0. From equation 1.5, the first Friedmann equation can be derived:

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8}{3}\pi G\rho - \frac{k}{a^2} + \frac{1}{3}\Lambda \tag{1.6}$$

According to General Relativity, k represents the curvature of space. If k is positive, then space is positively curved and the Universe is said to be 'closed'. In the absence of the final term of equation 1.6, the Universe will eventually begin to contract in this scenario. Conversely, if k is negative then the resulting 'open' Universe expands forever (again if $\Lambda = 0$). The value of k is believed to be close to zero implying that space is 'flat' with zero curvature.

The final term of equation 1.6 was introduced by Einstein in order to maintain a static Universe. However, following the evidence for an expanding Universe Einstein saw no further logical need for it. In fact, recent evidence suggests the need for such a term which we discuss in section 1.4.2; it is testament to the primacy of this theory that it has been able to accommodate these shifts in the cosmological paradigm. With the introduction of this cosmological constant, the spatial curvature and expansion history are no longer locked together; it is possible for instance for a closed Universe to expand forever.

As with other areas of astronomy, the use of Einstein's field equations as a description of the dynamics of the Universe assumes the validity of applying physics determined locally (in the solar system) to the farthest reaches of space and time. This assumption is not trivial since it has led to the invoking of two pieces of physics required in order to account for the missing mass and energy densities observed in recent experiments (see section 1.4), which themselves have not been discovered. However, until the existence of these dark components of the energy density are proved or otherwise, this theory provides an elegant framework which appears to offer a good approximation of reality.

1.1.4 THE BIG BANG

The concept that the Universe has expanded from a hot, dense state does not automatically spring from the idea of an expanding Universe; this theory has gained favour over the last fifty years only through the arrival of several pieces of observational evidence. The term 'big bang model' is now commonly used in reference to the standard cosmological model. The name is slightly misleading in that there is currently no observational evidence to suggest that there was a specific moment when the Universe began; only that the observable Universe was once much smaller than it is currently.

The most important evidence in support of the big bang model are the properties of the Cosmic Microwave Background (CMB; Penzias & Wilson 1965). This radiation is extremely isotropic, varying by only a few parts in 10^5 (e.g. Smoot et al. 1992), suggesting that the CMB uniformly fills space. The radiation has an almost perfect black body spectrum indicating that it was produced by a process in almost perfect thermodynamic equilibrium, and at an epoch at which the Universe was optically thick. A natural interpretation is that the CMB is radiation which became free to propagate through space after the recombination of electrons and protons, a process that was allowed to occur once the Universe, having been both very small and hot, expanded and cooled enough to allow matter and radiation to decouple.

In addition, the theory of Big Bang Nucleo-Synthesis (BBNS) predicts the relative quantities of the light elements, deuterium, helium and lithium, produced through nuclear fusion in the three minutes after the big bang (Alpher, Bethe & Gamow 1948). Current observations of the deuterium abundance at high redshifts point to a constraint of $\Omega_b h^2 \approx 0.014 \pm 0.009$ (e.g. Kurki-Suonio 2002), where Ω_b denotes the baryon density, a value which has been corroborated by independent constraints such as those arising from CMB observations (e.g. $\Omega_b h^2 = 0.0224 \pm 0.0009$, Bennett et al. 2003).

A further significant piece of evidence is derived from the observed ages of the currently known oldest objects which coincide with theoretical estimates for the age of the Universe. Taking definitions for the density of matter, the curvature parameter and the parameter associated with the cosmological constant respectively:

$$\Omega_m = \frac{8\pi G\rho_0}{3H_0^2}$$

$$\Omega_k = \frac{k}{(a_0 H_0)^2}$$

$$\Omega_\Lambda = \frac{\Lambda}{3H_0^2} \tag{1.7}$$

(where the subscript '0' denotes the current value as with the Hubble parameter), and since the mass density and scale factor vary as $\rho \propto a^{-3}$ and $a \propto (1+z)^{-1}$ if the material and radiation pressures are small, it follows from equation 1.6 that:

$$H_0 t_0 = \int_1^{\infty} \frac{dy}{y(\Omega_m y^3 + \Omega_k y^2 + \Omega_\Lambda)^{1/2}} \tag{1.8}$$

where $y = 1+z$ and t_0 represents the current age of the Universe. In the $k = 0$ Einstein-de Sitter limit, that is when the pressures due to matter and radiation are small compared to the mass density and the space curvature and cosmological constant are both negligible, equation 1.8 results in the expression $H_0 t_0 = \frac{2}{3}$. Current measurements indicating $h = 0.71 \pm 0.04$ (Bennett et al. 2003) therefore indicate that for an Einstein-de Sitter Universe the age is $t_0 \approx 9$ Gyr. More recent estimates including current constraints on the various components of the energy density suggest that the Universe is slightly older than this, with $t_0 = 13.7 \pm 0.2$ Gyr (Bennett et al. 2003). This constraint is consistent with the ages of globular clusters in the galactic halo derived from stellar evolution models (e.g. $t_0 = 16 \pm 2$ Gyr, Renzini et al. 1996).

1.2 Issues for the Big Bang Model

The Big Bang model presents the problem of how the various structures of galaxies that we observe in the local Universe formed. If we accept that the fluctuations observed in CMB experiments are the signature of perturbations in the mass distribution which are imprinted onto the photon field at the time of the decoupling of matter and radiation through pressure oscillations in the primordial soup of baryons and photons, then the growth of structure may be brought about from these seed fluctuations through the action of gravitational instability. The mass density may be expressed in the form:

$$\rho(\mathbf{r}, t) = \bar{\rho}(t)[1 + \delta(\mathbf{r}, t)] \tag{1.9}$$

where $\bar{\rho}$ represents the mean mass density at a particular moment in the history of the Universe, t, and $\delta = \delta\rho/\rho$ is the fractional departure from this mean at position $\mathbf{r}$ and time t. Considering only linear perturbation theory which assumes that the density fluctuations are small, $|\delta| \ll 1$, it can be shown that for an Einstein-de Sitter cosmology, $\delta \propto a \propto (z+1)^{-1}$. Gravity it seems may be able to account for the observed structure if the primordial density fluctuations are of a sufficient amplitude (given that the redshift

at recombination was $z_{rec} \approx 1000$), although reproducing the exact form of the large-scale structure at low redshifts also hinges on more subtle properties of the matter which are discussed in section 1.4.1. The problem remains as to what seeds these primordial fluctuations.

There are two further significant problems for the standard cosmological model. The first concerns the isotropy of the CMB; the observed uniformity is not expected as the CMB photons have been emitted from regions too widely separated to be causally connected; at the time of recombination, the horizon scale when projected onto the sky is $\approx 1°$. The second problem arises from the fact that observations indicate that the Universe is close to being flat with an energy density very near to the critical value, $\Omega_{tot} = 1$. However, the standard cosmological model predicts that the energy density rapidly evolves away from this value; the initial value of Ω_{tot} has to be fine-tuned such that it is equal to the critical value within 1 part in 10^{15}.

These issues confronting the standard cosmological model can be resolved by the theory of inflation (Guth 1981). In this scenario, a phase of accelerated expansion occurs in the second after the big bang, driven by a scalar ('inflaton') field. Quantum fluctuations in this field are amplified by the expansion yielding the primordial density fluctuations at the time of recombination which seed structure formation. In addition, the particle horizons in the observable Universe are eliminated therefore allowing regions previously beyond the horizon to be causally connected. The fine-tuning problem is also addressed since the expansion causes the curvature of space to become flatter, forcing Ω_{tot} towards the critical value whatever its initial value. The theory of inflation has gained wide acceptance due to the success with which it can eliminate the problems enumerated above. However, there are a huge variety of different models offering various mechanisms by which the expansion might occur. The overriding problem with this theory is that there are few observable constraints and even these are not able, even according to current theory, to eliminate many of the various models of inflation. This is a subject to which we return in chapter 5, where we attempt to constrain an observable of inflation, the distribution of primordial density fluctuations.

1.3 STATISTICAL COSMOLOGY

Constraining the parameters of the standard cosmological model, such as the density of dark matter, is a vital step in understanding the nature of the Universe as it allows us

to address fundamental issues such as the age of the Universe and its ultimate fate. In order to do this successfully, experiments which cover significant volumes are required in order to ensure that the results are not affected by cosmic variance as defined by the cosmological principle. With accordingly large surveys it is necessary to analyse the data in a statistical manner; the results of such surveys are reviewed in the next section. First, two commonly used statistics in the study of large-scale structure, the correlation function and power spectrum, are briefly described since this is a subject of particular relevance to this thesis. In chapters 4 and 5 the *angular* power spectrum and correlation functions are used for which a similar rationale applies; a detailed account of these statistics is given in sections 4.3 and 5.2 respectively.

The 2-point galaxy correlation function, $\xi(r)$, is defined by the joint probability of finding two galaxies in each of two volume elements, dV_1 and dV_2, separated by distance r:

$$dP = n^2 dV_1 dV_2 [1 + \xi(r)] \tag{1.10}$$

where n denotes the mean number of galaxies per unit volume. $\xi(r)$ thus defines the excess probability of finding two galaxies at separation r compared to a Gaussian random distribution; if there is any form of clustering present at scale r, $\xi(r)$ is positive as the probability of finding two galaxies with such separations increases.

Measuring the 2-point galaxy correlation function requires an accurate knowledge of the survey selection function from which the galaxy sample is drawn, so as to produce a distribution of randomly distributed points with the same spatial and angular limits, through which the excess probability of finding galaxy pairs can be determined. A commonly used estimator is:

$$\xi(r) = \frac{DD(r)}{RR(r)} \left(\frac{\bar{n}_R}{\bar{n}_D} \right)^2 - 1 \tag{1.11}$$

where DD and RR represent the number of galaxy and random pairs recorded as a function of separation, r, and $\bar{n}_D$ and $\bar{n}_R$ denote the density per unit volume of galaxy and random points respectively.

The correlation function has proved a powerful discriminant of cosmological models as we discuss in the next section. However, the power spectrum, the Fourier transform of the correlation function, has become a more favoured method of constraining the variance of galaxy density fluctuations and the parameters of the standard cosmological model. Both

statistics are affected by uncertainties in the mean density of galaxies, n; however, the correlation function is affected by these uncertainties on all scales whereas the power spectrum is affected only on large scales. In addition, the power spectrum is the quantity directly predicted by perturbation theory. The density contrast, δ, at position $\mathbf{r}$, described in equation 1.9, may be Fourier expanded:

$$\delta(\mathbf{r}) = \frac{1}{8\pi} \int \delta_k e^{-i\mathbf{k}.\mathbf{r}} \, d^3k. \tag{1.12}$$

The amplitude of the modes of the density field are given by the Fourier transform of this equation:

$$\delta_k = \int \delta(\mathbf{r}) e^{i\mathbf{k}.\mathbf{r}} \, d^3r. \tag{1.13}$$

The power spectrum, which defines the amplitude of density fluctuations as a function of wavenumber, k, is given by $P(k) = |\delta_k|^2$. Typically, the normalisation of the power spectrum is given in terms of the rms mass fluctuation on scales of $8\,h^{-1}\,\mathrm{Mpc}$, σ_8.

In chapter 4, we describe an angular power spectrum analysis which, rather than using a Fourier expansion as described here for the three-dimensional power spectrum, uses a spherical harmonic expansion of the angular density distribution of galaxies. This statistic is particularly relevant for the subject of statistical cosmology since powerful cosmological parameter constraints are derived from the angular power spectrum of temperature fluctuations in the CMB which are reviewed in the following section.

1.4 Components of the Standard Cosmological Model

1.4.1 Dark Matter

Considerable evidence has accumulated over the last eighty years suggesting that the dominant form of the mass component of the Universe does not reside in the baryonic matter visible in stars and galaxies. Hubble (1929) first estimated the visible mass density using galaxy number counts (see section 1.1.1) and the gravitational binding energies required to confine stars and gas to galaxies. Assuming that the predictions of Newtonian mechanics give a reasonable approximation of the dynamics of galaxies in clusters, this result was an order of magnitude lower than the subsequent mass estimates derived from velocity dispersion observations of galaxies in the Coma and Virgo clusters (Zwicky 1933). Support for the existence of this 'missing mass' was provided by observations indicating

a constant form to galaxy rotation curves out to larger radii than expected if the visible matter constituted the entire mass of the galaxy (e.g. Faber & Gallagher 1979). Additional evidence for dark matter has also arrived in the form of lensing experiments, the cluster abundance and also statistical analyses of large surveys of galaxies (which are reviewed in the rest of this section) and the CMB.

Since baryons might reside in diffuse clouds of gas or be locked in dim stellar remnants and planets as well as luminous stars, the missing mass might reasonably be presumed to be baryonic. However, BBNS constraints indicate that the mean baryon density is also an order of magnitude less than the cluster mass estimates of Zwicky (1933). This, and the fact that credible non-baryonic alternatives for the dark matter also emerged soon shifted the focus away from baryonic candidates. Dark matter candidates have since been recognised in two varieties: Warm Dark Matter (WDM), particles which have relativistic primeval velocities (e.g. neutrinos), and Cold Dark Matter (CDM) for which the particle energies are small enough at high redshifts to allow the dark matter to quickly become self-gravitating (e.g. Weakly Interacting Massive Particles).

This classification is necessary as the distinction has important consequences for the growth of structure. Since CDM is able to form structures on smaller scales and earlier times than WDM, a process of hierarchical structure formation ensues in which small dark matter halos merge to form more massive objects. In contrast, a mass density dominated by WDM results in large agglomerations of matter which subsequently fragment to form the smaller mass halos that exist at low redshifts. While neutrinos are an attractive candidate for the dark matter given that they are known to exist, CDM simulations indicate that the galaxy distribution at low redshifts is more consistent with observations than for WDM counterparts, assuming that the galaxies trace the underlying mass in some straightforward fashion.

1.4.2 Towards a Concordance Cosmology

Over the past two decades, the variance of galaxy density fluctuations in the local Universe has been measured to high accuracy at relatively small scales. This began through the use of the angular correlation function (e.g. Groth & Peebles 1977) derived from magnitude-limited surveys which at first contained information only about each galaxy's position (e.g. the Lick catalogue, Seldner et al. 1977). Following the development of models dominated by CDM (Peebles 1982, Bond & Szalay 1983), the results appeared to suggest an Ω_m=1 Einstein-de Sitter Universe with a relatively small baryon component as suggested by

BBNS constraints.

However, the widespread acceptance of this Standard Cold Dark Matter (SCDM) cosmology began to alter with the arrival of the APM galaxy survey (Maddox et al. 1990a). The associated angular correlation function was found to display a downturn at a similar scale to that of the Lick catalogue of $r \approx 10\,h^{-1}$ Mpc, but with a much shallower slope at larger scales, apparently inconsistent with a high mass density (Maddox et al. 1990b). Along with baryon fraction constraints from X-ray cluster observations indicating that $\Omega_b/\Omega_m \approx 0.15$, the projected correlation function of the APM galaxy survey suggested that the CDM component may be considerably lower than previously thought, with $\Omega_m \approx 0.3$ (Efstathiou et al. 1990, White et al. 1993).

Since then, numerous galaxy surveys have confirmed this picture (e.g. Tucker et al. 1997, Ratcliffe et al. 1998b, Guzzo et al. 2000). In particular, the precision with which the correlation function and power spectrum are measured has been revolutionised by two recent surveys, the 2dF Galaxy Redshift Survey (2dFGRS; Hawkins et al. 2003, Cole et al. 2005) and the Sloan Digital Sky Survey (SDSS; Zehavi et al. 2004, Tegmark et al. 2004). What marks these out from previous experiments is the combination between depth ($\bar{z} \approx 0.12$) and solid angle, leading to large surveyed volumes, as well as redshift data for an unprecedented number of galaxies. For the first time, it appears that the uncertainty in the associated cosmological parameters are dominated by systematic rather than statistical errors ($\Omega_m h = 0.168 \pm 0.016$, $\Omega_b/\Omega_m = 0.185 \pm 0.046$ in the case of the 2dFGRS (Cole et al. 2005), assuming h=0.72, a negligible neutrino density, Ω_ν=0, and a primordial slope of n_s=1). Both surveys provide considerable further support for a low-Ω_m cosmology.

Along with these developments, further support has been added to this picture by recent measurements of the CMB angular power spectrum by the Wilkinson Microwave Anisotropy Probe (WMAP; e.g. Spergel et al. 2003, Bennett et al. 2003). In combination with the surveys of local large-scale structure mentioned above, WMAP strongly supports the idea that the Universe has zero spatial curvature ($\Omega_{tot} = 1.02 \pm 0.02$), but also that the matter density is low ($\Omega_m = 0.27 \pm 0.04$, Bennett et al. 2003). This picture is consistent with recent observations of high redshift Type Ia supernovae (e.g. Perlmutter et al. 1997, Riess et al. 1998), which are also able to provide an explanation for the missing energy density. Type Ia supernovae are expected to be standard candles, and being extremely luminous are able to constrain the distance scale over a significant fraction of the observable Universe. Recent observations suggest that the Universe is in a phase of

accelerated expansion driven by a dark energy component (Λ) which dominates the energy density of the Universe, with $\Omega_\Lambda = 0.72 \pm 0.05$ (assuming a cosmological constant and spatially flat Universe, Tonry et al. 2003). However, concerns remain over the validity of using Type Ia supernovae as standard candles; the level to which systematics such as evolution and dust affect the results is currently poorly constrained.

1.5 Motivation and Thesis Aims

Significant problems with this ΛCDM cosmology (where $\Omega_m \approx 0.3$ and $\Omega_\Lambda \approx 0.7$) remain, primarily since the geometry of the Universe is dominated by two pieces of undiscovered physics. In addition, there are a number of outstanding observational problems for the ΛCDM model and several assumptions which might significantly affect the cosmological constraints; a number of these form the focus of this thesis:

One such problem is the issue of bright number counts as observed by the APM survey mentioned previously; the observed number of galaxies as a function of apparent magnitude appears to indicate an under-density in the local galaxy distribution (or 'local hole') which is at odds with the prediction of the ΛCDM model for the variance of density fluctuations on scales of $r \gtrsim 300\,h^{-1}$ Mpc (Frith et al. 2003, Busswell et al. 2004). This is an issue which is examined in detail in chapters 2 and 3 by determining the number counts from independent observations.

If the APM survey number counts were due exclusively to local large-scale structure as suggested by Frith et al. (2003) and Busswell et al. (2004), this should be observable as an excess of power over the ΛCDM prediction on large scales. Support for the ΛCDM model through measurements of the variance of local galaxy density fluctuations have only been obtained on relatively small scales ($r \lesssim 300\,h^{-1}$ Mpc). An inherent problem for such clustering analyses is the lack of understanding in the way in which the galaxy and underlying mass distributions are connected; it is assumed that on relatively large scales ($r \gtrsim 40\,h^{-1}$ Mpc) the galaxies trace the mass in a linear fashion. We attempt to address these issues in the near infrared through a determination of the K_s-band power spectrum on extremely large scales ($r \lesssim 1000\,h^{-1}$ Mpc) in chapter 4.

A common assumption when constraining cosmological parameters through observations of the local galaxy distribution, is that the primordial distribution of density fluctuations is Gaussian; such initial conditions are expected in standard inflationary models, however, primordial non-Gaussianity may be introduced in more complicated models.

If the latter were true, then the linear predictions which assume Gaussian intial conditions with which measurements of the variance of local galaxy density fluctuations are compared and the corresponding cosmological constraints would no longer be valid. In chapter 5, we probe the primordial density field by determining the higher order moments of the local galaxy distribution. With these statistics, we are also able to examine the way in which galaxies selected in the near infrared trace the underlying mass distribution in detail, determining constraints for high order, non-linear bias terms.

However, despite the outstanding issues for the study of local large-scale structure mentioned above, strong support for the concordance cosmology remains via CMB constraints and in particular from the accuracy of WMAP observations. In chapter 6, we address a possible contaminant of CMB data arising from the inverse Compton scattering of CMB photons by the hot gas present in large clusters of foreground galaxies, the thermal Sunyaev-Zeldovich effect. The level to which the associated cosmological constraints might be affected is also discussed.

CHAPTER 2

A LARGE LOCAL HOLE IN THE UNIVERSE?

2.1 INTRODUCTION

The counting of galaxies as a function of apparent magnitude is one of the most powerful tools in observational cosmology. Not only can this simple statistic form strong constraints on the level of evolution at the faint end, but also on the large-scale structure and the scales to which the cosmological principle can be said to hold from bright magnitude counts.

A recurring problem arising from the study of bright galaxy number counts has been the measured deficiency of galaxies around the Southern Galactic Pole (SGP). This was first examined in detail by Shanks (1990) and subsequently by the APM galaxy survey (Maddox et al. 1990a), which observed a large deficit in the number counts ($\approx$50 per cent at B =16, $\approx$30 per cent at B =17) over a $\approx$4000 deg^2 solid angle. If this anomaly was due solely to features in the galaxy distribution, this would be at odds with recent measurements of the variance of local galaxy density fluctuations (e.g. Hawkins et al. 2003, Cole et al. 2005) or the expected linear growth of density inhomogeneities at large scales.

Maddox et al. (1990c) examined possible causes of this deficiency. From redshift survey results over the APM survey area (Loveday et al. 1992), it was argued that a weak local under-density contributed to the observed deficiency at the $\lesssim$10 per cent level at $B \approx$17. Instead, Maddox et al. (1990c) suggested that strong low redshift galaxy evolution was the dominant contribution. This phenomenon has also been suggested as a possible explanation for large deficiencies in the Sloan Digital Sky Survey (SDSS) (Loveday 2004), although models without such strong low redshift evolution provide predictions consistent with observed redshift distributions (e.g. Broadhurst et al. 1988, Colless et al. 1990, Hawkins et al. 2003). In contrast, Shanks (1990) argued that evolution could not account for the observed slope and that large-scale structure was the principal cause of the deficiency in the counts.

However, another possible contribution to the low counts might be errors in the APM photometry. Comparing the photographic APM photometry with B-band CCD data, Metcalfe et al. (1995) detected a small residual scale error in the APM survey photometry for $B \gtrsim 17$. Correcting for this offset, the counts were now in good agreement with homogeneous predictions at faint magnitudes ($B \gtrsim 17.5$); however, the problematic deficiency at brighter magnitudes remained. More recently, Busswell et al. (2004) used B-band CCD data over $\approx$337 deg^2 within the APM survey area to provide the most accurate comparison to date with a sample of the APM survey photometry. The photometric zeropoint of this CCD data was in excellent agreement with the Millennium Galaxy Catalogue (Driver 2003) and the Sloan Digital Sky Survey Early Data Release (Yasuda et al. 2001). However, a comparison with the APM photometry suggested a large offset of 0.31 magnitudes for B <17.35. Applying this to the APM survey counts, a deficiency of $\approx$25 per cent remained at B =16; Busswell et al. (2004) determined that such a deficiency in the local galaxy distribution would still be at odds with a ΛCDM form to the galaxy correlation function and power spectrum at large scales.

The presence of a large under-density around the SGP has been confirmed by large galaxy redshift surveys. The 2dF Galaxy Redshift Survey (2DFGRS; Colless et al. 2003) has taken spectra for galaxies brightward of $b_J \approx 19.45$ over a solid angle of >600 deg^2 around the SGP. The $n(z)$ indicates remarkable structure in the local galaxy distribution with a large deficit to $z \approx 0.1$. Large deficiencies in the $n(z)$ are also indicated by other redshift surveys in the SGC (Shectman et al. 1996, Vettolani et al. 1997, Ratcliffe et al. 1998b).

In conjunction with the work of Busswell et al. (2004), Frith et al. (2003) investigated the number counts in the K_s-band from the 2 Micron All Sky Survey (2MASS) second incremental release (Jarrett et al. 2000). In the Northern and Southern 2dFGRS declination strips they found good agreement both with the corresponding optical number counts and the expected trend defined by the 2dFGRS $n(z)$, suggesting that the number counts in the 2dFGRS strips are caused by real features in the galaxy distribution, and that a significant effect from strong, low-redshift galaxy evolution is unlikely. However, due to incompleteness in the 2MASS second incremental release, a comparison with the optical APM survey counts could not be made, although the steeper than Euclidean counts over the partially surveyed galactic caps indicated the possible presence of large under-densities in both the Southern *and* Northern Galactic Caps to $\sim 300\,h^{-1}$ Mpc. This was supported by mapped 2MASS counts, showing the variations of counts over the sky in $5^\circ \times 5^\circ$ bins,

and indicating huge regions of under-density $\gtrsim 100°$ across in both hemispheres.

Here, we probe the large-scale structure of the local galaxy distribution with the recently completed 2MASS survey using three techniques. First, we compare the K_s-band 2MASS counts over the APM survey area with model predictions. We calculate the significance with respect to these models considering a ΛCDM form for the 2-point angular correlation function. Secondly, since the form of the local hole as suggested by Busswell et al. (2004) requires the presence of excess power at large scales over the ΛCDM prediction, we compute the 2MASS angular power spectrum and compare with a ΛCDM prediction constructed from the ΛCDM Hubble Volume simulation. Finally, we investigate the local galaxy distribution more directly by using a counts in cells analysis on the 2MASS and ΛCDM mock catalogues.

In section 2.2, we present details of the datasets used. In section 2.3, the number counts are presented. The clustering at large scales is investigated through a determination of the 2MASS and ΛCDM mock angular power spectra in section 2.4. In section 2.5 we present the counts in cells analysis. The discussion and conclusions follow in section 2.6.

2.2 DATA

2.2.1 THE 2MASS EXTENDED SOURCE CATALOGUE

The 2 Micron All Sky Survey (2MASS) extended source catalogue has now released K_s, H and J-band photometry for over 1.6×10^6 extended sources over the entire sky with high completeness to K_s=13.5 (Jarrett et al. 2000, Jarrett 2004). 2MASS is currently the largest existing all-sky galaxy survey and therefore represents a uniquely powerful probe of the local galaxy density field at large scales.

Previously, in order to estimate the total K_s-band magnitudes from the 2MASS second incremental release data, Cole et al. (2001) used the deeper J-band Kron magnitudes, colour-corrected to the K_s-band via the J and K_s default aperture magnitudes. The accuracy of this magnitude estimator was determined through a comparison with the K-band photometry of Loveday (2000); the Loveday photometry had better signal-to-noise and resolution than the 2MASS scans and so enabled more accurate 2MASS magnitudes to be determined.

The final release data uses revised magnitude estimates and the default aperture magnitudes used in Cole et al. (2001) have been abandoned (Jarrett - priv. comm.). In Fig. 2.1 we show a selection of 2MASS K_s-band magnitude estimates with the revised

2MASS photometry compared with the Loveday (2000) photometry used previously. In the place of the default aperture magnitudes used in Cole et al. (2001), we use fiducial elliptical Kron magnitudes in panels (b) and (d) to colour-correct the J-band magnitudes to the K_s-band. Of the many different magnitude estimates examined, the most accurate in terms of the scale error between the Loveday and 2MASS photometry and the zero-point offset uses the J-band extrapolated magnitude colour-corrected to the K_s-band as described above. Using the dust maps of Schlegel et al. (1998), we use extinction-corrected K_s-band magnitudes calculated in this way.

In order to verify the usefulness of the magnitude estimator used in this work as an estimate of the total K_s-band magnitude, we perform an internal check via a comparison with the magnitude estimates used in the 2MASS-selected 6dF Galaxy Survey (6dFGS). The 6dFGS K_s-band magnitudes are determined using a surface brightness correction to the K_s-band 20 mag. arcsec^{-2} isophotal elliptical aperture magnitude (Jones et al. 2004). We find excellent agreement with a slope of 1.022, an offset of 0.018 magnitudes and a spread of σ=0.048 magnitudes for $|b|$ >20° galaxies matched below $K_s = 13.5$.

The 2MASS dataset removes or flags sources identified as artefacts such as diffraction spikes and meteor streaks (Jarrett et al. 2000); we use the 2MASS cc_flag to remove such objects. We also employ a colour cut ($J - K_s$ <0.7 and $J - K_s$ >1.4) below K_s=12 in order to remove a small number of objects identified as non-extragalactic extended sources (Maller et al. 2003, Maller et al. 2005).

The resulting sample using the magnitude estimator described above provides a sample of 61 478 and 62 748 $b \leq -20°$ and $b \geq 20°$ galaxies below $K_s = 12.5$ respectively over a combined solid angle of 27 144 deg^2.

2.2.2 THE 2dF GALAXY REDSHIFT SURVEY

The 2dF Galaxy Redshift Survey (2dFGRS) is selected in the photographic b_J-band using the APM survey and subsequent alterations and extensions to it (Colless et al. 2003) for two declination strips in the northern and southern galactic caps, as well as 99 randomly selected 2° fields scattered over the APM survey area. The final release data has provided spectra for ≈220,000 galaxies (for a quality flag of Q≥3) over ≈1500 square degrees to an approximate magnitude limit of b_J=19.45.

In this chapter, we investigate the galaxy distribution in the K_s-band. We have therefore formed a 2MASS-2dFGRS catalogue, matched over the Northern and Southern 2dFGRS declination strips, using the K_s-band magnitude estimator described above.

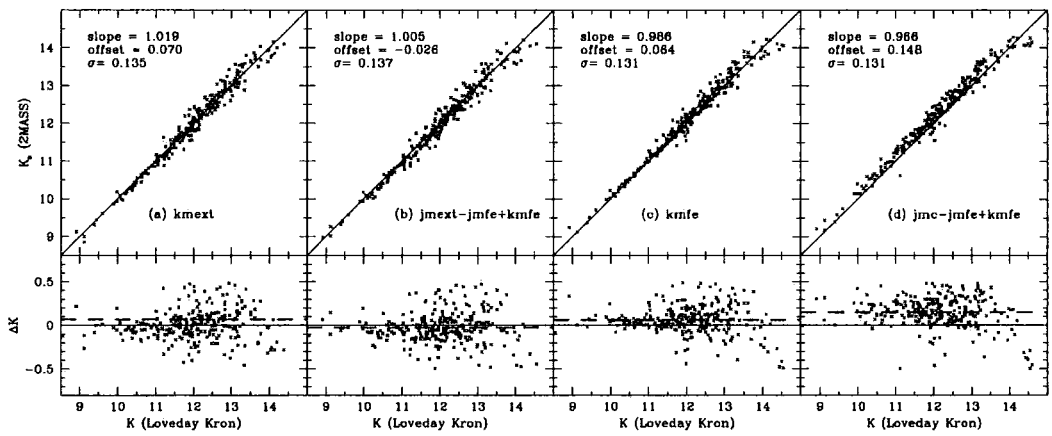


Figure 2.1: A selection of K_s-band magnitude estimates from the 2MASS full release compared with Loveday (2000) K-band photometry. In each case the lower panels display the residual. The $x = y$ slope is indicated by a solid line, while the mean offset is indicated in the lower panel by a dashed line. This offset (in magnitudes), the best fit slope determined from least squares fits and the rms scatter are indicated for each magnitude estimate. We estimate the magnitudes directly from the (a) K_s-band extrapolated and (c) K_s-band fiducial elliptical Kron magnitudes, and also from the (b) J-band extrapolated and (d) J-band circular Kron magnitudes colour-corrected to the K_s-band using the J and K_s-band fiducial elliptical Kron magnitudes.

Fig. 2.2 shows the K_s-band selected $n(z)$ and radial density functions, i.e. the observed $n(z)$ over the predicted, for the 2dFGRS declination strips with an applied magnitude limit of K_s=13.5.

2.2.3 THE ΛCDM HUBBLE VOLUME SIMULATION

The Hubble Volume catalogues represent one of the largest volume N-body simulations of the Universe to date. The ΛCDM simulation follows the evolution of 10^9 dark matter particles from $z \approx 50$ over a volume of $3000^3\ h^{-3}\mathrm{Mpc}^3$. The associated cosmological parameters are Ω_m=0.3, Ω_b=0.04, h=0.7, σ_8=0.9 (Jenkins et al. 1998).

In this work, we construct mock 2MASS catalogues from the $z = 0$ ΛCDM Hubble Volume simulation dark matter particles. We divide the total volume into 27 virtually independent spherical volumes of $r = 500\,h^{-1}$ Mpc. These are subjected to the 2MASS selection function:

$$n(z) = \frac{3z^2}{2(\bar{z}/1.412)^3} \exp\left(-\left(\frac{1.412z}{\bar{z}}\right)^{3/2}\right) \tag{2.1}$$

(Baugh & Efstathiou 1993, Maller et al. 2005) where $\bar{z}$ is determined from the 2MASS-2dFGRS matched sample described in section 2.2; for K_s <12.5, $\bar{z}$=0.05. Equation 2.1 is normalised to match the total number of observed 2MASS galaxies for $|b|$ >20°. These mocks are mainly used in the angular power spectrum analysis in section 4.

For the counts in cells analysis in section 5, it is necessary to more accurately mimic the galaxy sample. In addition to the 2MASS selection function used in equation 2.1 we therefore use a bias prescription:

$$P(\nu) = \begin{cases} \exp(\alpha\nu + \beta\nu^{3/2}) & \text{for } \nu \geq 0 \\ \exp(\alpha\nu) & \text{for } \nu < 0, \end{cases} \tag{2.2}$$

(Cole et al. 1998) where the bias probability, $P(\nu)$, represents the probability of placing a galaxy onto a dark matter particle and is based on the density field at the epoch at which the particles are selected rather than the initial density field, and ν is the number of standard deviations of the density away from the mean. For the counts in cells analysis we use a magnitude limit of K_s =12.5. We use parameters of α=0.6 and β=-0.15 to match the K_s <12.5 angular correlation function at small scales ($\theta \lesssim 3°$). The resulting correlation function amplitude is ≈1.5 higher at these scales than for the unbiased mocks described above.

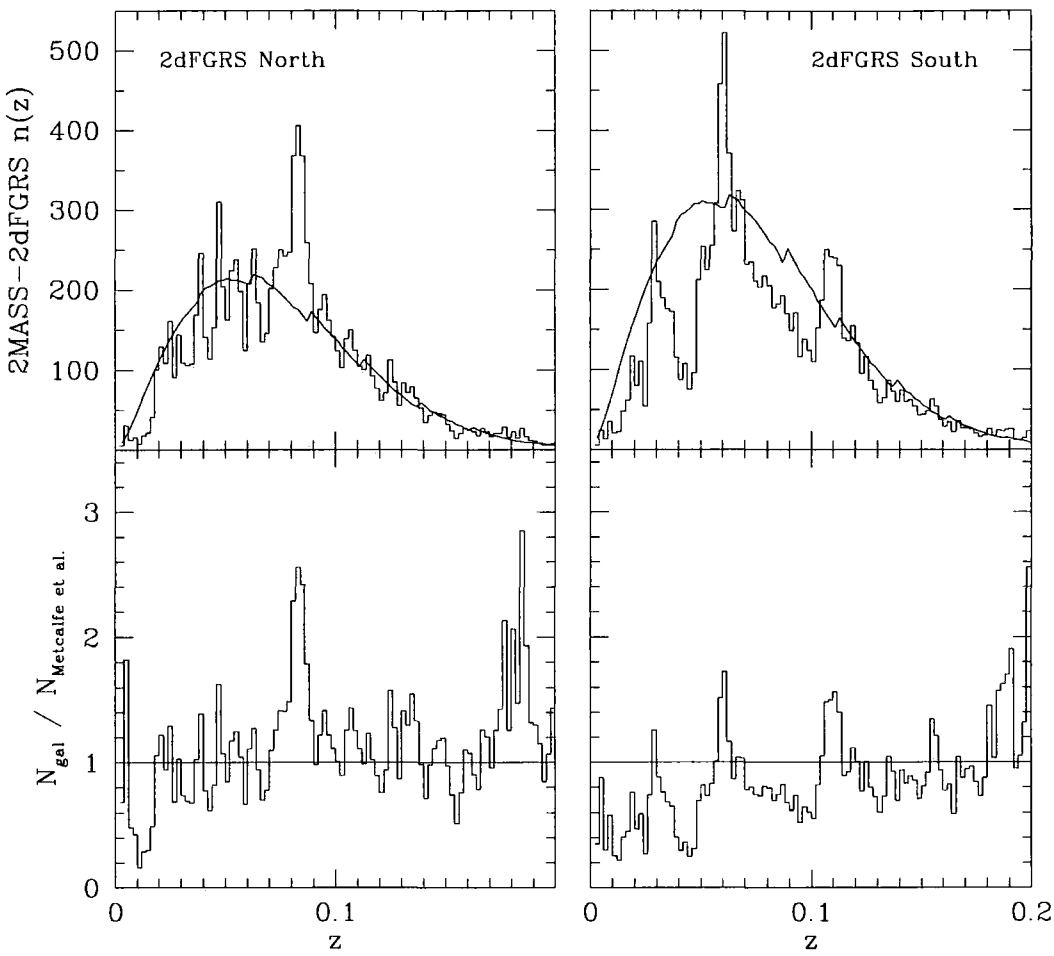


Figure 2.2: Number redshift histograms for the 2MASS-2dFGRS matched sample described in section 2.2 for the Northern (left hand panels) and Southern 2dFGRS declination strips (right hand panels). In each case the solid line indicates a homogeneous prediction constructed from the K-band luminosity function parameters of Metcalfe et al. (2001) and K-corrections of Bruzual & Charlot (1993). The upper panels show the $n(z)$ for K_s <13.5 galaxies. The lower panels show the corresponding radial density functions, i.e. the $n(z)$ divided through by the Metcalfe et al. (2001) homogeneous prediction.

Morphological Type	ϕ^* ($\times 10^{-4}\, h^3\, \mathrm{Mpc}^{-3}$)	$M_K^* - 5\log h$	α
E/S0	7.42	-21.82	-0.7
Sab	3.70	-21.89	-0.7
Sbc	4.96	-22.16	-1.1
Scd	2.18	-21.99	-1.5
Sdm	1.09	-21.45	-1.5

Table 2.1: K-band luminosity function parameters from Metcalfe et al. (2001) as a function of morphological type. These were initially determined via bright B-band CCD data; the M_B^* parameters were subsequently colour-corrected to the K_s-band for each morphological type.

2.3 NUMBER COUNTS

2.3.1 MODEL NORMALISATION

The issue of the local hole in the APM survey area rests critically on the model number count normalisation. In the K-band, the number count predictions are in remarkable agreement with the observations to $K \approx 23$ (McCracken et al. 2000). Also, the predicted number below $K \approx 18$ is fairly insensitive to the evolutionary model or the assumed cosmology. Therefore, the K-band number counts are a particularly useful probe of the local Universe since the model predictions can be constrained at fainter magnitudes with few concerns over uncertainties in the amount of evolution or the cosmology.

In this work, we use a non-evolving K-band model computed from the luminosity function parameters of Metcalfe et al. (2001) (see table 2.1 and Fig. 2.3) and the K-corrections of Bruzual & Charlot (1993). Comparing the number count predictions in the fitting range $14 < K < 18$ to the faint counts used in McCracken et al. (2000) as well as other data (Vaisanen et al. 2000, Szokoly et al. 1998, Huang et al. 1997, Huang et al. 2001, Kummel & Wagner 2000, Martini 2000, McLeod et al. 1995), we find that the Metcalfe et al. (2001) model provides a good fit to the observations with $Y=0.96\pm0.06$ (where $Y=1$ represents the Metcalfe et al. (2001) normalisation). This fitting range is chosen to avoid systematic effects from large-scale structure at the bright end and uncertain sensitivity to evolution at fainter magnitudes.

In order to test for the presence of strong evolution at low redshifts and zero-point offsets in the 2MASS data, we also construct model predictions for the number counts from the Metcalfe et al. (2001) homogeneous prediction described above and the observed $n(z)$. These are determined by varying the luminosity function normalisation as a function

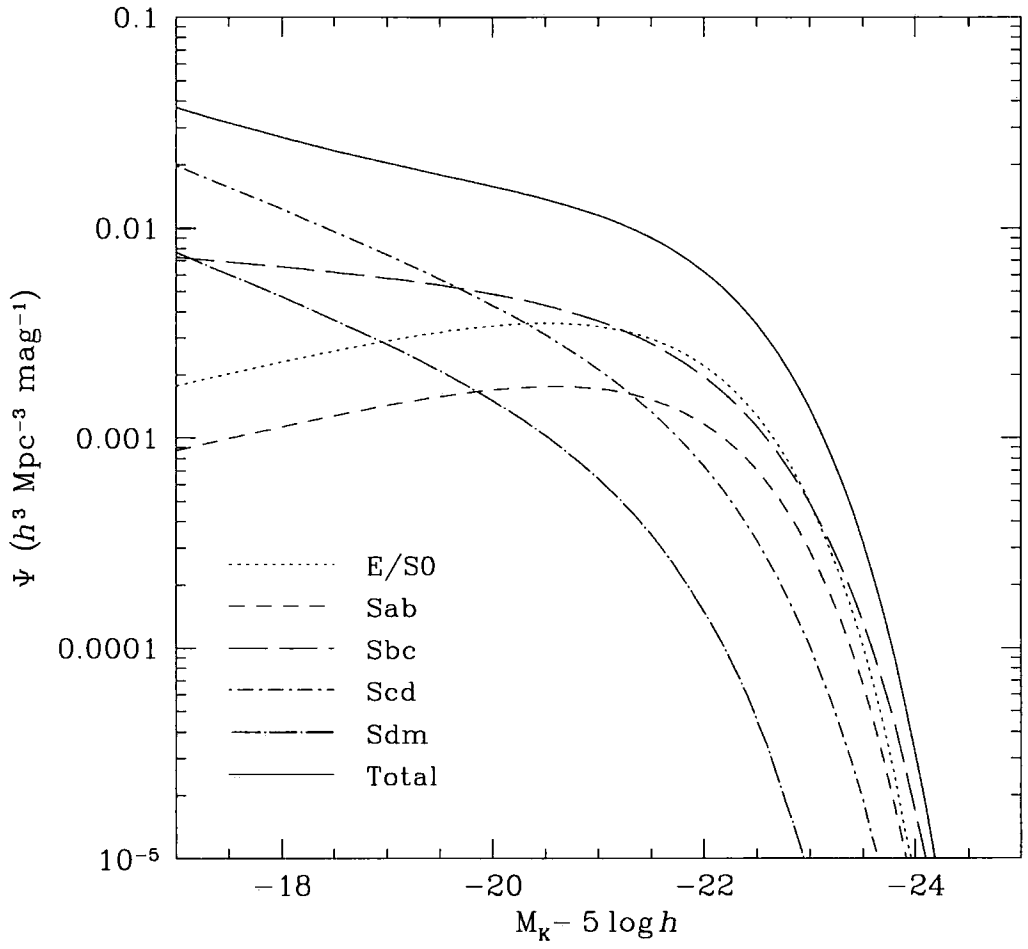


Figure 2.3: K-band luminosity functions from Metcalfe et al. (2001) as a function of morphological type using the parameters shown in table 2.1. The solid line shows the total K-band luminosity function.

of the redshift; the luminosity function parameter ϕ^* is multiplied by the relative density (Fig. 2.2, lower panels). These variable ϕ^* models (Frith et al. 2003) therefore provide a simple prediction for the number counts associated with a given galaxy distribution, assuming that there is no significant effect from strong galaxy evolution at low redshifts.

2.3.2 RESULTS

In order to verify the consistency of bright number counts with the corresponding $n(z)$ previously noted by Frith et al. (2003), we compare the K_s-band number counts extracted from the 2dFGRS strips with the corresponding variable ϕ^* models constructed from the Metcalfe et al. (2001) model, using the newly-completed 2dFGRS and 2MASS catalogues (Fig. 2.4). The K_s-band number counts extracted from the 2dFGRS fields are in reasonable agreement with the corresponding variable ϕ^* models. This suggests that the form of the bright number counts is exclusively due to features in the local galaxy distribution and that a significant effect from strong, low-redshift galaxy evolution is unlikely. The agreement between the counts and the variable ϕ^* models is independent of the model normalisation, since any change in the number count models also alters the $n(z)$ model normalisation and therefore the implied deficiency to the same degree.

Having confirmed the consistency between the number counts and the underlying large-scale structure in the 2dFGRS strips, we are now in a position to examine the number counts over the APM survey area. Fig. 2.5 shows the K_s-band 2MASS counts extracted for the $\approx$4000 deg^2 field. We also show the 2dFGRS variable ϕ^* models (as in Fig. 2.4) for reference.

There is a surprisingly good agreement between the K_s-band counts in the APM survey area and the 2dFGRS Southern variable ϕ^* model. This suggests that the local galaxy distribution in the APM survey area may be similar to that of the much smaller 2dFGRS Southern strip. This is supported by similar deficiencies observed in the Durham-UKST redshift survey (Ratcliffe et al. 1998b), the Las Campanas Redshift Survey (Shectman et al. 1996) and the ESO Slice Project (Vettolani et al. 1997), which are all situated within the APM survey area, and also the optical number counts and corresponding variable ϕ^* models (Frith et al. 2004, Frith et al. 2003, Busswell et al. 2004). The degree of under-density may therefore be inferred from the observed 2MASS-2dFGRS redshift distribution and the model normalisation. Taking the Metcalfe et al. (2001) K-band model normalisation and the 2MASS-2dFGRS matched $n(z)$ (Fig. 2.2) implies a 23 per cent under-density to z=0.1. If we take the best fit K-band model normalisation detailed in section 3.1 (i.e.

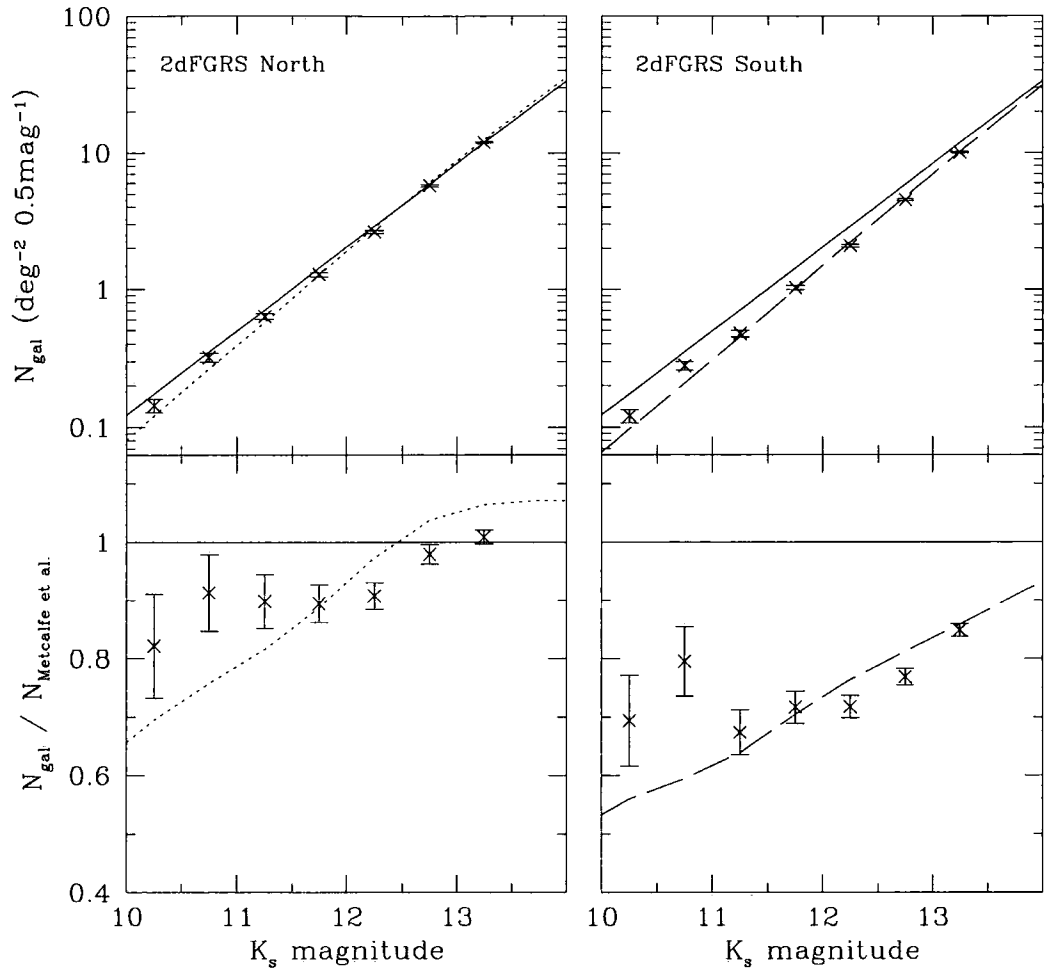


Figure 2.4: K_s-band 2MASS number counts extracted for the 2dFGRS Northern (left hand panels) and Southern (right hand panels) declination strips. The upper panels show the counts on a logarithmic plot, while the lower panels show the residual, i.e. the number count divided through by the Metcalfe et al. (2001) homogeneous prediction. The homogeneous model is indicated by a solid line, with the 2dFGRS Northern and Southern variable ϕ^* models shown by the dotted and dashed lines respectively. The errorbars indicate the expected 1σ Poisson fluctuation; these are indicative of the error from bin to bin only.

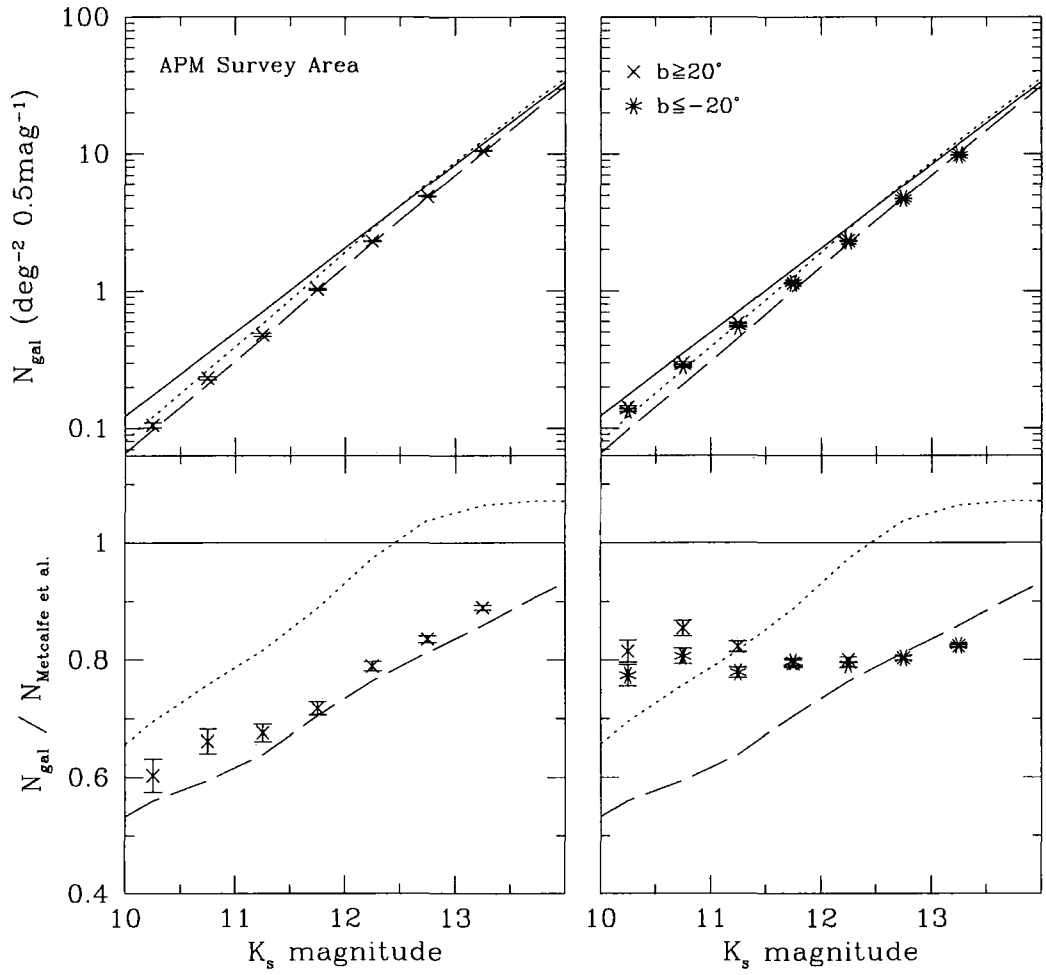


Figure 2.5: K_s-band 2MASS number counts extracted for the ≈ 4000 deg^2 APM survey area and the $|b| \geq 20°$ galactic caps, presented as in Fig. 2.4. As before the errorbars indicate the Poisson error only; an analysis of the expected dispersion in these counts due to large-scale structure follows in section 2.3.3.

Model Normalisation	Y	Deficiency	Significance
Metcalfe et al.	1.00	26.9 per cent	5.0σ
Best fit	0.96	23.9 per cent	4.5σ
$\bar{n}_{\,2MASS}$	0.80	9.7 per cent	1.8σ

Table 2.2: Significances for the observed deficiency in the K_s-band counts extracted for the APM survey area for K_s <12.5 considering various model normalisations. The first two rows show the significance calculation for the Metcalfe et al. and best fit model normalisations; the lowest row indicates the significance if the model is lowered to match the mean 2MASS number density; the relative normalisations are indicated by Y. In each case a ΛCDM correlation function is considered of A=0.28, $1-\gamma=-0.71$ for θ <5.0°, A=5.3, $1-\gamma=-2.5$ for $\theta \geq$5.0° where $\omega = A\theta^{1-\gamma}$. The best fit parameters at small scales are taken from the 2MASS correlation function analysis of Maller et al. (2005), while the large scale fit is determined from the K_s <12.5 correlation function (presented in Fig. 2.7 for $\theta \geq$10°) constructed from the 27 mock ΛCDM 2MASS mocks with an applied linear bias of b_K=1.1 (Maller et al. 2005), providing good agreement with the Maller et al. (2005) power law best fit at $\theta \approx$1°.

Y=0.96), the implied under-density over the APM survey area to z=0.1 is 20 per cent.

In order to see whether this structure persists to even larger scales than the APM survey area, we have determined the 2MASS K_s-band counts over the entire $|b| \geq$20° galactic caps (Fig. 2.5, right hand panels). The two sets of counts are very low with respect to the Metcalfe et al. (2001) model but are in good agreement with each other. This may indicate that the normalisation of the Metcalfe et al. (2001) model should be even lower, 2.5σ below the best fit to the K-band counts at 14< K <18 (i.e. Y=0.96). Alternatively, it might also indicate the presence of a zero-point offset between the 2MASS photometry and the K-band model of ≈0.15 magnitudes; any alteration to the 2MASS zero-point would compromise the agreement between the K_s-band 2dFGRS strip counts and the corresponding variable ϕ^* models, and also with the Loveday (2000) photometry zero-point with which the 2MASS magnitudes are compared. If the low counts over the galactic caps were due to real features in the galaxy distribution, this would imply that the local Universe is globally under-dense, and that even surveying over ≈27,000 deg^2 to $r \gtrsim 150\,h^{-1}$ Mpc does not constitute a fair sample of the Universe.

2.3.3 Determining the Significance

Using assumed forms to the galaxy correlation function at large scales, it is possible to determine the associated significance of features in the galaxy distribution. Busswell et

al. (2004) considered various forms to the real-space correlation function to calculate the significance of an assumed 3-dimensional form to the local hole. Here, we use the 2-dimensional analogue to determine the significance implied by the *observed* deficiency in the number counts with respect to various normalisations of the Metcalfe et al. (2001) model using the angular correlation function:

$$\left(\frac{\delta N}{\bar{N}}\right)^2 = \frac{1}{\bar{N}} + \frac{1}{\Omega^2}\int d\Omega_1 d\Omega_2 \omega(\theta_{12}) \tag{2.3}$$

(Peebles 1980) where $\omega(\theta_{12})$ is the value of the 2-point angular correlation function between two area elements $d\Omega_1$ and $d\Omega_2$. Ω is the total solid angle of the survey, n is the mean galaxy density such that $\bar{N} = n\Omega$ is the total number of galaxies in the survey area. A power law form of the correlation function is used such that $\omega = A\theta^{1-\gamma}$.

In Table 2.2, we apply this technique to the observed deficiency over the APM survey area in the 2MASS K_s-band counts for K_s <12.5 with respect to the Metcalfe et al. (2001) model, the best fit normalisation of this model described in section 3.1 and the observed mean 2MASS number count for $|b| \geq 20°$ (see Fig. 2.5). This magnitude limit is chosen such that the peak in the selection function lies within the redshift range of interest; at K_s <12.5 the $n(z)$ peaks at $z \approx 0.05$, and so this magnitude limit preferentially samples the mean depth of the possible local hole in the APM survey area. We consider a ΛCDM form for the angular correlation function determined from fits to the mock 2MASS catalogues (see Fig. 2.7), with parameters A=0.28, $1-\gamma = -0.71$ for θ <5.0° and A=5.3, $1-\gamma = -2.5$ for $\theta \geq 5.0°$. From this we determine the expected 1σ dispersion in the number of galaxies below K_s=12.5 over APM-like solid angles as 5.4 per cent.

Both the Metcalfe et al. (2001) model and the best fit normalisation appear to imply that the observed 2MASS number counts over the APM survey area represent either an extremely rare fluctuation in the galaxy density or a challenge to the large-scale clustering predicted by the ΛCDM Hubble Volume simulation, assuming a linear biasing scheme on these scales. Only if the best fit model is lowered by 2.5σ to the mean 2MASS number density at K_s <12.5 for $|b| \geq 20°$ ($\bar{n}_{\,2MASS}$), do the counts become more consistent with the ΛCDM prediction. To remove any discrepancy in the K_s <12.5 counts over the APM survey area requires the best fit model normalisation to be lowered by 3.8σ. Combining the uncertainty on the normalisation and the significance estimate for the best fit normalisation yields a significance of 2.9σ. Clearly, unless the faint K_s-band counts are systematically too high, then the 2MASS counts over the APM survey area suggest the presence of excess power at large scales over the ΛCDM prediction, assuming a linear

biasing scheme on these scales.

2.4 THE 2MASS ANGULAR POWER SPECTRUM

The large local hole in the APM survey area, as suggested by Busswell et al. (2004) (a $\approx$25 per cent deficiency to z=0.1 over $\approx$4000 deg^2) and the K_s <12.5 2MASS number counts, appear to imply the presence of excess power at large scales over the ΛCDM prediction. In order to determine whether any such excess power exists at large scales, we compute the angular power spectrum for $|b| \geq 20°$, K_s <12.5 2MASS galaxies and 27 mock 2MASS catalogues constructed from the ΛCDM Hubble Volume simulation (see section 2.2.3).

Following the usual method (e.g. Peebles 1973, Peebles & Hauser 1973, Peebles & Hauser 1974, Scharf et al. 1992), the angular power is estimated through a spherical harmonic expansion of the surface density of galaxies. This method is described in detail in chapter 4; we also determine the associated cosmological constraints and possible sources of systematic error.

2.4.1 RESULTS

The 2MASS angular power spectrum for 124 264 K_s < 12.5, $|b| \geq 20°$ galaxies is presented in Fig. 2.6. We also show the mean and 1σ spread determined from 27 unbiased mock 2MASS catalogues described in section 2.3. The linear model corresponding to the ΛCDM Hubble Volume mock catalogue input parameters constructed from transfer function fitting formulae of Eisenstein & Hu (1998) is indicated by the dotted line; this model is detailed further in Chapter 4.

The 2MASS angular power spectrum is in reasonable agreement with the mock 2MASS angular power spectra although the 2MASS slope is steeper and there is some discrepancy with the mock catalogue in an unbiased scenario. Therefore it appears that either there is an excess of power in the 2MASS catalogue over the ΛCDM Hubble Volume or there exists a scale-dependent bias within the scales shown. Clearly, the issue of bias is critical in determining the level of disagreement at large scales. Taking a reasonable value of the K_s-band bias of b_K=1.1 (Maller et al. 2005), the disagreement at large scales ($l \leq 30$; this corresponds to $r \gtrsim 30\,h^{-1}$ Mpc at the mean depth of the K_s <12.5 sample) is at the $\approx 3\sigma$ level.

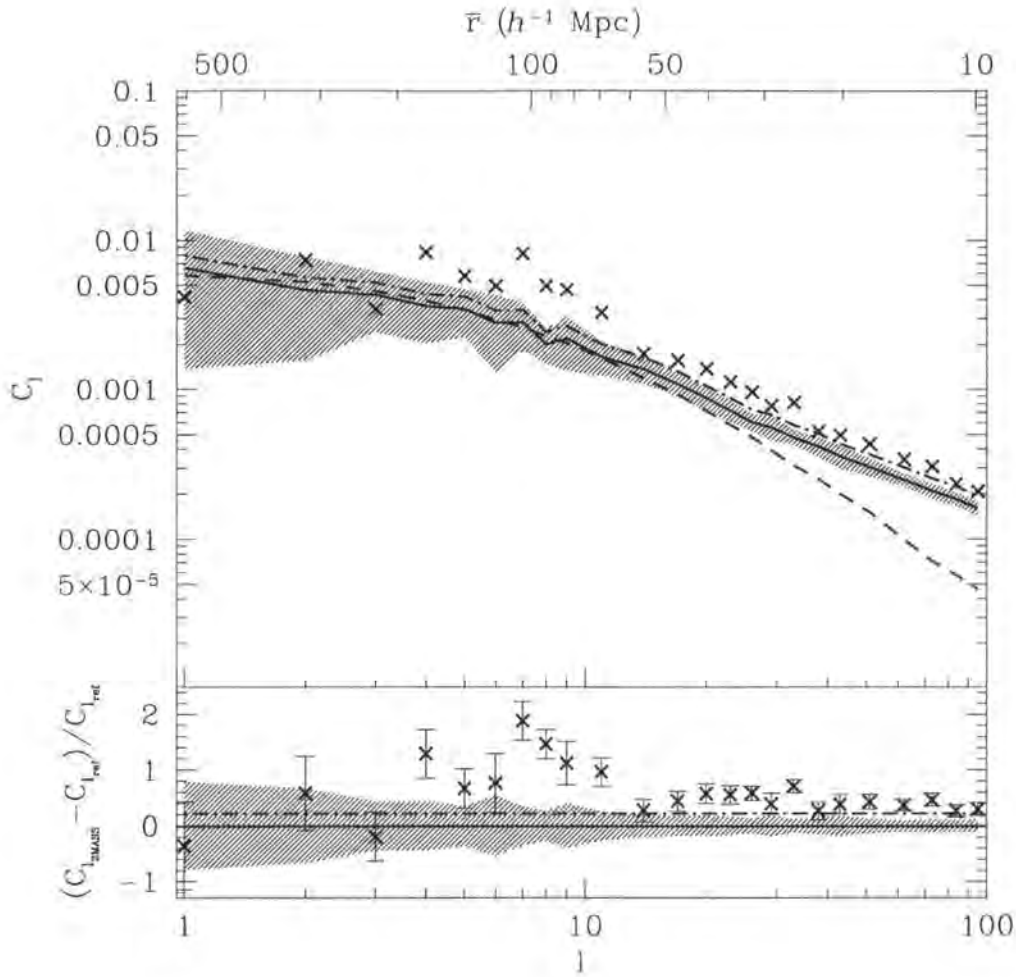


Figure 2.6: The $|b| \geq 20°$ 2MASS angular power spectrum for 124 264 $K_s < 12.5$ galaxies. The crosses indicate the 2MASS datapoints with the shaded region and solid line indicating the 1σ spread and mean power spectrum of 27 unbiased mock 2MASS catalogues constructed from the ΛCDM Hubble Volume simulation. The linear model corresponding to the Hubble Volume mock catalogue input parameters of Ω_m=0.3, Ω_b=0.04, h=0.7 and σ_8=0.9 is indicated by the dashed line. In the lower panel we show the 2MASS, mock and model power spectra expressed as the fractional deviation from the mean mock power spectrum. In addition we indicate the approximate mean distance scale probed by the data on the top x-axis. For reference, the mean mock power spectrum is also shown, using the linear bias of b=1.1 used subsequently, by the dot-dashed line in both panels.

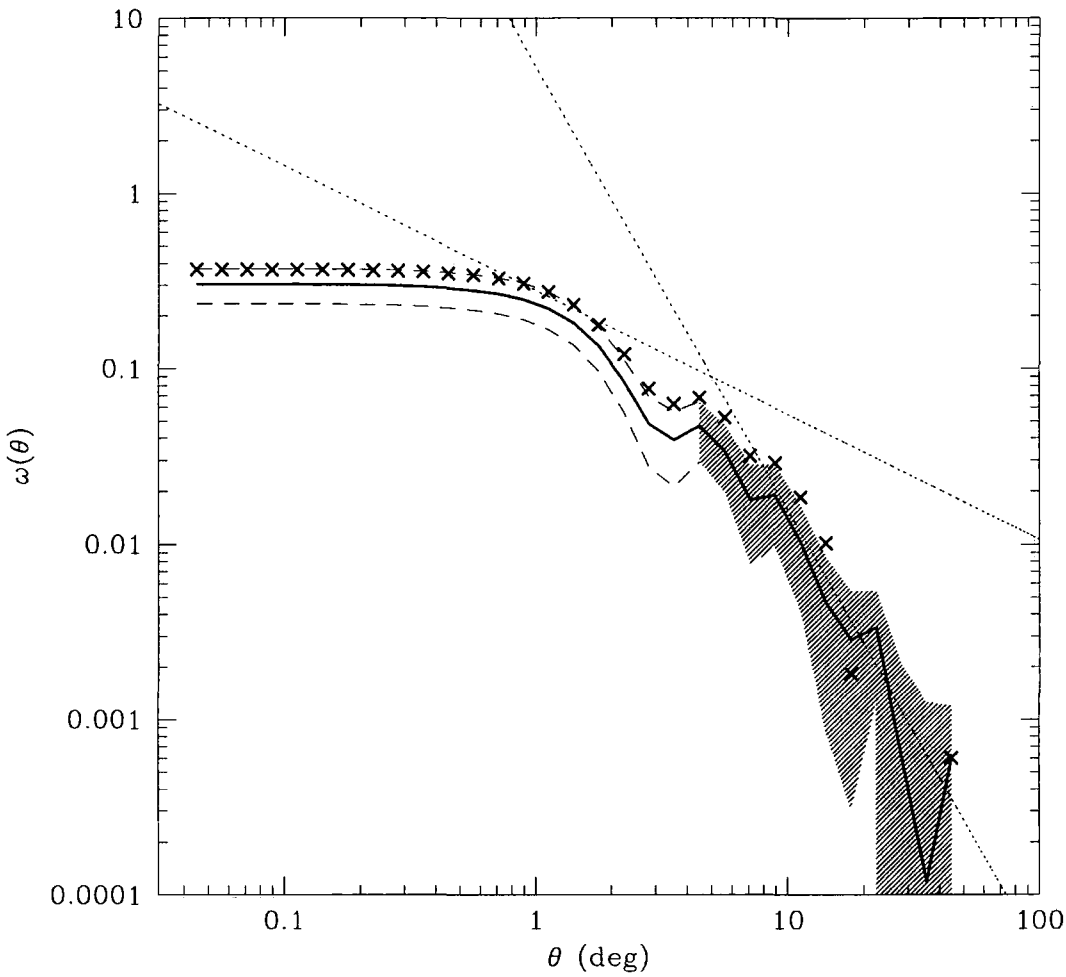


Figure 2.7: The K_s <12.5, $|b|$ ≥20° 2MASS (crosses) and mock 2MASS (solid line) 2-point angular correlation functions determined via a Bessel function transform of the K_s <12.5 angular power spectra for 1≤ l ≤100. We have applied a bias factor of b_K=1.1 to the (originally unbiased) mock 2MASS angular correlation function. The dotted lines indicate power law fits of A=0.28, $1-\gamma = -0.71$ for θ <5.0°, A=5.3, $1-\gamma = -2.5$ for θ ≥5.0° where $\omega = A\theta^{1-\gamma}$. The small scale fit is taken directly from Maller et al. (2005); at large scales the fit is determined for the θ ≥10° mock 2MASS datapoints; these fits are those used in the significance calculation of section 2.3.3. The shaded region indicates the 1σ spread determined from the 27 mock 2MASS catalogues on scales for which the transformation shown in equation 2.4 is robust, given that the sum ranges only between $l = 1$ and $l = 100$; for this reason the correlation function on scales smaller than $\theta \approx 4°$ is unreliable due to a lack of small scale information from the power spectrum in Fig. 2.6; on these scales the 1σ spread is indicated by dashed lines.

2.4.2 THE RELEVANCE FOR A LARGE LOCAL HOLE

Using equation 2.3, we can determine whether the possible excess of power observed at large scales in the 2MASS angular power spectrum over the ΛCDM prediction can account for a large local hole in the APM survey area, via a transform of the angular power spectrum to the angular correlation function. For this we use a Bessel function transform (Efstathiou 1990):

$$\omega(\theta) \approx \frac{1}{2\pi} \sum_{l} l\ C_l\ J_0(l\theta) \tag{2.4}$$

In Fig. 2.7, we present the transformation of the $K_s < 12.5$ 2MASS and ΛCDM mock 2MASS power spectra shown in Fig. 2.6 to the angular correlation function, together with the best fit power laws used in section 3.3. To the mock 2MASS angular correlation function determined from the unbiased ΛCDM Hubble Volume simulation we have applied a bias factor of b_K=1.1 (Maller et al. 2005). Both the 2MASS and mock 2MASS angular correlation functions are in good agreement with the Maller et al. (2005) best fit slope at $\approx 1°$; at smaller scales, the transformation becomes unreliable due to the lack of angular power spectrum information beyond l=100.

The 2MASS angular correlation function is in good agreement with the mock 2MASS catalogues at large scales, although there is a small difference in slope. We now wish to examine the effect of this difference to the form of the angular correlation function at large scales in order to verify whether the significance estimates determined in section 2.3.3 are robust to such a change. Computing the significance as in Equation 2.3 using a new power law fit to the 2MASS angular correlation function at large scales (A=11.1, $1-\gamma = -2.7$ for $\theta \geq 6.4°$, determined using $\theta \geq 10°$ 2MASS datapoints) does reduce the significance estimates shown in table 2.2; these are reduced by $\lesssim 0.5\sigma$ in each case. For example, the previous significance estimate with respect to the best fit model normalisation of 4.5σ is reduced to 4.0σ. (Including the uncertainty in the normalisation this corresponds to a reduction from 2.9σ to 2.8σ).

Thus the K_s <12.5 2MASS data may show increased power at large scales over what is expected from the Hubble Volume ΛCDM model. However, this excess power is still not able to account completely for the large local hole. The possible ways out are that we are detecting non-Gaussianity at large scales, or that the whole volume sampled by 2MASS is under-dense, or that the relationship between 2MASS galaxies and the underlying dark matter is non-linear on large scales, or that the uncertainty in the number count model

normalisation may be bigger than estimated. The first possibility is tested for in section 2.5 and we discuss the other possibilities in section 2.6.

2.5 COUNTS IN CELLS

Our third technique to investigate the local large-scale structure is a counts in cells analysis of the $K_s < 12.5$ 2MASS sample. At one level this is simply a check of the significance calculation used in section 2.3.3. However, it is also interesting to investigate whether the Hubble Volume simulation is able to reproduce the observed structure over cosmologically significant volumes.

2.5.1 METHOD

We sample the 2MASS survey area with a large number of randomly placed circular cells with an angular radius r. The Count Probability Distribution Function (CPDF; e.g. Croton et al. 2004) is defined as the probability of finding an exact number of galaxies, N, in a particular cell as a function of the cell size (a more detailed definition of this statistic is given in section 5.2.1). The CPDF is normalised such that the integral with respect to N for a particular cell size is unity. Here, we determine the CPDFs for N=0, 2, 6 and 20, sampling scales of $\lesssim 3°$. We therefore limit the area over which the cell centres are placed to $|b| \geq 20°$ and the 2MASS K_s <12.5 sample to $|b| \geq 10°$ to remove spurious edge effects. This magnitude limit is chosen in order that the peak in the selection function lies within the redshift range of interest; at K_s <12.5 the $n(z)$ peaks at $z \approx 0.05$, and so this magnitude limit preferentially samples the mean depth of the possible local hole in the APM survey area.

In order to probe much larger scales, of interest in this work, it is more useful to probe the density of galaxies rather than the absolute number. We define a Density Probability Distribution Function (DPDF) as the probability of finding a cell of given density (determined with respect to the mean 2MASS density) as a function of cell size. Since we wish to probe the galaxy distribution at large angular scales, we limit the cell centres to $|b| \geq 42°$ and the 2MASS K_s <12.5 sample to $|b| \geq 10°$ as before. For both the CPDFs and DPDFs, we mimic the 2MASS sample with the 27 biased and unbiased mock 2MASS catalogues described in section 2.2.3.

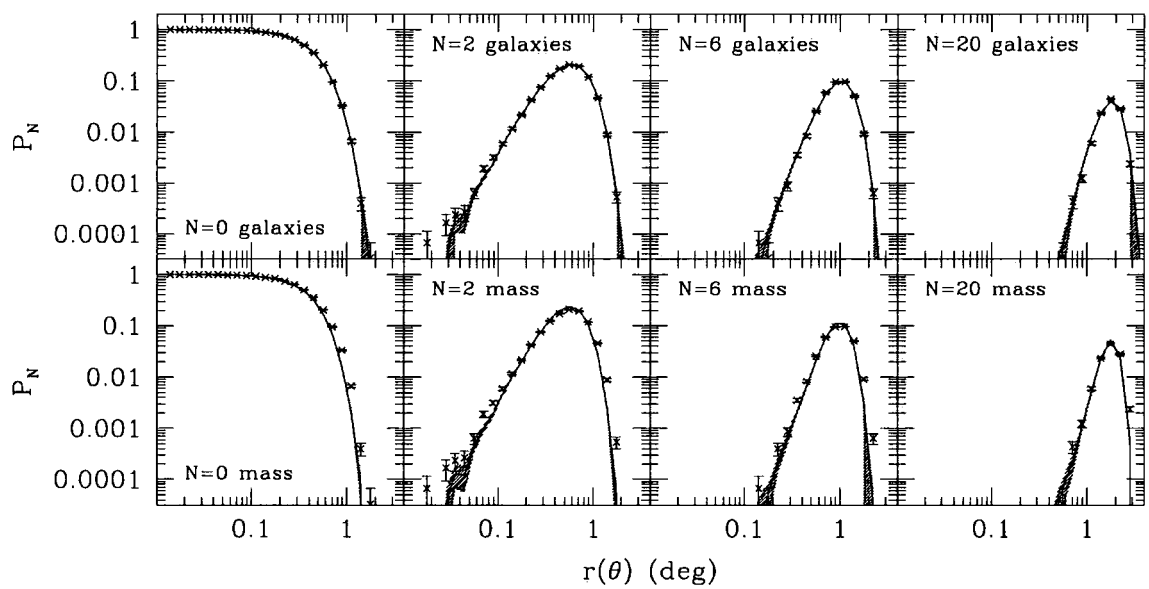


Figure 2.8: Count Probability Distribution Functions (CPDF) for N=0, 2, 6 and 20 for K_s <12.5 2MASS galaxies (crosses). The mean CPDFs (solid line) and 1σ spread from the 27 mock 2MASS catalogues are also shown, for unbiased (lower panels) and biased (upper panels) particles. The errorbars show the Poisson error for the number of cells recorded in each bin.

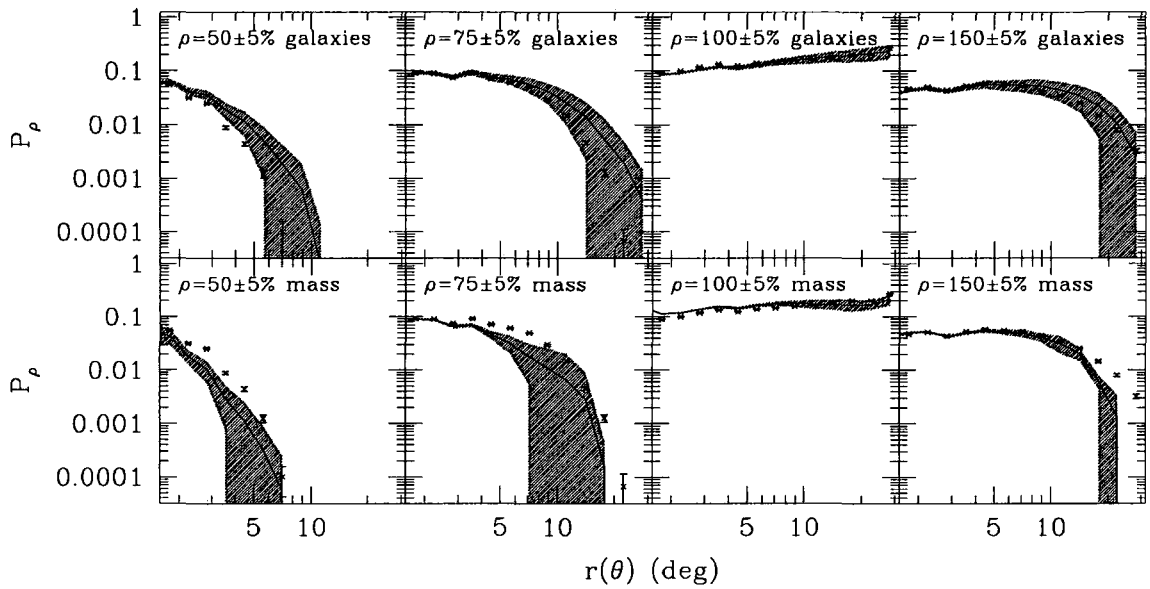


Figure 2.9: Density Probability Distribution Functions (DPDF) for densities of ρ=50±5 per cent, 75±5 per cent, 100±5 per cent and 150±5 per cent for K_s <12.5 2MASS galaxies (crosses). The mean DPDFs (solid line) and 1σ spread from the 27 mock 2MASS catalogues are also shown, for unbiased (lower panels) and biased (upper panels) particles. The errorbars show the Poisson error for the number of cells recorded in each bin.

2.5.2 RESULTS

In Fig. 2.8, we show the CPDFs for N=0, 2, 6 and 20 determined for the 2MASS and mock 2MASS samples. The agreement between the 2MASS and the mock galaxy samples is excellent, although this is not unexpected since the mock galaxy sample is produced to match the observed $\omega(\theta)$ at small scales.

Moving to larger scales, Fig. 2.9 shows the DPDFs for the 2MASS and mock galaxy samples for a range of observed densities, determined with respect to the mean 2MASS density. Again, the agreement with the biased ΛCDM prediction is excellent to extremely large scales ($\theta \lesssim 30°$). It is also interesting to note that the introduction of bias has an extremely significant effect on the resulting DPDF. While the solid angles probed in the largest bin are slightly smaller than the APM survey area, the good agreement between the biased mocks and the 2MASS sample (especially in the ρ = 75 per cent case of interest for the possible local hole) confirms the significance calculation shown in table 2.2 which suggests that, when compared to the mean $|b| \geq 20°$ 2MASS number density, the observed deficiency in the APM survey area is not significant when compared to the ΛCDM prediction; equally under-dense areas occur in the biased mock simulations. There is therefore no need from this analysis to invoke non-Gaussianity, for example, to explain the under-density in the APM survey area. However, the mock 2MASS catalogues are normalised to the mean number count observed by 2MASS and so any increase in the global galaxy density, as might be suggested by the faint ($14< K <18$) K-band counts compiled from the literature (see section 3.1), increases the significance of a large local hole.

2.6 DISCUSSION & CONCLUSIONS

Recent evidence (Busswell et al. 2004, Frith et al. 2003) has suggested that while the optical number counts over the APM survey area may be significantly less deficient than originally proposed, the resulting under-density might still present a challenge to the form of clustering predicted by ΛCDM at large scales. In this chapter, we have presented three different methods of analysis to probe the possible existence of this large local hole in the galaxy distribution around the SGP using the recently completed 2MASS survey.

First, we determined the 2MASS K_s-band number counts over the APM survey area. In order to probe the underlying galaxy distribution, we compared these counts with a model guided by the Southern 2dFGRS $n(z)$; this variable ϕ^* model provides a good

agreement with the corresponding K_s-band counts extracted for the Southern 2dFGRS strip. The agreement between the Southern 2dFGRS variable ϕ^* model and the K_s-band counts over the APM survey area is remarkable, and suggests that the galaxy distribution over this $\approx$4000 deg^2 area may be similar to that of the much smaller 2dFGRS Southern strip. Using the 2MASS-2dFGRS matched $n(z)$, this would imply a mean deficiency in the galaxy distribution of 23 per cent to $z = 0.1$ with respect to the Metcalfe et al. (2001) model used in this paper; taking the best fit normalisation of this model to faint K-band data compiled from the literature in the range 14$<K<$18 implies a mean deficiency of 20 per cent to z=0.1.

The issue is complicated by the $b \geq 20°$ and $b \leq -20°$ 2MASS counts which are in good agreement with each other but are significantly below the Metcalfe et al. (2001) homogeneous prediction. There are three possible interpretations. The first is that the model normalisation is too high. However, if the model were scaled down in order to agree with the mean 2MASS number density, this would require a 2.5σ deviation from the faint (14$<K<$18) K-band data compiled from the literature; to account for the low APM counts entirely through a change in the normalisation would require a 3.8σ deviation. Secondly, the low cap counts might indicate a zero-point offset between the Metcalfe et al. (2001) model and the 2MASS data. If one were to invoke the $\approx$0.15 magnitude shift necessary to bring the cap counts and the model into line, this would compromise the good agreement between the 2MASS and Loveday (2000) photometry and also between the 2dFGRS K_s-band counts and the corresponding variable ϕ^* models. Thirdly, the low 2MASS cap counts might indicate that the entire local galaxy distribution is globally under-dense. While this might appear to be unlikely, the observed 2MASS counts over the APM survey area suggest that large inhomogeneities exist in the galaxy distribution over extremely large volumes, and so perhaps only a few such features are necessary to bias the entire local sample. It is therefore not possible to rule out this final possibility without further analysis. However, since the $b \geq 20°$ and $b \leq -20°$ counts are similar, this position requires the coincidence that we are positioned in the centre of this local under-density (see Loveday 2004 for further discussion on this point).

In order to determine the significance of the observed K_s-band counts over the APM survey area, we calculated the expected dispersion in the K_s <12.5 counts over $\approx$4000 deg^2 considering a ΛCDM form of the 2-point angular correlation function at large scales. The observed deficiency is calculated with respect to the Metcalfe et al. (2001) model, the best fit normalisation of this model described in section 2.3.1, and the mean 2MASS

$|b| \geq 20°$ number density ($\bar{n}_{2MASS}$). In the first two cases, the observed counts represent an extremely rare fluctuation from that expected in a ΛCDM cosmology. If the Metcalfe et al. (2001) model is effectively lowered by 2.5σ with respect to the faint K-band data to the mean 2MASS number density, then the observed counts begin to become more consistent with a ΛCDM form of the correlation function. Therefore, unless the faint K-band data are systematically too high or the galaxy distribution in the SGC is an extremely rare fluctuation in the galaxy density, then the K_s-band counts over the APM survey area appear to imply an excess of power at large scales over the ΛCDM prediction.

Our second technique was therefore to examine the large-scale power of the 2MASS galaxy sample through a determination of the $|b| \geq 20°$, $K_s < 12.5$ angular power spectrum. We compare this to a ΛCDM prediction determined from the Hubble Volume simulation. The two are in reasonable agreement although there is some discrepancy in the slopes; taking a K_s-band bias of $b_K = 1.1$ (Maller et al. 2005) results in a 3σ excess over the mean ΛCDM angular power spectrum at large scales ($l \leq 30$). In order to determine the effect of this apparent excess on the significance estimates used previously, we transform the angular power spectrum to the angular correlation function via a Bessel function transform. The corresponding best fit to the 2MASS angular correlation function at large scales ($\theta \geq 10°$) decreases the significance estimates by $\lesssim 0.5\sigma$ compared to the ΛCDM angular correlation function used previously. Therefore while there appears to be an excess of power at large scales in the 2MASS catalogue over the ΛCDM Hubble Volume simulation, it is not enough to account for the observed deficiency in the APM survey area. One caveat to this is that if the local galaxy distribution was globally under-dense with respect to the faint K-band counts, as might be inferred from the low $|b| \geq 20°$ counts in Fig. 2.5, then the corresponding significance estimates would change due to the fact that the 2MASS correlation function is not drawn from a fair sample of the Universe.

Our third technique was to use a counts in cells analysis of the $K_s < 12.5$ 2MASS galaxy sample and also the unbiased and biased mock 2MASS catalogue constructed from the ΛCDM Hubble Volume simulation (again normalised to the mean $|b| \geq 20°$ 2MASS galaxy density). This enables us to verify the significance estimates used previously, and also that the form of real features in the galaxy distribution are reproduced by the ΛCDM Hubble Volume simulation at large scales. The biased mock catalogues reproduce the observed galaxy distribution to very large scales ($\lesssim 30°$). This supports the significance calculation determined previously (with respect to the mean 2MASS number density) which indicates that the local hole is not a challenge to ΛCDM if the $K_s < 12.5$ 2MASS

catalogue is a fair sample of the Universe.

In conclusion, the issue of the large local hole in the local galaxy distribution has yet to be resolved. The 2MASS K_s-band number counts extracted for the $\approx$4000 deg^2 APM survey area indicate a clear deficiency in the local galaxy distribution consistent with the form of the $n(z)$ in the much smaller 2dFGRS Southern strip. However, to determine the level of the deficiency in the galaxy distribution requires an accurate normalisation of the K-band number count model. Using the Metcalfe et al. (2001) model, which provides an excellent fit to faint K-band data compiled from the literature in the range $14< K <18$, implies a deficiency over the APM survey area which is at odds with ΛCDM and a local galaxy distribution which is globally under-dense. Only if the model is lowered by 3.8σ below the faint K-band data can the normalisation account for the low counts over the APM survey area.

In the next chapter, we aim to better constrain possible sources of systematic error, particularly the issue of the model normalisation, using new deep H-band photometry.

CHAPTER 3

H-BAND GALAXY NUMBER COUNTS

3.1 INTRODUCTION

In the previous chapter, the issue of the deficiency in the bright number counts around the Southern Galactic Pole (SGP) was investigated in the near infrared K_s-band. To recap, the K_s-band counts extracted from the 2 Micron All Sky Survey (2MASS) over the $\approx$ 4000 deg^2 APM survey area were first determined; a similar deficiency was observed to the optical APM survey counts (with the zeropoint offset determined by Busswell et al. (2004) applied), with a $\approx$25 per cent deficit at K_s =12 compared to the no evolution model of Metcalfe et al. (2001). Using a ΛCDM form for the angular correlation function at large scales, and assuming that the observed counts were solely due to features in the local galaxy distribution, the observed counts represented a 5σ fluctuation. However, this result was complicated by the fact that the 2MASS K_s-band number counts for almost the entire survey ($|b|$ >20°, covering $\approx$27 000 deg^2) were also low, with a constant deficiency of $\approx$20 per cent between K_s =10 and K_s =13.5.

Does this surprising result perhaps indicate that the K_s-band Metcalfe et al. (2001) model normalisation is too high? Or, as suggested previously, could low redshift luminosity evolution significantly affect the bright counts? These issues were also addressed in the previous chapter: First, the Metcalfe et al. (2001) model was compared with faint K-band data collated from the literature. Fitting in the magnitude range 14< K <18 it was found that the best fit model normalisation was slightly too high, although not significantly (this magnitude range was used so as to avoid fluctuations in the counts arising from large-scale structure at bright magnitudes and significant effects from galaxy evolution at the faint end). Accounting for the normalisation uncertainty (of $\pm$6 per cent) the observed deficiency in the K_s-band counts over the APM survey area still represented a $\approx 3\sigma$ fluctuation. Second, the issue of low redshift luminosity evolution was also addressed: 2MASS galaxies below K_s = 13.5 were matched with the Northern and Southern areas of the 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2003). The resulting

$n(z)$, covering > 1000 deg^2 in total, was consistent with the no evolution model of Metcalfe et al. (2001). In addition, these K_s-band redshift distributions were used to form predictions for the number counts over the Northern and Southern 2dFGRS areas respectively. This was done by multiplying the luminosity function parameter ϕ^* (which governs the model normalisation) used in the Metcalfe et al. (2001) model by the relative density observed in the K_s-band $n(z)$ as a function of redshift. These 'variable ϕ^* models' were then compared with 2MASS counts extracted for the 2dFGRS areas in order to determine whether the observed counts were consistent with being due solely to features in the local galaxy distribution; the variable ϕ^* models were in good agreement with the number counts, indicating that low redshift luminosity evolution is unlikely to have a significant impact on the observed deficiency in the counts, in the K_s-band at least.

In this chapter we aim to address the issue of low, bright number counts in the near infrared H-band. In particular we wish to address a drawback to the K_s-band analysis in the previous chapter - the issue of the number count model normalisation; while the K_s-band model used was compared with faint data and was found to be in good agreement, the level to which systematic effects, arising perhaps via zeropoint offsets between the bright and faint data or cosmic variance in the faint data, might affect the conclusions were uncertain. We address this issue in the H-band using new faint data covering 0.3 deg^2 to $H = 18$, calibrated to match the 2MASS zeropoint. In section 3.2, we first verify that the H-band counts provide number counts over the APM survey area which are consistent with the previous results in the B-band (Busswell et al. 2004) and K_s-band (previous chapter), and that the form of the counts is not significantly affected by low redshift luminosity evolution through comparisons with variable ϕ^* models (described above). In section 3.3, we provide details of the data reduction of the new faint H-band photometry. The associated counts are presented in section 3.4. In section 3.5 we discuss possible systematics affecting the bright number counts including the model normalisation and incompleteness. The conclusions follow in sections 3.6.

3.2 BRIGHT H-BAND COUNTS FROM 2MASS

We wish to examine the form of bright number counts in the H-band in order to verify that the counts over the APM survey area ($\approx$4000 deg^2 around the SGP) are comparable to those measured previously in the optical B-band (Busswell et al. 2004) and near infrared K_s-band (previous chapter). The near infrared has the advantage of being sensitive to

the underlying stellar mass and is much less affected by recent star formation history than optical wavelengths. For this reason, number count predictions in the near infrared are insensitive to the evolutionary model at bright magnitudes. In Fig. 3.1 we show faint H-band data collated from the literature along with bright counts extracted from 2MASS over $\approx$27 000 deg^2. The 2MASS magnitudes are determined via the 2MASS H-band extrapolated magnitude; this form of magnitude estimator has previously been shown to be an excellent estimate of the total flux in the K_s-band (see previous chapter) through comparison with the total magnitude estimates of Jones et al. (2004) and the K-band photometry of Loveday (2000). Throughout this work we use 2MASS H-band counts determined via this magnitude estimator. We also show two models in Fig. 3.1 corresponding to homogeneous predictions assuming no evolution and pure luminosity evolution models. These are constructed from the H-band luminosity function parameters listed in Metcalfe et al. (2005) and the $K+E$-corrections of Bruzual & Charlot (1993). At bright magnitudes the two are indistinguishable; only at $H \gtrsim 18$ do the model predictions begin to separate. The faint data is in good agreement with both the no evolution and pure luminosity evolution predictions to $H \approx 26$.

Before examining the H-band counts over the APM survey area, we first verify that the bright counts are consistent with relatively insignificant levels of low redshift luminosity evolution (i.e. less than a few tenths of a magnitude for $z < 0.2$) in the manner carried out in the previous chapter for the K_s-band counts. In the upper panels of Fig. 3.2 we show H-band $n(z)$ to the 2MASS limiting magnitude of $H = 14$, determined through matched 2MASS and 2dFGRS galaxies over the 2dFGRS Northern (left hand) and Southern (right hand panels) declination strips (see section 2.2.2 for further details of the matching technique). The solid lines indicate the expected homogeneous distribution constructed from the pure luminosity evolution predictions of Metcalfe et al. (2005) (there is no discernible difference between this and the no evolution prediction). In the lower panels we divide through by this prediction; these panels show the relative density as a function of redshift. The observed $n(z)$ are consistent with the expected trends, with relatively homogeneous distributions beyond $z = 0.1$ (1 per cent and 8 per cent over-dense in the North and South respectively for $0.1 \leq z \leq 0.2$). For this reason, Fig. 3.2 suggests that the level of luminosity evolution is relatively insignificant at low redshifts in the H-band; strong luminosity evolution produces an extended tail in the predicted $n(z)$ which is not observed in the data.

As a further check against strong low redshift luminosity evolution, we can use the ob-

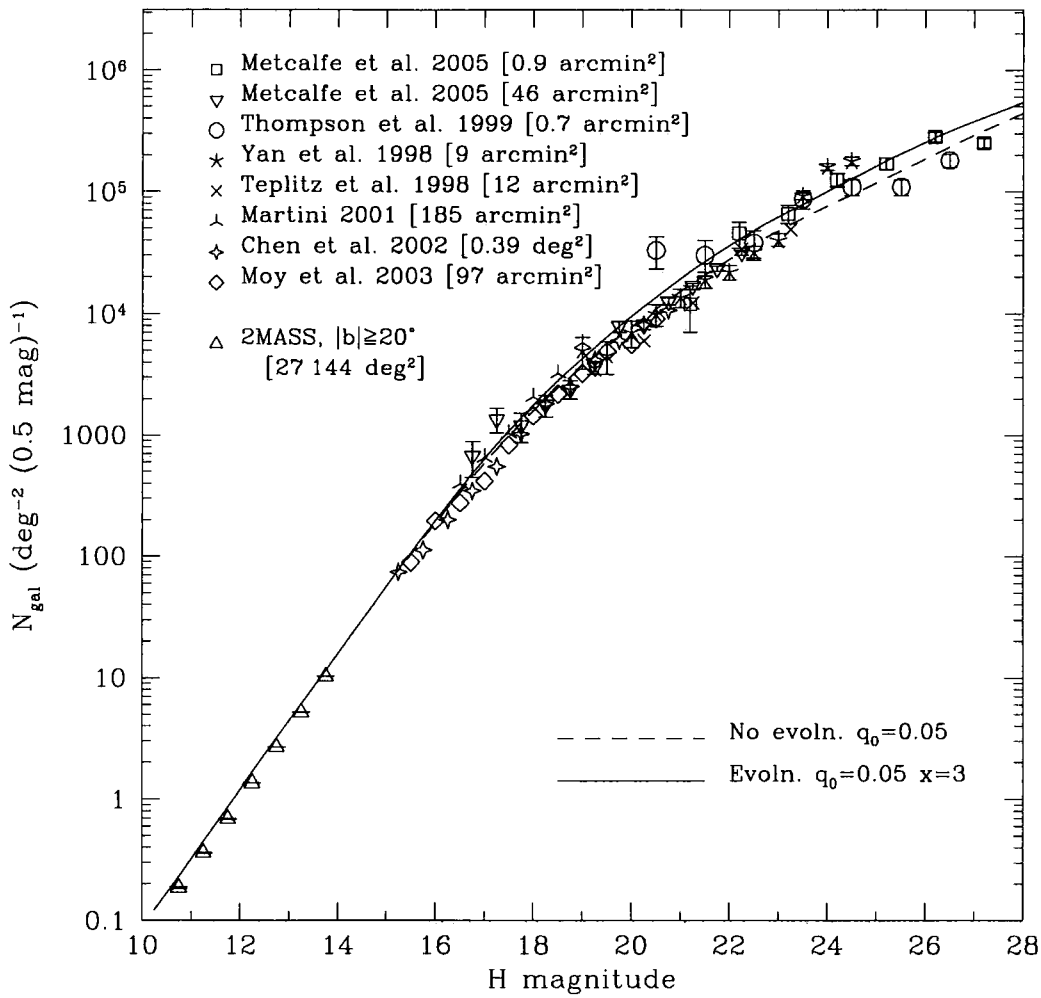


Figure 3.1: H-band galaxy number counts collated from the literature. The dashed and solid lines indicate the no evolution and pure luminosity evolution predictions respectively, described in section 3.2. We also show bright H-band counts extracted from the 2MASS extended source catalogue for $|b|$ >20°. For each dataset, we indicate the associated observed solid angle in square brackets.

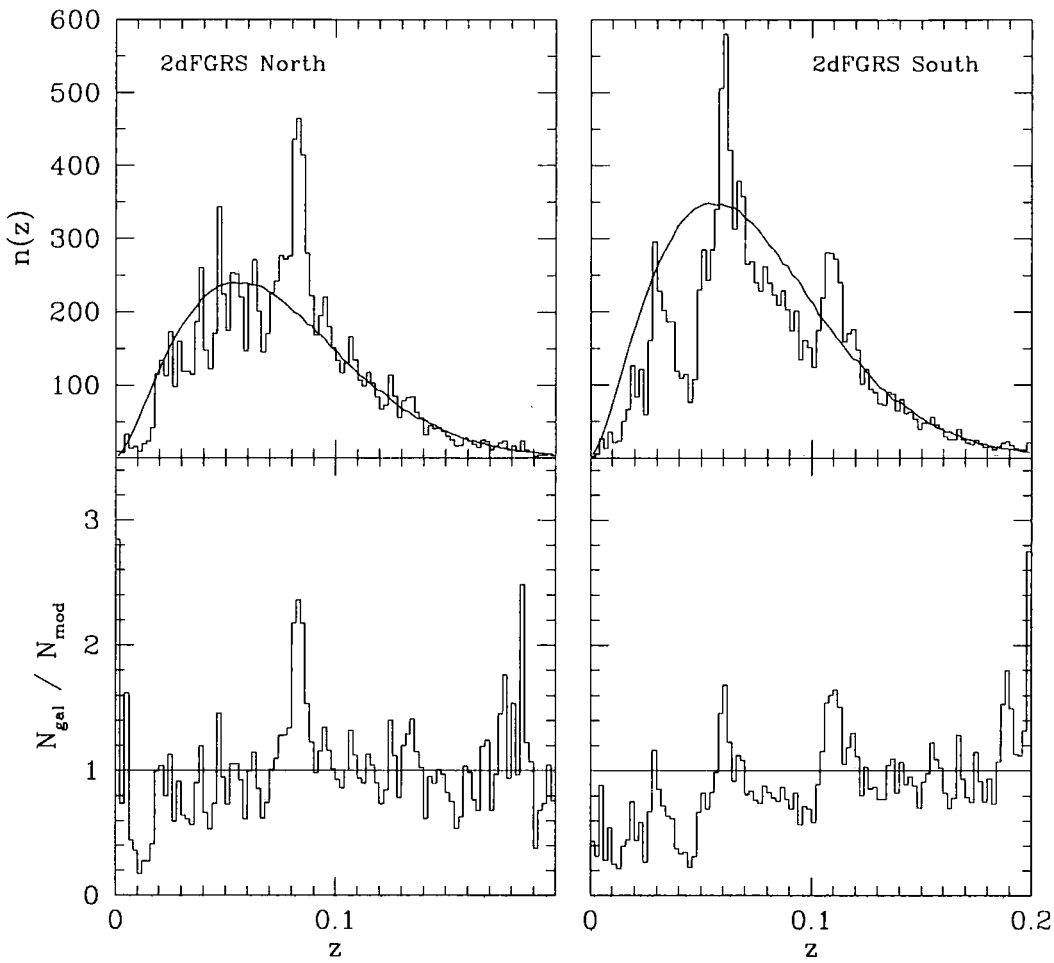


Figure 3.2: Number redshift histograms for 11 501 and 13 687 H <14 2MASS galaxies matched with the 446 deg^2 Northern (left hand) and 647 deg^2 Southern (right hand panels) 2dFGRS declination strips respectively. In each case the solid lines indicate the passive pure luminosity evolution prediction for a homogeneous distribution described in section 3.2 normalised by the respective solid angles. We also indicate the relative density in the lower panels, dividing the observed $n(z)$ by the homogeneous prediction.

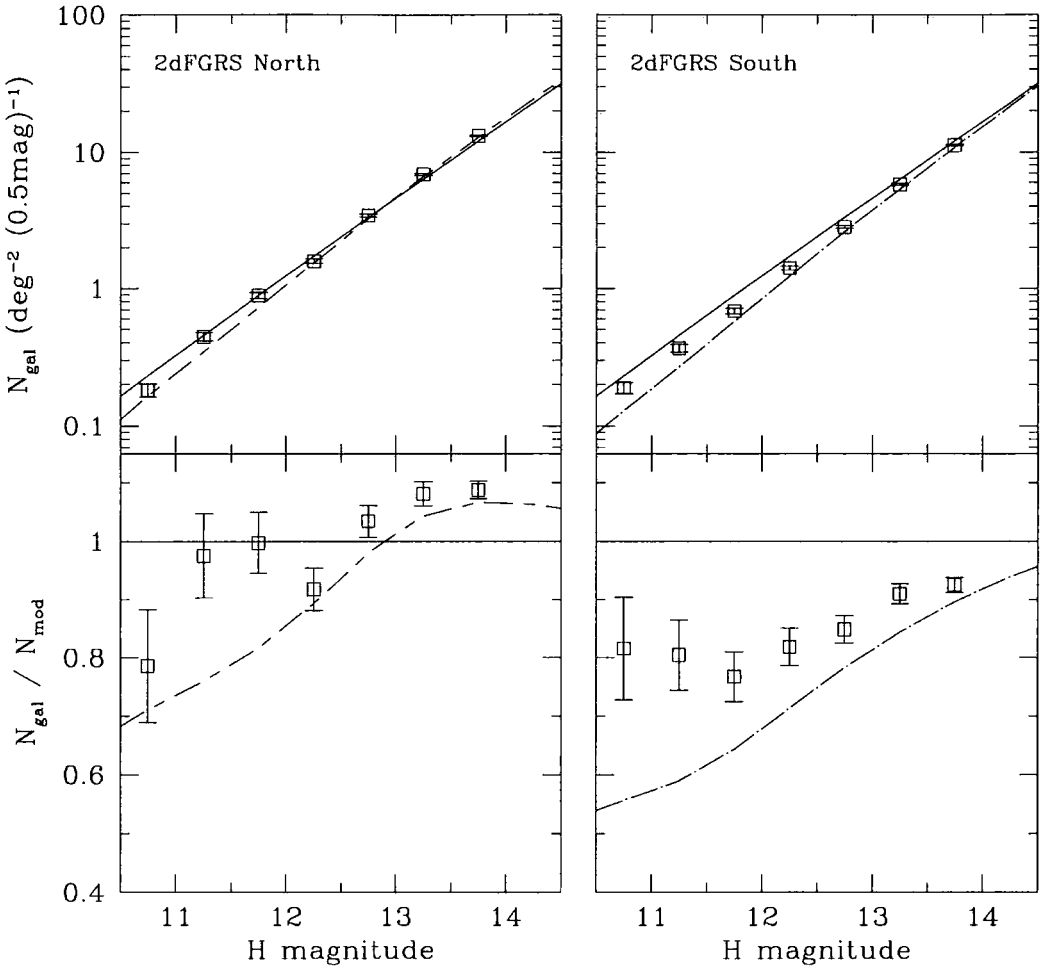


Figure 3.3: H-band 2MASS galaxy number counts extracted from the Northern (left hand) and Southern (right hand) 2dFGRS declination strips. The solid line indicates the homogeneous pure luminosity evolution prediction described in section 2 (this and the no evolution prediction are indistinguishable at these magnitudes). The dashed and dot-dashed lines indicate the variable ϕ^* models for the Northern and Southern 2dFGRS strips respectively; these indicate the expected number counts given the observed $n(z)$ (Fig. 3.2). In the lower panels we divide through by the homogeneous prediction. In each case the errorbars indicate the Poisson uncertainty in each bin.

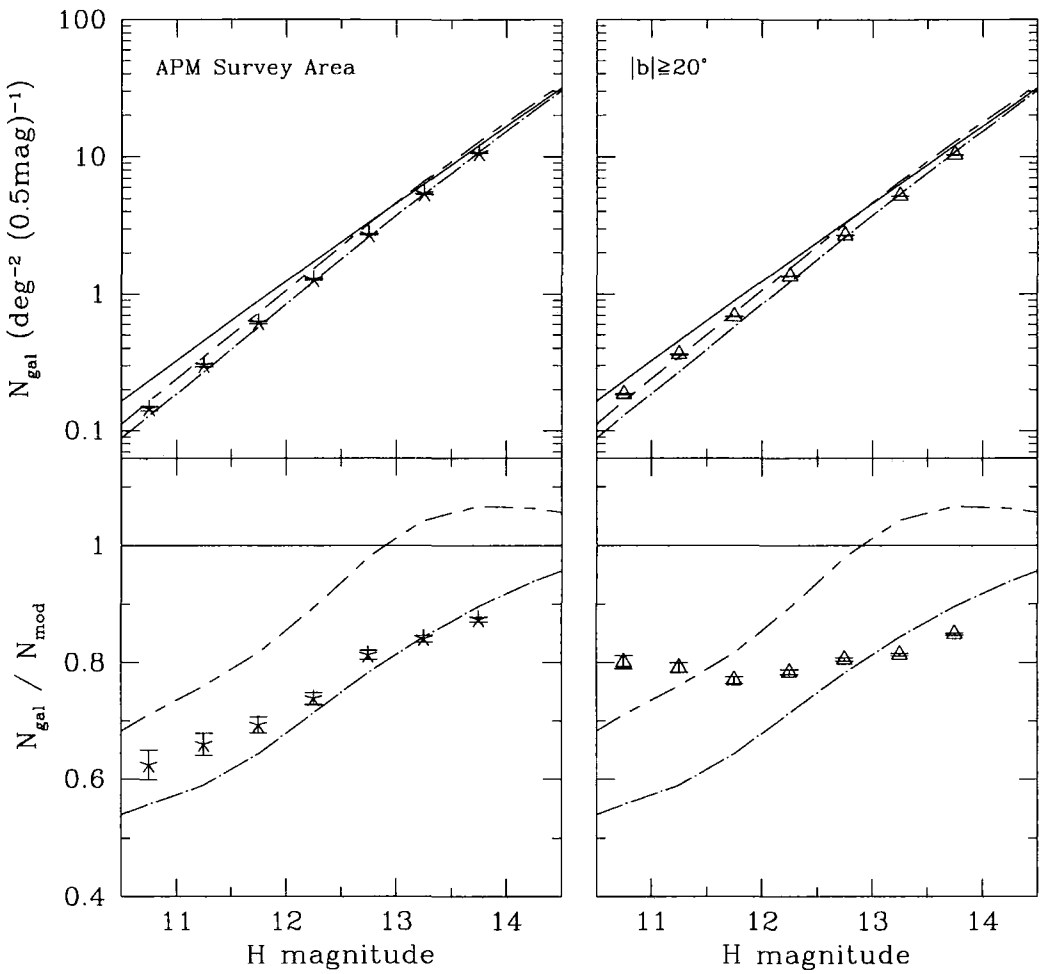


Figure 3.4: H-band 2MASS galaxy number counts extracted for the APM survey area ($\approx$4 000 deg^2) and for $|b|$ >20° ($\approx$27 000 deg^2), shown in the left and right hand panels respectively. As in Fig 3.2, we show the homogeneous pure luminosity prediction (solid line), and the Northern (dashed) and Southern (dot-dashed) variable ϕ^* models, indicating the expected number counts for the redshift distributions shown in Fig. 3.2. As before, in the lower panels we divide through by the homogeneous prediction. In each case the errorbars indicate the Poisson uncertainty in each bin.

served $n(z)$ to predict the expected H-band number counts over the 2dFGRS declination strips. This technique is described in detail in Frith et al. (2003). To recap, we use the observed density (Fig. 3.2, lower panels), to vary the luminosity function normalisation (ϕ^*) used in the Metcalfe et al. (2005) model as a function of redshift (for $z \leq 0.2$). We show these 'variable ϕ^* models' along with the 2MASS H-band counts extracted for the 2dFGRS strips in Fig. 3.3. In each case, the upper panels indicate the number count on a logarithmic scale; in the lower panels we divide through by the homogeneous prediction. In both the Northern and Southern 2dFGRS areas, the counts are in good agreement with the expected trend, defined by the corresponding variable ϕ^* model. This indicates that real features in the local galaxy distribution are the dominant factor in the form of the observed H-band number counts, and that strong low redshift luminosity evolution is unlikely to have a significant role in any under-density observed in the APM survey area; any strong evolution model would have to be fine-tuned in order to maintain this agreement and fit the observed $n(z)$.

We are now in a position to examine the number counts over the APM survey area. In Fig. 3.4 we show counts extracted for the $\approx$4000 deg^2 field along with the homogeneous and the Northern and Southern 2dFGRS variable ϕ^* models shown in Fig. 3.2. The form of the counts is in good agreement with the B-band (Busswell et al. 2004) and K_s-band (previous chapter) bright number counts measured over the APM survey area, with a deficiency of $\approx$25 per cent below $H = 13$. In addition, the form of the counts is similar to that of the counts extracted from the 2dFGRS Southern declination strip and the corresponding variable ϕ^* model (this is also observed in the B and K_s-band); this perhaps indicates that the form of the local galaxy distribution in the $\approx$600 deg^2 2dFGRS Southern declination strip is similar to that of the much larger APM survey area, with an under-density of $\approx$25 per cent to z =0.1. However, the 2MASS H-band counts over almost the entire survey ($|b|$ >20°, $\approx$27 000 deg^2) are also deficient (as are the K_s-band counts), with a relatively constant deficit of $\approx$15-20 per cent to $H = 14$ (Fig 3.3, right hand panels).

The low $|b|$ >20° counts raise the question as to whether systematic effects are significant, or whether these counts are due to real features in the local galaxy distribution, as suggested by the agreement between the variable ϕ^* models and corresponding counts in Fig. 3.3. If the latter is true, then the size of the local hole would not only be much larger than previously suggested but would also represent an even more significant departure from the form of clustering at large scales expected in a ΛCDM cosmology. In the

following two sections we address a possible source of systematic error using new faint H-band photometry - the model normalisation. Other possible causes for the low counts are also discussed in section 3.5.

3.3 NEW FAINT H-BAND DATA

3.3.1 OBSERVATIONS

The new H-band data was taken between the 1^{st} and 3^{rd} of September 2004 with Nigel Metcalfe on the Calar Alto 3.5m telescope using the 15'.4×15'.4 Omega 2000 camera. The primary objective of the observing run was to observe the William Herschel Deep Field (WHDF) in the H and K-bands to faint magnitudes ($H \approx 23$). However, we were also able to observed a larger solid angle to shallower depth in order to examine the normalisation of the H-band model mentioned previously at intermediate magnitudes. We took 15 minute exposures for 6 pointings; the combined solid angle of this 'Calar Alto (CA) field' was 0.24 deg^2, excising unreliable regions at the edge of the exposure. The approximate magnitude limit was $H \gtrsim 19$ for most of this area; due to the non-photometric conditions however, for 0.024 deg^2 of this the magnitude limit was considerably lower, $H \gtrsim 17.5$. Including the observations of the WHDF, we use new faint H-band data covering a combined solid angle of 0.30 deg^2 to $H = 17.5$ and 0.27 deg^2 to $H = 19.0$.

3.3.2 CALIBRATION

Photometric calibration of the H-band images is obtained through comparison with the 2MASS point source catalogue. Fig. 3.5 shows the 2MASS magnitudes compared with our data for 393 matched point sources over the CA field and the WHDF. The zeropoint of our data is chosen to match that of the 2MASS objects and is accurate to ±0.01 magnitudes. The large datapoints and errorbars indicate the mean offset and *rms* dispersion as a function of magnitude. When comparing this data to the 2MASS number counts at bright magnitudes it is important to note that the 2MASS point source catalogue includes a maximum bias in the photometric zeropoint of <2 per cent around the sky (see the 2MASS website).

3.3.3 STAR/GALAXY SEPARATION

We use the Sextractor software to separate objects below $H = 18$; for this magnitude limit, the associated STAR_CLASS parameter provides a reliable indicator of stars and

galaxies. We identify 30.0 per cent as galaxies (CLASS_STAR<0.1), 58.9 per cent as stars (CLASS_STAR>0.9), leaving 11.1 per cent as unclassified.

3.4 FAINT H-BAND COUNTS

3.4.1 COMPARISON WITH THE LCIRS

Before determining number counts for the new H-band data described in the previous section, we first examine the photometry of the Las Campanas Infra-Red Survey (LCIRS; Chen et al. 2002). The published data covers 847 arcmin^2 in the Hubble Deep Field South (HDFS) and 561 arcmin^2 in the Chandra Deep Field South (CDFS); the combined solid angle (0.39 deg^2) represents the largest H-band dataset for $14 \lesssim H \lesssim 20$. The associated number counts are $\approx$15 per cent below the homogeneous Metcalfe et al. (2005) predictions at $H = 18$ (see Fig. 3.1). This is significant, as if the model normalisation was altered to fit, the deficiency in the 2MASS counts at bright magnitudes (Fig. 3.3) would become much less severe. However, various other surveys show higher counts, although over much smaller solid angles. With the LCIRS data in particular therefore, it is vital to ensure that the photometric zeropoint is consistent with the 2MASS data at bright magnitudes.

In Fig. 3.6 we compare the LCIRS and 2MASS H-band photometry for 438 points sources matched over the HDFS and CDFS fields. There appears to be a large offset which is approximately constant for K >12. Using point sources matched at all magnitudes, we determine a mean offset of -0.28$\pm$0.01 magnitudes; this is robust to changes in the magnitude range and is consistent over both the HDFS and CDFS fields.

3.4.2 NEW H-BAND COUNTS

In Fig. 3.7 we show counts determined for the new H-band data described in section 3.3, the 0.27 deg^2 CA field and the 0.06 deg^2 WHDF (see also table 3.1). Both sets of counts are in excellent agreement with the pure luminosity evolution and no evolution homogeneous predictions of Metcalfe et al. (2005). In addition we show LCIRS counts determined in the 0.24 deg^2 HDFS and 0.16 deg^2 CDFS, applying the 0.28 magnitude zeropoint offset determined with respect to 2MASS in section 3.4.1. The associated counts are also in excellent agreement with the Metcalfe et al. (2005) models at all magnitudes.

In Fig. 3.8, we show counts determined from our data and the LCIRS combined, with a consistent zeropoint applied as in Fig. 3.7. We estimate the uncertainty arising from cosmic variance using field-to-field errors, weighted by the solid angle of each field.

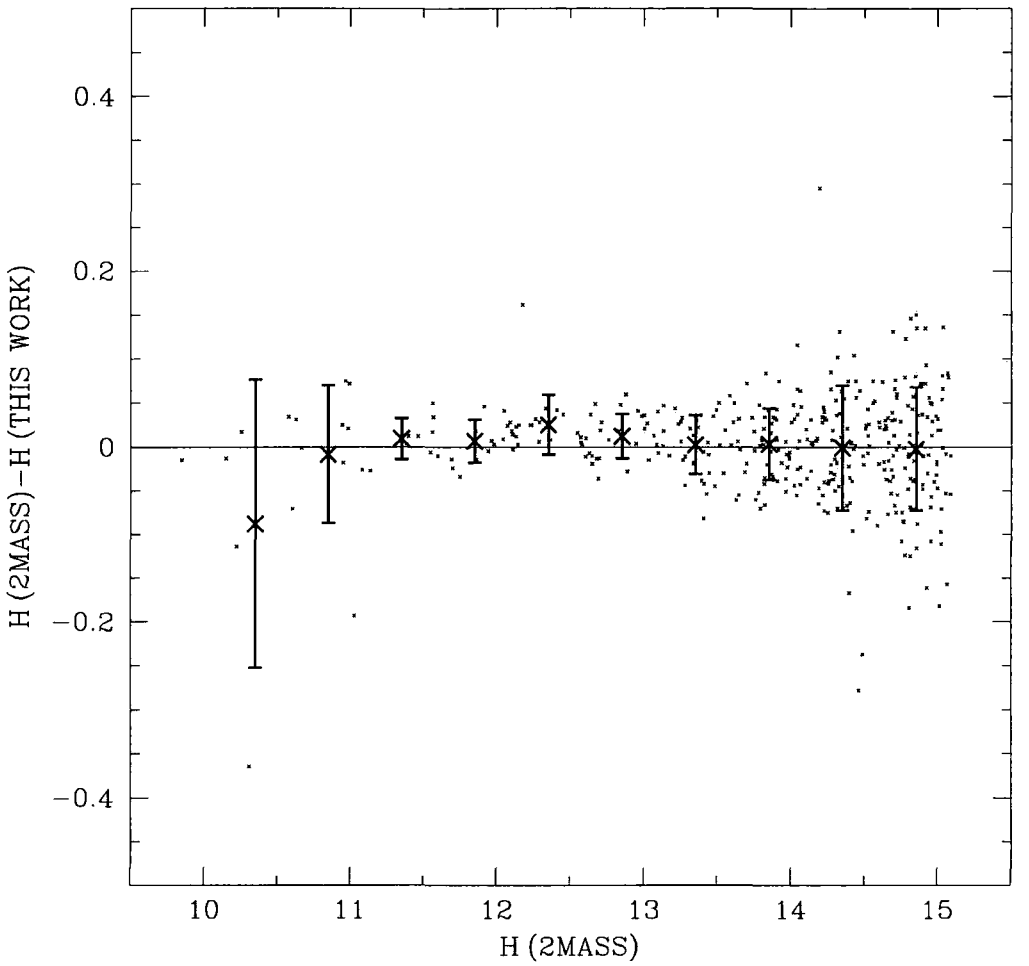


Figure 3.5: Here we examine the uncertainty in our photometric calibration with 2MASS. The H-band magnitudes determined by 2MASS and the residual with our photometry are indicated for 393 point sources below $H = 15.1$. The large datapoints indicate the mean offset and rms dispersion as a function of magnitude. The zeropoint used is indicated by the solid line and is accurate to ± 0.01 magnitudes at 1σ confidence.

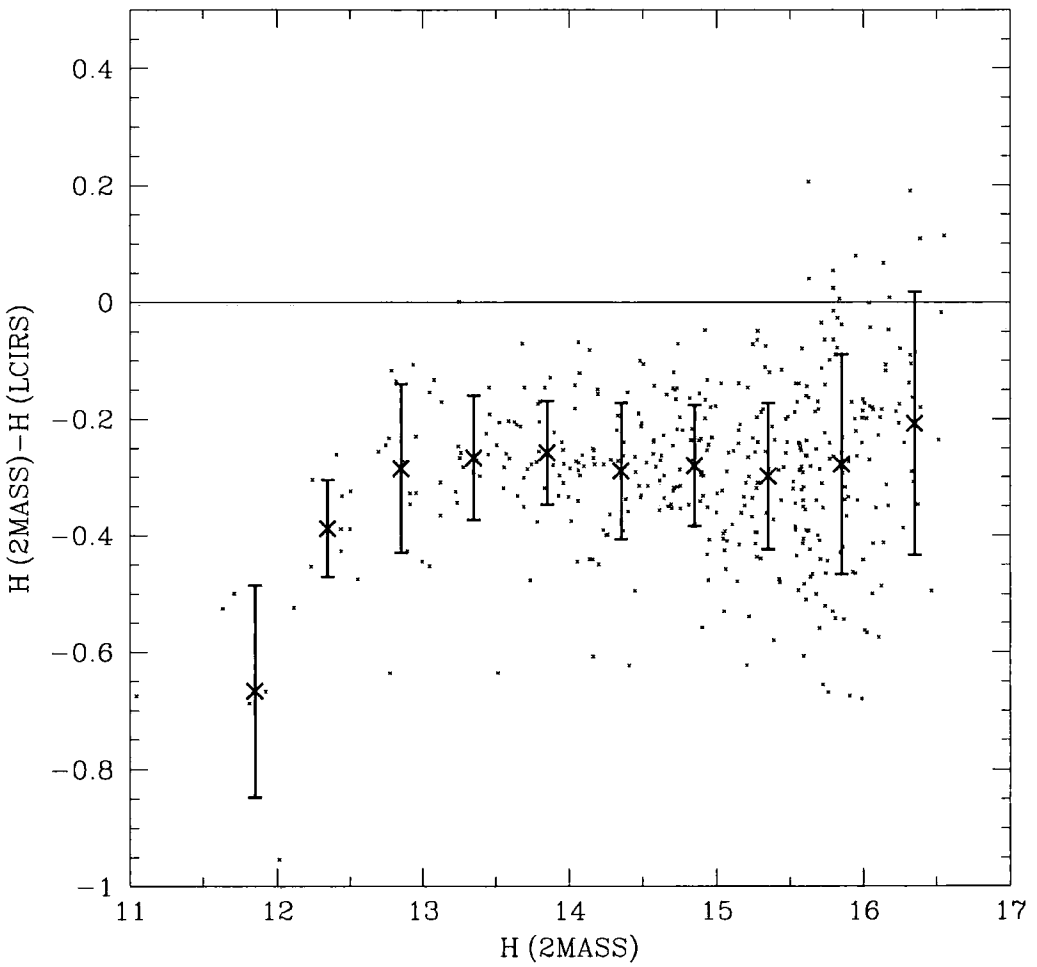


Figure 3.6: We compare the H-band photometry of the LCIRS (Chen et al. 2002) with 2MASS using 438 points sources. As in Fig. 3.5, the large datapoints indicate the mean offset and rms dispersion as a function of magnitude. The mean offset is -0.28±0.01 magnitudes at 1σ confidence. The zeropoint used in the new data presented in this work is indicated by the solid line.

H	$N_{CA\ field}$	N_{WHDF}	N_{HDFS}	N_{CDFS}	N_{tot} (deg^{-2})	N_{mod} (deg^{-2})
14.25	10	4	6	8	40.8	23.0
14.75	17	5	12	8	61.1	43.5
15.25	21	9	23	23	110	81.9
15.75	41	14	31	43	188	153
16.25	55	15	77	51	288	280
16.75	133	39	163	73	594	500
17.25	217	58	238	135	943	861
17.75	283	77	337	256	1.44×10^3	1.43×10^3

Table 3.1: The raw number counts per half magnitude are shown for the new H-band data described in section 3.3 - the CA field (0.27 deg^2) and WHDF (0.06 deg^2) in columns 2 and 3. In addition, we show the counts for the LCIRS fields, the HDFS (0.24 deg^2) and CDFS (0.16 deg^2) in columns 4 and 5, applying the zeropoint offset determined with respect to 2MASS in section 3.4.1. The total number count per deg^2 for all fields combined (0.69 deg^2) is shown in column 6 along with the homogeneous pure luminosity evolution prediction of Metcalfe et al. (2005) in column 7. The faintest magnitude bin for the CA field is slightly smaller (0.21 deg^2) than at brighter magnitudes; the combined solid angle for the faintest bin in column 6 is therefore 0.66 deg^2.

These combined counts are in good agreement with the Metcalfe et al. (2005) models, particularly at fainter magnitudes where the dispersion in the counts arising from cosmic variance appears to be small. We perform least squares fits between these counts and the pure luminosity evolution model; in the magnitude range $14< H <18$ we find a best fit normalisation of $1.095^{+0.035}_{-0.034}$, where 1.0 corresponds to the Metcalfe et al. (2005) normalisation shown in Fig. 3.8. Varying the fitting range does slightly alter the result; in the range $16< H <18$ we find a best fit normalisation of $1.061^{+0.048}_{-0.033}$ for example.

3.5 DISCUSSION

In the previous sections, bright H-band number counts from 2MASS were determined over the APM survey area ($\approx$ 4000 deg^2) and almost the entire survey ($|b| > 20°$, $\approx$27 000 deg^2), along with faint counts to $H = 18$ over a combined solid angle of 0.69 deg^2 applying a zeropoint consistent with 2MASS. The bright H-band number counts over the APM

survey area are extremely low ($\approx$ 25 per cent at $H = 13$) with respect to homogeneous predictions, and reproduce the form of the bright counts observed in the optical B-band (Busswell et al. 2004) and the near infrared K_s-band (previous chapter). Previous work has suggested that if due solely to local large-scale structure, these low counts would be at odds with the form of clustering expected in a ΛCDM cosmology. In addition, the bright H-band $|b| > 20°$ counts were also found to be low. In the following section, various possible causes for these low counts are examined.

3.5.1 Model Normalisation

The normalisation of number count models may be determined by fixing the predicted to the observed number of galaxies at faint magnitudes. Ideally, the magnitude range at which this is done should be bright enough to avoid large uncertainties in the evolutionary model while faint enough such that large fluctuations in the counts arising from cosmic variance are expected to be small. Near infrared wavelengths are expected to be insensitive to luminosity evolution at bright magnitudes, making the H-band particularly useful for such analysis. Of vital importance when determining the model normalisation is that when making comparisons between faint and bright counts, the zeropoints are consistent; an offset of a few tenths of a magnitude between the two, for example, would be enough to remove the observed anomaly in the bright counts over the APM survey area.

Applying the 2MASS zeropoint to the faint H-band data presented in this work and the LCIRS data (Chen et al. 2002), covering a combined solid angle of 0.69 deg^2, it is clear that a discrepancy between the bright and faint counts exists; the model normalisation used previously, which indicates low counts below $H = 14$ over the APM survey area (and for $|b| > 20°$), provides good agreement with the faint data. In fact, fixing the model to the faint counts implies a slightly higher normalisation. This agreement, as indicated by the errorbars in Fig. 3.8, suggests that the discrepancy between the bright and faint counts is not due to cosmic variance in the faint data. To remove the observed deficit in the APM survey area counts below $H = 14$ by renormalising the model, requires a deviation from the faint counts of 7.0σ using the best fit normalisation of $1.095^{+0.035}_{-0.034}$ (determined for $14 < H < 18$). Similarly, renormalising to the $|b| > 20°$ counts would require a deviation of 7.2σ from the faint data.

In addition, the model normalisation may also be scrutinised through comparison with redshift distributions. Fig. 3.2 shows the Metcalfe et al. (2005) pure luminosity evolution model compared with H-band $n(z)$ determined through a match between 2MASS and the

2dFGRS Northern and Southern declination strips. The model predictions appear to be consistent with the observations, with relatively homogeneous distributions beyond $z = 0.1$ (1 per cent and 8 per cent over-dense in the North and South respectively). Lowering the model normalisation to fit the bright 2MASS number counts would compromise this agreement and imply large over-densities beyond $z = 0.1$ (19 per cent and 27 per cent in the North and South respectively).

3.5.2 Galaxy Evolution

A change in amplitude therefore, cannot easily account for the discrepancy in the number counts at bright magnitudes. However, could an unexpected change in the slope of the number count model contribute? In section 3.2, we examined the consistency of the number counts at bright magnitudes with the underlying redshift distribution, assuming a model with insignificant levels of luminosity evolution at low redshift. The predictions derived from the observed $n(z)$ were in good agreement with the observed number counts indicating that luminosity evolution at low redshift is unlikely to have a significant impact on the form of the counts at bright magnitudes. This is supported by the consistency of the pure luminosity evolution model with the observed redshift distributions (Fig. 3.2); strong low redshift luminosity evolution produces a tail in the $n(z)$ which would imply large deficiencies at high redshift.

Could unexpectedly high levels of luminosity evolution at higher redshifts affect our interpretation of the bright counts? If the slope of the homogeneous prediction were to increase significantly above $H \approx 14$ from the evolutionary models considered in this paper, then the model normalisation could effectively be lowered into agreement with the bright counts. The problem with this is that the number counts beyond $H \approx 14$ are consistent with low levels of luminosity evolution to extremely faint magnitudes ($H \approx 26$). Models with significantly higher levels of luminosity evolution above $H \approx 14$ would therefore compromise this agreement.

Therefore, it appears that relatively low levels of luminosity evolution are consistent with number count observations to high redshifts. Also, recent evidence from the COMBO-17 survey, examining the evolution of early-type galaxies using nearly 5000 objects to $z \approx 1$ (Bell et al. 2004), suggests that density evolution will also not contribute; ϕ^* appears to decrease with redshift indicating that the number of objects on the red sequence increases with time, and so acts contrary to the low counts observed at bright magnitudes. This picture is supported by the K20 survey (Cimatti et al. 2002), which

includes redshifts for 480 galaxies to a mean depth of $\bar{z} \approx 0.7$ and a magnitude limit of $K_s = 20$ with high completeness. The resulting redshift distribution is consistent with low levels of luminosity *and* density evolution (Metcalfe et al. 2005).

In summary, significant levels of evolution are not expected in passive or star forming pure luminosity evolution models, although could occur through dynamical evolution. However, the pure luinosity evolution models of Metcalfe et al. (2005) fit the observed $H < 14$ $n(z)$ at $z > 0.1$; it is at lower redshifts that there are fluctuations. In addition, these models continue to fit the observed $n(z)$ at very high redshift and the number counts to extremely faint magnitudes ($K \approx 23$), suggesting that there is little need for evolution at $z \approx 1$, far less $z \lesssim 0.1$. Some combination of dynamical and luminosity evolution might be able to account for these observations; however it would require fine-tuning in order to fit both the steep counts at bright magnitudes and the unevolved $n(z)$ at low and high redshifts.

3.5.3 Photometry Issues & Completeness

The number counts shown in Figs. 3.7 and 3.8 show bright and faint counts with a consistent zeropoint applied. Photometry comparisons have been made using several hundred points sources matched at bright magnitudes. In order to check that the applied zeropoints are consistent with the *galaxy* samples, we also compare the 2MASS photometry with 24 matched galaxies in the CA field and WHDF and 16 in the LCIRS samples; we find that the mean offsets are -0.01 ± 0.04 and -0.32 ± 0.06, consistent with the zeropoints determined via the 2MASS point sources. The comparisons with the 2MASS point source catalogue (Figs. 3.5 and 3.6) also indicates that there is no evidence of scale error in either of the faint samples to $H \approx 16$.

Could the discrepancy between the bright and faint counts arise from an underestimation of the total flux of the galaxies? Recall that we make no correction to total magnitude for the faint data presented in this work; however, under-estimating the total flux in the faint data would only increase the observed deficit in the counts at bright magnitudes, if the model normalisation is adjusted to fit the faint counts. The good agreement between the point source and galaxy zeropoints suggests that the estimate for the total galaxy flux is comparable in the bright and faint data. At bright magnitudes, the 2MASS extrapolated H-band magnitudes are used. In the K_s-band, this magnitude estimator has been shown to be an excellent estimate of the total flux, through comparisons with the total K_s-band magnitude estimator of Jones et al. (2004) and the K-band photometry of

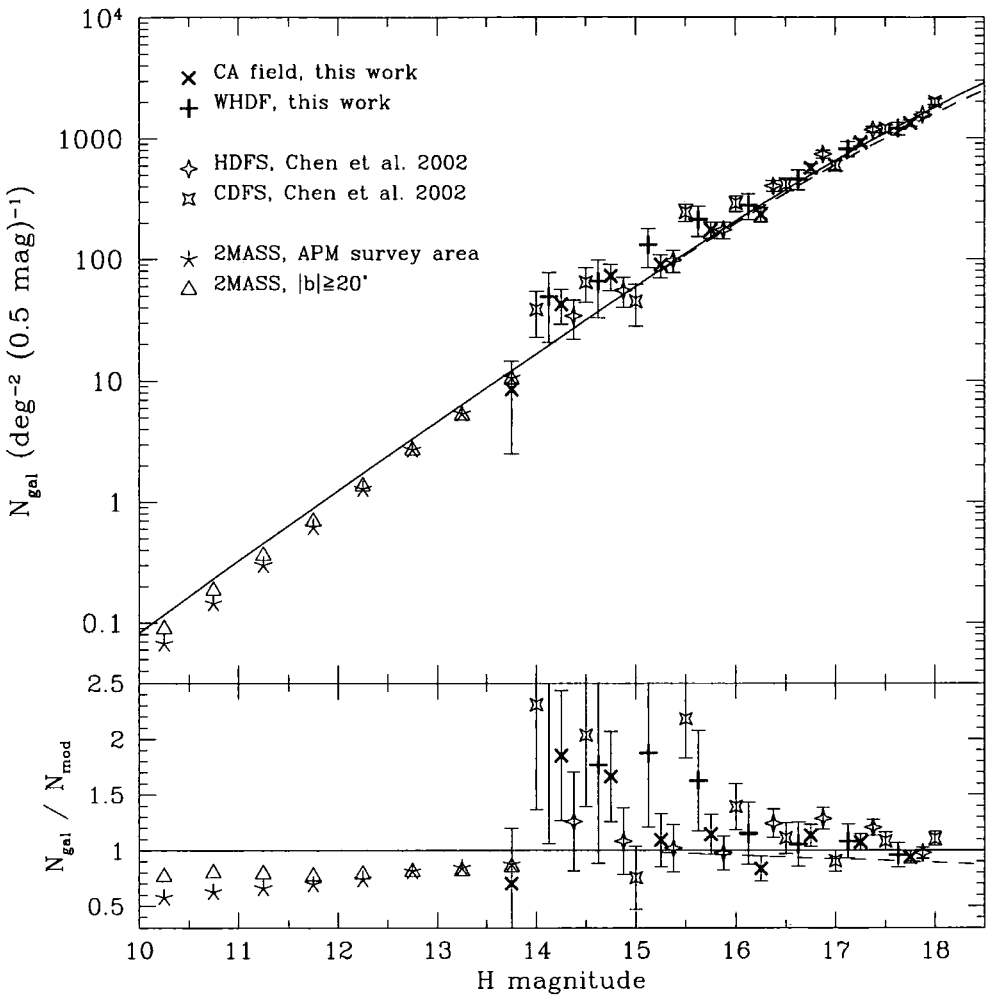


Figure 3.7: Here we show H-band galaxy number counts for the two separate fields observed in this work, the Calar Alto field (CA field; 0.27 deg^2) and the William Herschel Deep field (WHDF; 0.06 deg^2). We also show number counts determined for the two separate fields of the LCIRS (Chen et al. 2002) situated in the Hubble Deep Field South (HDFS; 0.24 deg^2) and Chandra Deep Field South (CDFS; 0.16 deg^2), subtracting 0.28 magnitudes in each case in order to bring the LCIRS and 2MASS zeropoints (and hence also the CA field and WHDF zeropoints) into agreement. We also show bright number counts extracted from 2MASS for the APM survey area and for $|b|$ >20° as shown in Fig. 3.3. The models are indicated as in Fig. 3.1. In the lower panel, we divide through by the pure luminosity evolution homogeneous prediction as in Figs. 3.3 and 3.4. At faint magnitudes, we indicate the Poisson uncertainty in each bin. We omit Poisson errors on the bright counts for clarity (see Fig. 3.4 for these). We discuss the uncertainty in the counts arising from cosmic variance in section 3.5.

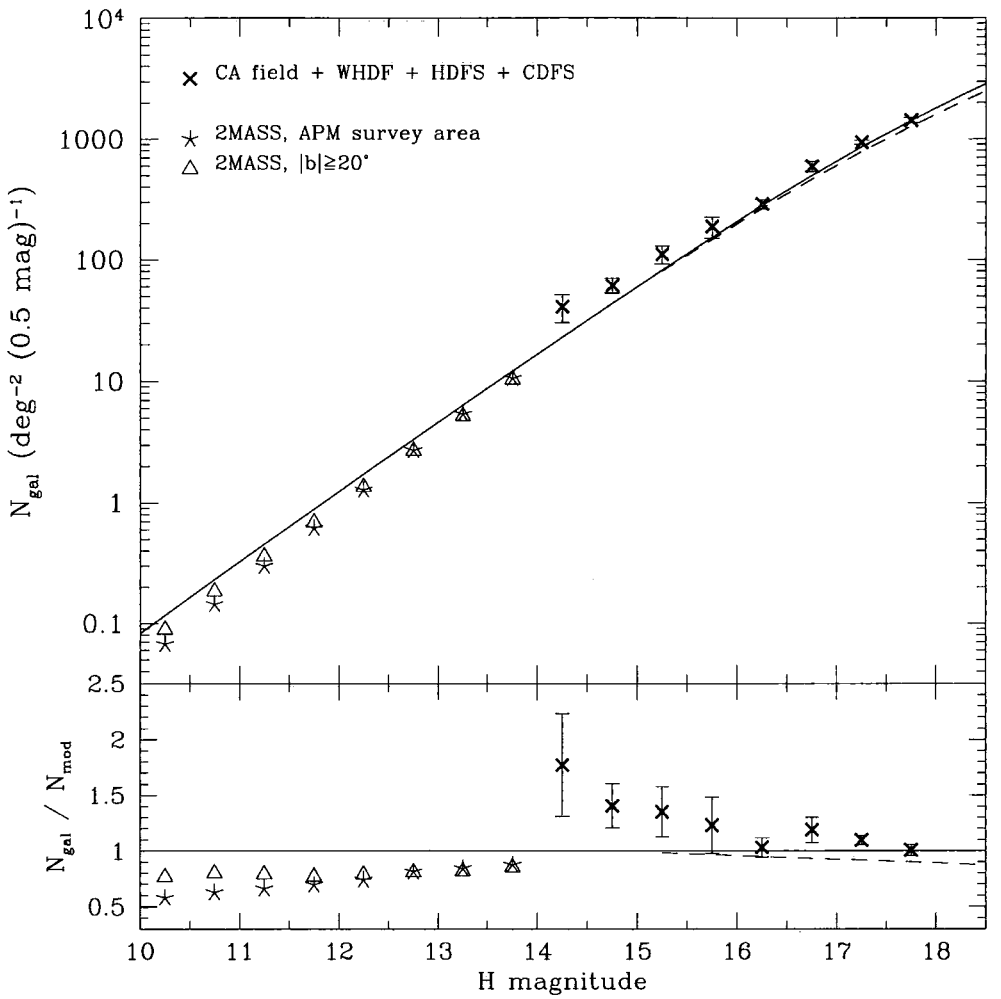


Figure 3.8: Here we show the faint H-band data from the two fields presented in this work (CA field and WHDF) and the two fields published by the LCIRS (HDFS and CDFS; Chen et al. 2002), applying a zeropoint to the LCIRS data consistent with the bright H-band 2MASS data (and hence the CA field and WHDF also), as shown in Fig. 3.7. The errorbars at faint magnitudes indicate the field-to-field error, weighted in order to account for the different solid angles of each field. Bright H-band counts extracted from 2MASS for the APM survey area and for $|b|$ >20° are shown as previously. In the lower panel, the counts are divided through by the pure luminosity evolution homogeneous prediction as before.

Loveday (2000).

Another possible contribution to the low counts could be high levels of incompleteness in the 2MASS survey. As with the possible systematic effects described previously, it is differing levels of completeness in the faint and bright data which would be important. The 2MASS literature quotes the extended source catalogue completeness as > 90 per cent (see the 2MASS website for example). Independently, Bell et al. (2003) suggest that the level of completeness is high ($\approx$99 per cent), determined via comparisons with the SDSS Early Data Release spectroscopic data and the 2dFGRS. The faint data presented in this work and the LCIRS data is likely to suffer less from incompleteness, as we cut well below the magnitude limit, are subject to lower levels of stellar confusion and suffer less from low resolution effects. Incompleteness in 2MASS will therefore affect the observed deficit in the bright counts at the < 10 per cent level, although the effect is likely to be lower (by a few per cent) due to incompleteness also in the faint catalogues and suggestions that the 2MASS extended source catalogue is fairly complete.

3.5.4 Large-Scale Structure

It appears therefore, that the observed deficiency in the bright counts may be significantly affected by incompleteness in the 2MASS extended source catalogue. However, the level to which other systematic effects such as the model normalisation, luminosity evolution and photometry issues appears to be small. The question then is – accounting for these various sources of error, are the deficiencies in the bright H-band counts over the APM survey area and for $|b| > 20^\circ$ still at odds with the expected fluctuations in the counts arising from local large-scale structure in a ΛCDM cosmology, as suggested in previous work (Busswell et al. 2004, Frith et al. 2003)?

We determine the expected fluctuations due to cosmic variance via ΛCDM mock 2MASS counts. These ΛCDM mock 2MASS catalogues are described in more detail in Frith et al. (2005a). To recap, we apply the 2MASS selection function to 27 virtually independent volumes of $r = 500\,h^{-1}$ Mpc formed from the $3000^3 h^{-3}\mathrm{Mpc}^3$ ΛCDM Hubble Volume simulation. This simulation has input parameters of $\Omega_m = 0.3$, $\Omega_b = 0.04$, $h = 0.7$ and $\sigma_8 = 0.9$ (Jenkins et al. 1998). The mean number density of the counts at the magnitude limit is set to that of the observed 2MASS density.

We are now in a position to estimate the significance of the observed bright H-band counts. We use the 1σ fluctuation in the counts expected in a ΛCDM cosmology (determined using the 2MASS mocks described above), which for the APM survey area is

Field	H_{lim}	Significance (no incompleteness correction)	Significance (assuming 10 per cent incompleteness)
APM	13.0	3.7σ	2.5σ
APM	14.0	4.2σ	2.4σ
$\|b\| > 20°$	13.0	6.1σ	3.8σ
$\|b\| > 20°$	14.0	6.8σ	4.0σ

Table 3.2: Here we show the significance of the H-band 2MASS counts extracted for the $\approx$4000 deg^2 APM survey area and for $|b| > 20°$, for $H < 13$ and $H < 14$. In each case we determine the expected cosmic variance using a ΛCDM form to the large-scale power determined via mocks constructed from the Hubble Volume simulation. In addition we use the best fit normalisation of the Metcalfe et al. (2005) pure luminosity evolution model determined at faint magnitudes of 1.095 and add the uncertainty on this ($\pm$3.1 per cent) in quadrature to the expected cosmic variance. In the third column use the observed counts as shown in Figs. 3.4, 3.7 and 3.8; in the fourth column we account for an upper limit on the incompleteness in the 2MASS extended source catalogue of 10 per cent; the level to which this will affect the significance is likely to be lower due to incompleteness in the faint data.

7.63 per cent (for $H < 13$) and 4.79 per cent (for $H < 14$), and for $|b| > 20°$ is 3.25 per cent (for $H < 13$) and 1.90 per cent (for $H < 14$). In addition we also take into account the uncertainty in the model normalisation; we use the best fit normalisation of the Metcalfe et al. (2005) pure luminosity evolution model (a factor of 1.095 above the Metcalfe et al. model) and add the uncertainty of ± 3.1 per cent derived from the faint H-band counts (presented in Fig. 3.8) in quadrature. Regarding the possible effect arising from survey incompleteness, we first assume that the level of incompleteness is comparable in the faint and bright data; the resulting significance for the APM survey area and $|b| > 20°$ bright counts are shown in column 3 of table 3.2. This represents an upper limit on the significance since we have effectively assumed that there is no difference in the incompleteness between the bright and faint datasets. In column 4 of table 3.2, we assume that there is a difference in the completeness levels in the faint and bright data of 10 per cent. This represents a lower limit on the significance (assuming that there are no further significant systematic effects), since we assume that the completeness of the 2MASS extended source catalogue is 90 per cent (the lower limit) and that there is no incompleteness in the faint data.

Therefore, assuming a ΛCDM cosmology, it appears that the observed counts over the APM survey area might be in line with a rare fluctuation in the local galaxy distribution. However, the counts over 66 per cent of the sky ($|b| > 20°$) suggest a deficiency in the counts that are at odds with ΛCDM, even accounting for a 10 per cent incompleteness effect and the measured uncertainty in the best fit model normalisation.

3.6 CONCLUSIONS

We have presented new H-band photometry over two fields with a combined solid angle of 0.30 deg^2 to $H \approx 19$. The zeropoint is chosen to match that of the 2MASS photometry at the bright end and is accurate to ± 0.01 magnitudes. In addition we have examined the faint H-band data of the LCIRS (Chen et al. 2002) which covers two fields with a combined solid angle of 0.39 deg^2 to $H \approx 20$. The zeropoint of this data appears to be offset from the 2MASS photometry by 0.28 ± 0.01 magnitudes. Applying a consistent zeropoint, the faint counts determined from the new data presented in this work and the LCIRS are in good agreement with the pure luminosity evolution model of Metcalfe et al. (2005), although with a best fit normalisation a factor of $1.095^{+0.035}_{-0.034}$ higher.

In contrast, the bright H-band counts extracted from 2MASS over the ≈ 4000 deg^2

APM survey area are low with respect to this model, in good agreement with previous results in the optical B-band (Busswell et al. 2004) and near infrared K_s-band (previous chapter). In addition, the counts extracted for almost the entire survey, covering 66 per cent of ths sky, are also low with a deficit of $15-20$ per cent to $H=14$. Importantly, this descrepancy does not appear to be due to zeropoint differences between the faint and bright data or uncertainty in the model normalisation set by the faint counts.

We have investigated various possible sources of systematic error which might affect this result: The counts are consistent with low levels of luminosity and density evolution, as predicted by the pure luminosity evolution model of Metcalfe et al. (2005), to extremely faint magnitudes (see Fig. 3.1). Also, the photometry appears to be consistent between the faint and bright galaxy data with the 2MASS zeropoint applied from comparisons between point sources. However, differing incompleteness in the bright and faint galaxy samples might have a significant impact; incompleteness in the 2MASS extended source catalogue is <10 per cent.

Finally, we determine the expected cosmic variance in bright number counts from ΛCDM mock 2MASS catalogues. Allowing for the model normalisation uncertainty determined from the faint counts, and using an upper limit on the effects arising from incompleteness in the 2MASS galaxy sample, the deficiency in the counts over the APM survey area represents a rare ($\approx$1 in 100) fluctuation in a ΛCDM cosmology. However, the low H-band counts for $|b|>20^{\circ}$ suggest that this deficiency might extend over the entire local galaxy distribution; allowing for incompleteness and the model normalisation uncertainty as before, this would represent a 4σ fluctuation ($<$1 in 10 000) in the local galaxy distribution, and therefore be at odds with the expected form of clustering at large scales expected in a ΛCDM cosmology.

In the next chapter, we investigate the form of galaxy clustering at extremely large scales; an excess of power over the ΛCDM prediction has previously been investigated in chapter 2 where the 2MASS angular power spectrum was first presented. We next investigate this statistic in greater detail and examine the implied cosmological constraints.

CHAPTER 4

THE 2MASS GALAXY ANGULAR POWER SPECTRUM

In the previous two chapters, the issue of the possible deficiency around the Southern Galactic Pole has been examined in the near infrared K_s and H-bands, and the level to which possible systematic effects might affect these counts has been constrained. It appears that the low counts cannot easily be accounted for by common sources of systematic error and that the dominant contribution to the observed deficiency is likely to arise from local large-scale structure. However, if this is the case, it would be at odds with the expected form of clustering at large scales in a ΛCDM cosmology. In chapter 2, the 2 Micron All Sky Survey (2MASS) angular power spectrum was presented for $K_s < 12.5$. We now examine the K_s-band angular power spectrum in greater detail and use it to constrain various cosmological parameters.

4.1 INTRODUCTION

The nature of galaxy fluctuations at extremely large scales ($r \lesssim 1000\ h^{-1}$ Mpc) is poorly constrained. Over the last decade, large galaxy surveys have constrained the form of the galaxy density field to a few hundred Megaparsecs. In recent years, large redshift surveys of both galaxies (e.g. Percival et al. 2001, Tegmark et al. 2004, Cole et al. 2005) and QSOs (e.g. Outram et al. 2003) have determined $P(k)$ at relatively small scales. Using the 2dF Galaxy Redshift Survey (2dFGRS), Cole et al. (2005) have constrained the form of galaxy density fluctuations to scales of $r \approx 300\, h^{-1}$ Mpc and the associated cosmological parameters to $\Omega_m h$=0.168±0.016 and Ω_b/Ω_m=0.185±0.046 (assuming h=0.72, a negligible neutrino mass, and a primordial $n_s = 1$ spectrum). However, determining the power spectrum through such redshift surveys suffers from large statistical uncertainty at large scales due to the relatively few objects available, as well as uncertainties arising from cosmic variance due to the relatively small volumes surveyed.

Using imaging surveys as opposed to redshift surveys provides a greater number of objects over larger solid angles. With angular power spectrum analysis of such surveys it is therefore possible to constrain the form of galaxy fluctuations to extremely large scales. However, since the clustering signal from a particular scale in real space is smeared over a range of angular scales, cosmological constraints through comparisons with linear theory predictions at smaller scales cannot be made; the departure from linearity at scales of $r \lesssim 40\,h^{-1}\,\mathrm{Mpc}$ (Percival et al. 2001) affects the clustering signal in the angular power spectrum over a wide range of scales. Nevertheless at large scales, where this effect is insignificant, angular power spectrum analysis represents one of the most effective probes of local large-scale structure.

Previously, the galaxy angular power spectrum has been determined for the Sloan Digital Sky Survey Early Data Release, the Edinburgh-Durham Southern Galaxy Catalogue, and a sample of IRAS galaxies (Tegmark et al. 2002, Huterer, Knox & Nichol 2001, Scharf et al. 1992, respectively), which along with the recent analyses of redshift surveys has constrained the form of galaxy fluctuations to $r \approx 300\,h^{-1}\,\mathrm{Mpc}$.

2MASS has recently been completed and provides near infrared photometry for over a million galaxies over the entire sky (Jarrett 2004, Jarrett et al. 2000); at the time of writing this dataset represents the largest all sky galaxy survey. 2MASS therefore represents a uniquely powerful probe of the local galaxy density field at large scales; the sample used in this analysis probes a volume approximately 5 times larger than the final 2dFGRS volume.

In this chapter, we use data from the 2MASS final release extended source catalogue to determine the K_s-band galaxy angular power spectrum with the aim of determining the form of the clustering of galaxies at extremely large scales, and constraining the shape and normalisation of the power spectrum. In section 4.2, we describe the 2MASS dataset and the magnitude estimator used. The method of analysis is outlined and the 2MASS angular power spectrum is determined and compared to mock power spectra in section 4.3. In section 4.4, we investigate various sources of systematic error. We determine constraints for various cosmological parameters in section 4.5. The conclusions follow in section 4.6.

4.2 DATA

We use the galaxy sample described in chapter 2. However, in this analysis the main sample is limited at K_s =13.5 (rather than K_s =12.5) and includes 518 576 galaxies above a galactic latitude of $|b|$ =20°; this represents a factor of ≈5 increase over the volume used in the final 2dFGRS power spectrum analysis (Cole et al. 2005). To illustrate the usefulness of the 2MASS data as a probe of the variance of local galaxy density fluctuations, in Fig. 4.1 we show the number of galaxy pairs as a function of separation for the final 2dFGRS release data and for a mock $|b|$ >20° 2MASS random catalogue constructed using the 2MASS selection function detailed in equation 2.1. It is clear that while the 2dFGRS contains pairs at larger separations, there are significantly more galaxy pairs to scales of $\approx 800\,h^{-1}$ Mpc in the 2MASS sample used in this work.

For reference later, the surface density of the 2MASS galaxy sample used is 19.1 deg^{-2}. We also use the shallower sample limited at K_s=12.5 and $|b|$ >20° as in chapter 2 which includes 124 264 galaxies, for which the surface density is 4.58 deg^{-2}.

Similarly, we use the unbiased K_s <12.5 mock 2MASS catalogues derived from the ΛCDM Hubble Volume simulation described in the previous chapter. In addition we use a deeper set of K_s <13.5 mock catalogues with mean redshift $\bar{z}$=0.074, where $\bar{z}$ is determined from the 2MASS-2dFGRS matched sample described in chapter 2. As before, these mocks are produced such that the number density of objects matches that of the 2MASS $|b|$ >20° galaxy sample. Due to the volume of the 27 mock 2MASS catalogues, the selection function is artificially truncated for the K_s <13.5 mocks at z ≈0.156. However, this has a negligible effect on the work in this chapter; at this redshift, ≈95 per cent of the galaxies are sampled for K_s <13.5.

4.3 THE 2MASS ANGULAR POWER SPECTRUM

4.3.1 ESTIMATING THE POWER SPECTRUM

Following the usual method (e.g. Peebles 1973, Peebles & Hauser 1973, Peebles & Hauser 1974, Scharf et al. 1992), the angular power is estimated through a spherical harmonic expansion of the surface density of galaxies. The coefficients of this expansion are determined over the observed solid angle Ω_{obs}:

$$a_l^m = \sum_{N_{gal}} Y_l^m(\theta,\phi) - \mathcal{N}\int_{\Omega_{obs}} Y_l^m(\theta,\phi)d\Omega \tag{4.1}$$

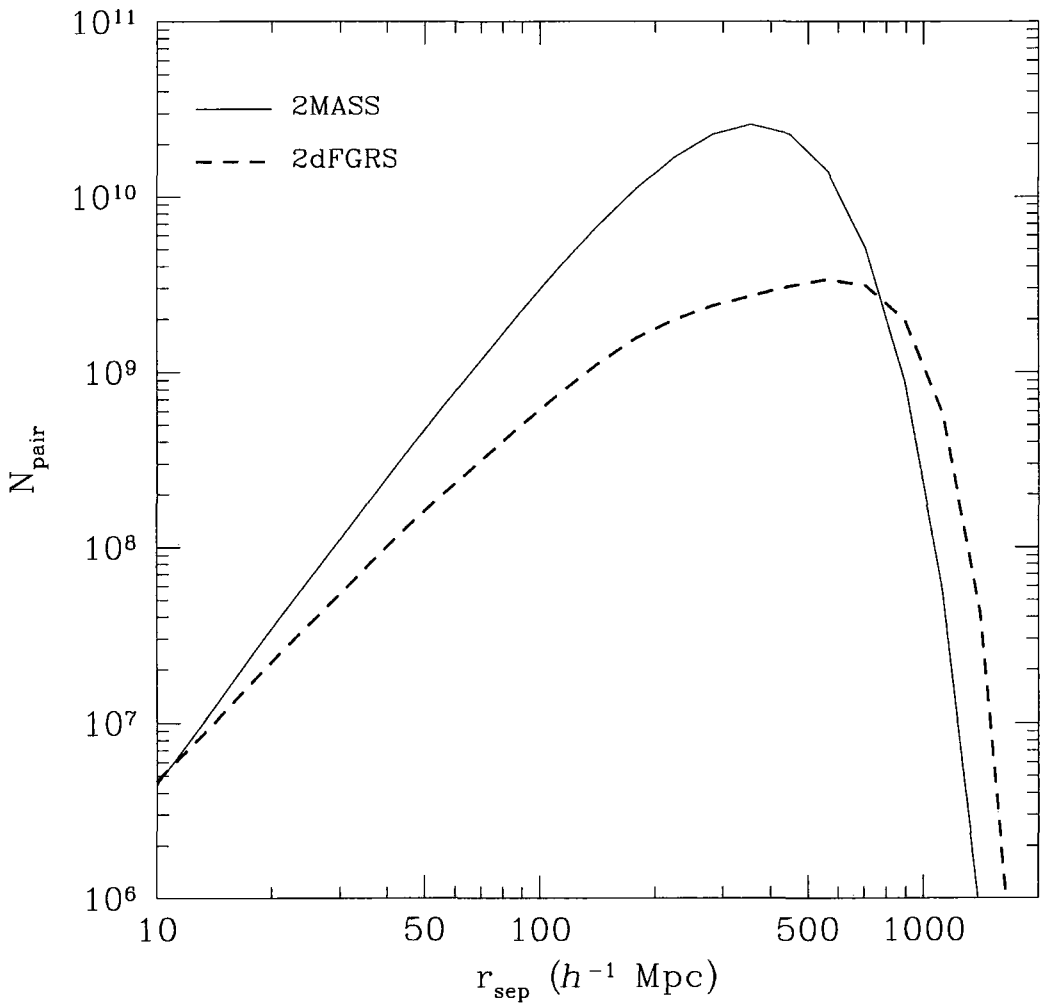


Figure 4.1: Number of galaxy pairs as a function of separation for a mock $|b|$ >20° 2MASS random catalogue constructed using the 2MASS selection function detailed in equation 2.1 and the final 2dFGRS release data. The counts are divided into 10 bins per decade.

where $\mathcal{N}=N_{gal}/\Omega_{obs}$ is the observed number of galaxies per steradian. The angular power is then determined:

$$\mathcal{Z}_l = \frac{1}{2l+1}\sum_m \frac{|a_l^m|^2}{\mathcal{J}_l^m} \tag{4.2}$$

where,

$$\mathcal{J}_l^m = \int_{\Omega_{obs}} |Y_l^m(\theta,\phi)|^2 d\Omega. \tag{4.3}$$

This term accounts for the survey window; throughout most of this analysis we use a $|b|$ >20° cut. The angular power is then normalised, subtracting the expected shot noise contribution:

$$C_l = \frac{\mathcal{Z}_l}{\mathcal{N}} - 1 \tag{4.4}$$

such that C_l=0 corresponds to a random distribution.

4.3.2 Fitting to the Power Spectrum

In order to compare the angular power spectrum with cosmological predictions, we determine an expected form for the angular power spectrum for various cosmological parameters using the relation between the three and two-dimensional power spectra:

$$|a_l^m|^2 = \frac{2}{\pi}\int \left(\int r^2\Phi(r) j_l(kr) dr\right)^2 k^2 P(k) dk + \mathcal{N} \tag{4.5}$$

(Scharf et al. 1992, Tegmark et al. 2002, Huterer, Knox & Nichol 2001), which we normalise as before. Here, $\Phi(r)$ is the 2MASS selection function, and j_l is a spherical Bessel function. The 2MASS selection function is determined using equation 2.1.

We use the transfer function fitting formulae of Eisenstein & Hu (1998) to obtain a linear theory prediction for the power spectrum of matter, $P(k)$, with input parameters for the matter, vacuum, baryon and neutrino densities (Ω_m, Ω_Λ, Ω_b and Ω_ν), h (such that $H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) and matter power spectrum normalisation (σ_8). We also employ a linear biasing scheme such that $P_{gal}(k)=b^2 P_{matter}(k)$ to provide a linear prediction for the galaxy $P(k)$. This is then transformed to a galaxy angular power spectrum prediction using the spherical Bessel function transform in equation 4.5.

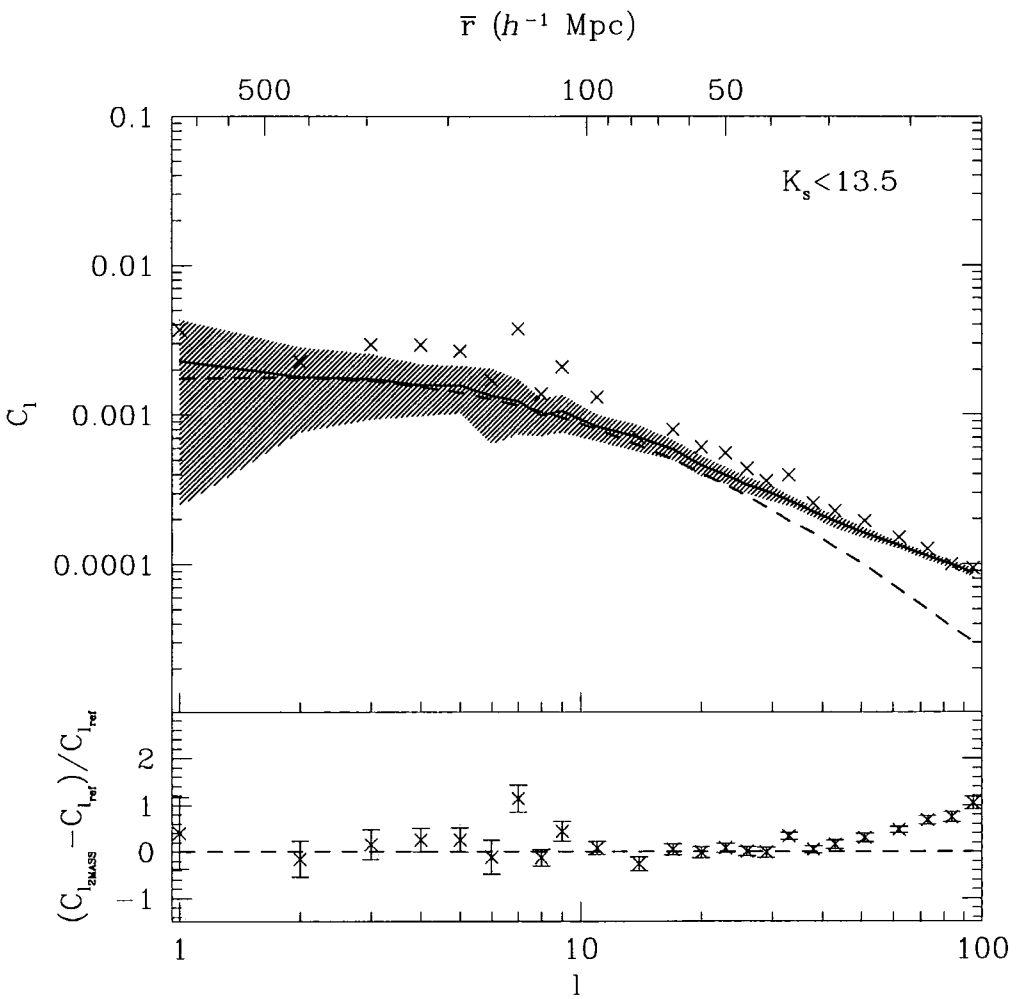


Figure 4.2: The $|b|$ >20° 2MASS galaxy angular power spectrum for 518 576 K_s < 13.5 galaxies. The crosses indicate the 2MASS datapoints with the shaded region and solid line indicating the 1σ spread and mean power spectrum of the 27 mock unbiased 2MASS catalogues constructed from the ΛCDM Hubble Volume mock catalogue as described in section 4.2. An unbiased linear theory model corresponding to the Hubble Volume mock catalogue input parameters of Ω_m=0.3, Ω_b=0.04, h=0.7 and σ_8=0.9 is indicated by the dashed line. In the lower panel we show the fractional deviation of the 2MASS power spectrum from this model applying the best fit power spectrum normalisation, $(\sigma_8 b_K)^2$=1.36, (determined in section 4.5 for the K_s <13.5 sample) to the linear prediction, with errors taken from the mock 2MASS 1σ spread. In addition we indicate the approximate mean distance scale probed by the data on the top x-axis.

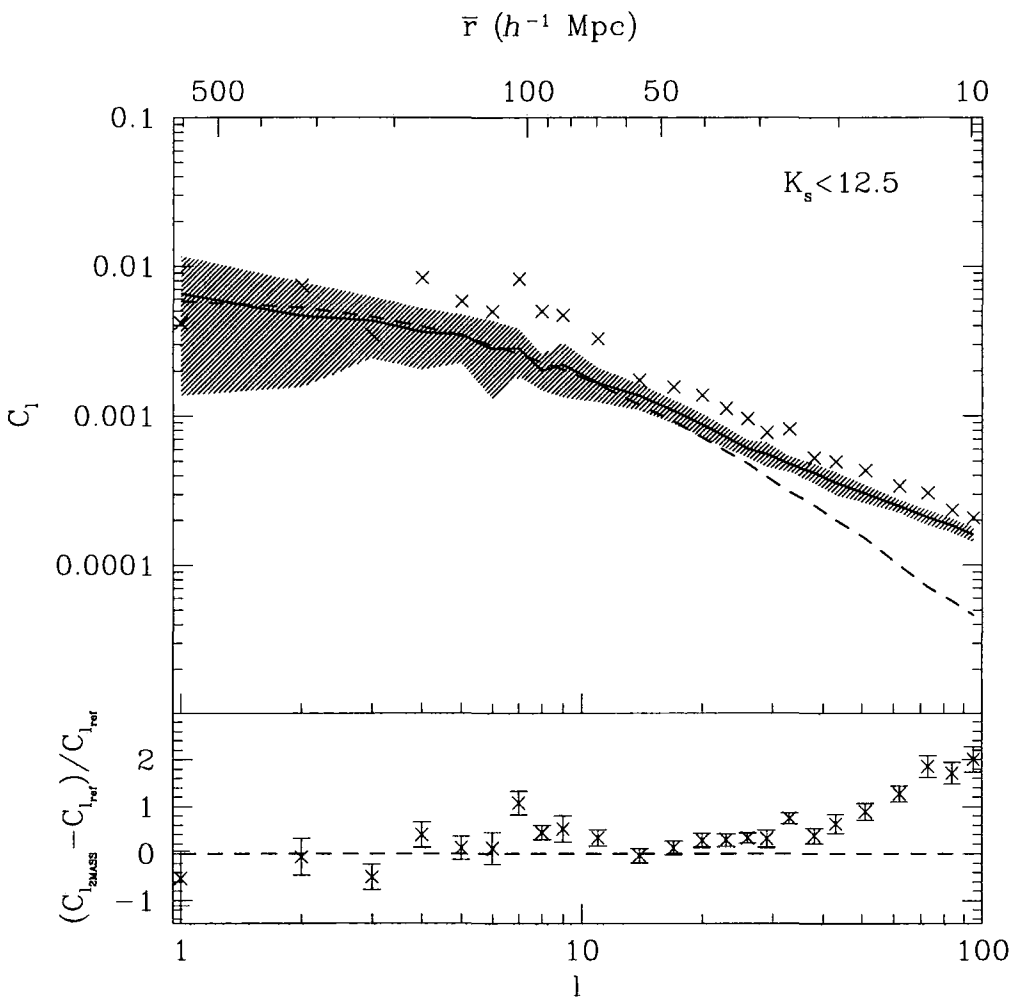


Figure 4.3: The $|b|$ >20° 2MASS galaxy angular power spectrum for 124 264 K_s < 12.5 galaxies. As in Fig. 4.2, the crosses indicate the 2MASS datapoints with the shaded region and solid line indicating the 1σ spread and mean power spectrum of the 27 mock unbiased 2MASS catalogues. We also show an unbiased linear theory model corresponding to the Hubble Volume mock catalogue input parameters. The lower panel show the fractional deviation of the K_s < 12.5 2MASS power spectrum from the linear prediction (applying the best fit power spectrum normalisation of $(\sigma_8 b_K)^2$=1.36 as before).

4.3.3 RESULTS

The angular power spectrum for 518 576 $K_s < 13.5$, $|b| > 20°$ 2MASS galaxies is presented in Fig. 4.2, determined through a spherical harmonic expansion of the galaxy number density as described in section 4.3.1. In order to determine the expected scatter due to cosmic variance we determine the angular power spectrum for the 27 unbiased mock 2MASS catalogues constructed from the ΛCDM Hubble Volume simulation described in section 2.2; the mean and 1σ spread are indicated by the solid line and shaded region. On the top x-axis we also indicate the approximate distance scale probed by the angular power spectrum at the mean depth of the sample determined from the 27 mock 2MASS catalogues. At the very smallest l-modes, the K_s <13.5 power spectrum probes scales of $\gtrsim 500\,h^{-1}$ Mpc.

We have also calculated the linear prediction corresponding to the ΛCDM Hubble Volume input parameters (Ω_m=0.3, Ω_Λ=0.7, Ω_b=0.04, h=0.7, σ_8=0.9 and Ω_ν=0) through a spherical Bessel function transform of the three-dimensional power spectrum as described in section 4.2; this is indicated for a bias of 1.0 by the dashed line. The linear model and the mean mock 2MASS power spectrum are in good agreement at large scales. At smaller angular scales (l >30) the effects of non-linear clustering become significant.

In order to verify whether the form and scatter of the mock power spectra, which we later use to estimate the error on the observed angular power spectrum, is consistent with the data, we perform a χ^2 fit between the two. We marginalise over the normalisation of the mean mock angular power spectrum and use the binning as shown in order to reduce the covariance to insignificant levels. We assume that the spread in the mock power spectra is independent of normalisation, i.e. we apply the same spread determined for the unbiased mock power spectra to the observed angular power spectrum. In this particular case, this is likely to provide an optimistic view of the observed errors since we are not shot noise limited. In this scenario, the errors are likely to be independent of the power spectrum amplitude; on the other hand, if the observed power spectrum is cosmic variance limited the errors scale with model normalisation (see Feldman, Kaiser & Peacock (1994) for further discussion on this point). We investigate the impact of this assumption on the associated cosmological constraints in section 4.5. First, we perform a χ^2 fit over the full angular range $1 \leq l \leq 100$ between the K_s <13.5 2MASS galaxy angular power spectrum and the mean mock 2MASS power spectrum; we find that $\chi^2/d.o.f.$=3.0. Limiting the angular range to scales which are not significantly affected by non-linear

clustering ($l \geq 30$), the form of the mock power spectra are in better agreement with the observed 2MASS galaxy angular power spectrum, with $\chi^2/d.o.f.=2.0$.

The form of the 2MASS angular power spectrum is therefore in good agreement with the ΛCDM prediction in the linear regime, although it is clear from Fig. 4.2 that there is some difference in slope at small scales. Assuming the validity of the prediction, this is due either to scale-dependent bias in the non-linear regime or resolution effects in the Hubble Volume simulation. Consistency with the ΛCDM prediction in the linear regime, of interest in this work, is confirmed through a comparison (in the lower panel) with the linear prediction for the ΛCDM Hubble Volume simulation input parameters applying a scale-independent bias to match the normalisation of the observed power spectrum at large scales (see section 4.5).

4.4 Systematic Errors

4.4.1 Magnitude Limits

Before turning to the cosmological constraints inferred from the 2MASS galaxy angular power spectrum it is important to verify that the results are robust and not significantly affected by potential sources of systematic error. While the 2MASS catalogue is >98 per cent reliable for $|b| > 20°$, $K_s < 13.5$ galaxies (Jarrett et al. 2000) and 99 per cent complete for $|b| > 30°$, $12.0 < K_s < 13.7$ galaxies (Maller et al. 2005), we wish to verify that the angular power spectrum is robust to changes in the magnitude limit, and is not adversely affected by variable incompleteness or reliability at faint magnitudes or scale errors in the photometry.

Figs. 4.2 and 4.3 show the 2MASS galaxy angular power spectrum as a function of imposed magnitude limit. The shape and normalisation of the power spectrum, with respect to both the linear model and the mean mock 2MASS power spectrum, are remarkably robust to changes in the magnitude limit. The departure of the linear model from the observed power spectrum occurs at larger angular scales with the shallower magnitude limit due to the reduced mean depth of the sample. For this reason also, the mock 2MASS power spectrum is more significantly distorted at the very smallest scales by resolution affects resulting in a slightly steeper slope at $l \gtrsim 70$.

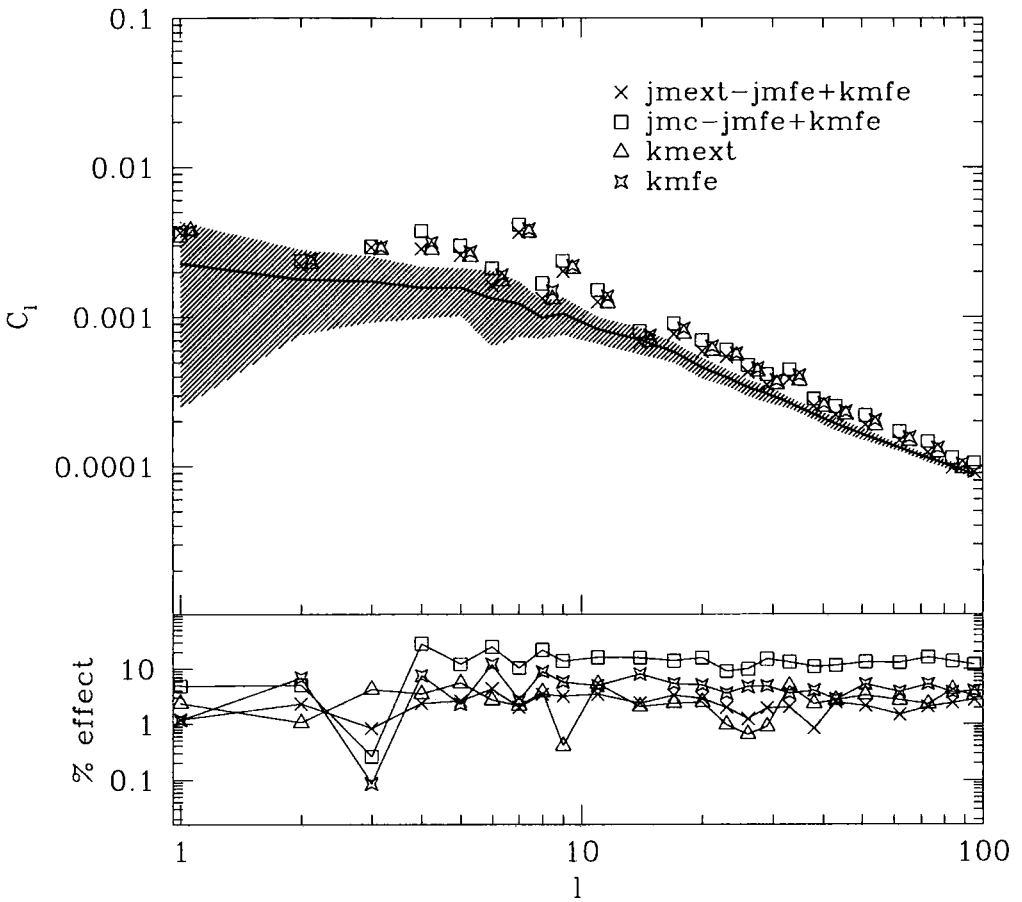


Figure 4.4: The $|b|$ >20°, K_s <13.5 extinction-corrected 2MASS galaxy angular power spectra for the four magnitude estimators shown in Fig. 2.1 of chapter 2 using the raw magnitude estimate (i.e. without a zero-point correction to the Loveday (2000) photometry). The 1σ spread and mean 2MASS mock power spectrum are shown as in Fig. 4.2. The lower panels indicate the effect of each magnitude estimator on the resulting power spectrum compared to the colour-corrected J-band extrapolated magnitude estimator (with the zero-point correction) used in Fig. 4.2 In the upper panels we have displaced the *kmext* and $kmfe$ datapoints for clarity.

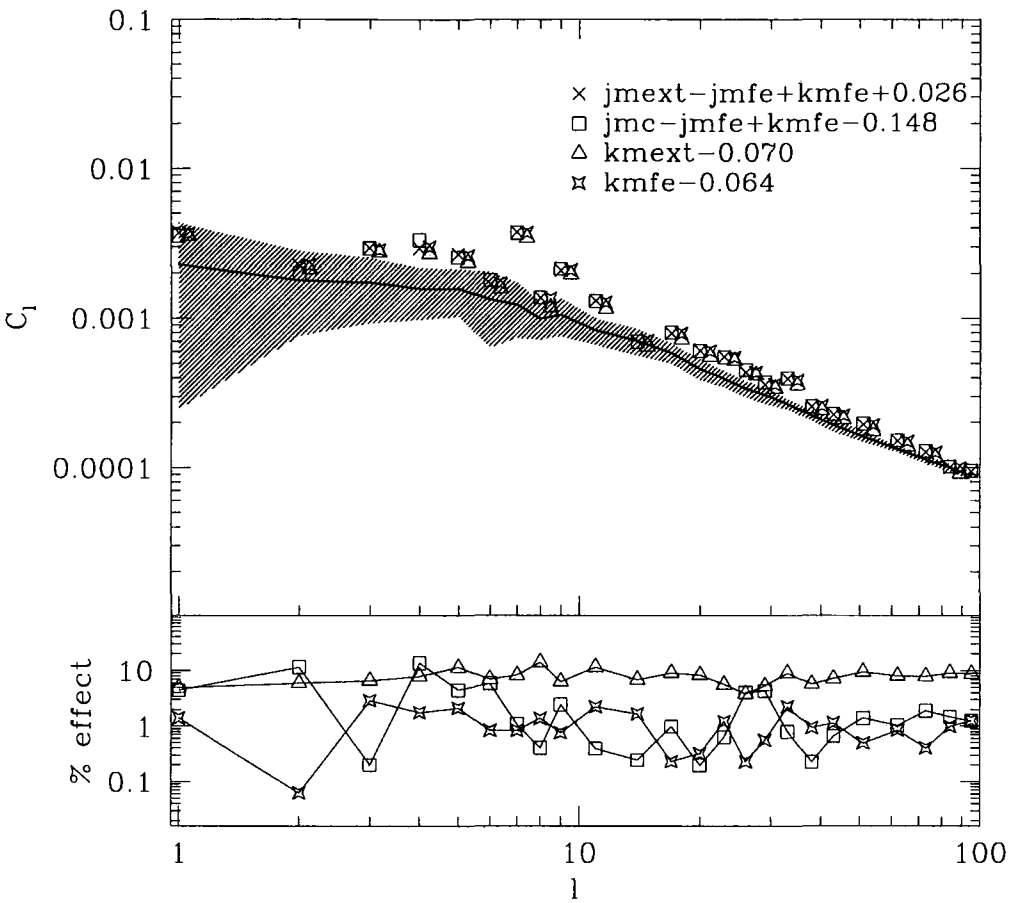


Figure 4.5: The $|b|$ >20°, K_s <13.5 extinction-corrected 2MASS galaxy angular power spectra for the four magnitude estimators shown in Fig. 2.1 of chapter 2, this time using a zero-point correction to account for the offset determined with respect to the Loveday (2000) photometry. The 1σ spread and mean 2MASS mock power spectrum are shown as in Fig. 4.2. The lower panels indicate the effect of each magnitude estimator on the resulting power spectrum compared to the colour-corrected J-band extrapolated magnitude estimator (with the zero-point correction) used in Fig. 4.2 and also indicated here by the black crosses. In the upper panels we have displaced the *kmext* and *kmfe* datapoints for clarity.

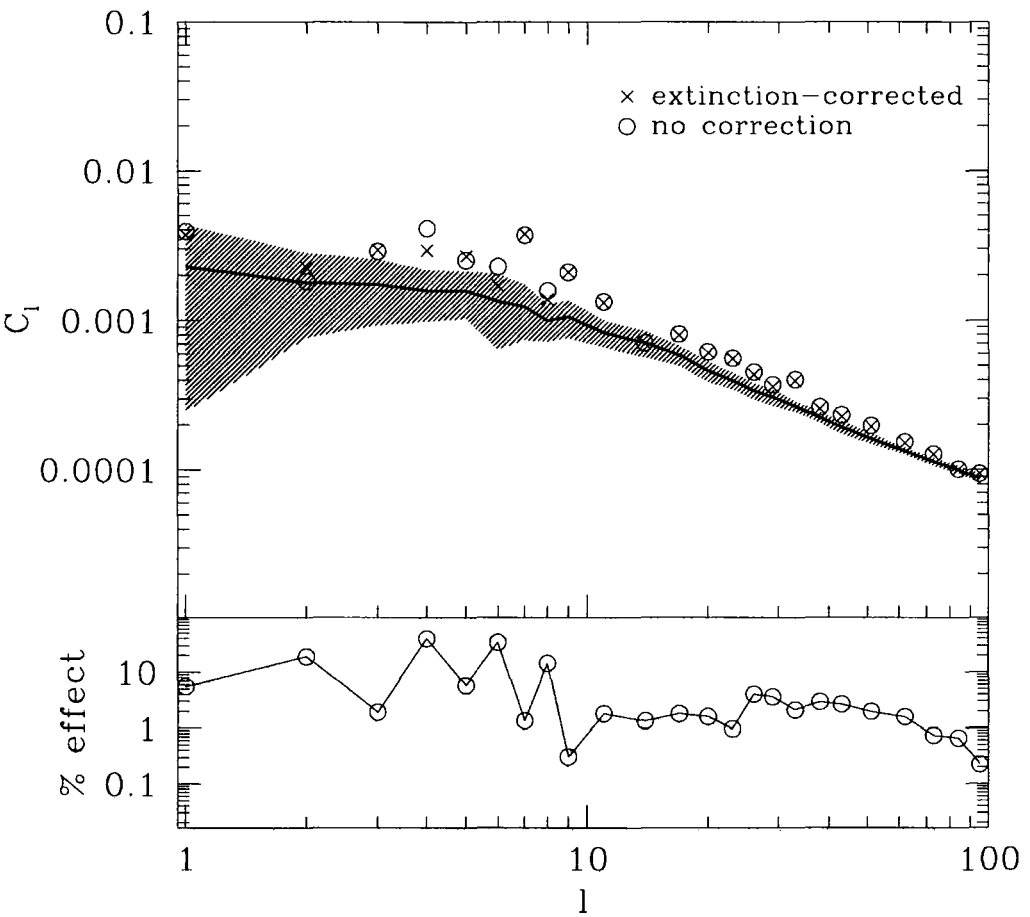


Figure 4.6: The $|b|$ >20°, K_s <13.5 2MASS galaxy angular power spectra including no extinction correction, and as previously an extinction correction derived from the Schlegel et al. (1998) dust maps. The mock 2MASS mean angular power spectrum and 1σ spread are shown as before. In the lower panel we indicate the effect of this correction on the power spectrum through a comparison with the corrected sample (indicated by the crosses in the upper panel and as shown in Fig. 4.2).

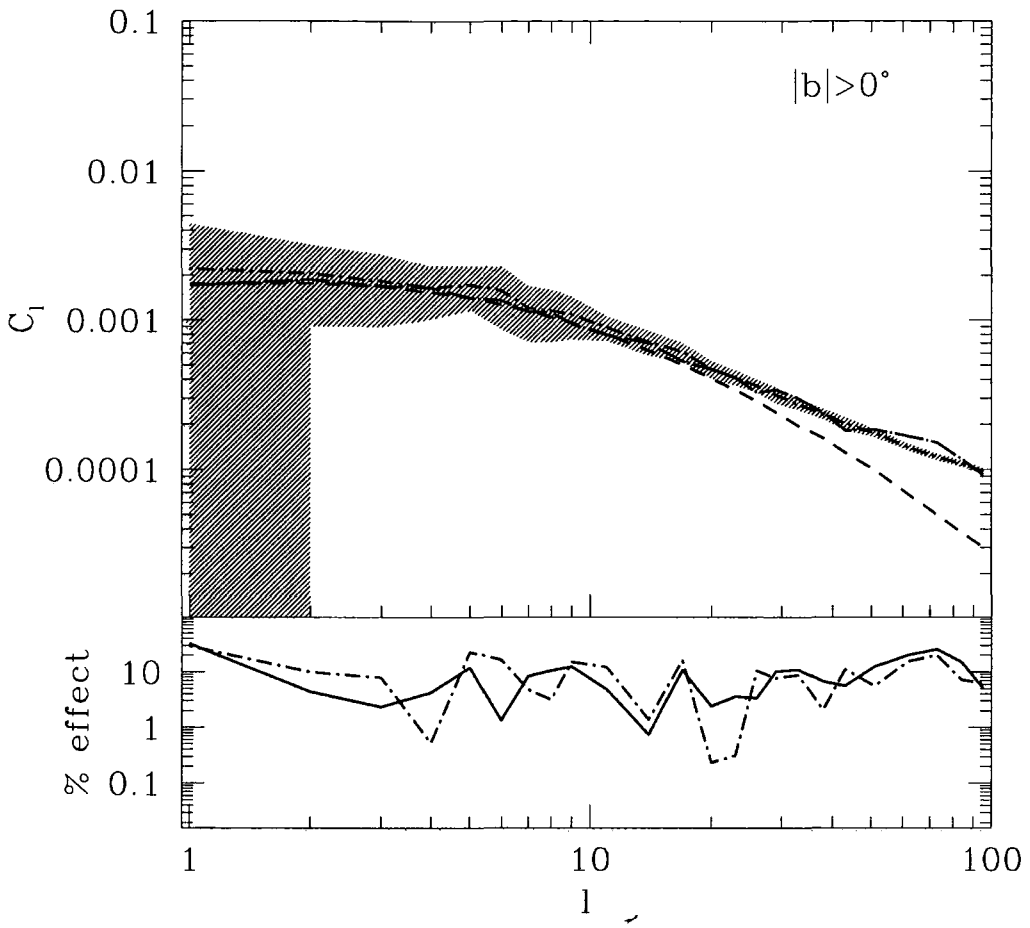


Figure 4.7: The $|b|$ >0°, K_s <13.5 mean power spectrum and 1σ spread determined from the 27 mock 2MASS catalogues (solid line and shaded region). As in Fig. 4.2 the dashed line indicates the expected linear trend for the ΛCDM Hubble Volume mock input parameters of Ω_m=0.3, Ω_b=0.04, h=0.7 and σ_8=0.9. As a consistency check, we also show the ΛCDM Hubble Volume angular power spectrum (large dot-dashed line) calculated via the directly determined ΛCDM Hubble Volume $P(k)$ (Carlton Baugh - priv. comm.) transformed to the angular power spectrum as described in section 4.2. In the lower panel we compare this prediction with the $|b|$ >0° (dot-dashed line) and $|b|$ >20° (solid line) mean mock 2MASS power spectra.

4.4.2 MAGNITUDE ESTIMATOR

Throughout this work, we estimate the K_s-band magnitudes using the J-band extrapolated magnitudes colour-corrected using the K_s and J-band fiducial elliptical Kron magnitudes, as this results in a smaller zero-point offset and scale error when compared to the more accurate K-band photometry of Loveday (2000), as in chapter 2. We wish to investigate the effect on the power spectrum by the choice of magnitude estimator; in Figs. 4.4 and 4.5 we compare the power spectra for the four magnitude estimators presented in Fig. 2.1 of chapter 2 with and without respectively the associated correction to the Loveday (2000) zero-point.

The power spectrum is robust to changes in the magnitude estimate and zero-point at the $\lesssim 10$ per cent level. This is due to the fact that the change in the depth of the survey due to differences in the magnitude limit and scale error effects are insignificant. In any case, this potential systematic does not significantly affect the associated cosmological constraints examined in section 4.5, since any shift in the zeropoint is modelled in the linear theory prediction through a change in the associated mean redshift of the galaxy sample (see equations 4.5 and 2.1).

4.4.3 EXTINCTION

While the level of extinction in the K_s-band is low and the 2MASS magnitudes have been corrected using the Schlegel et al. (1998) dust maps, it is useful to examine the potential level of systematic error introduced by extinction. Fig. 4.6 shows the 2MASS galaxy angular power spectrum with and without correction for extinction. In this extreme case, the effect of removing the dust correction to the magnitude estimate is at the $\lesssim 10$ per cent level at large scales and $\lesssim 1$ per cent for $l \gtrsim 10$. The effect on the recovered cosmological parameters of Γ_{eff} and $(\sigma_8 b_K)^2$ for instance (see section 4.5) is -0.007 and $+0.007$ respectively, considerably less than a 1σ shift.

4.4.4 THE WINDOW FUNCTION

Throughout this paper a $|b| > 20°$ galactic latitude cut is applied in order to avoid the high levels of extinction and stellar contamination in the zone of avoidance. We wish to determine the level of any systematic effect on the spread of the Hubble Volume mock power spectra (and therefore our interpretation of the statistical uncertainty) introduced by the window function. In Fig. 4.7 the 27 mock 2MASS power spectra and corresponding

linear theory model for the ΛCDM Hubble Volume input parameters are shown with no galactic latitude cut. Neither the shape nor the spread of the power spectra are significantly altered. The effect of the window function on the angular power spectrum is $\lesssim 5$ per cent at all scales.

In order to check the consistency of our results we provide a further verification of the mock 2MASS power spectrum results through a comparison with the transform of the directly determined ΛCDM Hubble Volume simulation $P(k)$ (Carlton Baugh - priv. comm.). There is excellent agreement with both the $|b|$ >0° and $|b|$ >20° mean mock 2MASS power spectra.

4.5 COSMOLOGICAL CONSTRAINTS

Using the 2MASS galaxy angular power spectrum we have determined the form of the galaxy density field at extremely large scales and verified that it is not significantly affected by common sources of systematic error; the $< 1\sigma$ shift on the recovered cosmological parameters arising from the extreme case of removing the extinction correction to the 2MASS K_s-band magnitudes represents the largest effect from the systematics examined in section 4.4. We now wish to determine the associated cosmological constraints.

Using the Eisenstein & Hu (1998) transfer function fitting formulae we have determined linear theory predictions for the three-dimensional power spectrum of matter, $P(k)$, using input parameters of Ω_m, Ω_Λ, Ω_b, h and matter power spectrum normalisation, σ_8; in the subsequent analysis we assume a negligible neutrino mass density, a primordial n_s=1 spectrum and $\Omega_\Lambda = 1 - \Omega_m$. We form galaxy angular power spectrum predictions using the spherical Bessel function transform described in section 4.2 and a linear biasing scheme.

First, we perform fits to the galaxy power spectrum shape and normalisation. Assuming a CDM cosmology, the power spectrum may be defined through a parameterisation of the shape

$$\Gamma_{\rm eff} = \Omega_m h \exp(-\Omega_b(1 + \sqrt{2h}/\Omega_m)) \tag{4.6}$$

(Sugiyama 1995), and a normalisation, which for galaxy power spectra may be parameterised through the galaxy bias and σ_8. Varying the parameters Ω_m, Ω_b and h in this equation, we constrain $\Gamma_{\rm eff}$ and $(\sigma_8 b_K)^2$ using a grid of 200×800 models between $0.1 \leq \Gamma_{\rm eff} \leq 0.3$ and $0.0 \leq (\sigma_8 b_K)^2 \leq 8.0$ respectively. We perform least squares fits to the

$|b|$ >20°, K_s <13.5 angular power spectrum as shown in Fig. 4.2 at scales of $l \leq 30$ (binned as shown to reduce the covariance to insignificant levels); beyond $l \approx 30$ the angular power spectrum begins to be significantly affected by non-linear effects.

We take the spread determined from the 27 mock 2MASS angular power spectra in order to estimate the errors on the 2MASS datapoints, assuming that the uncertainty remains the same for a biased as for an unbiased distribution (as in section 4.3). In doing this, we assume that the ΛCDM Hubble Volume mock catalogue provides an accurate description of the local galaxy distribution at large scales and that the associated uncertainty in the datapoints is realistic. However, since these errors are valid only in an unbiased ΛCDM cosmology we are required to make assumptions as to the nature of the cosmic variance in the various other cosmologies scrutinised in these fits. Here we assume that the errors are independent of cosmology and power spectrum normalisation; the likely impact of this assumption is examined below. We find that:

$$\Gamma_{\rm eff} = 0.14 \pm 0.02$$

and

$$(\sigma_8 b_K)^2 = 1.36 \pm 0.10$$

marginalising over the normalisation and power spectrum shape respectively. The associated confidence regions are indicated by the filled contours in Fig. 4.8.

This value of $\Gamma_{\rm eff}$ is in excellent agreement with the 2dFGRS fit (Percival et al. 2001) of $\Gamma_{\rm eff} = 0.18 \pm 0.04$ (for h=0.7) and the WMAP value (Spergel et al. 2003) of $\Gamma_{\rm eff} = 0.15 \pm 0.01$ (for n_s=0.99). However, our value is slightly higher than the Maller et al. (2005) result which constrains $\Gamma_{\rm eff} = 0.116 \pm 0.009$ at 95 per cent confidence using a measure of the three dimensional K_s-band galaxy power spectrum via an inversion of the 2MASS angular correlation function.

Our constraint on the K_s-band galaxy power spectrum normalisation of $(\sigma_8 b_K)^2 = 1.36 \pm 0.10$ is also slightly higher than the Maller et al. (2005) result of $\sigma_8 b_K$=1.0±0.1. Using the WMAP-2dFGRS best fit matter power spectrum normalisation of $\sigma_8 = 0.84 \pm 0.04$ (Bennett et al. 2003), we constrain the K_s-band bias to $b_K = 1.39 \pm 0.12$, in reasonable agreement with previous measurements determined from the 2MASS clustering dipole of $b_K = 1.37 \pm 0.34$ (Maller et al. 2003) and the 2MASS angular correlation function analysis

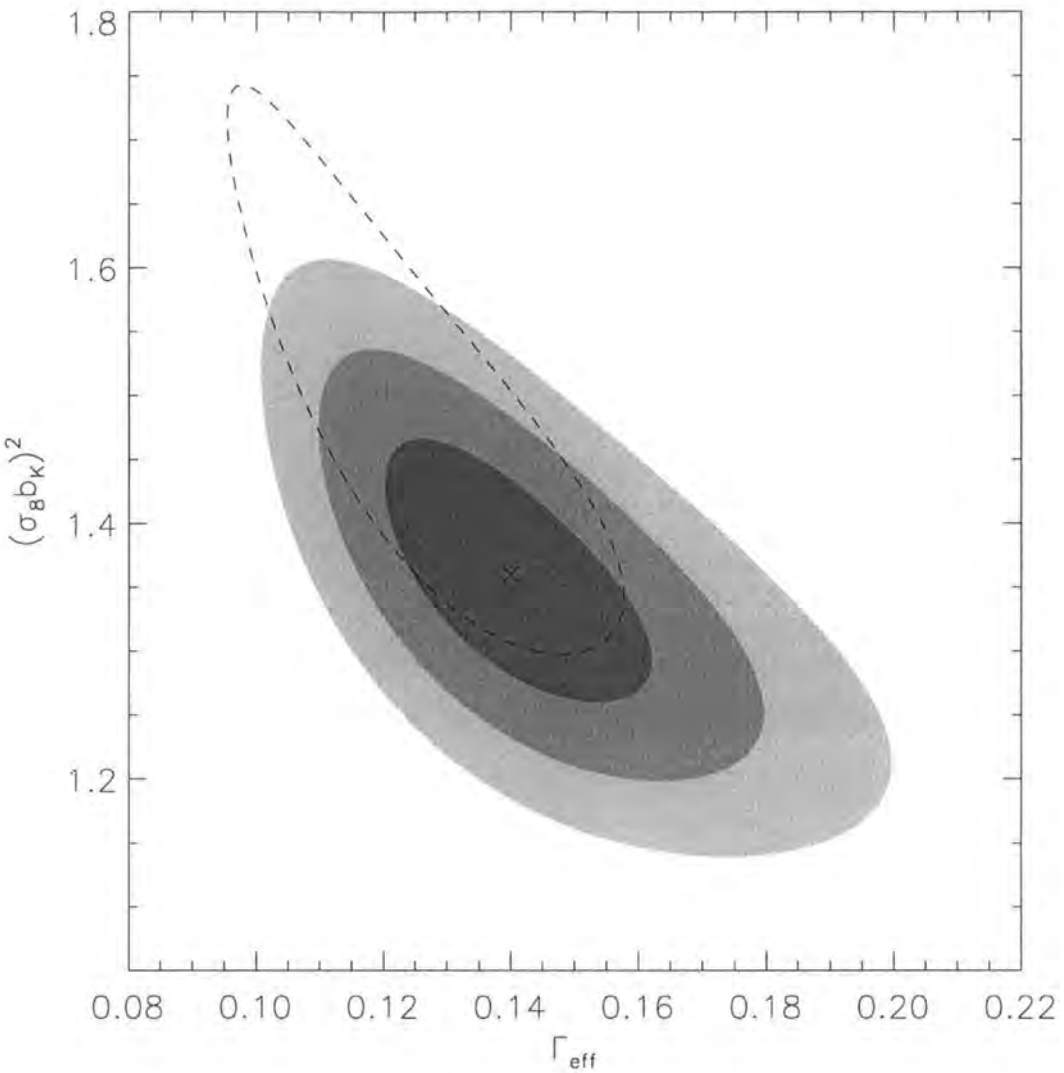


Figure 4.8: Filled contours representing the 1σ, 2σ and 3σ confidence regions for the galaxy power spectrum shape and normalisation determined from χ^2 fits to the 2MASS $|b| > 20°$ $K_s < 13.5$ galaxy angular power spectrum in the range $l \leq 30$. The cross indicates the best fit parameters of $\Gamma_{\rm eff}=0.14$ and $(\sigma_8 b_K)^2=1.36$. We also show the 1σ confidence region for the 2MASS result as above where we use errors which scale with the model power spectrum normalisation (dashed line).

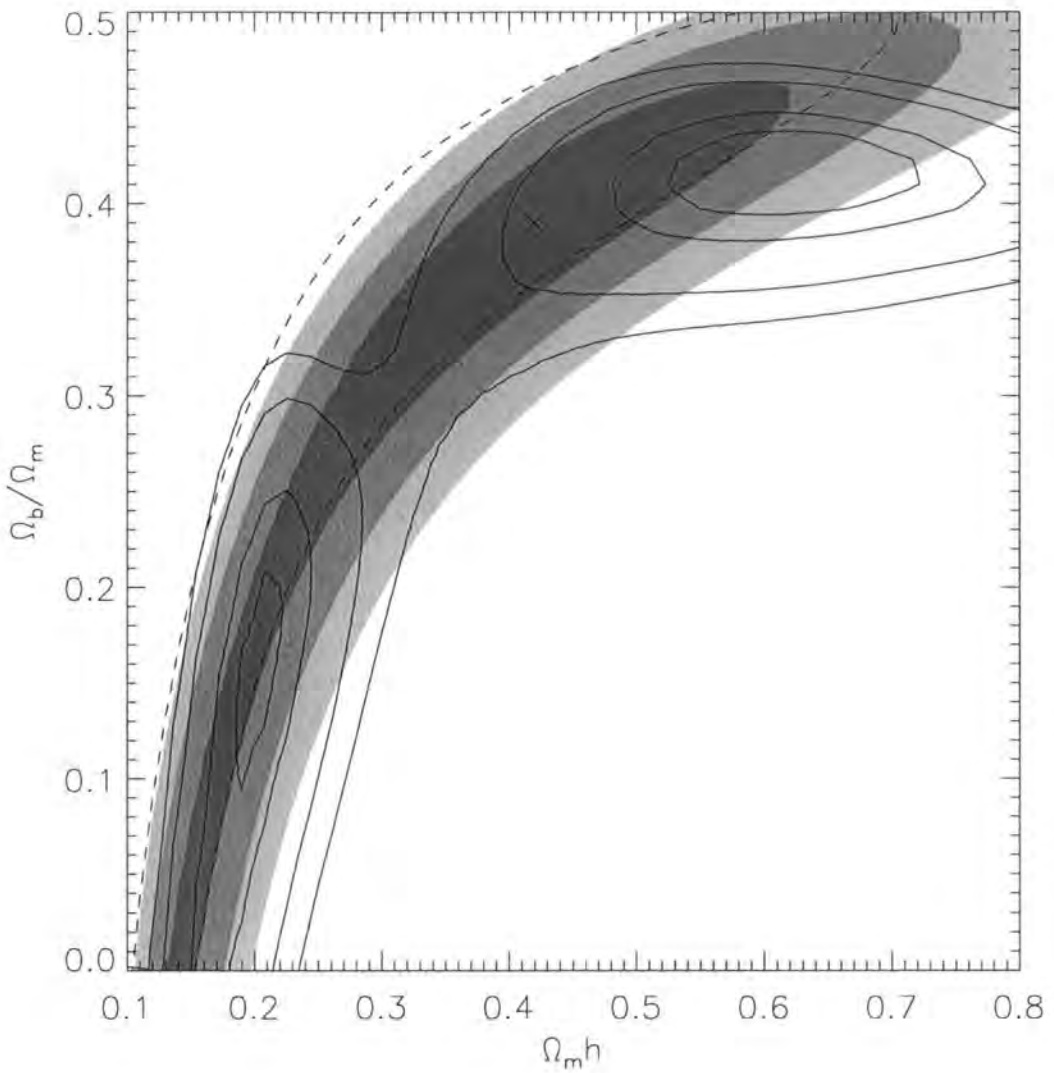


Figure 4.9: Contours of decreasing likelihood in the $\Omega_m h$ - Ω_b/Ω_m plane for the best-fitting angular power spectrum in the range $l \leq 30$. The filled contours indicate the 1σ, 2σ and 3σ confidence regions for the 2MASS $|b| > 20^\circ$ $K_s < 13.5$ galaxy angular power spectrum, determined from simple χ^2 fits, marginalising over the normalisation and h. The solid contours indicate the 1σ, 2σ, 3σ and 4σ confidence regions determined from the 2dFGRS 100k release $P(k)$ (Percival et al. 2001). The cross marks the best fit model to the 2MASS data of $\Omega_m h$=0.42 and Ω_b/Ω_m=0.39. As in Fig. 4.8 we also show the 1σ confidence region for the 2MASS result as above where we use errors which scale with the model power spectrum normalisation (dashed line).

of $b_K = 1.1 \pm 0.1$ (Maller et al. 2005). The constraint on the bias derived in this work rejects $b_K = 1$ at $> 3\sigma$; it appears therefore that galaxies selected in the K_s-band are clustered more strongly than both the underlying mass distribution and galaxies selected in optical wavebands for which $b \approx 1$ (e.g. Verde et al. 2002, Gaztañaga 1994).

We are also able to provide constraints on other cosmological parameters. We fit to $\Omega_m h$ and Ω_b/Ω_m since these primarily determine the shape of the input $P(k)$ and the size of the baryon oscillations. We determine model angular power spectra in a $71\times51\times11$ grid between $0.1\leq \Omega_m h \leq 0.9$, $0.0\leq \Omega_b/\Omega_m \leq 0.5$ and $0.4\leq h \leq 0.9$ (the effect of h on the angular power spectrum is fairly small and we therefore use a lower resolution), marginalising over the model normalisation. We perform least squares fits to the K_s <13.5, $|b|$ >20° angular power spectrum at scales of $l \leq 30$, using errors determined for the 2MASS datapoints as before which are independent of power spectrum normalisation.

The filled contours in Fig. 4.9 show the associated confidence regions marginalising over the normalisation. We are able to provide weak constraints on the cosmology of $\Omega_m h < 0.62$ and $\Omega_b/\Omega_m < 0.46$ (at 1σ confidence). Unlike previous redshift survey analyses, these constraints are particularly insensitive to the baryon density since the associated acoustic oscillations are smoothed over a wide range of angular scales. However our constraints are in good agreement with the previous results at smaller scales from the 2dFGRS $P(k)$ (Percival et al. 2001, Cole et al. 2005). As an example of how our results can differentiate between different cosmological models we show the 2MASS galaxy angular power spectrum compared with ΛCDM and SCDM predictions in Fig. 4.10.

We also wish to examine our assumption, used throughout this work, that the uncertainty due to cosmic variance determined from the 27 ΛCDM mock 2MASS catalogues is independent of the power spectrum normalisation. To do this, we instead assume that the errors determined from the ΛCDM mock catalogues simply scale with the model power spectrum normalisation as would be the case in the cosmic variance limited scenario, and compare the two cases. In Figs. 4.8 and 4.9 we show the associated 1σ confidence regions by the dashed lines, marginalising over the power spectrum normalisation. We find best fit parameters of $\Gamma_{\rm eff}=0.125\pm0.030$, $(\sigma_8 b_K)^2=1.47^{+0.27}_{-0.17}$, $\Omega_b/\Omega_m < 0.52$ and $\Omega_m h < 0.71$. This constraint on the galaxy power spectrum normalisation implies a K_s-band bias of $b_K=1.47^{+0.27}_{-0.17}$ (using the WMAP-2dFGRS constraint on σ_8 as before). It is clear that while the associated confidence regions for each parameter are slightly larger the results are in fair agreement whichever error analysis is used. However, it is clear from Fig. 4.8 that using this alternative assumption about the errors it is more difficult to reject combi-

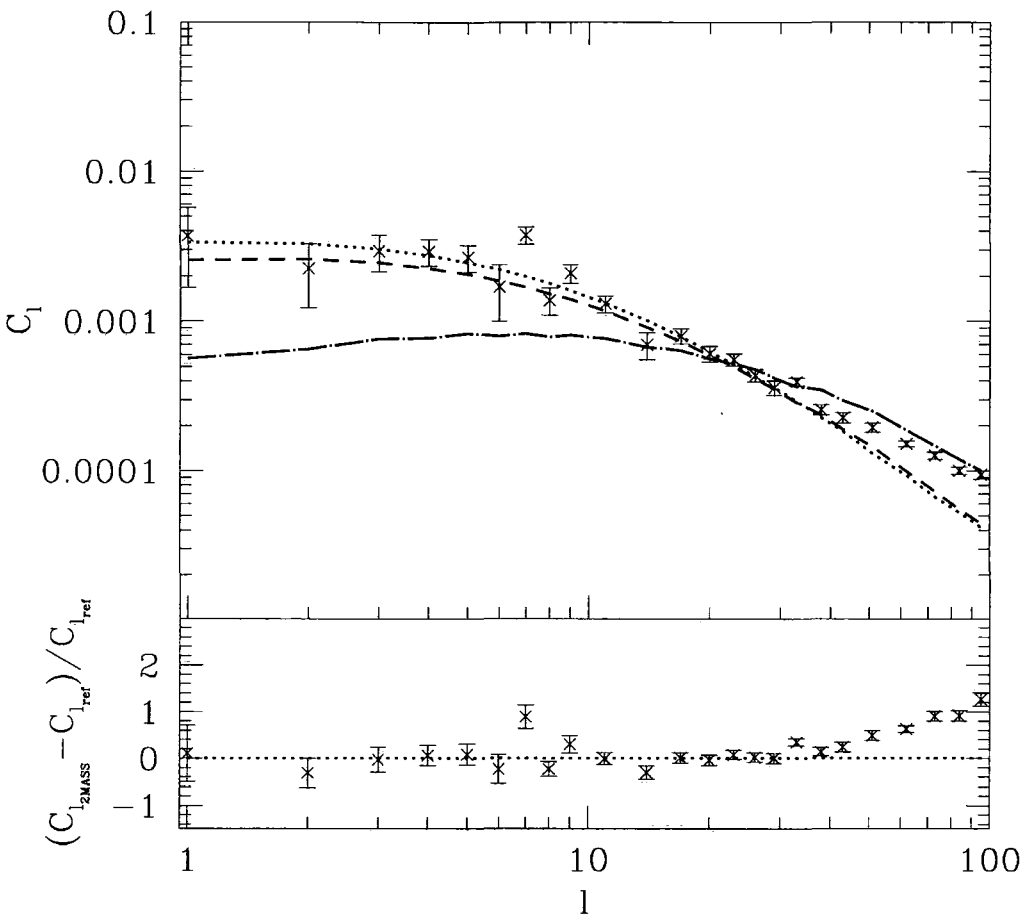


Figure 4.10: The angular power spectrum for $|b|$ >20° K_s <13.5 2MASS galaxies (as in Fig. 4.2) is compared to a linear theory SCDM prediction using input parameters of Ω_m=1.0, Ω_b=0.04, h=0.50 (dot-dashed line), a ΛCDM prediction using the Hubble Volume input parameters as before (dashed line), and the best fit power spectrum shape (for $l \leq$30) of $\Gamma_{\rm eff}$=0.14 (dotted line). In each case we use the best fit normalisation of $(\sigma_8 b_K)^2$=1.36. The errorbars indicate the 1σ spread determined from the 27 mock 2MASS power spectra. In the lower panel we show the fractional deviation from the best fit $\Gamma_{\rm eff}$=0.14 prediction.

nations of high bias and steeper $\Gamma_{\rm eff}$ slopes. For example, $\Gamma_{\rm eff}$=0.05 would only be rejected at 2.5σ. More simulations of other cosmologies are needed to check whether these errors or the errors used elsewhere in this paper are most likely to be correct.

4.6 CONCLUSIONS

We have used 518 576 $K_s < 13.5$, $|b| > 20^\circ$ galaxies selected from the 2MASS full release extended source catalogue to determine the associated angular power spectrum and constrain the form of galaxy fluctuations to Gigaparsec scales. We have compared this to a ΛCDM N-body mock prediction constructed from the Hubble Volume simulation; it is in reasonable agreement although there is a discrepancy in the slopes at l >30 in that the 2MASS result is significantly steeper than the mock prediction. We compare these to a linear theory prediction using the ΛCDM Hubble Volume simulation input parameters; there is good agreement with the mock prediction at scales where non-linear effects are insignificant ($l \lesssim 30$).

Possible sources of systematic error were investigated. We first examined the effect of imposed magnitude limit; the 2MASS angular power spectrum slope was robust with respect to the 2MASS mock and model predictions. The 2MASS galaxy angular power spectrum is also robust to different magnitude estimators and zero-point corrections (imposed to agree with the Loveday (2000) photometry) at the $\approx$10 per cent level. We correct for extinction using the Schlegel et al. (1998) dust maps; the effect on the angular power spectrum is $\approx$10 per cent at $l \lesssim 10$, and $\approx$1 per cent at smaller scales. Our results are also robust to window function effects; the effect of a $|b| > 20^\circ$ cut is $\lesssim 5$ per cent at all scales.

Finally, we have used linear theory predictions for the 2MASS galaxy angular power spectrum formed from the transfer function fitting formulae of Eisenstein & Hu (1998) to determine constraints on $\Omega_m h$ and Ω_b/Ω_m assuming a flat CDM cosmology, a primordial n_s=1 spectrum and a negligible neutrino mass. Our results are in agreement with the 2dFGRS $P(k)$ constraints (Percival et al. 2001), and we are able to provide weak constraints of $\Omega_m h$ <0.62 and Ω_b/Ω_m <0.46 (at 1σ confidence). Angular power spectrum analysis is particularly insensitive to the baryon density since any associated baryon oscillations are likely to be smoothed over a wide range of angular scales. However, given the huge volume probed ($\approx$5 times the final 2dFGRS volume) the associated constraints on the power spectrum shape and normalisation are more significant. We also determine constraints for the galaxy power spectrum shape, $\Gamma_{\rm eff}$, and normalisation, $(\sigma_8 b_K)^2$. In

agreement with the 2dFGRS and WMAP values, we find that $\Gamma_{\rm eff} = 0.14 \pm 0.02$. This is slightly higher than an alternative value found by Maller et al. (2005) using the 2MASS dataset of $\Gamma_{\rm eff} = 0.116 \pm 0.009$, determined through an inversion of the angular correlation function. We also tightly constrain the K_s-band galaxy power spectrum normalisation to $(\sigma_8 b_K)^2 = 1.36 \pm 0.10$. Using the WMAP-2dFGRS value of $\sigma_8 = 0.84 \pm 0.04$ (Bennett et al. 2003), this implies a K_s-band bias of $b_K = 1.39 \pm 0.12$.

We also investigated the likely impact on our assumption that the errors which we use to constrain various cosmological parameters, determined from the unbiased ΛCDM mocks, are independent of cosmology and power spectrum normalisation by instead assuming that these errors simply scale with the power spectrum normalisation as would be the case in the cosmic variance limited scenario. We find that while the associated confidence regions are slightly larger the results are in fair agreement. However it becomes less easy to reject models with lower $\Gamma_{\rm eff}$; therefore although the data appears to prefer a ΛCDM power spectrum slope, it may still not be possible to rule out a significantly steeper $\Gamma_{\rm eff}$.

CHAPTER 5

HIGH-ORDER 2MASS GALAXY CORRELATION FUNCTIONS

In the previous chapter, we determined the variance of the local galaxy density field as observed by the 2 Micron All Sky Survey (2MASS); we now examine higher order moments. While this analysis has no direct bearing on the possible local hole discussed in chapters 2 and 3, it has important consequences for two issues mentioned in chapter 1; the nature of the primordial density field and the way in which galaxies trace the underlying mass distribution.

5.1 INTRODUCTION

The variance of local galaxy density fluctuations has been measured to high accuracy over the last decade, both through the 2-point correlation function (e.g. Zehavi et al. 2004, Maller et al. 2005, Hawkins et al. 2003) and its Fourier transform, the power spectrum (e.g. Cole et al. 2005). For a perfectly Gaussian density field, the 2-point statistic forms a complete description of the galaxy distribution as all higher order connected (i.e. not including any contribution from lower orders) moments are zero.

Assuming a Gaussian form for the primordial distribution of density fluctuations, perturbation theory predicts non-zero higher order correlation functions of the local galaxy density field; as the initial inhomogeneities grow gravitationally, the density distribution becomes asymmetric, developing non-zero skewness and kurtosis etc. However, it is possible to construct models of inflation, and also models which contain non-linear structures in the primordial density field, such that the initial conditions themselves are non-Gaussian (e.g. Silk & Juszkiewicz 1991, Weinberg & Cole 1992). Therefore, if non-zero high-order moments of the local galaxy density field are detected, these arise either via the gravitational collapse of initially Gaussian density fluctuations or from some degree of primordial non-Gaussianity.

Examining the high-order correlation functions of the local galaxy distribution there-

fore represents a particularly powerful approach to understanding the nature of the primordial density field. It is now well established that under the assumption of Gaussian initial conditions, these high-order moments are expected to display a hierarchical scaling such that the p-order cumulants of the local density field $< \delta^p >_c$ (where the subscript c indicates the connected moment) can be expressed in terms of the variance of the distribution, such that $< \delta^p >_c = S_p < \delta^2 >^{p-1}$ (e.g. Peebles 1980, Fry 1984, Bouchet et al. 1992, Bernardeau 1992, Bernardeau 1994a, Gaztañaga & Baugh 1995, Baugh et al. 2004).

These S_p coefficients, which quantify the departure from Gaussian behaviour, are therefore expected to be constant with scale and are insensitive to cosmic epoch or assumed cosmology at scales where the growth of the density field is linear or quasi-linear. Departures from the hierarchical scaling of the high-order galaxy correlation functions on these scales might be expected only in the case of strongly non-Gaussian initial conditions or some form of scale-dependency at large scales in the bias between the galaxy distribution and the underlying mass distribution.

Several surveys have made a significant contribution to our understanding of this issue. It has long been known that high-order correlation functions are non-zero on small scales (e.g. Groth & Peebles 1977, Saunders et al. 1991, Bouchet et al. 1993, Gaztañaga 1994, Szapudi et al. 2000, Hoyle et al. 2000). More recently, Croton et al. (2004) used volume-limited catalogues drawn from the 2dF Galaxy Redshift Survey (2dFGRS) to constrain S_p in redshift space (for $p \leq 5$) to $r \approx 10\,h^{-1}$ Mpc. For $r \lesssim 4\,h^{-1}$ Mpc the hierarchical scaling parameters, S_p, were found to be approximately constant with respect to scale. However, at larger values of r these coefficients increased with scale for some luminosities, consistent with some models of structure formation with strongly non-Gaussian initial density fields (Gaztañaga & Mahonen 1996, Gaztañaga & Fosalba 1998, White 1999, Bernardeau et al. 2002). However, the results were complicated by the presence of two massive superclusters which, when removed from the analysis (corresponding to a reduction in the volume of $\approx$2 per cent), resulted in constant scaling parameters on all scales to $r \approx 10\,h^{-1}$ Mpc, consistent with Gaussian initial conditions. It appears therefore that the 2dFGRS does not probe a large enough volume to constitute a fair sample of the local galaxy distribution for high-order correlation functions.

The form of the scaling parameters on larger scales are also poorly constrained. Previously, Gaztañaga (1994) used the APM galaxy survey to $b_J = 20$ to constrain high-order ($p \leq 9$) angular correlation functions and the associated angular scaling parameters (see

equation 5.14), s_p, to $\theta \approx 7°$ (corresponding to $r \lesssim 30\,h^{-1}\,\mathrm{Mpc}$). Despite the fact that the APM galaxy survey covers a $\approx 4\times$ larger solid angle than the 2dFGRS, a similar increase in s_3 on large scales was found as seen for S_3 observed in the 2dFGRS. At higher orders, further departures from the expected hierarchical scaling were also observed, although the scales probed were limited to $\theta \lesssim 3°$ and the statistics became increasingly uncertain. These features have also been measured in the smaller Edinburgh-Durham Galaxy Catalogue (Szapudi & Gaztañaga 1998). Of course, the increase in s_3 at large scales detected by Gaztañaga (1994) may also be due to the same massive supercluster observed in the Southern 2dFGRS field, as the 2dFGRS is drawn from the larger APM galaxy sample. However, it would be surprising if the volume probed by the APM galaxy survey still did not constitute a fair sample of the Universe.

High-order correlation functions and the amplitude of the associated scaling parameters also provide us with a powerful probe of the way in which galaxies trace the underlying mass distribution. Recent measurements of the linear bias, that is the bias associated with the variance of the density field such that $\bar{\xi}_{2,gal} = b_1^2 \bar{\xi}_{2,DM}$, indicate that in the near infra-red $b_1 > 1$ (see chapter 4) whereas for optically-selected galaxies $b_1 \approx 1$ (e.g. Verde et al. 2002, Gaztañaga et al. 2005). Whether there exist non-linear contributions to the galaxy bias, such that the bias is a function of the density field, can be examined using high-order moments of the galaxy density field (e.g. Fry & Gaztañaga 1993).

Previous high-order clustering analyses have universally found that the observed skewness etc. are significantly lower than the expected values in a ΛCDM cosmology. This discrepancy has been interpreted as evidence for a negative, non-linear contribution to the galaxy bias. Most recently, Gaztañaga et al. (2005) used volume-limited samples of the 2dFGRS to determine the redshift space 3-point correlation function and constrained b_J-band bias parameters to $b_1 = 0.94^{+0.13}_{-0.11}$ and a non-linear, quadratic bias of $c_2 = b_2/b_1 = -0.36^{+0.13}_{-0.09}$, although as with previous 2dFGRS analyses these results are significantly affected by the presence of two massive superclusters. Independently, Pan & Szapudi (2005) determined the monopole contribution to the 2dFGRS 3-point correlation function and determined a similar b_J-band linear bias to Gaztañaga et al. (2005) and a negative c_2 parameter as well, although considerably smaller and at a reduced significance to the other 2dFGRS constraint. Computing the bispectrum for the PSCz catalogue (which is selected from the 60μm IRAS galaxy sample), Feldman et al. (2001) also constrain infra-red bias parameters of $b_1 = 0.83 \pm 0.13$ and $c_2 = -0.50 \pm 0.48$.

The 2 Micron All Sky Survey (2MASS) has recently been completed and provides K_s,

H and J-band photometry for 1.6×10^6 extended sources over the entire sky to $K_s\gtrsim13.5$. 2MASS is the largest existing all-sky galaxy survey and therefore represents a uniquely powerful probe of the local galaxy density field at large scales; the solid angle of the 2MASS $|b|\geq10^\circ$ sample used in this paper represents an order of magnitude increase over the APM galaxy survey, meaning that clustering statistics determined from the 2MASS galaxy sample will suffer less from projection effects while probing a comparable volume. The 2MASS survey also represents an order of magnitude increase in volume over the largest volume-limited 2dFGRS sample (assuming an equal weighting scheme for each galaxy); it is therefore possible to probe much larger scales than the 2dFGRS although with the added complication of projection effects. A further advantage of 2MASS over previous datasets is that the photometry is extremely accurate with high completeness for $|b|\geq10^\circ$; correcting for the variable completeness over the 2dFGRS survey area for instance, complicates the analysis and increases the uncertainty from possible systematic effects. The main drawback to high-order clustering analysis of the 2MASS data (as with the APM galaxy survey) is the lack of available three-dimensional information; the clustering signal from a particular scale in real space is smeared over a range of angular scales. For this reason detailed features in the real space correlation function, such as the shoulder at $r\approx10\,h^{-1}$ Mpc (Baugh 1996, Hawkins et al. 2003, Zehavi et al. 2004), may not be detected by the 2MASS projected correlation function.

In this chapter, we aim to determine the high-order angular correlation functions and the associated scaling parameters of the local galaxy density field to high precision and large scales, using the final 2MASS extended source catalogue. In section 5.2, the details of the method for estimating high-order correlation functions are given. The galaxy sample used and the error analysis are described in section 5.3. We present the p-point angular correlation functions (for $p\leq9$) and the hierarchical scaling relations in section 5.4. We also examine possible systematic effects arising from extreme fluctuations in the observed galaxy density field. The implications for the form of primordial density fluctuations and non-linear galaxy bias are discussed in section 5.5. The conclusions follow in section 5.6.

5.2 Method of Estimation

5.2.1 The p-point Correlation Function

The p-point galaxy correlation function estimates the joint probability that p galaxies are separated by a certain scale and can be defined through considering fluctuations in

the galaxy density field. The connected or reduced part of this statistic corresponds to the contribution to this probability which does not include any conditional probability on lower orders:

$$\omega_p(\theta_1, ..., \theta_p) \equiv < \delta_1, ..., \delta_p >_c, \tag{5.1}$$

where δ denotes the density fluctuation; for $p \leq 3$, the unreduced and reduced correlation functions are the same. In this analysis, we work with the reduced p-point correlation function only.

The 2-point angular galaxy correlation function, $\omega_2(\theta)$, is given in terms of the probability of finding two galaxies in area elements $d\Omega_1$ and $d\Omega_2$ separated by angle θ:

$$dP_2 = \mathcal{N}^2 \, [1 + \omega_2(\theta)] \, d\Omega_1 d\Omega_2, \tag{5.2}$$

where $\mathcal{N}$ is the mean number of galaxies per unit solid angle (e.g. Groth & Peebles 1977, Peebles 1980). Similarly, the 3-point function, $\omega_3(\theta)$, is defined by the joint probability of finding galaxies in each of three area elements:

$$dP_3 = \mathcal{N}^3 \, [1 + \omega_2(\theta_{12}) + \omega_2(\theta_{23}) + \omega_2(\theta_{13})$$

$$+ \omega_3(\theta_{12}, \theta_{23}, \theta_{31})] \, d\Omega_1 d\Omega_2 d\Omega_3 \tag{5.3}$$

The first term in equations 5.2 and 5.3 include contributions from galaxy pairs or triplets respectively which are accidentally seen as close together in projection but are at very different radial distances. Similarly, the following three terms in equation 5.3 describe the contributions from one correlated pair and a third uncorrelated galaxy which forms a triplet by chance line-of-sight clustering. This leaves the final term which defines the contribution from the real clustering of triplets.

A simple way in which to estimate the high-order correlation functions is through the area-averaged correlation function, $\bar{\omega}_p$, defined as:

$$\bar{\omega}_p(\theta) = \frac{1}{\Omega^p} \int_\Omega \omega_p(\theta_1, ..., \theta_p) \, d\Omega_1, ..., d\Omega_p, \tag{5.4}$$

where Ω is the solid angle of the cone defined by its angular radius θ. The area-averaged correlation function, $\bar{\omega}_p$, is estimated by considering the central moments of the angular counts:

$$m_p(\theta) =< (N - \bar{N})^p >= \sum_{N=0}^{N=\infty} (N - \bar{N})^p P_N(\theta), \tag{5.5}$$

where $P_N(\theta)$ denotes the count probability distribution function and is calculated by placing circular cells of angular radius θ over the survey area and determining the number of cells containing exactly N galaxies:

$$P_N(\theta) = \frac{N_N}{N_C}, \tag{5.6}$$

where N_N and N_C denote the number of cells containing N galaxies and the total number of cells respectively. $\bar{N}$ in equation 5.5 is the mean number of galaxies in a cell of angular radius θ and may be determined directly from the count probability distribution function:

$$\bar{N} = \sum_{N=0}^{N=\infty} N P_N(\theta) \tag{5.7}$$

The moments of the count probability distribution function determined via equation 5.5 yield the unreduced correlation function through the relation $m_p =< \delta^p > \bar{N}^p$. In order to obtain the reduced correlation function, the connected moments μ_p are determined:

$$\mu_2 = m_2$$

$$\mu_3 = m_3$$

$$\mu_4 = m_4 - 3m_2^2, \tag{5.8}$$

(see Gaztañaga (1994) for higher order relations). In addition we apply a shot noise correction (Gaztañaga 1994, Baugh et al. 1995) such that:

$$k_2 = \mu_2 - \bar{N}$$

$$k_3 = \mu_3 - 3k_2 - \bar{N}$$

$$k_4 = \mu_4 - 7k_2 - 6k_3 - \bar{N} \tag{5.9}$$

The area-averaged, reduced angular correlation function is then determined from the relation $\bar{\omega}_p =< \delta^p >_c = k_p/\bar{N}^p$ (e.g. Gaztañaga 1994).

5.2.2 HIERARCHICAL SCALING

In perturbation theory, the density field, evolved by gravity from an initially Gaussian distribution, leads to a hierarchical clustering pattern, such that all high-order correlations can be expressed in terms of the 2-point correlation function:

$$\bar{\xi}_{p,\mathrm{DM}} = S_{p,\mathrm{DM}}\, \bar{\xi}_{2,\mathrm{DM}}^{p-1}, \tag{5.10}$$

where $\bar{\xi}_{p,DM}$ is the volume-averaged p-point dark matter correlation function. Importantly, this relation is preserved for the galaxy density field such that $\bar{\xi}_{p,gal} = S_{p,gal}\, \bar{\xi}_{2,gal}^{p-1}$ (Fry & Gaztañaga 1993), assuming that the relation between the dark matter and galaxy density fields may be expressed through a Taylor expansion of the dark matter density contrast:

$$\delta_{\mathrm{gal}} = \sum_{p=0}^{\infty} \frac{b_p}{p!} (\delta_{\mathrm{DM}})^p. \tag{5.11}$$

For the skewness, it can be shown that (Fry & Gaztañaga 1993):

$$S_{3,\mathrm{gal}} = \frac{1}{b_1}(S_{3,\mathrm{DM}} + 3c_2), \tag{5.12}$$

where b_1 is the linear bias such that $\bar{\xi}_{2,gal} = b_1^2 \bar{\xi}_{2,DM}$ on scales where the variance of the density field is small, and $c_2 = b_2/b_1$ quantifies the second-order contribution to the galaxy bias.

Using perturbation theory, it is possible to determine precise quantitative predictions for the $S_{p,DM}$ parameters. Assuming a power law form for the three-dimensional power spectrum of density fluctuations, $P(k) \propto k^n$, and a spherical top-hat window function (i.e. a square window function that is constant within, and vanishes outside, a spherical survey volume), the skewness of the matter distribution, $S_{3,DM}$ may be determined (Juszkiewicz et al. 1993):

$$S_{3,\mathrm{DM}} = \frac{34}{7} - (n+3) \tag{5.13}$$

We later use these expressions to determine constraints on non-linear bias through comparisons with $S_{3,gal}$ in the linear and quasi-linear regimes. We assume a power spectrum slope of $n = -2$ (Percival et al. 2001, Cole et al. 2005). The uncertainties on n are small (Δn <0.1) compared to the sampling errors in the measurement of $S_{3,gal}$; we therefore assume the concordance value cited above to generate predictions for $S_{3,DM}$ and neglect small uncertainties in this parameter.

5.2.3 TRANSFORMATION TO THREE DIMENSIONS

For a projected galaxy distribution, a hierarchical scaling relation may also be defined in terms of the area-averaged p-point galaxy correlation functions and angular scaling coefficients, s_p:

$$\bar{\omega}_p = s_p \, \bar{\omega}_2^{p-1} \tag{5.14}$$

We wish to transform these angular scaling parameters, s_p, to the three-dimensional coefficients, S_p, in order to make comparisons with perturbation theory and constrain non-linear galaxy bias. Following the method of Gaztañaga (1994), we transform to three dimensions via the relation:

$$S_p(\bar{r}) \approx \frac{s_p(\theta) B_p(\gamma)}{r_p(\gamma) C_p(\gamma)}, \tag{5.15}$$

where $\bar{r} = \theta \mathcal{D}$ is the mean scale probed at an angular scale θ for a survey of median depth $\mathcal{D}$, γ is the slope of the 2-point real space correlation function, and B_p and C_p are related to the number of different configurations of the three-dimensional and angular hierarchical tree graphs respectively (these show the hierarchy structure from the smallest groups or pairs of galaxies to the largest superclusters; see Gaztañaga (1994) for further details). Here we use the form of the real space correlation function determined from the 2dFGRS (Hawkins et al. 2003). The r_p factor is related to the selection function Ψ:

$$r_p = \frac{I_1^{p-2} I_p}{I_2^{p-1}} \tag{5.16}$$

where

$$I_j = \int_0^\infty \Psi^j x^{(3-\gamma)(j-1)} (1+z)^{(3+\epsilon-\gamma)(1-j)} F(x) x^2 dx \tag{5.17}$$

where ϵ describes the evolution of clustering with redshift (Groth & Peebles 1977) and is taken to be ϵ=0 in good agreement with recent observational and theoretical considerations (Hamilton et al. 1991, Peacock & Dodds 1994, Carlberg et al. 2000, Wilson 2003). Here, x denotes the comoving distance and $F(x)$ a correction for curvature such that $F(x) = [1 - (H_0 x/c)^2(\Omega_m - 1)]^{1/2}$. In this work we use the concordance value of Ω_m=0.3. For the selection function we use a parameterised form for the $n(z)$ such that:

$$n(z) = \frac{3z^2}{2(\bar{z}/1.412)^3} \exp\left(-\left(\frac{1.412z}{\bar{z}}\right)^{3/2}\right) \tag{5.18}$$

(Baugh & Efstathiou 1993, Maller et al. 2005) where $\bar{z}$ is determined from the 2MASS-2dFGRS matched sample described in section 2.2.2. For reference $\bar{z}$=0.074 for K_s <13.5. In this case, the normalisation of the selection function factors out (see equations 5.16 and 5.17).

The transformation described in equation 5.15 is robust to reasonable changes in the selection function and choice of cosmological parameters (Gaztañaga 1994). However, this relation becomes uncertain on large angular scales, $\theta > 2°$ (Gaztañaga 1994), due to the fact that the power law form to the 2-point correlation function and the value of γ are not well constrained on large scales.

5.3 Analysis of the 2MASS Data

5.3.1 The 2MASS Extended Source Catalogue

We use the dataset described in section 2.2.1, selecting objects from the 2MASS final release extended source catalogue (Jarrett 2004) above a galactic latitude of $|b| = 10°$ in order to remove regions of high extinction and stellar contamination (see Fig. 5.3) and below an extinction-corrected magnitude limit of K_s=13.5 using the dust maps of Scharf et al. (1992). The subsequent sample of 650 745 galaxies probes to a median depth of $\mathcal{D} \approx 220\,h^{-1}$ Mpc and covers 83 per cent of the entire sky ($\approx$34 000 deg^2).

In order to calculate the high-order correlation functions of this 2MASS sample, we determine the count probability distribution function detailed in equation 5.6 by randomly placing N_C=10^6 cell centres over the survey area. Each cell is then allowed to grow and the number of galaxies as a function of the angular radius is recorded. The size of the cells is limited by the galactic latitude limit of $|b| = 10°$ in the 2MASS sample. We replace cells which are lost as they encroach the boundary of the galaxy sample such that N_C=10^6, independent of the angular scale probed. We use cell radii in the range 0°.01 to 25°.1, equivalent to a range in the mean scale probed of $0.04\,h^{-1}$ Mpc to $104\,h^{-1}$ Mpc.

5.3.2 Error Estimation

The statistical uncertainty associated with the correlation function and angular scaling parameters in this work are determined using bootstrap estimates. The full 2MASS sample is split into 20 equal area regions of $\approx$1700 deg^2; 20 of these sub-areas are selected at random (with repeats) and the associated clustering statistic is determined. As with the analysis of the full sample, cells are replaced as they encroach the boundary of each

sub-area. We repeat this 1000 times and determine the standard deviation; these are indicated by the errorbars in Figs. 5.1 and 5.2. The size of these errors is not significantly altered if we vary the number of realisations or sub-areas within reasonable limits.

Once we have determined the angular correlation functions and scaling parameters and the associated errors in this way, we wish to carry out comparisons with predictions from perturbation theory. Since correlation function estimates determined at different cell radii are highly correlated being integral quantities, it is necessary to account for the covariance between each datapoint when performing fits to the data. We determine the covariance matrix (C_{ij}) using the bootstrap method above. We then determine χ^2:

$$\chi^2 = \sum_i \sum_j \Delta_i C_{ij}^{-1} \Delta_j, \tag{5.19}$$

where $\Delta_i = s_p^{obs}(i) - s_p^{mod}(i)$ for example. We compute the inverse covariance matrix, C_{ij}^{-1}, using the Numerical Recipes Singular Value Decomposition algorithm (Press et al. 1988). As noted by Croton et al. (2004), previous constraints on the high-order scaling parameters (except for the 2dFGRS results) ignored the correlations between different bins leading to unrealistically small errors in the fitted values.

5.4 RESULTS

5.4.1 AREA-AVERAGED CORRELATION FUNCTIONS

In the first column of Fig. 5.1 we plot the area-averaged correlation functions for $p \leq 9$ determined for the full $|b| \geq 10°$ 2MASS galaxy sample described in section 5.3.1. In each case the errors are determined via the bootstrap technique described in section 5.3.2. For reference we compare our result for $\bar{\omega}_2$ with the best fit to the 2MASS K_s <13.5 2-point correlation function of Maller et al. (2005) (A=0.10±0.01, $1-\gamma = -0.79 \pm 0.02$, determined on scales of $\theta \leq 2°.5$, where $\omega_2 = A\theta^{1-\gamma}$).

The p-point correlation functions are consistent with power laws over several orders of magnitude in angular scale. In particular, we note that the area-averaged $\bar{\omega}_2$ determined in this work is in good agreement with the directly-determined result of Maller et al. (2005) at small scales (we find $A = 0.13 \pm 0.01$, $1-\gamma = -0.77 \pm 0.04$ for $\theta \leq 2°.5$). This agreement is robust to changes in the magnitude estimator and galactic latitude cut; Maller et al. (2005) use the K_s-band elliptical Kron magnitude estimator and a $|b| \geq 20°$ cut.

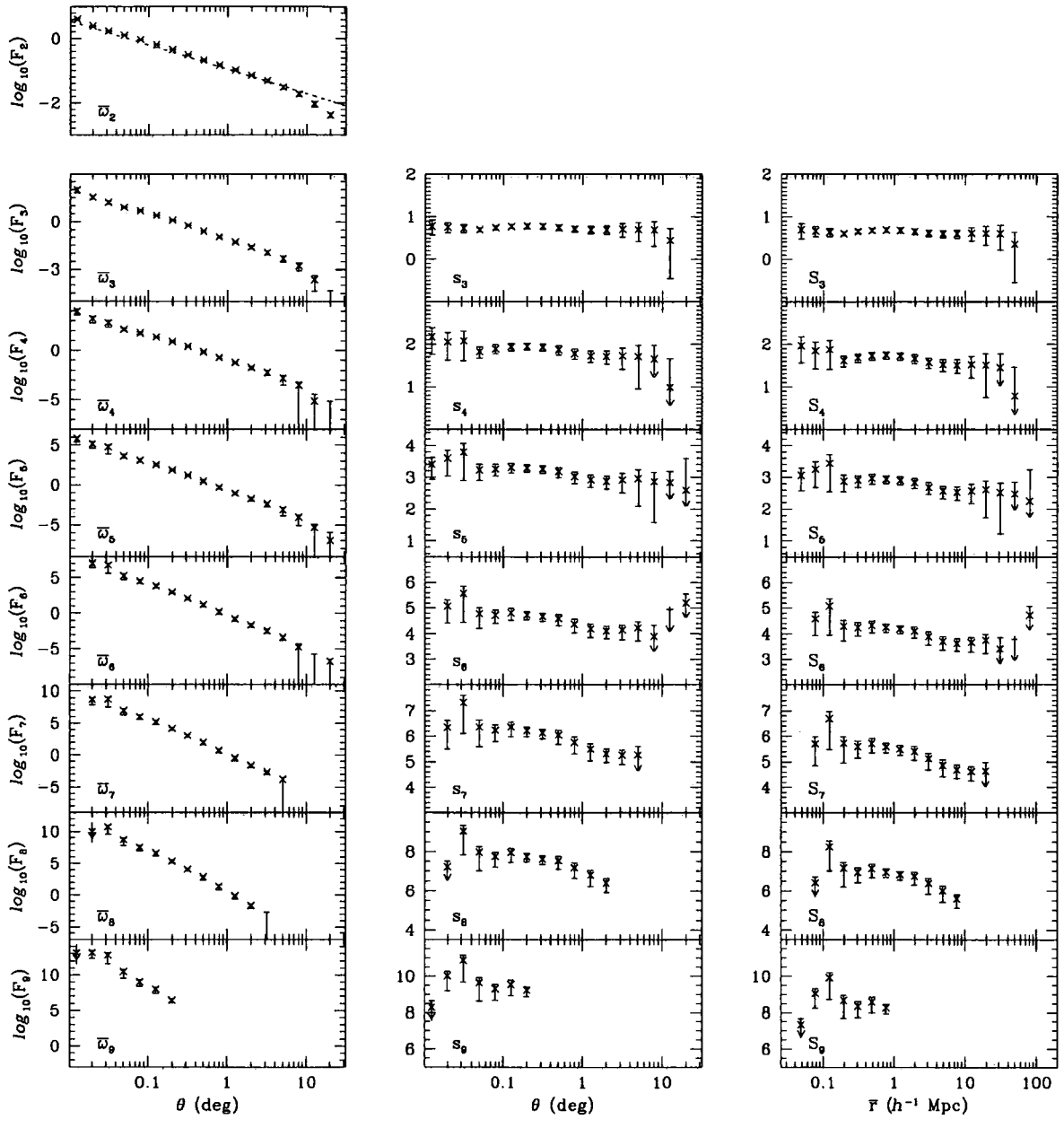

log10(F2)
ω2
log10(F3)
ω3
log10(F4)
ω4
log10(F5)
ω5
log10(F6)
ω6
log10(F7)
ω7
log10(F8)
ω8
log10(F9)
ω9
θ (deg)
S3
S4
S5
S6
S7
S8
S9
θ (deg)
r̄ (h⁻¹ Mpc)

Figure 5.1: Moments of the 2MASS galaxy density field for the full $|b| \geq 10°$ K_s <13.5 sample (all on logarithmic scales). Each row corresponds to a different moment of the galaxy density field. Since in each column we plot a different statistic (ω_p, s_p or S_p) we denote the y-axis label as F_p. In each case the errors are taken from bootstrap estimates described in section 5.3.2. Datapoints with extremely large errorbars are omitted for clarity. The columns are set out as follows:

First column: Area-averaged correlation functions for 650 745 K_s <13.5 2MASS galaxies. In the $\bar{\omega}_2$ panel, the best fit result at small scales from Maller et al. (2005) is indicated by the dotted line.

Second column: The angular scaling parameters (s_p) determined via equation 5.14 for the full sample.

Third column: The real space scaling parameters (S_p) determined via equation 5.15 for the full sample.

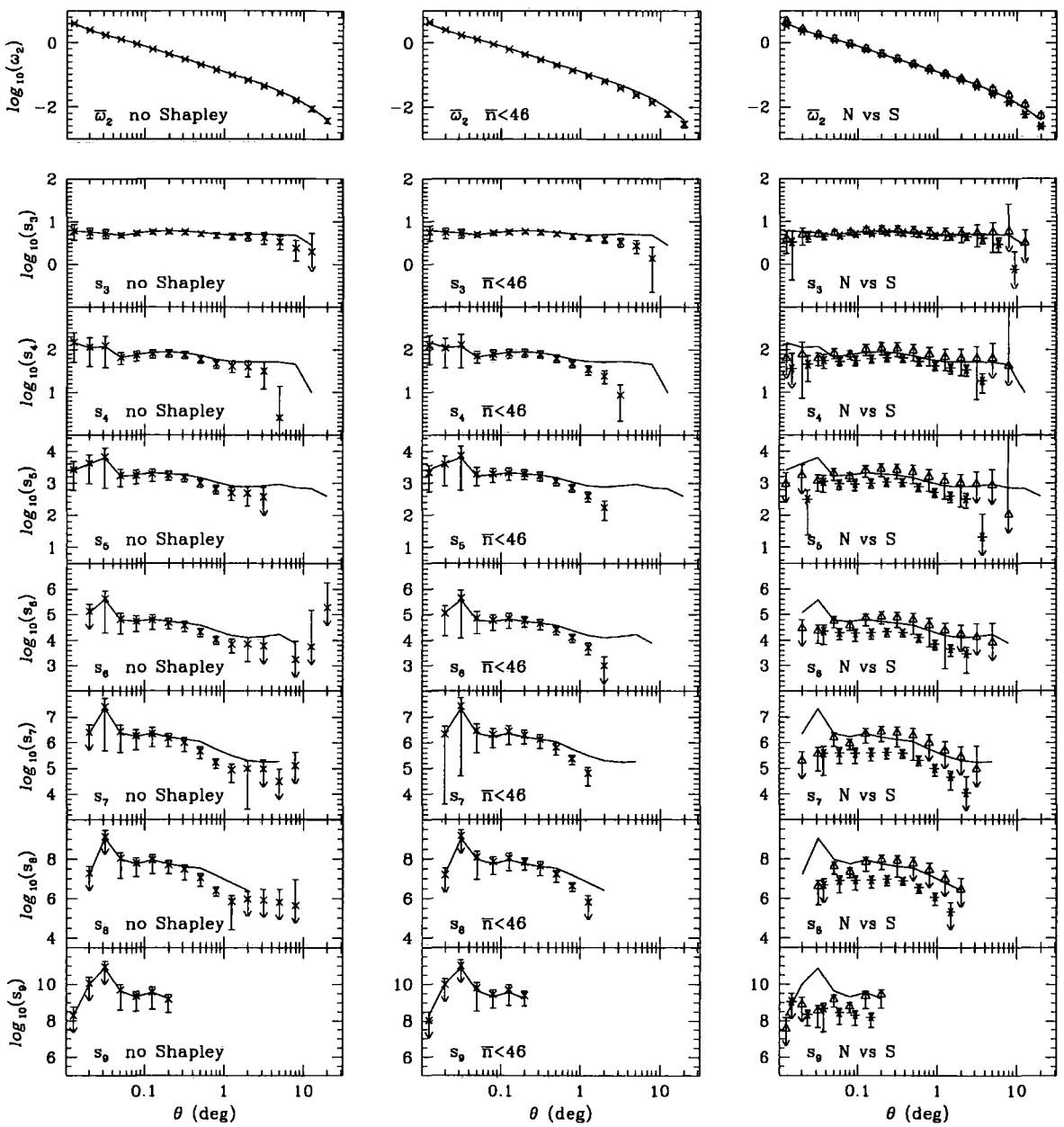

log10(ω2)
ω̄2 no Shapley
ω̄2 n̄<46
ω̄2 N vs S
log10(s3)
s3 no Shapley
s3 n̄<46
s3 N vs S
log10(s4)
s4 no Shapley
s4 n̄<46
s4 N vs S
log10(s5)
s5 no Shapley
s5 n̄<46
s5 N vs S
log10(s6)
s6 no Shapley
s6 n̄<46
s6 N vs S
log10(s7)
s7 no Shapley
s7 n̄<46
s7 N vs S
log10(s8)
s8 no Shapley
s8 n̄<46
s8 N vs S
log10(s9)
s9 no Shapley
s9 n̄<46
s9 N vs S
0.1
1
10
θ (deg)

Figure 5.2: Moments of the 2MASS galaxy density field making various cuts to the full $|b| \geq 10^\circ$ K_s <13.5 sample. As in Fig. 5.1 each row corresponds to a different moment and the errors are determined via bootstrap estimates as described in section 5.3.2. For p=2 (the top row) we show the area-averaged correlation function; for higher order moments we display the angular scaling parameters, s_p. In each case we indicate the corresponding results from the full sample shown in Fig. 5.1 by a solid line. Datapoints with extremely large errorbars are omitted for clarity. The columns are set out as follows:

First column: We show the 2-point function and the higher-order angular scaling parameters having removed a region of radius 6° centred on the largest over-density in the sample, the Shapley supercluster (see section 5.4.2). This corresponds to a removal of 1.1 per cent of the galaxies and 0.33 per cent of the solid angle of the full sample.
Second column: We apply a more stringent cut such that areas of radius 6° centred on the ten most over-dense pixels in Fig. 5.3 are removed (see table 5.3) corresponding to a removal of 6.3 per cent of the galaxies and 2.6 per cent of the solid angle of the full sample. This sample is equivalent to including only those cells with a number density of $\bar{n} < 46\,\mathrm{deg}^{-2}$ (see Fig. 5.3).
Third column: Here, we split the full sample into north (triangles) and south (stars). For clarity, we have displaced the southern angular scaling parameter datapoints to the right.

p	s_p	S_p
3	5.28 ± 0.45	4.29 ± 0.39
4	57.7 ± 9.2	36.0 ± 5.7
5	1510 ± 507	678 ± 228
6	(3.68 ± 2.08)×10^4	(1.20 ± 0.68)×10^4
7	(9.74 ± 8.72)×10^5	(2.25 ± 2.01)×10^5
8	(2.64 ± 3.69)×10^7	(4.19 ± 5.86)×10^6
9	(2.03 ± 1.56)×10^9	(2.26 ± 1.74)×10^8

Table 5.1: Small scale fits to the high-order scaling parameters (assuming constant values) determined for the full 2MASS $|b| \geq 10°$, K_s <13.5 sample (see Fig. 5.1). In the second column the best fit angular scaling parameters (for $3 \leq p \leq 9$) are shown, determined from χ^2 fits in the range $0°.04 < \theta < 1°.0$ for $p \leq 8$ and $0°.04 < \theta < 0°.25$ for p=9. Similarly, in the third column we show the best fit real space scaling parameters, S_p, fitted over an equivalent range of scales, $0.15 < \bar{r} < 4.0\,h^{-1}$ Mpc for $p \leq 8$ and $0.15 < \bar{r} < 1.0\,h^{-1}$ Mpc for p=9. The errors in each case take into account the covariance matrix determined from the bootstrap estimates described in section 5.3.2.

In order to determine whether these results are consistent with the hierarchical scaling described in section 5.2.2 we compute the angular scaling parameters, s_p (see equation 5.14); these are shown in the second column of Fig. 5.1. The angular scaling parameters are transformed into the real space S_p parameters in the third column of Fig. 5.1 (see equation 5.15). The s_p and S_p coefficients are consistent with constant values to large scales ($\theta \lesssim 20°$, $\bar{r} \lesssim 100\,h^{-1}$ Mpc), although there may be a decrease in amplitude between the non-linear and quasi-linear regime ($\theta \approx 1°$, $r \approx 4\,h^{-1}$ Mpc). One caveat to this is that while these coefficients are consistent with a slope of zero, the constraints become increasingly weak at higher orders such that in the range $1°.0 < \theta < 10°$ the slopes are constrained to $\gamma(s_3) = 0.01^{+0.34}_{-0.43}$, $\gamma(s_4) = 0.02^{+0.45}_{-0.84}$, $\gamma(s_5) = -0.01^{+0.63}_{-0.94}$ and $\gamma(s_6) = -0.39^{+0.88}_{-0.91}$ at 1σ confidence (where $s_p \propto \theta^\gamma$, marginalising over the normalisation); we investigate the constraint on the slope of S_3 and the level to which this can reject primordial non-Gaussianity in section 5.5.1.

Since we probe well into the linear and quasi-linear regimes we are able to make comparisons with predictions from perturbation theory. We perform χ^2 fits to these scaling parameters considering the covariance in the datapoints (see equation 5.19). We consider small and large scales separately due to the fact that, despite the consistency of

p	s_p	S_p
3	4.91 ± 0.60	4.00 ± 0.49
4	54.2 ± 11.1	33.8 ± 6.9
5	740 ± 240	332 ± 108
6	(1.01 ± 0.67)×10^4	(3.30 ± 2.19)×10^3

Table 5.2: Large scale fits to the high-order scaling parameters (assuming constant values) determined for the full 2MASS $|b| \geq 10°$, K_s <13.5 sample (see Fig. 5.1). In the second column the best fit angular scaling parameters (for $3 \leq p \leq 6$) are shown, determined from χ^2 fits in the range $1°.0 < \theta < 10°$. Similarly, in the third column we show the best fit real space scaling parameters, S_p, fitted over an equivalent range of scales, 4.0< $\bar{r}$ <40 h^{-1} Mpc. The errors are determined as in Table 5.1.

the scaling parameters over all scales considered, perturbation theory is not expected to be valid on small scales (Bernardeau et al. 2002); approximately, the scales used represent fits in the non-linear and quasi-linear or linear regimes. Also, it is important to remember that the conversion from angular to real space scaling parameters becomes increasingly uncertain at large angular scales ($\theta > 2°$). The best fit scaling parameters are shown in Table 5.1 for small scales and Table 5.2 for large scales.

5.4.2 Fair Sample Issues

It was noted in section 5.1, that a considerable problem in previous high-order clustering analyses is the presence of extreme fluctuations in the galaxy samples which have a significant effect on the observed scaling parameters. The APM and 2dF Galaxy Redshift Surveys observe rising S_p values for $r \gtrsim 4\,h^{-1}$ Mpc for example, consistent with some models of structure formation with strongly non-Gaussian initial conditions (Gaztañaga & Fosalba 1998, Gaztañaga & Mahonen 1996, Bernardeau et al. 2002); with the 2dFGRS at least the clustering signal is significantly altered when two superclusters are removed from the sample (corresponding to a reduction in the volume of ≈2 per cent) yielding constant S_p parameters for $p \leq 5$ to scales of $r \approx 10\,h^{-1}$ Mpc (Baugh et al. 2004, Croton et al. 2004).

We therefore wish to examine the level of such systematic effects by removing large over-densities from the 2MASS galaxy sample and examining the effect on the observed scaling parameters. In Fig. 5.3, we show the pixelated 2MASS K_s <13.5 galaxy density distribution smoothed on ≈7° scales (each pixel has a solid angle of 13.5 deg^2); clusters

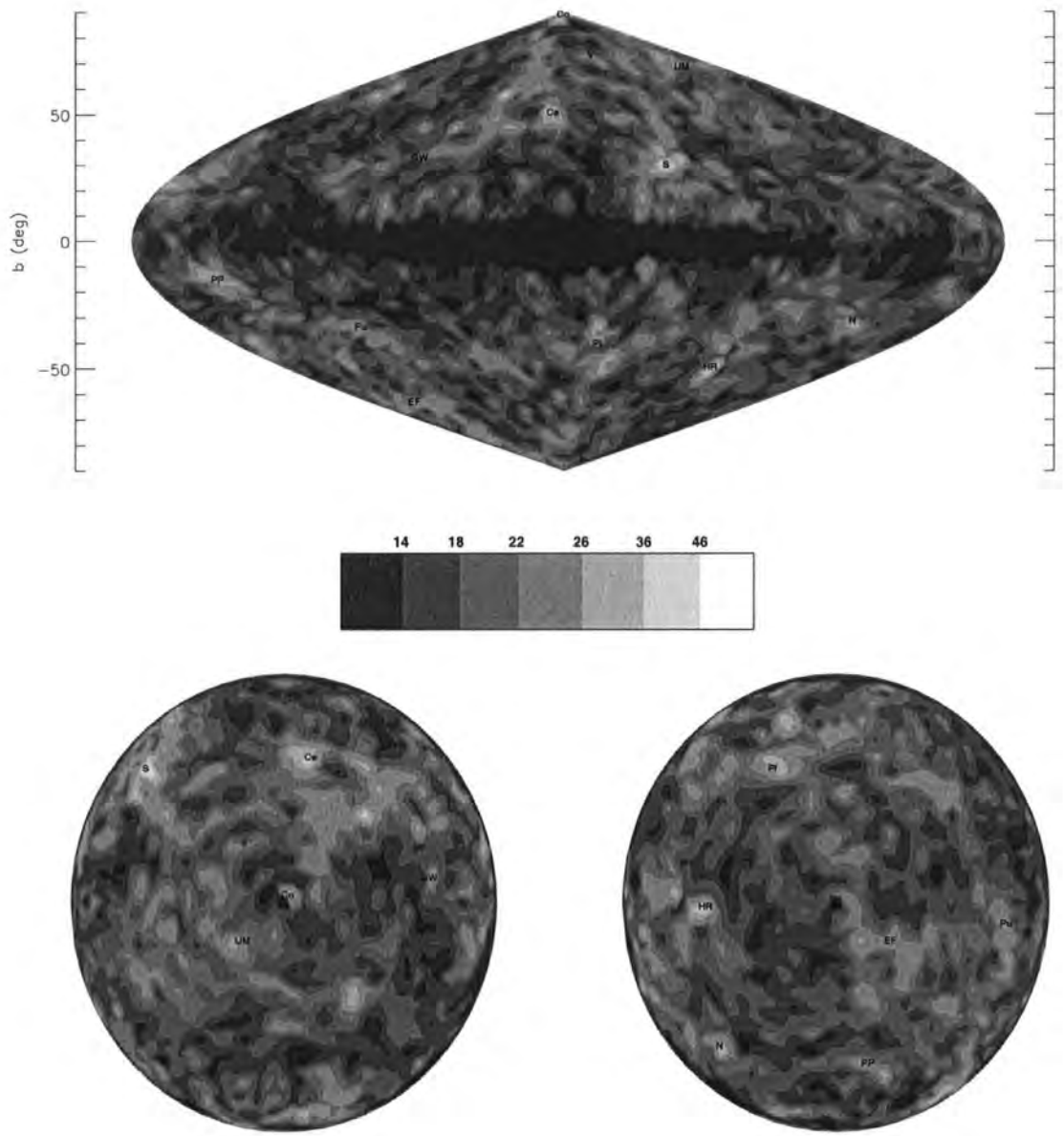


Figure 5.3: The number density (in $\deg^{-2}$) of 2MASS K_s <13.5 galaxies binned in 13.5 $\deg^2$ pixels; under-dense regions are indicated by the dark filled contours; areas of over-density by lighter filled contours. For reference, the mean $|b| \geq 10°$ number density is 19.1 $\deg^{-2}$. In the upper plot we show the entire sky in projection; in the lower panels we show the $b \geq 0°$ (left-hand plot) and $b \leq 0°$ (right-hand plot) hemispheres, such that the galactic poles are positioned in the centres of each. Prominent clusters are indicated as follows: Co - Coma cluster, S - Shapley supercluster, V - Virgo supercluster, UM - Ursa Major cloud, GW - Great Wall, Ce - Centaurus, HR - Horologium-Reticulum, PP - Perseus-Pisces chain, PI - Pavo-Indus wall, N - NGC 1600 Group, Pu - Puppis, EF - Eridanus-Fornax.

	l (°)	b (°)	$\bar{n}$ (deg^{-2})	Cluster
1	312.2	30.0	69.1	Shapley supercluster
2	266.5	-51.3	54.4	Horologium-Reticulum
3	45.0	57.4	52.4	Centaurus
4	10.4	51.3	52.2	Centaurus
5	303.8	32.8	49.4	Shapley supercluster
6	266.8	-48.1	47.5	Horologium-Reticulum
7	219.4	-32.8	47.5	NGC 1600 Group
8	343.1	-32.8	46.8	Pavo-Indus wall
9	312.2	35.7	46.2	Shapley supercluster
10	9.6	48.1	46.1	Centaurus

Table 5.3: The positions and galaxy densities of the ten most over-dense pixels of the smoothed $K_s < 13.5$ 2MASS galaxy distribution shown in Fig. 5.3. We also note the cluster with which each pixel is associated.

are indicated by lighter filled contours and regions of under-density by darker shades. In the upper panel we show the entire K_s <13.5 2MASS sample in Aitoff projection; in the lower panels the galaxy density field for each galactic hemisphere is plotted in projection separately. We also list the ten most over-dense pixels (corresponding to a limit in the number density of $\bar{n}$ >46 deg^{-2}) in table 5.3.

In Fig. 5.2, we plot $\bar{\omega}_2$ and the high-order angular scaling parameters with various cuts to the full galaxy sample used in Fig. 5.1. In each case, the corresponding result for the full sample is indicated by a solid line. First, we omit the largest supercluster only; in the first column, we have removed a circular region with an angular radius of 6° centred on the most over-dense pixel in Fig. 5.3 sampling the Shapley supercluster. This corresponds to a removal of 1.1 per cent of the galaxies and 0.33 per cent of the solid angle of the full $|b| \geq 10°$ sample. The effect of this removal on the 2-point correlation function is insignificant. The form of s_3 remains consistent with the result for the full sample, although the best fit slope at large scales ($1°.0 < \theta < 10°$) changes to $\gamma = -0.27^{+0.25}_{-0.29}$ (compared to $\gamma = 0.01^{+0.34}_{-0.43}$ for the full sample). At higher orders, there is also no significant effect although the statistical uncertainty increases at large scales (datapoints with extremely large errorbars are omitted for clarity).

In the second column in Fig. 5.2 we omit all galaxies within 6° of the ten most over-

dense pixel centres (see table 5.3). This corresponds to a removal of 6.3 per cent of the galaxies and 2.6 per cent of the solid angle of the full sample. There is a small effect on the 2-point correlation function at large scales ($\theta \gtrsim 10°$). The effect on the higher order angular scaling parameters also becomes more significant. However the effect on s_3 at least is limited; the best fit slope at large scales ($1°.0 < \theta < 10°$) is $\gamma = -0.33^{+0.26}_{-0.29}$.

It has previously been observed that the form of the galaxy distribution in the northern and southern galactic caps are significantly different in many respects; Maller et al. (2005) detected a difference in the $b \geq 20°$ and $b \leq -20°$ 2MASS 2-point angular correlation functions at large scales for example. It has also been suggested in chapters 2 and 3 that the southern galactic cap may contain a large local under-density in the galaxy distribution covering $\gtrsim 4000$ deg^2 around the Southern Galactic Pole to $r \approx 300\,h^{-1}$ Mpc, which may be at odds with the form of the ΛCDM $P(k)$ at large scales (see also Frith et al. 2003, Busswell et al. 2004). It is therefore interesting to compare the form of high-order clustering statistics in the galactic caps.

We plot the 2-point angular correlation functions and high-order scaling parameters for $b \geq 10°$ and $b \leq -10°$ $K_s < 13.5$ 2MASS galaxies in the third column of Fig. 5.2. We observe a similar discrepancy in the 2-point function as determined previously by Maller et al. (2005) with a steeper slope in the southern $\bar{\omega}_2$ at large scales. However, the high-order scaling parameters in the northern and southern local galaxy distributions, for $p \leq 4$ at least, are consistent with each other and constant values over ≈ 3 orders of magnitude of angular scales to $\theta \lesssim 20°$. At higher orders there are differences between the two which increase with p. Whether these differences are due to real north-south differences in the local galaxy distribution, or simply due to other systematic effects is unclear.

5.5 Discussion

In the previous section, high-order angular and real space hierarchical scaling parameters were determined to extremely large scales ($r \gtrsim 40\,h^{-1}$ Mpc) from a sample of 650 745 $K_s < 13.5$ 2MASS galaxies. Unlike previous high-order clustering analyses, the scaling relations for $p \lesssim 4$ are relatively robust even when the most prominent clusters are removed from the sample.

As described in sections 5.1 and 5.2, high-order clustering statistics have significant implications for the nature of the primordial density distribution and also the way in which galaxies trace the underlying mass distribution. In the following section we examine the

consistency of our results with Gaussian initial conditions and hierarchical clustering and discuss the implications for non-Gaussian models of the primordial density distribution. Assuming the former, we also infer constraints on non-linear galaxy bias.

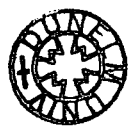
5.5.1 Implications for the Primordial Density Field

5.5.1.1 Consistency with Gaussian Initial Conditions

The nature of the primordial distribution of density fluctuations is predicted to be close to or exactly Gaussian in standard inflationary models (Falk et al. 1993, Gangui et al. 1994, Lesgourgues 1997, Wang & Kamionkowski 2000, Maldacena 2002, Acquaviva et al. 2003). This occurs as a consequence of the slow-roll conditions on the inflation potential which require the potential energy of the field to dominate over the kinetic energy in order to produce a phase of accelerated expansion which lasts for a sufficiently long period of time. Deviations from Gaussianity under these assumptions depend on the inflationary model but are generally extremely small.

Under the assumption of Gaussian initial conditions, a hierarchical scaling (see equation 5.10) of the high-order moments of the dark matter and galaxy density fields (Fry & Gaztañaga 1993) is expected through the evolution under gravitational instability of the primordial density fluctuations (e.g. Peebles 1980, Fry 1984, Bouchet et al. 1992, Bernardeau 1992, Bernardeau 1994a, Gaztañaga & Baugh 1995, Baugh et al. 2004). This holds on scales where the gravitational collapse of dark matter halos evolves linearly; it is important therefore to examine high-order moments in the linear and quasi-linear regime.

It is clear from Fig. 5.1 that our results provide an excellent agreement with the expected scaling relations at large scales, with s_p and S_p parameters consistent with constant values over several orders of magnitude in scale for $p \leq 7$. As noted previously by the 2dFGRS (Baugh et al. 2004), this hierarchical scaling extends to smaller scales than expected by perturbation theory predictions (Bernardeau et al. 2002). This consistency with Gaussian initial conditions is also apparent from Tables 5.1 and 5.2 where we perform fits to the scaling parameters to unprecedented scales; there is good agreement between the small and large scale fits (although there is a small correlation between the two). Alternatively, we can examine the measured slope of the S_p parameter at the scales of interest; in Fig. 5.4 we determine confidence limits for the slope and amplitude of the S_3 parameter at large scales ($4.0 < \bar{r} < 40\,h^{-1}\,\mathrm{Mpc}$) as shown in Fig. 5.1. The constraints provide excellent agreement with the expected constant value as a function of scale, with a



best fit slope of $\gamma_S = -0.02^{+0.34}_{-0.44}$ (where $S_3 \propto r^{\gamma_S}$, marginalising over the normalisation).

This consistency with hierarchical scaling represents a departure from recent constraints on high-order correlation functions. The APM (Gaztañaga 1994) and Edinburgh-Durham Southern Galaxy Catalogue (Szapudi & Gaztañaga 1998) surveys observe significant upturns in the angular skewness at scales of $\theta \gtrsim 1^\circ$ (corresponding to $r \gtrsim 4\,h^{-1}\,\mathrm{Mpc}$). Also, analysis of the 2dFGRS (Croton et al. 2004, Baugh et al. 2004) indicates rising S_p parameters. The issue is complicated by the fact that these measurements are not independent as there is significant overlap between the observed survey fields. Coupled with this, the 2dFGRS identify two superclusters which significantly alter their results; when removed there is good agreement with the expected hierarchical scaling. However, due to the survey volume the scaling parameters are determined only below $r \approx 10\,h^{-1}\,\mathrm{Mpc}$ and do not probe the larger scales of interest here.

5.5.1.2 Constraints on Non-Gaussianity

While the assumption of Gaussian initial conditions is acceptable assuming the validity of the simplest inflationary models, there also exist alternative models of inflation in which deviations from a Gaussian form for the primordial density field are predicted. For example, it is possible to introduce non-Gaussianity if the scalar field driving inflation has more than one component (e.g. Kofman & Pogosyan 1988, Bernardeau & Uzan 2003); strongly non-Gaussian initial conditions are also a feature of models in which the inflaton is not a slowly rolling scalar field but a fast moving ghost condensate (Arkani-Hamed et al. 2004). Additionally, it is possible to construct models in which the initial conditions are non-Gaussian as a result of non-linear structures, such as cosmic strings or global textures, within the primordial density field (e.g. Gaztañaga & Mahonen 1996, Avelino et al. 1998, Gaztañaga & Fosalba 1998).

When considering the effect of non-Gaussian initial conditions on moments of the local galaxy density field, it is convenient to consider the departure from primordial Gaussianity in two ways. Here, we consider the strongly non-Gaussian regime as might arise from cosmic strings or textures; weak departures from Gaussianity as might arise from the various models of inflation described above are more difficult to constrain since the effect can be a shift in the s_p and S_p parameter amplitudes rather than a change in slope. In the case of strong departures from Gaussian initial conditions, the scaling of high-order moments of the density field is expected:

$$\bar{\xi}_p = B_p \, \bar{\xi}_2^{p/2}, \tag{5.20}$$

where for non-Gaussian models seeded by topological defects the B_p scaling parameters (not to be confused with B_p used in equation 5.15) are expected to be constant at large scales and of order unity (Turok & Spergel 1991, Gaztañaga & Mahonen 1996, Bernardeau et al. 2002). Equivalently, the typical signature of strong non-Gaussianity in the S_p scaling parameters is a slope of $\gamma_S \gtrsim 0.6$ (where $S_p \propto r^{\gamma_S}$) at large scales, with a characteristic minimum at $r \approx 10\,h^{-1}\,\mathrm{Mpc}$ (Gaztañaga & Mahonen 1996).

In Fig. 5.5 we constrain the slope and amplitude of the B_3 parameter through χ^2 fits to the 2MASS results (measuring the projected $B_3 = \bar{\omega}_3/\bar{\omega}_2^{3/2}$ and using the transformation to real space described in equation 5.15) at large scales ($4.0 < \bar{r} < 40\,h^{-1}\,\mathrm{Mpc}$). We are able to reject the scaling expected in strongly non-Gaussian models (i.e. constant B_p parameters, $\gamma_B = 0$) as described in equation 5.20, and therefore non-Gaussian models seeded by topological defects, at the $\approx 2.5\sigma$ confidence level.

5.5.2 Non-Linear Galaxy Bias

High-order clustering analysis also represents a powerful probe of the way in which galaxies trace the underlying mass distribution. In chapter 4, we constrained the galaxy bias associated with the variance of the density field to $b_1 > 1$ in the K_s-band at the $> 3\sigma$ level; we also noted in section 5.1 that $b_1 \approx 1$ for optically-selected galaxies. It is unclear whether deviations from the linear bias model exist, and if they do, how this coincides with current theories of galaxy evolution.

Assuming Gaussian initial conditions, we use predictions from perturbation theory for the dark matter skewness and the relationship between the dark matter and galaxy skewness (equations 5.12 and 5.13) to compute the K_s-band non-linear bias coefficient $c_2 = b_2/b_1$ implied by our results. We use an $n = -2$ power spectrum slope (Percival et al. 2001, Cole et al. 2005) and a K_s-band linear bias of $b_1 = 1.39 \pm 0.12$ measured from the $K_s < 13.5$ 2MASS galaxy angular power spectrum analysis in chapter 4. Since these predictions from perturbation theory are valid only in the quasi-linear and linear regimes (Bernardeau 1994b), we use only the fits on large scales. First, we use the best fit galaxy skewness in the range $4.0 < \bar{r} < 40\,h^{-1}\,\mathrm{Mpc}$ (see Table 5.2); we find that $c_2 = 0.57 \pm 0.41$. Since the conversion from s_3 to S_3 becomes increasingly uncertain at scales of $\theta > 2^\circ$ (Gaztañaga 1994), we also use a narrower range of $4.0 < \bar{r} < 10\,h^{-1}\,\mathrm{Mpc}$ (equivalent to

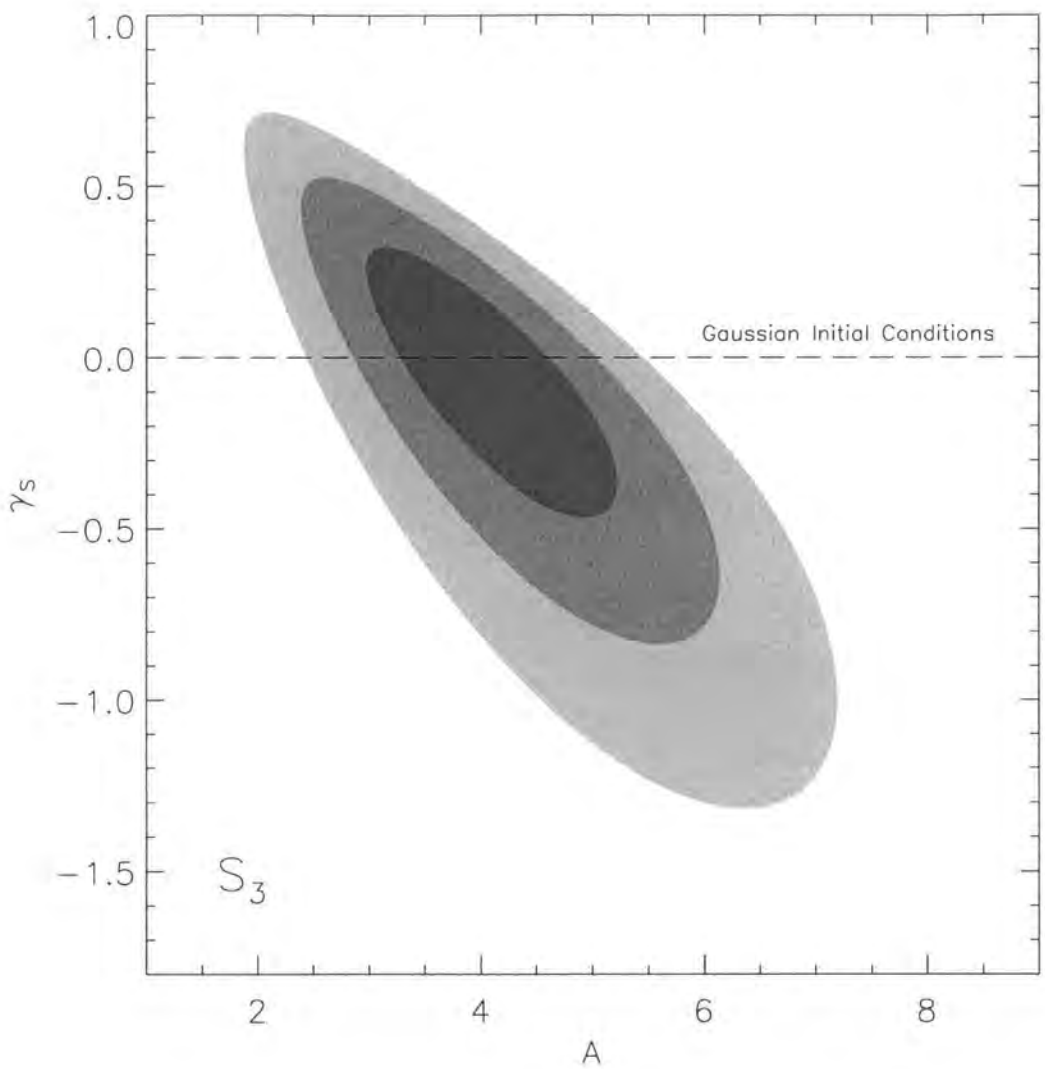


Figure 5.4: Filled contours representing the 1σ, 2σ and 3σ confidence regions for the real space galaxy skewness slope and amplitude determined from χ^2 fits (accounting for the covariance) to the 2MASS $|b| \geq 10°$ K_s <13.5 S_3 parameter (as shown in Fig. 5.1) at large scales ($4.0 < \bar{r} < 40\,h^{-1}$ Mpc). The best fit parameters are A=4.0 and γ_S=-0.02, where we model the skewness using $S_3 = A\, r^{\gamma_S}$. The dashed line indicates the predicted constant form (i.e. γ_S=0) for S_3 in the case of a Gaussian distribution of primordial density fluctuations.

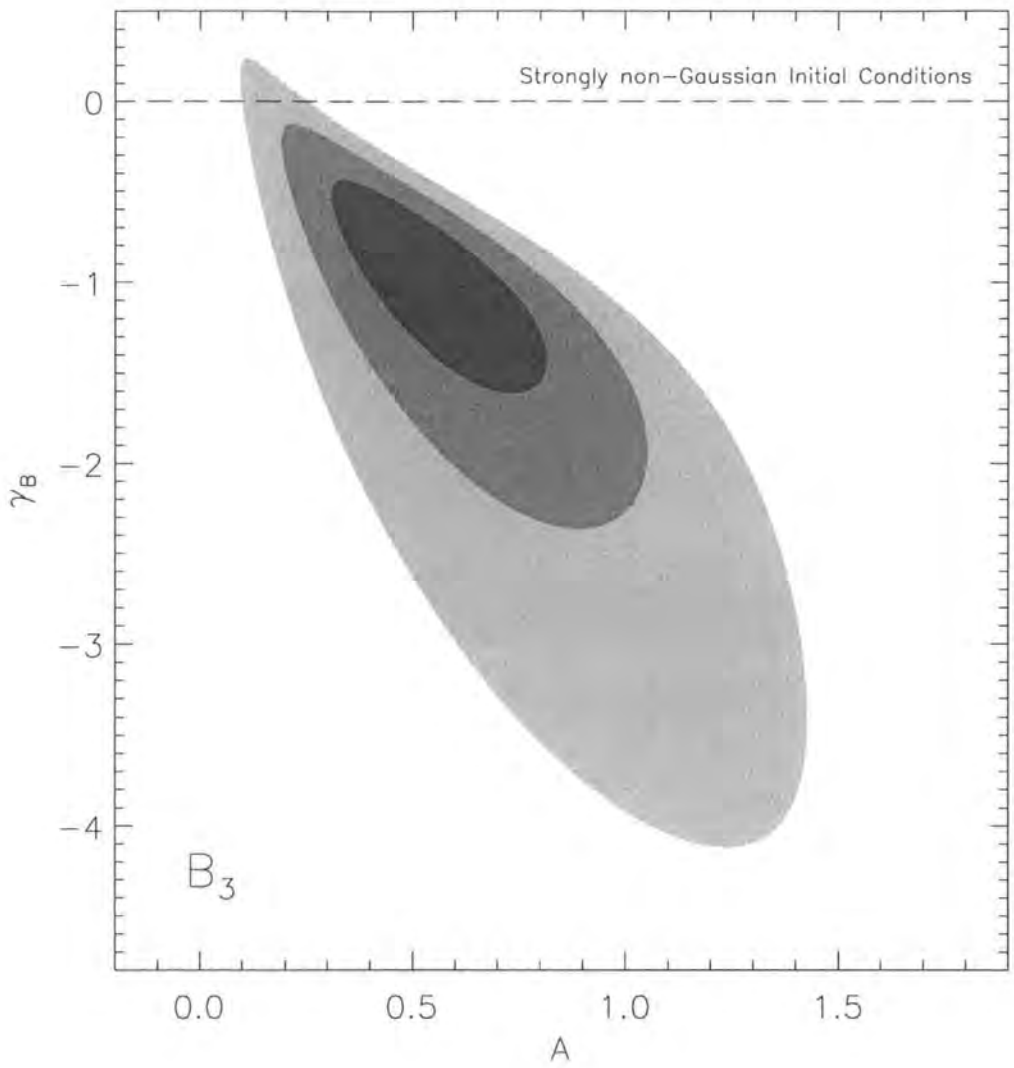


Figure 5.5: Filled contours representing the 1σ, 2σ and 3σ confidence regions for the real space B_3 parameter (see equation 5.20) slope and amplitude determined from χ^2 fits as in Fig. 5.4. The best fit parameters are A=0.53 and γ_B=-0.93, where we use $B_3 = A\, r^{\gamma_B}$ as before. The dashed line indicates the predicted constant form for B_3 in the case of a strongly non-Gaussian distribution of primordial density fluctuations.

$1°.0 < \theta < 2°.5$). Using the best fit galaxy skewness in this range of S_3=4.01±0.34, we find that $c_2 = 0.57 \pm 0.33$.

We therefore detect a *positive* quadratic contribution to the K_s-band galaxy bias at the $\approx 2\sigma$ level. This means that the density of 2MASS galaxies rises more quickly than the mass density contrast. This differs from all previous constraints on the c_2 non-linear bias parameter which have been *negative*, most recently with constraints from the 2dFGRS which limit $b_1 = 0.94^{+0.13}_{-0.11}$ and $c_2 = -0.36^{+0.13}_{-0.09}$ in the optical b_J-band (Gaztañaga et al. 2005), and also from the IRAS PSCz catalogue of $b_1 = 0.83 \pm 0.13$ and $c_2 = -0.50 \pm 0.48$ in the infra-red (Feldman et al. 2001). This compares to near infra-red K_s-band constraints of $b_1 = 1.39 \pm 0.12$ (chapter 4) and the optimal constraint in this analysis of $c_2 = 0.57 \pm 0.33$.

It is possible to understand these results by examining the analytic predictions of Mo et al. (1997) for the high-order bias coefficients, formed via the Press & Schechter (1974) formalism and an initially Gaussian density field. From this the first two terms in the Taylor expansion of equation 5.11 are predicted to be:

$$b_1 = 1 + \frac{\nu^2 - 1}{\delta_c} \tag{5.21}$$

$$b_2 = 2\left(1 - \frac{17}{21}\right)\frac{\nu^2 - 1}{\delta_c} + \left(\frac{\nu}{\delta_c}\right)^2(\nu^2 - 3), \tag{5.22}$$

where $\nu = \delta_c/\sigma(M)$, $\sigma(M)$ denotes the linear *rms* fluctuation on the mass scale of a dark matter halo of mass M and δ_c is the linear theory over-density at the time of collapse (for reference δ_c=1.686 for Ω=1).

This relatively simplistic scenario is not able to provide accurate quantitative predictions which match the observational results above. However, by considering halos of differing mass it is possible, qualitatively at least, to understand these apparently contradictory constraints on c_2. For instance, if we consider the most massive halos for which $\nu^2 \gtrsim 3$, Mo et al. (1997) predict $b_1 > 1$ and $c_2 > 0$. In contrast, for typical mass halos for which $\nu^2 \approx 1$, bias parameters of $b_1 \approx 1$ and $c_2 < 0$ might be expected. With this in mind, it is possible to understand why a near infra-red survey, which is more sensitive to early-type galaxies than optical or infra-red surveys (e.g. Jarrett 2004, Croton et al. 2005, Oliver et al. 1996, respectively), produces galaxy samples with higher values for the linear and quadratic bias parameters.

5.6 CONCLUSIONS

We have measured reduced angular correlation functions, $\bar{\omega}_p$, to ninth order using 650 745 galaxies selected from the 2MASS extended source catalogue. From our estimates for the angular correlation functions, we have determined the projected and real space hierarchical scaling parameters, s_p and S_p respectively. The prime motivation for such analysis is to test the hierarchical scaling hypothesis which predicts these parameters to be constant in the linear and quasi-linear regimes (e.g. Peebles 1980, Fry 1984, Bouchet et al. 1992, Bernardeau 1992). As such we are able to probe the primordial density field and constrain various models of inflation and structure formation. High-order clustering analysis also allows us to probe the way in which galaxies trace the underlying mass distribution; a negative offset between ΛCDM predictions and observations by the 2dFGRS have recently been interpreted as evidence for a quadratic contribution to the galaxy bias, although these conclusions are based on constraints in the weakly non-linear regime (Gaztañaga et al. 2005).

The most comparable recent work are the analyses of the 2dFGRS (Croton et al. 2004, Baugh et al. 2004) and APM surveys (Gaztañaga 1994); the galaxy sample used in this work represents an order of magnitude increase in volume and solid angle over each respectively. Previous analyses of high-order clustering statistics have proved extremely challenging; due to the relatively small volumes probed in 3-dimensional analyses for instance, the results have been sensitive to the presence of clusters and superclusters within the galaxy sample. Not only this, but direct comparisons with perturbation theory have also proved difficult since the statistical uncertainty at large scales is considerable, and frequently the covariance in the statistics has been ignored leading to unrealistically small errors. In addition, the results for the hierarchical scaling parameters have consistently displayed a puzzling upturn in values at $r \approx 4\,h^{-1}\,\mathrm{Mpc}$, consistent with some models of structure formation with strongly non-Gaussian initial conditions (Gaztañaga & Mahonen 1996, Gaztañaga & Fosalba 1998, White 1999, Bernardeau et al. 2002).

Here, we are able to determine the scaling parameters to high accuracy to unprecedented scales, $r \lesssim 100\,h^{-1}\,\mathrm{Mpc}$. We also carry out a full covariance analysis in order to take account of correlations in the datapoints at different cell radii. We are therefore in a position to make direct comparisons with the predictions of perturbation theory since we probe well into the quasi-linear and linear regimes and we have a good understanding of the statistical uncertainty. However, since we are working with a projected galaxy

sample, we have to convert the associated angular scaling parameters to real space via a transformation which becomes increasingly uncertain on large scales ($\theta > 2°$). Nevertheless, this work currently provides the best estimates of high-order clustering statistics at large scales. We are able to reach a number of conclusions:

(i) Our results are in line with the expected hierarchical scaling relation, with s_p and S_p parameters consistent with constant values over $\approx$3 orders of magnitude in scale to $r \approx 40\,h^{-1}\,$Mpc for $p \leq 7$; we constrain the slope of S_3 to $\gamma_S = -0.02^{+0.34}_{-0.44}$ (where $S_p \propto r^{\gamma}$). Such a scaling pattern is expected if an initially Gaussian density field evolves under the action of gravitational instability (e.g. Peebles 1980, Fry 1984, Bouchet et al. 1992, Bernardeau 1992). This result is in contrast to recent results drawn from the 2dFGRS, APM and EDSGC surveys (Gaztañaga 1994, Szapudi & Gaztañaga 1998, Baugh et al. 2004, Croton et al. 2004) which display rising scaling parameters at large scales.

(ii) The scaling parameters are relatively robust to the removal of the largest over-density, the Shapley supercluster, although the best fit slope of s_3 at large scales ($1°.0 < \theta < 10°$) becomes steeper, yet is still consistent with a constant value ($\gamma = -0.27^{+0.25}_{-0.29}$). We also use a more stringent cut by removing the ten most over-dense regions (see Fig. 5.3 and Table 5.3) corresponding to a cut to the main sample of 6.3 per cent of the galaxies and 2.6 per cent of the total solid angle; the resulting s_3 parameter remains broadly consistent with the result for the main sample, although with a slightly steeper best fit slope at large scales ($\gamma = -0.33^{+0.26}_{-0.29}$).

(iii) Since strong non-Gaussianity in the primordial density field, as might be expected in models seeded by topological defects such as cosmic strings or global textures (Avelino et al. 1998, Gaztañaga & Fosalba 1998, Gaztañaga & Mahonen 1996), results in a strong upturn in the hierarchical scaling parameters at large scales (e.g. Bernardeau et al. 2002), we are able to reject strongly non-Gaussian initial conditions, producing relations of the form $\bar{\xi}_p \propto \bar{\xi}_2^{p/2}$, at the $\approx 2.5\sigma$ confidence level.

(iv) We compare our constraints on S_3 at large scales (where we have assumed a constant value) to predictions from perturbation theory. We detect a significant deviation consistent with a non-linear, quadratic contribution to the K_s-band galaxy bias, parameterised as $c_2 = 0.57 \pm 0.33$ (derived from fits in the range $4.0 < \bar{r} < 10\,h^{-1}\,$Mpc), implying that the 2MASS galaxy density rises more quickly than the mass density contrast. This *positive* result represents a significant difference from the negative values found previously; constraints on c_2 from the optically-selected 2dFGRS and the infra-red PSCz samples yield negative values. We explain these apparently contradictory results through

an examination of the model of Mo et al. (1997), which predicts bias parameters of $b_1 > 1$ (the linear bias) and $c_2 > 0$ (the quadratic bias) if the surveyed galaxies typically reside in large mass halos. Similarly, as observed previously, we might expect $b_1 \approx 1$ and $c_2 < 0$ from galaxy samples in which late-type galaxies are over-represented.

CHAPTER 6

AN EXTENDED SZ EFFECT IN WMAP?

In the previous four chapters, we have examined the clustering of galaxies on large scales directly, investigating a large inhomogeneity in the bright number counts and the variance and higher order moments of the local galaxy density field. In this chapter, we now use catalogues of large clusters of galaxies to focus on another potential problem for modern cosmology mentioned in chapter 1; the extent to which Cosmic Microwave Background (CMB) observations and the associated cosmological constraints are affected by foreground contamination. In particular, we examine the Sunyaev-Zeldovich (SZ) effect, the inverse Compton scattering of CMB electrons by the hot gas in galaxy clusters, and its effect on the Wilkinson Microwave Anisotropy Probe (WMAP) results.

6.1 INTRODUCTION

The first year data release of the WMAP CMB anisotropy experiment has provided accurate confirmation that the first acoustic peak in the temperature power spectrum occurs at $l = 220 \pm 10$ and further constrained the form of primordial temperature fluctuations to $l \approx 500$ (Hinshaw et al. 2003). These results appear to provide further support for the standard ΛCDM cosmology (Spergel et al. 2003). However, an intriguing result from WMAP is the detection of polarisation at large scales arising from an epoch of reionisation at $10 < z < 20$ (Kogut et al. 2003). This might significantly reduce the amplitude of the acoustic peaks in the CMB power spectrum through Thomson scattering. In addition, WMAP also detected a low quadrupole in the power spectrum (Hinshaw et al. 2003); this is not expected if the Integrated Sachs-Wolfe (ISW) effect, arising from the relatively recent domination of dark energy effects, is present. It is clear therefore, that while WMAP provides strong support for the concordance model, there are a number of significant complications and strong evidence that the data is affected by cosmic foregrounds.

The detection of foreground contamination in the first year WMAP data has been investigated by a number of authors. The WMAP team (Bennett et al. 2003) list 208

point sources detected at $> 5\sigma$ in the WMAP data and identify them as radio galaxies and quasars. Giommi & Colafrancesco (2003) also claim to detect significant blazar contamination in WMAP and Boomerang data. The search for the ISW effect has proved fruitful: Boughn & Crittenden (2004) and Nolta et al. (2003) have claimed evidence for the ISW effect in the WMAP data from cross-correlation with the NVSS catalogue. Diego, Silk & Sliwa (2003) have cross-correlated X-ray data and the WMAP data. Similarly Fosalba & Gaztañaga (2003) have cross-correlated APM galaxies and WMAP data; they find a marginal detection of the ISW effect at 5-10 deg scales and suggest that the lack of a detection at smaller scales may be due to cancellation with other foreground effects.

Here, we investigate another possible contaminant of the WMAP data arising from low redshift processes; the thermal SZ effect. Various authors have made model-dependent predictions for the level of contamination in WMAP data by the SZ effect and usually conclude that the contaminating effects are small (Refregier et al. 2000a, Refregier et al. 2000b, Komatsu & Kitayama 1999). Afshordi, Loh & Strauss (2003) have claimed the detection of SZ, ISW and point sources in a power spectrum analysis of WMAP data and the 2MASS galaxy catalogue. The WMAP team (Bennett et al. 2003) find a 2.5σ detection of the SZ effect in the XBACS sample of 242 X-ray bright Abell clusters (Ebeling et al. 1996). However, they only looked for SZ decrements on the scale of the WMAP beam and did not explore any larger scales. Hernandez-Monteagudo & Rubino-Martin (2003) have obtained upper limits on the SZ effect from superclusters from failing to find any cross-correlation with the Abell-Corwin-Olowin (ACO) and other cluster catalogues but they did find significant correlations from individual clusters.

The chapter is set out as follows: In section 6.2 we give details of the CMB data and foreground galaxy cluster catalogues used. The cross-correlation technique and error analyses are outlined in section 6.3, the results of which are presented in section 6.4. Possible implications for our understanding of cosmological parameters and CMB power spectrum fits are discussed in section 6.5. The conclusions follow in section 6.6.

6.2 DATA

6.2.1 THE WILKINSON MICROWAVE ANISOTROPY PROBE DATA

WMAP has collected and published data in five bands: W (94 GHz), V (61 GHz), Q (41 GHz), Ka (33 GHz) and K (23 GHz). In this work we shall principally use the W band data because of its relatively high resolution compared to the other bands (the

FWHM is 12′.6 compared to 19'.8, 29'.4, 37'.2 and 49'.2 respectively), and also since it is most sensitive to the SZ effect. Although, none of the beams is exactly Gaussian (see Fig. 2 of Page et al. (2003)) we have found that Gaussians of the above FWHM are good approximate fits to cross-correlation results between faint radio point sources and WMAP data. We will also use the Internal Linear Combination (ILC) map produced by the WMAP team, which combines each of the five bands with constant weights. The data is published in HEALPix format, with equal area pixels of $49\,\mathrm{arcmin}^2$. Where necessary we use the Kp0 WMAP mask (Bennett et al. 2003) which mainly masks Galactic contamination; its effect in this work is small as we shall generally be working at Galactic latitudes of $|b|$ >40°. The maps all use thermodynamic temperature and the cosmological dipole has already been subtracted from the data by the WMAP team.

6.2.2 Galaxy Cluster Catalogues

We use galaxy group and cluster data selected from three catalogues. First, we use the ACO catalogue of Abell, Corwin & Olowin (1989) which lists clusters with 30 or more members within a $1.5h^{-1}\,\mathrm{Mpc}$ radius within 2 magnitudes of the 3rd brightest cluster member. A richness class is assigned from this ($0 \leq R \leq 5$) with Coma classed as $R = 3$. The Northern catalogue with $b > 40°$ lists 2489 clusters with $R \geq 0$; the Southern catalogue lists 1346 for $b < -40°$. For reference, the sky density in the North is therefore $0.52\,\mathrm{deg}^{-2}$ and in the South it is $0.28\,\mathrm{deg}^{-2}$. Of importance later, for $R \geq 2$ clusters, the sky density is $0.063\,\mathrm{deg}^{-2}$ with an average redshift of $z = 0.15$.

Second, we use the galaxy group and cluster catalogues of Myers et al. (2003), derived from the APM Galaxy Survey (Maddox et al. 1990a). The survey area covers $\delta < -2.5\,°$ and $b < -40\,°$. The groups and clusters were identified using an algorithm in which the radius of a circle, laid down around each APM galaxy group or cluster (with $B < 20.5$), is increased until the overdensity (β) falls to $\beta = 8$; those galaxies whose circles overlap are called groups. Here, we include groups and clusters of minimum memberships $m \geq 7$ and $m \geq 15$ respectively, where m denotes the number of galaxies within the overlap region. These membership limits define minimum group effective 'radii' of 1.′2 and 1.′7, since the APM galaxy surface density is $N \approx 750\,\mathrm{deg}^{-2}$ at $B < 20.5$. We take an average redshift of $z = 0.1$ for both APM samples. For reference, the sky density of groups and clusters is $3.5\,\mathrm{deg}^{-2}$ and $0.35\,\mathrm{deg}^{-2}$ respectively. At $m \geq 15$ there are considerable differences between the ACO and APM catalogues. For instance, an $R = 0$ Abell cluster at $z = 0.15$ may contain a galaxy sky density of $260\,\mathrm{deg}^{-2}$ within its 11.′5 Abell radius compared to a

minimum sky density of $5250\,\mathrm{deg}^{-2}$ for galaxies within the APM groups. For this reason, the $R \leq 1$ ACO clusters may accommodate much lower galaxy densities than even the $m \geq 7$ APM groups which are guaranteed to sample higher density regions, albeit over smaller areas.

Our third galaxy cluster sample is derived from the final data release of the 2MASS extended source catalogue (Jarrett 2004) as described in chapter 2. We use the algorithm mentioned above on $K_s \leq 13.5$ galaxies; we detect 500 groups and clusters with $m \geq 35$ members at the density contrast $\beta = 8$ for $|b| \geq 10\,^\circ$ with average redshift $z \approx 0.07$. K-selected galaxy samples are dominated by early-type galaxies because of their red colours and early-type galaxies are the most common galaxy-type found in rich galaxy clusters. Therefore the 2MASS survey provides an excellent tracer of the high density parts of the Universe out to $z < 0.15$ and so provides a further test for the existence of the SZ effect.

6.3 Method

6.3.1 Cross-Correlation Technique

In order to search for correlations between the WMAP data and the foreground cluster catalogues listed in section 6.2.2., we use a simple cross-correlation technique. The average temperature difference (with respect to the mean), ΔT, of the CMB is determined in annuli as a function of the angular distance θ between each cluster or group and each HEALPix pixel centre. The mean temperature difference is then computed as a function of angular scale and compared to the mean WMAP ΔT over the foreground survey solid angle. Since the mean temperature difference of the WMAP data is not necessarily zero even over large areas, we plot the mean ΔT as an effective zero level (solid lines).

6.3.2 Error Analysis

Due to the presence of clustering in the cluster catalogues, monte-carlo errors using randomly distributed sources are likely to under-estimate the errors at large scales for the galaxy cluster samples. This is due to the preferential sampling of particular regions of sky in a clustered catalogue; a greater spread is thus expected for monte-carlo errors using clustered mock catalogues. Clustered mock catalogues are created using a new technique; points are first randomly distributed over a region larger than the survey solid angle. Each point or mock galaxy cluster is assigned a weight, $(1 + w)$, derived from every other mock cluster, where w is the value of an input 2-point correlation function. The assigned

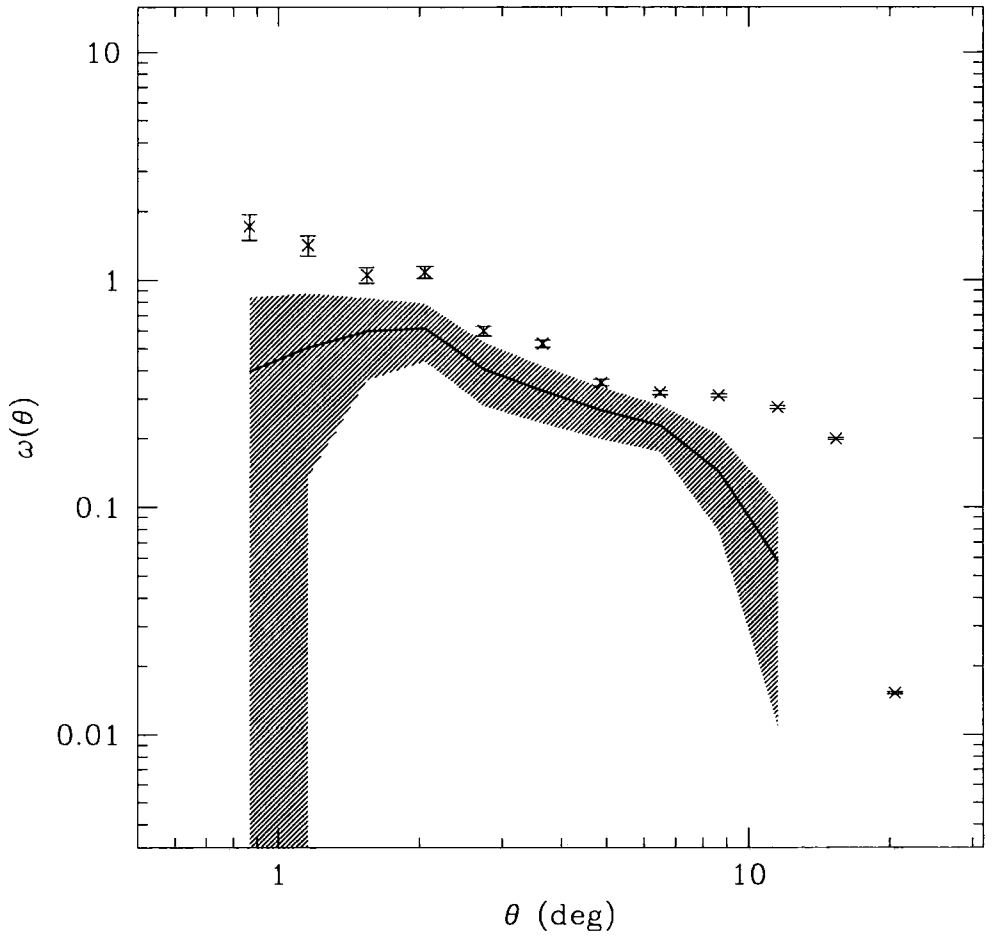


Figure 6.1: The angular correlation function of the 377 Southern ACO clusters (crosses) and the mean angular correlation function and 1σ spread for the corresponding one hundred mock catalogues (solid line and shaded region). The errorbars show the Poisson uncertainty in each bin.

weight is therefore dependent on an input slope, γ, and amplitude s_0. Having assigned weights to each point, the $(n-1)^{th}$ root is taken for each point (where n is the number of points laid down) and these are rescaled such that each new value lies between 0 and 1. Each mock cluster is then rejected with a certain probability (its rescaled weight); any remaining outside the survey field are excluded and outstanding excesses are then removed randomly. The mock cluster catalogue that remains has the same number of points as the real cluster catalogue by design. It is also necessary to ensure that the initial random distribution of mock cluster positions is laid out over a larger area than the survey solid angle, as mock clusters are preferentially down-weighted at the edge of the field due to fewer clusters in close proximity. Poorly clustered mock catalogues are then rejected using a χ^2 analysis, determined between the angular correlation functions of the mock catalogue and the data. In each case one hundred independent mock catalogues are used.

In practice, the weighting system needs an extremely large input slope and normalisation, as well as additional tweaking such that the weights are increased at small relative to large scales using a simple step function. A fairly consistent 2-point correlation function for the ACO mock catalogues for instance is then produced (see Fig. 6.1). However, problems were encountered when using this clustering method for the APM m$\geq$15 and m$\geq$7 sample, due to the large numbers of clusters involved and the consequential smaller differences between the largest and smallest assigned weights. As a result, we estimate clustered monte-carlo errors for the APM samples by first using random, unclustered mocks; the standard deviation calculated from these is scaled up using the difference between the clustered and unclustered ACO errors for each bin. These scaled errors should provide an over-estimate of the error, as the clustering is considerably weaker in the APM m$\geq$15 and m$\geq$7 samples than in the ACO cluster catalogue used.

As an additional powerful constraint on the systematic errors, the ACO clusters were also rotated around the galactic poles. The variation in the mean ΔT around each cluster as a function of scale and rotation in galactic longitude provides an additional estimate on the error.

6.4 RESULTS

In Fig. 6.2a and b we show the cross-correlations of the APM clusters and groups respectively with the WMAP W band data. Both datasets indicate an anti-correlation with

respect to the mean ΔT over the survey area (solid line), although there is considerable covariance between each bin. Accounting for this, the integrated significances are shown in the first two columns of table 6.1. The APM cluster sample ($m \geq 7$) appears to indicate a significant anti-correlation signal within the beam (12'.6). However, for the APM galaxy groups, while the amplitude of the signal is slightly lower than for the clusters ($\Delta T \approx -0.008\,\mathrm{mK}$ compared to $\Delta T \approx -0.01\,\mathrm{mK}$), the anti-correlation appears to extend beyond the beam, only reducing to 1.5σ at $\theta \approx 60'$. In general, the amplitude of the temperature decrement with both APM datasets is consistent with the expected signal at small scales with the thermal SZ effect (Refregier et al. 2000a). However, it would be surprising both if this signal extended to scales of $\approx 1^\circ$ and if it persisted with galaxy groups with a sky density of $3.5\,\mathrm{deg}^{-2}$, a factor $\approx 30\times$ higher than that of the Abell clusters.

In Fig. 6.2c we examine the cross-correlation of ACO $R \geq 2$ clusters with the WMAP W band. This sample includes 229 clusters with $b > 40^\circ$ and 377 clusters with $b < -40^\circ$. Again, as with the APM data an anti-correlation is measured with respect to the mean ΔT over the associated area of sky. Accounting for the covariance from bin to bin, the significance calculated as a function of scales is shown in the third column of table 6.1. As with the APM groups, the signal is significant to extremely large scales; $\approx 2.2\sigma$ level at $\theta < 1^\circ$ (equivalent to $\approx 7.5\,h^{-1}\,\mathrm{Mpc}$). This analysis was also carried out for ACO clusters of all richnesses; an insignificant anti-correlation was found. This may be due to the fact that clusters with $R \leq 1$ have too low densities to produce a strong SZ signal. However, it appears that the ACO $R \geq 2$ cluster sample confirms the extended detection made with the APM group and cluster catalogues; given the amplitude of this signal at small scales, it is possible that the observed anti-correlations at large scales are caused also by the SZ effect.

In Fig. 6.2d we cross-correlate the 500 $m \geq 35$ groups drawn from the 2MASS extended source catalogue with the WMAP W band. As with the APM and ACO datasets an anti-correlation is observed to scales of $\approx 1^\circ$, although with marginal significance. At the mean 2MASS group redshift, scales of $\theta < 1^\circ$ corresponds to $r < 3\,h^{-1}\,\mathrm{Mpc}$; this result provides further evidence for possible extended SZ decrements to $> 1\,h^{-1}\,\mathrm{Mpc}$ scales.

We now wish to investigate possible correlations between foreground clusters and the WMAP data as a function of frequency. In Fig. 6.3a-f we show the measured cross-correlation between the ACO $R \geq 2$ clusters and the WMAP W, V, Q, Ka and K bands and also the ILC. As with the W band (Figs. 6.2 and 6.3a) there appears to be anti-correlations at all other frequencies observed by WMAP. Given the W band result this

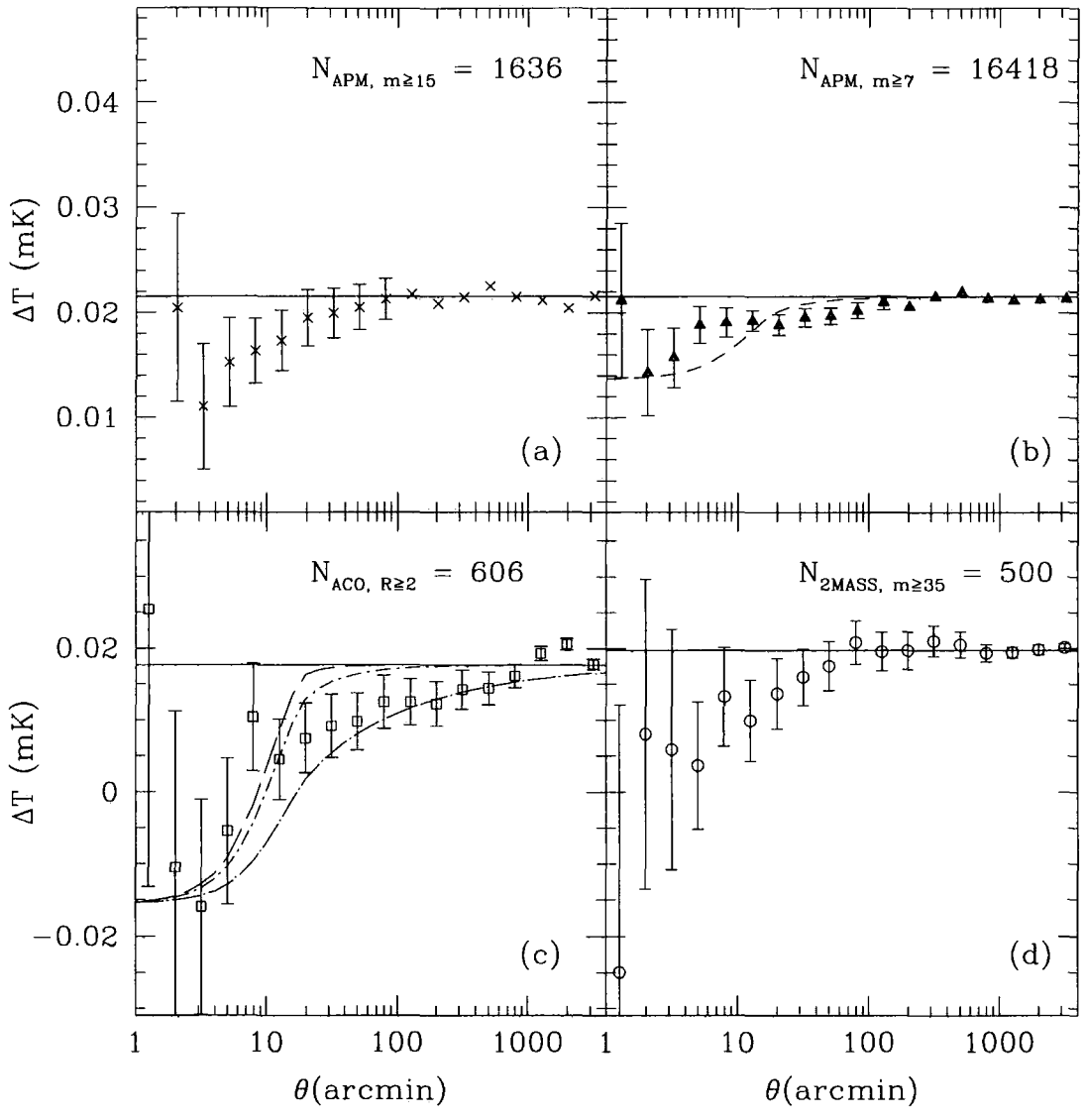


Figure 6.2: Cross-correlation of the WMAP 94 GHz W band data with (a) $m \geq 15$ APM groups and clusters, (b) $m \geq 7$ APM groups and clusters, (c) ACO $R \geq 2$ clusters for $|b| > 40°$ and (d) $m \geq 35$ 2MASS groups and clusters. In each case the solid line indicates the effective zero level, the mean ΔT over the appropriate survey area. In panels (b) and (c) we show isothermal SZ models convolved with a Gaussian beam profile of width $\sigma = 5.'25$. In panel (b) the dashed line includes parameters $\Delta T(0) = 0.015\,\mathrm{mK}$, $\theta_c = 2.'3$ and $\beta = 0.75$. In panel (c) we use parameters of $\Delta T(0) = 0.083\,\mathrm{mK}$, $\beta = 0.75$, $\theta_c = 1.'5$ (dashed line); $\Delta T(0) = 0.050\,\mathrm{mK}$, $\beta = 0.5$, $\theta_c = 1.'5$ (dotted line); and $\Delta T(0) = 0.12\,\mathrm{mK}$, $\beta = 1.0$, $\theta_c = 1.'5$ (dot-dash line). In all cases the Monte-Carlo error estimates are shown indicating the effect from the clustering of clusters.

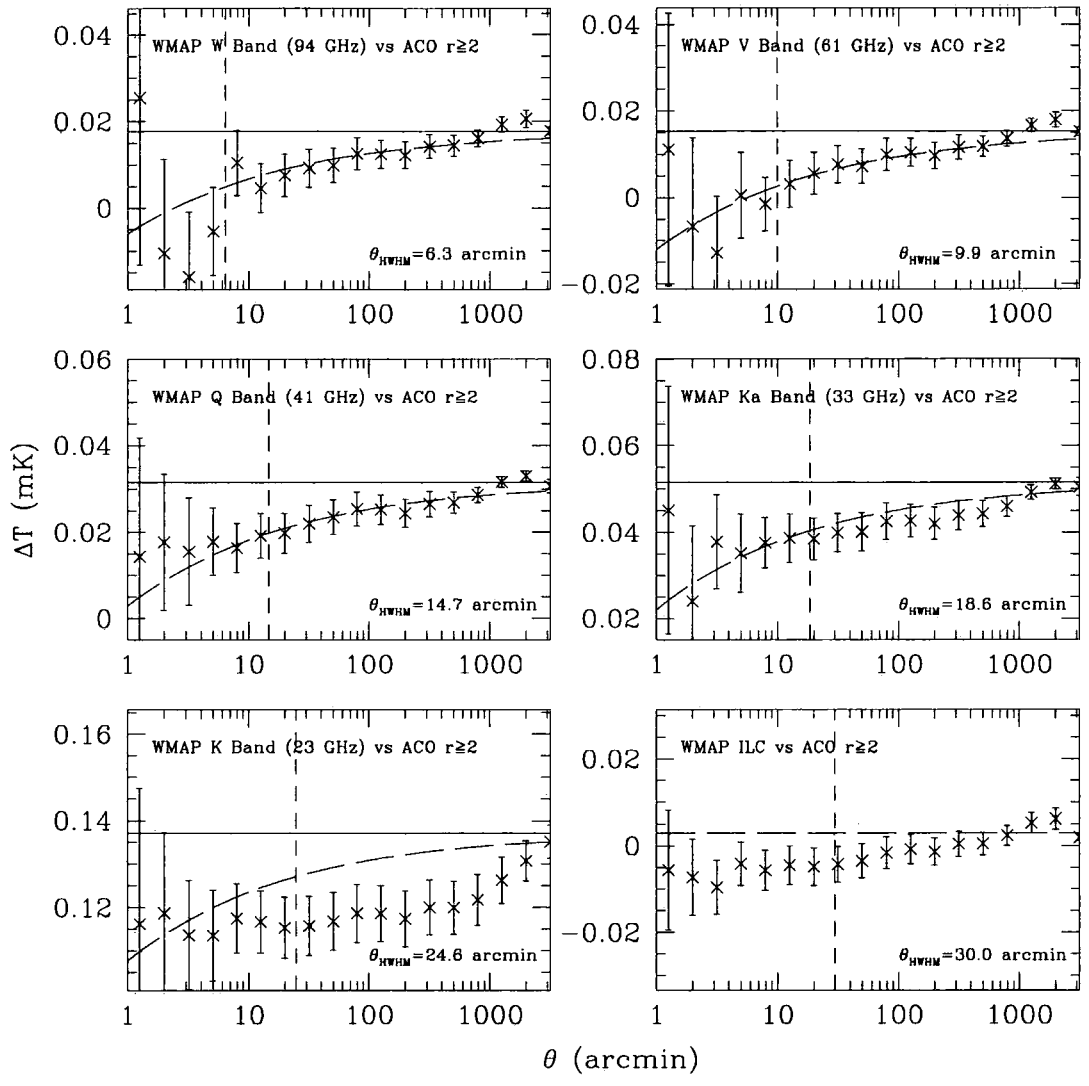


Figure 6.3: Cross-correlation of WMAP data in the W, V, Q, Ka, K and ILC bands with 606 $|b| > 40$ deg ACO clusters. As before, the solid line is the simple average ΔT over the area surveyed. The vertical short-dashed line indicates the beam half-maximum in each band. The long-dashed line shows a $\theta^{-1/3}$ fit to the W band, scaled for the SZ frequency dependence in the other bands; this can be used as a reference line for the frequency dependence.

θ	APM $m \geq 15$	APM $m \geq 7$	ACO R$\geq$2
<6'	1.8σ	2.1σ	2.7σ
<10'	1.7σ	2.0σ	2.1σ
<25'	1.2σ	2.0σ	2.4σ
<40'	0.8σ	1.6σ	2.2σ
<65'	0.6σ	1.5σ	2.2σ
<100'	0.3σ	1.3σ	1.8σ

Table 6.1: Significances of the anti-correlations seen in Fig. 6.2a-c. These are derived from the clustered Monte-Carlo error estimates for the ACO dataset and from scaled clustered errors in the case of the APM cluster catalogues. As such, the APM significances are likely to be slightly under-estimated.

is not unexpected; according to equations (11) and (13) of Refregier et al. (2000a) the SZ decrement relative to the W band should increase by factors of 1.16, 1.21, 1.25 and 1.25 for the V, Q, Ka and K bands respectively. The frequency dependence of the anti-correlation appears therefore to be consistent with the SZ prediction (compare with the dashed lines) in all but the K band; this anomaly is likely to be due to a combination of poor resolution and residual Galactic contamination. However, the errors are too large to discriminate between the SZ and CMB spectral indices. This spectral similarity makes SZ contamination difficult to quantify in the WMAP data. We have also examined the frequency dependence of the APM anti-correlations; there is a significant signal in the V and Q bands but less so in the K and Ka bands, probably due to lower resolution.

In order to check that the observed temperature decrements are not due to residual Galactic foreground contamination, we cross-correlate the ACO clusters with the WMAP foreground maps of Bennett et al. (2003); we find no indication of any strong systematic effect due to Galactic foregrounds. As a further check we have also cross-correlated the APM and ACO group and cluster samples with the cleaned map of Tegmark et al. (2003); we measure an effect consistent with that found with the WMAP ILC. This further supports the idea that the observed anti-correlations are not due to foreground systematics.

As an additional check against systematic effects caused by foreground contamination, we rotate the ACO $R \geq 2$ clusters around the galactic poles and cross-correlate with the high resolution WMAP W and V band data; we therefore determine temperature decrements as a function of the rotating angle. We show the cross-correlation signals

relative to the mean W and V band ΔT over the survey area integrated to the beam half-maximum, $\theta < 6.'3$ for W, $\theta < 9.'9$ for V (open circles), $\theta < 60'$ (crosses) and $\theta < 500'$ (filled circles), determined at intervals in galactic longitude of $\Delta l = 10°$. In order to test whether the temperature decrement within the beam size is significant assuming that the anti-correlation at large scales is due to systematic effects we examine the difference between the beam half-maximum and the $\theta < 500'$ signals (solid line); the beam half-maximum decrement is significant whether marginalising over the large scale signal in this way (2.8σ for W, 3.1σ for V) or comparing to the overall mean ΔT (3.1σ for W, 3.6σ for V). In addition, the $\theta < 60'$ signal also appears to be significant whether measured relative to the large scale $\theta < 500'$ results (2.0σ for W, 1.7σ for V) or the overall mean (2.3σ for W, 2.2σ for V). However, these latter results should be treated with caution as the rotated cross-correlation results are more correlated at $\theta < 60'$ than at beam half-maximum. Nevertheless, the $\Delta l = 0°$ result displays the lowest temperature decrement in the case of the beam half-maximum and $\theta < 60'$ results. In the case of the $\theta < 500'$ points there are several rotated points which show lower temperature decrements; it is likely that there may be systematics which are beginning to dominate over any real SZ signal at these scales. However, this technique supports the suggestion that there is a significant anti-correlation to scales of $\theta < 60'$ and that these signals are robust to possible systematics on $\theta < 500'$ scales.

6.5 Discussion

We have found significant evidence for temperature decrements in the first year WMAP data extending to large scales, $\theta \approx 1°$, with marginal evidence that the signal may extend to even larger scales. Previously, cross-correlation of the ACO catalogue with the Rosat All-Sky Survey indicated diffuse X-ray emission associated with $R \geq 1$ Abell Clusters, extending to scales of $\theta \approx 2°$ from the cluster centres (Soltan, Freyberg & Hasinger 2002); this is comparable to the scale of the anti-correlation measured (see Fig. 6.2c).

One possible cause of the extended temperature decrement might be the cumulative effect of beam size SZ profiles arising from overlapping clusters. The angular correlation function of ACO clusters (Bahcall & Soneira 1983) suggests that the mean number of excess clusters in the range $20' < \theta < 100'$ from an average cluster is ≈ 1.3. We estimate that this cluster excess should contribute $\Delta T = 0.5\,\mu$K at $20' < \theta < 100'$; this compare to the observed decrement of $\Delta T = 6\,\mu$K. Therefore it does not seem possible that a

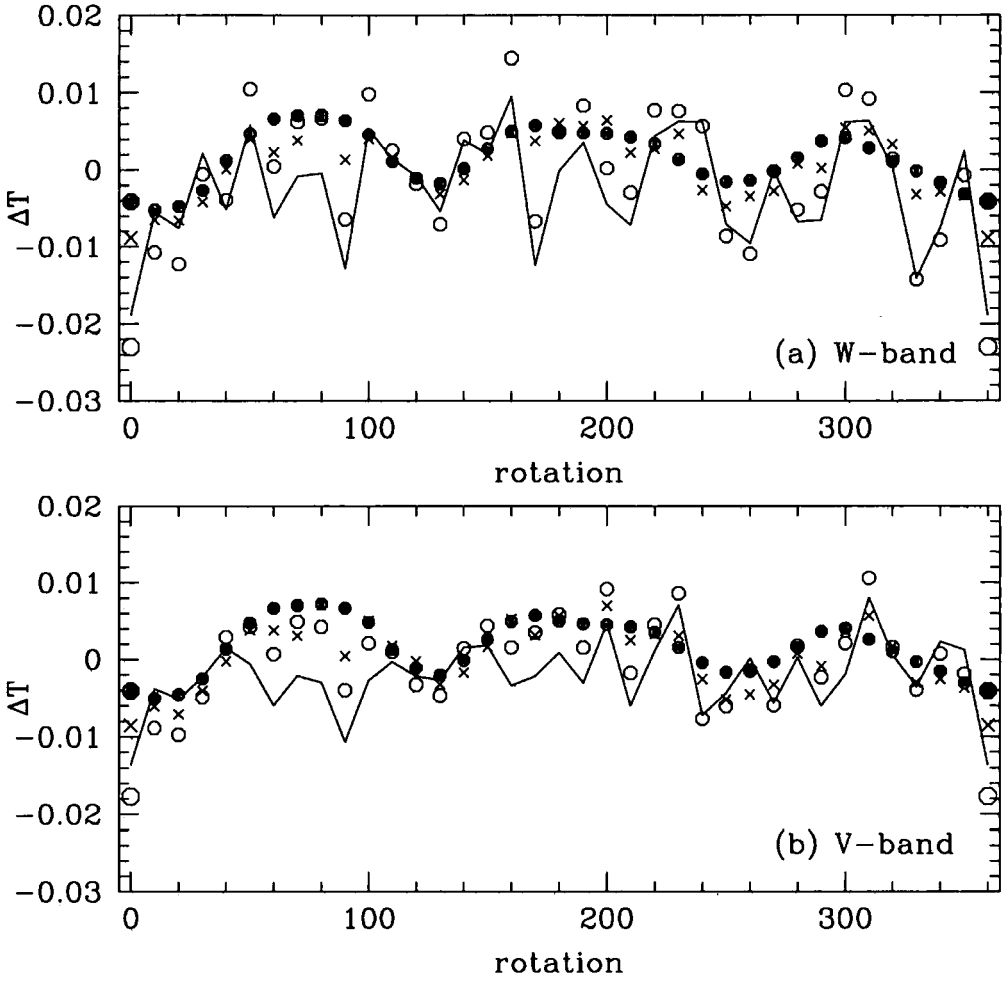


Figure 6.4: Here we cross-correlate the 606 ACO clusters used previously with the (a) W band and (b) V band WMAP data as a function of rotation angle, Δl, around the galactic poles. The open circles represent the mean ΔT within the beam half-maximum ($\theta < 6.'3$ for the W band and $\theta < 9.'9$ for the V band). The crosses represent the average ΔT for $\theta < 60'$. The filled circles represent the average ΔT for $\theta < 500'$. In each case the mean ΔT over the surveyed area has been subtracted. The solid line represents the difference between the beam half-maximum and $\theta < 500'$ results; this indicates the effect on the beam-size result if the anti-correlation observed at $\theta < 500'$ is assumed to be due entirely to systematics.

cumulative SZ effect arising from the clustering of clusters can explain the observed form to the temperature decrement. Instead, an extended ($\theta \geq 1°$) gas halo around individual clusters appears to be needed.

Assuming that the observed anti-correlation originated from such diffuse supercluster gas, we make a first order estimate of the Compton parameter, y. We use the relation $\Delta T_{SZ}/T_0 = yj(x)$ (Refregier et al. 2000a) where T_0 is the CMB temperature, $x = h\nu/kT_0$ and $j(x)$ is a spectral function which takes the value $j(x) = -1.56$ at 94 GHz. We use the APM $m \geq 7$ result since these groups have the biggest space density and so are the most representative of average sightlines. The mean sky density is $3.5\,\mathrm{deg}^{-2}$ and the mean SZ decrement extends to $\theta > 0.5°$; the sky coverage is therefore approximately unity. In the range $0.1 < \theta < 0.5°$, the temperature decrement is relatively constant (see Fig. 6.2b); effects due to the WMAP beam may be small at these scales. We measure $\Delta T_{SZ} = -3.0 \pm 1.6\,\mu\mathrm{K}$ in this range. We use the above relation; using $T_0 = 2.726\,\mathrm{K}$ our result for the APM $m \geq 7$ groups implies a Compton parameter of $y(z < 0.2) = 7 \pm 3.8 \times 10^{-7}$. Refregier et al. (2000a) note that a 40 per cent contribution to y originates at $z < 0.2$ in CDM models (Scaramella, Cen & Ostriker 1993, Persi et al. 1995). Assuming this we estimate $y(z < \infty) = 1.8 \pm 1.0 \times 10^{-6}$. This compares to the 3σ upper limit on the total integrated y parameter from the COBE-FIRAS measurement of the spectral distortion of the CMB of $y(z < \infty) = 2.2 \times 10^{-5}$ (Fixsen et al. 1996); the 3σ upper limit from cross-correlating COBE DMR and FIRAS is $y(z < \infty) = 4.5 \times 10^{-6}$ (Fixsen et al. 1997). In addition, Banday et al. (1996) find a 3σ upper limit of $y(z < 0.2) < 1.5 \times 10^{-6}$ by cross-correlating COBE DMR with the ACO cluster catalogue. Therefore, our result is consistent with previous observational upper limits. We also note that our estimate of y is $2 - 3\times$ higher than that predicted in the SCDM model of Scaramella, Cen & Ostriker (1993). It is also similar to the ΛCDM predicted value of Persi et al. (1995).

Next, we fit isothermal models to the ACO R≥2 anti-correlation profiles. We use the model from equation 15 of Refregier et al. (2000a):

$$\Delta T_{SZ}(\theta) = \Delta T_{SZ}(0)\left(1 + \left(\frac{\theta}{\theta_c}\right)^2\right)^{3\beta/2+1/2} \tag{6.1}$$

where β denotes the slope of the SZ profile, θ_c is the projected core radius of the cluster and $\Delta T(0)$ denotes the normalisation and indicates the strength of the effect at the centre of the cluster. This is then convolved with a Gaussian of width $\sigma = 5.'25$ to mimic the effect of the beam in the W band. We assume the value of $\beta = 0.75$ quoted for Coma,

and $\theta_c = 1.'5$ which represents the Coma value scales to the mean redshift of the ACO sample ($z = 0.15$). We measure $\Delta T(0)$=0.083 mK from our cross-correlation result; this compares to $\Delta T(0)$=0.5 mK quoted for Coma. This isothermal model is displayed in Fig. 6.2b and c. It appears that the decrement observed in the data extends to greater scales than this model. Instead a lower value of β might be implied as illustrated by the $\beta = 0.5$ model also shown.

We now turn to the gas mass implied by our results. We use the equations of Refregier et al. (2000a), first to determine the central electron density:

$$\left(\frac{n_0}{10^{-3}\mathrm{cm}^{-3}}\right) = \left(\frac{\Delta T_{SZ}(0)}{1K}\right)\left(\frac{1keV}{kT_e}\right)\left(\frac{1\mathrm{Mpc}}{r_c}\right)\left(\frac{1\mu K}{-38.8\mu K}\right)\left(\frac{-2}{j(x)}\right)\left(\frac{\Gamma(3\beta/2)}{\Gamma(3\beta/2-1/2)}\right) \tag{6.2}$$

where kT_e is the isothermal cluster temperature, r_c is the cluster core radius and $j(x)$ is the spectral function used previously. The total number of electrons per cluster may then be determined:

$$N_e = n_0 \int_0^{rsz} \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3\beta/2} \tag{6.3}$$

where r_{SZ} is the radial distance to which an SZ effect is assumed. Assuming a gas composition completely of hydrogen and a spatial cluster number density ($\rho_{cluster}$), and using a derived value for the critical density ($\rho_{crit} = \frac{3H_0^2}{8\pi G}$), the gas mass density may be calculated:

$$\Omega_{gas} = \frac{N_e m_p \rho_{cluster}}{\rho_{crit}} \tag{6.4}$$

where m_p is the proton mass.

First we consider the ACO $R \geq 2$ cluster sample. We assume values of $kT = 5\,\mathrm{keV}$ and $r_c = 0.2\,h^{-1}\,\mathrm{Mpc}$, and use the measured W band value of $\Delta T(0) = 0.083\,\mathrm{mK}$; the implied central electron density is $n_0 = 1.8h \times 10^{-3}\,\mathrm{cm}^{-3}$. Assuming $\beta = 0.75$ to $r < 1.75\,h^{-1}\,\mathrm{Mpc}$ ($\approx 13'$), this gives a gas mass of $M \approx 3 \times 10^{13} h^{-2} M_\odot$, which is reasonable considering the X-ray gas mass of $M \approx 1 \times 10^{14} h^{-2.5} M_\odot$ detected within a similar radius in the Coma cluster (Lea et al. 1973).

Next we consider the mean gas mass associated with the APM $m \geq 7$ group and cluster sample. We assume $r_c = 0.2\,h^{-1}\,\mathrm{Mpc}$ and use the observed decrement in the W band of $\Delta T(0) = 0.015\,\mathrm{mK}$; the implied central electron density is $n_0 = 1.6h \times 10^{-3}\,\mathrm{cm}^{-3}(kT/1\,\mathrm{keV})$. Using $\beta = 0.75$ to $r < 1.75\,h^{-1}\,\mathrm{Mpc}$ ($\approx 20'$) as before then gives

a gas mass of $M \approx 3 \times 10^{13} h^{-2} (kT/1\,\mathrm{keV}) M_{\odot}$. These masses are reasonable given the observed relation between the gas temperature and X-ray gas mass (e.g. Fabian et al. 2001); these estimates of the gas temperature and mass would suggest an X-ray luminosity of $L_X \gtrsim 10^{45}\,\mathrm{erg\,s^{-1}}$, in line with observations of rich clusters (Donahue et al. 1998).

In order to determine the gas mass density, ${\Omega_0}^{gas}$ (equation 6.4), we take the APM $m \geq 7$ sample since the associated groups and clusters are more numerous than the ACO clusters. Using an APM group density of $3 \times 10^{-4} h^3\,\mathrm{Mpc^{-3}}$ (Croom & Shanks 1999, Myers et al. 2003), we find ${\Omega_0}^{gas} \approx 0.03 h^{-1} (1\,\mathrm{keV}/kT)(\theta_{max}/20')^{0.75}$. We then assume parameters of $h = 0.7$, $kT = 1\,\mathrm{keV}$ and $\theta_{max} = 20'$; we determine ${\Omega_0}^{gas} \approx 0.04$. This is in line with the WMAP result for the baryon density of ${\Omega_0}^{baryon} = 0.044 \pm 0.004$ (Spergel et al. 2003). However, following the evidence for a more extended decrement ($\theta_{max} \approx 60'$, see Fig 6.2b), the gas masses might rise by a factor of ≈ 2 implying ${\Omega_0}^{gas} \approx 0.1$, now a factor of $\approx 2\times$ higher than the standard value for ${\Omega_0}^{baryon}$. Considering baryons present in groups less massive than the catalogue limit, the implied baryon fraction will increase above the values estimated in this analysis.

A tentative result of this analysis is therefore a possible alteration to the widely accepted cosmological parameters. Is there any possibility that SZ contamination in the WMAP data could affect the CMB power spectrum and the acoustic peak measurements? This question arises since we observe significant decrements to $\theta \approx 1°$, similar to the measured scale of the first CMB acoustic peak (l=220). In order to address this question we have formed simple models in which mock clusters are randomly distributed over simulated CMB maps determined via CMBFAST power spectra (Zaldarriaga & Seljak 2000). If the SZ clusters have the same sky density as the APM $m \geq 7$ groups ($3.5\,\mathrm{deg^{-2}}$) and we assume an SZ profile which extends only to $\theta < 0°.5$ with a temperature decrement of $\Delta T_{SZ} = -3\,\mu\mathrm{K}$ then the amplitude of the associated SZ power spectrum at l=220 is $\approx$2 orders of magnitude below the measured WMAP amplitude. If we assume that these groups and clusters extend unevolved in their gas content past the $z < 0.2$ APM limit to $z < 0.5$ then the associated sky density rises to $\approx 50\,\mathrm{deg^{-2}}$. If we further assume values for the SZ profile more in line with what is observed here ($\theta < 1°$ and $\Delta T_{SZ} = 5\,\mu\mathrm{K}$) then the amplitude of the SZ power spectrum at $l = 220$ rises to $\approx$30 per cent of the measured WMAP amplitude. At higher multipoles, the effects on the second and third acoustic peaks are likely to be more significant. However, the above assumptions might run up against the upper limit of $y(z < \infty) = 4.5 \times 10^{-6}$ (Fixsen et al. 1997), and thus the effect on the WMAP power spectrum may be less significant than suggested for the

most extreme scenario. Nevertheless, it appears that the question of whether the SZ effect may significantly contaminate the primordial power spectrum is reopened by the spatial extent of the SZ signal found in our results. Higher resolution CMB data and deeper group and cluster catalogues are needed to constrain the SZ contribution from $z > 0.2$ clusters in order to address this issue.

6.6 CONCLUSIONS

We find evidence for an anti-correlation between WMAP W band data and galaxy cluster catalogues derived from the ACO, APM and 2MASS surveys. We interpret this temperature decrement as caused by the SZ effect, inverse Compton scattering of CMB photons by hot gas in galaxy groups and clusters. We find significant evidence that this signal extends to $\approx 1^\circ$ scales around ACO $R \geq 2$ clusters, implying that they have extended gaseous halos which may also constitute a diffuse gas component in superclusters. We use the higher sky density sample of APM $m \geq 7$ groups and clusters to estimate the Compton parameter (y); we find $y(z < 0.2) = 7 \pm 3.8 \times 10^{-7}$. This is in line with previous observational upper limits and CDM predictions. We also estimate the average gas mass associated with ACO and APM clusters. For the ACO sample we estimate $M \approx 3 \times 10^{13} h^{-2} M_\odot$, assuming $kT = 5\,\mathrm{keV}$ and β=0.75 for $r < 1.75\,h^{-1}\,\mathrm{Mpc}$. This is in reasonable agreement with X-ray observations of ACO clusters within the central radius. For the APM $m \geq 7$ sample we estimate a similar value for the average gas mass, assuming $kT = 1\,\mathrm{keV}$ and and SZ profile as before. Using the concordance value of h=0.7 this value implies a gas density of ${\Omega_0}^{gas} \approx 0.04$, in line to the standard value of ${\Omega_0}^{baryon} = 0.044$ (Spergel et al. 2003). However, since it is likely that the X-ray temperatures are below 1 keV and also our cross-correlation results indicate that the SZ decrements may extend to scales beyond the beam size (i.e. $r > 1.75\,h^{-1}\,\mathrm{Mpc}$), the estimated value of ${\Omega_0}^{gas}$ may well increase.

We examine briefly the possible effect of an extended SZ effect on the measured WMAP power spectrum. If the temperature decrements extend to $\theta < 0^\circ.5$ over the redshift range probed by the APM groups, then there is little effect on the WMAP power spectrum. However, if the SZ signal extends to $\theta \approx 1^\circ$ scales as suggested by our results, and to redshifts of $z \approx 0.5$, then the power spectrum fits may be compromised, even at the position of the first acoustic peak. Further analysis of deeper cluster catalogues and high resolution CMB data is needed to determine the level of potential SZ contamination.

CHAPTER 7 | CONCLUSIONS

7.1 SUMMARY

Over the past decade the ΛCDM model has gained wide acceptance within the astronomical community following a number of surveys which provide considerable support for this cosmology. In particular, large surveys of the local galaxy distribution ($z \lesssim 0.3$), namely the 2dF Galaxy Redshift Survey (2dFGRS) and Sloan Digital Sky Survey (SDSS), and observations of the Cosmic Microwave Background (CMB) by the Wilkinson Microwave Anisotropy Probe (WMAP), have provided constraints on cosmological parameters for which the uncertainties appear for the first time to be dominated by systematic rather than statistical errors. Possible systematics have been investigated by the survey teams; for the respective power spectrum estimates they suggest that these effects are likely to be small and are unlikely to significantly affect the cosmological parameter estimates (Bennett et al. 2003, Tegmark et al. 2004, Cole et al. 2005). In combination, the WMAP and 2dFGRS constraints appear to indicate that the Universe is flat ($\Omega = 1.02 \pm 0.02$), and dominated by matter and dark energy components ($\Omega_m = 0.27 \pm 0.04$, $\Omega_\Lambda = 0.73 \pm 0.04$), with a small baryon component of $\Omega_b = 0.044 \pm 0.004$ (Bennett et al. 2003).

However, a number of issues confront a ΛCDM cosmology; using the recently completed 2 Micron All Sky Survey (2MASS), the largest all sky survey of the local Universe at the time of writing, the aim of this thesis was to investigate the following:

- The problematic deficiency of galaxies in the APM survey number counts: The analysis of Frith et al. (2003) indicated that the near infrared counts were similar to the corrected APM counts in the optical B-band (Busswell et al. 2004), although the 2MASS second incremental release data used only partially surveyed the APM survey area, and the optical APM survey photometry is uncertain at bright magnitudes. Busswell et al. suggested that if the deficiency observed in the corrected APM survey counts were due to large-scale structure alone, then this would be at odds with the expected form of clustering on large scales in a ΛCDM cosmology. It was therefore an aim of this thesis to investigate this issue in more detail, using the

full 2MASS release data and constraining possible systematic contributions to the low counts.

- Since the results of Frith et al. (2003) and Busswell et al. (2004) suggested that an excess of power over the ΛCDM prediction might exist, the second aim of this thesis was to determine the form of clustering at extremely large scales through an angular power spectrum analysis of 2MASS. This survey provides a unique tool with which to study the local Universe, probing a volume approximately 5 times larger than that of the final 2dFGRS. From this statistic, cosmological constraints could also be determined.

- A common assumption in power spectrum analyses is that the distribution of primordial density fluctuations, from which present day large-scale structure forms through gravitational instability, is Gaussian. It is possible to investigate this issue using high-order correlation functions of the local galaxy distribution. Previous such analyses have been plagued by systematic effects arising from rare structures in the relatively small volumes probed. Also, as a result of the small surveyed volumes, no reliable estimates of such statistics have been made at linear scales where comparisons can be made with perturbation theory. We aimed to address this issue, determining high-order correlation functions to large scales from the 2MASS final release, and examining their consistency with Gaussian initial conditions and hierarchical structure formation. As an interesting aside, these statistics are also able to probe whether non-linear contributions to the galaxy bias exist; we also aimed to investigate this, for the first time in the near infrared.

- A critical issue for the cosmological interpretation of CMB observations is the level to which foregrounds contaminate the signal. The final aim of this thesis was to investigate the extent to which the hot gas in nearby clusters of galaxies affects WMAP measurements through the thermal Sunyaev-Zeldovich (SZ) effect; the angular extent of the associated temperature decrement is of particular interest, since it is possible that the cosmological fits of the CMB power spectrum might be affected. In addition therefore, we aimed to examine the potential impact of any detected SZ signal.

7.2 MAIN RESULTS

Having addressed the potential problems for a ΛCDM cosmology enumerated in the previous section, we find that:

- The near infrared K_s and H-band bright galaxy number counts extracted for the ≈ 4000 deg^2 APM survey area are similar to those of the optical B-band (Busswell et al. 2004). Using predictions for the counts constructed from the 2dFGRS $n(z)$, it appears that the form of the redshift distribution over the APM survey area may be similar to that of the Southern 2dFGRS strip, with a deficiency of ≈ 25 per cent to $z = 0.1$. However, a surprising result is that in both these passbands, the counts over almost *the entire sky* ($|b| > 20°$, $\approx 27\,000$ deg^2) display a relatively constant deficiency of $15 - 20$ per cent.

- We investigate various possible causes for the observed low counts:

  - *Model normalisation*: In the K_s-band the model normalisation used is in good agreement with K-band data collated from the literature, with a best fit of $Y = 0.96 \pm 0.06$ (where $Y = 1$ corresponds to the original normalisation). In the H-band analysis in chapter 3, we presented new faint data which is matched to the 2MASS zeropoint at bright magnitudes; using additional faint data with a consistent zeropoint applied, we tightly constrain the H-band model normalisation to $Y = 1.095^{+0.035}_{-0.034}$.
  - *Evolution*: At low redshifts, we compared the homogeneous models used with near infrared selected $n(z)$ using the 2dFGRS data; there is excellent agreement at relatively high redshifts ($0.1 < z < 0.2$). In addition, we use these redshift distributions to construct predictions for the corresponding number counts over the 2dFGRS strips; there is reasonable agreement suggesting that real features in the local galaxy distribution are the dominant cause of variations in the number counts. As regards unexpectedly high levels of evolution at higher redshifts, which might affect the model normalisation, the passive evolution models used provide good agreement with counts collated from the literature to extremely faint magnitudes ($K \approx 23$, $H \approx 24$). This picture is consistent with results from independent results from the COMBO-17 and K20 surveys.
  - *Photometry and completeness*: These issues were examined in the H-band analysis of chapter 3. The zeropoints of the bright and faint data were matched

through comparisons of several hundred point sources; matched galaxy samples were also used and found to give consistent results. Completeness could have a significant effect however; the level of incompleteness in the faint data is uncertain but likely to be at the few per cent level; in the 2MASS catalogue the completeness is given as > 90 per cent. The effect of incompleteness on the observed deficiency is therefore < 10 per cent.

- *Large-scale structure*: First, in the K_s-band analysis of chapter 2 we used a ΛCDM form to the angular correlation function at large scales. The observed counts over the APM survey area represent a 5.0σ fluctuation. Using the best fit normalisation ($Y = 0.96 \pm 0.06$) and incorporating the uncertainty, this is reduced to 2.9σ. Second, in the H-band (chapter 3) we equivalently used mock ΛCDM 2MASS catalogues to determine the significance; using the best fit normalisation ($Y = 1.095^{+0.035}_{-0.034}$) and incorporating the associated uncertainty, and assuming an upper limit to the effect arising from incompleteness (10 per cent), the H-band APM survey area counts represent a rare fluctuation in the galaxy distribution ($\approx 2.5\sigma$, ≈ 1 in 100); the $|b| > 20°$ counts however suggest a challenge to the form of clustering predicted by ΛCDM on large scales ($\approx 4.0\sigma$, < 1 in 10 000).

- In order to detect any possible excess of power over the ΛCDM prediction, we used the $|b| > 20°$ 2MASS catalogue to determine the associated angular power spectrum and the form of galaxy clustering at extremely large scales ($r \lesssim 1000\,h^{-1}$ Mpc). In chapter 2, we limited at $K_s < 12.5$ as the peak in the associated selection function lies in the redshift range of interest ($z^* = 0.05$). We detected a 3σ excess at large scales ($l \leq 30$, $r \gtrsim 30\,h^{-1}$ Mpc) in the 2MASS angular power spectrum over that of initially unbiased ΛCDM mock 2MASS catalogues, applying a bias of $b_K = 1.1$ (Maller et al. 2005) to the resulting power spectrum. However, this apparent excess has little impact on the significance calculations mentioned above.

- We used the 2MASS angular power spectrum to determine the associated cosmological constraints in chapter 4, this time limiting to $K_s < 13.5$. Using errors determined from the ΛCDM mock catalogues that are independent of cosmology and normalisation, and assuming a flat geometry with negligible neutrino density and a primordial $n_s = 1$ spectrum, we find best fit parameters for the shape and normalisation of the power spectrum of $\Gamma_{eff} = 0.14 \pm 0.02$ and $(\sigma_8 b_K)^2 = 1.36 \pm 0.10$, marginal-

ising over the normalisation and shape respectively. Taking the WMAP-2dFGRS constraint of $\sigma_8 = 0.84 \pm 0.04$ (Bennett et al. 2003), this implies a K_s-band bias of $b_K = 1.39 \pm 0.12$. If we instead assumed that the errors used simply scale linearly with power then the constraints weaken slightly but remain in fair agreement.

- The higher order moments of the 2MASS galaxy density field were found to be consistent with hierarchical clustering and Gaussian initial conditions, with constant skewness and kurtosis measured on linear and quasi-linear scales ($10 \lesssim r \lesssim 100\,h^{-1}$ Mpc). Unlike previous such analyses, this agreement is relatively robust to the removal of the largest superclusters. We are able to reject a strongly non-Gaussian form to the distribution of primordial density fluctuations, as might be seeded by topological defects such as cosmic strings or global textures, at the $\approx 2.5\sigma$ confidence level. In contrast to all previous observations, we also measure a *positive* quadratic bias of $c_2 = b_2/b_1 = 0.57 \pm 0.33$ (assuming a power spectrum slope of $n = -2$).

- We detected significant evidence for temperature decrements in the first year WMAP data associated with galaxy cluster samples selected from the ACO, APM survey and 2MASS galaxy catalogues. These we interpret as due to the Sunyaev-Zeldovich (SZ) effect arising from hot inter-cluster gas. With the ACO and APM samples, the decrement appears to extend significantly beyond the WMAP beam size to scales of $\theta \gtrsim 1°$. Assuming reasonable values for the gas temperature ($kT = 1keV$), the cluster core radius ($r = 0.2\,h^{-1}$ Mpc) and $h = 0.7$, a beam size SZ effect in the APM group and cluster sample extending to scales of $\theta = 20'$ implies a gas mass density of $\Omega_{gas} \approx 0.04$, in line with WMAP constraints on the baryon density. However, we find evidence indicating that the SZ decrement extends to significantly larger scales, suggesting much larger values for the gas mass density; $\Omega_{gas} \approx 0.1$ for a $\theta = 1°$ SZ effect for example. Such a signal would in turn compromise the fits to the WMAP power spectrum.

7.3 FINAL CONCLUSIONS

The issue of the large local hole around the SGP has yet to be resolved; in this thesis, we have succeeded in constraining common sources of systematic error which might contribute to the measured deficiency in the number counts and determined that they cannot reconcile the observations with the expected form of clustering predicted in a ΛCDM cosmology. The counts suggest a deficiency in the galaxy distribution of ≈25 per cent over the APM survey area covering ≈ 4000 deg^2 to $z \approx 0.1$; perhaps a deficiency even extends over the entire local galaxy distribution. However, no corresponding excess is detected in the 2MASS angular power spectrum on large scales sufficient to account for the observed number counts; in fact the 2MASS C_ls provide strong support for ΛCDM with $\Gamma_{eff} = 0.14 \pm 0.02$. It is possible that the power spectrum may be affected by cosmic variance at large scales as suggested by the $|b| > 20°$ counts. However it is difficult to reconcile such an idea with measurements of the power spectrum at $z \approx 0.2$ and $z \approx 1.4$ from recent surveys of galaxies and QSOs which provide strong support for a low-Ω_m cosmology. The large increase in faint near infrared data from the UK Infrared Deep Sky Survey (UKIDSS) should help to resolve this issue.

The high-order clustering analysis also provides support for the standard paradigm, with results consistent with expectations for a Gaussian distribution of primordial density fluctuations that have evolved under the action of gravitational instability. The results are able to effectively rule out strongly non-Gaussian initial conditions. However, further support for this picture is needed from large-scale structure observations to complement those of WMAP. Another interesting result of this analysis is the measurement of a positive non-linear bias which differs significantly from the negative values determined previously. Qualitatively, this is reasonable given analytic predictions, although the theory needs further refinement in order to understand how this result fits in with semi-analytic models and theories of galaxy formation.

The detection of evidence for an extended SZ effect in WMAP data may provide a challenge to the use of the CMB power spectrum as a reliable probe of the geometry of the Universe. If such an effect exists on $\gtrsim 1°$ scales, it appears that not only would a large gas mass fraction be implied which would be in contradiction to the concordance value, but the CMB power spectrum fits would also be compromised. Further tests, using larger cluster catalogues and less noisy CMB data, are needed to resolve this issue.

7.4 FUTURE PROSPECTS

The advances made over the last decade in our understanding of the parameters of the standard cosmological model have been enormous. Widespread acceptance in the concept that the Universe has zero curvature, is dominated by a dark energy component and has a low matter density has largely been brought about through the recent results of ambitious surveys measuring the local large-scale structure and CMB to unprecedented accuracy.

However, as examined in this thesis, a number of outstanding issues remain for a ΛCDM cosmology, and a number of observational tests are still required in order to confirm the measurements made by recent surveys. In the next decade, these may be satisfied to some extent by the completion of surveys such as the Sloan Digital Sky Survey (SDSS), Planck (e.g. Lamarre et al. 2003) and the Supernova/Acceleration Probe (SNAP; e.g. Linder et al. 2003).

Whether or not the support for the ΛCDM model offered by recent experiments is confirmed by forthcoming surveys, the future prospects for observational cosmology are numerous and exciting. Assuming the validity of the ΛCDM model, the most pressing issue remains the lack of understanding of the CDM and dark energy components of the energy density. Can these phenomena be explained by existing physics within the framework of the standard cosmological model, or is some more radical explanation needed?

One such interesting alternative is the Cardassian model (Freese & Lewis 2002), which utilises a modification to the Friedmann equation (see equation 1.6) such that $H^2 = A\rho + B\rho^n$, where $A = 8\pi G/3$ and B are constants, H is the Hubble parameter, ρ is the energy density and n is a parameter of the Cardassian model. This additional second term can give rise to a cosmology which is consistent with the recent evidence for an increasing expansion rate, but is both flat and matter dominated. The need for a dark energy component is therefore removed, and the expansion is driven solely by the new ρ^n term in the modified Friedmann equation. The theoretical motivation for this Cardassian model of cosmology (see Freese & Lewis 2002) is fairly speculative however, and the implications for Einstein's equations of General Relativity are currently unexplored. The cosmological constraints for this model have been examined in Frith (2004). Using measurements for the locations of the first, second and third Doppler peaks in the CMB power spectrum from WMAP (Hinshaw et al. 2003) and Archeops data (Hu et al. 2001), and a sample of type 1a supernovae (Tonry et al. 2003), constraints on the matter density and the Cardassian parameter n were determined. While the supernovae and CMB constraints

favour slightly differing Cardassian cosmologies, this alternative to the concordance model cannot yet be ruled out.

Whether the CDM and dark energy components are explained by alterations to the standard cosmological model, such as with the Cardassian model, may be determined with the next generation of instruments such as the Square Kilometre Array (SKA; e.g. Blake et al. 2004), the Large Synoptic Survey Telescope (LSST; e.g. Stubbs et al. 2004), and PANSTARRS (e.g. Kaiser et al. 2005). With such surveys we will be able to constrain the equation of state of the dark energy and discriminate between a number of inflationary models, as well as map the galaxy distribution to high redshifts in exquisite detail. These and forthcoming surveys such as UKIDSS, Planck and SNAP should resolve the issues raised in this thesis, giving us greater understanding of the various components of the energy density, and perhaps allowing us to determine the ultimate fate of the Universe.

Bibliography

The following abbreviations are used in this bibliography:

AAS: American Astronomical Society

A&A: Astronomy and Astrophysics

AJ: The Astronomical Journal

ApJ: The Astrophysical Journal

ApJL: The Astrophysical Journal Letters

ApJS: The Astrophysical Journal Supplement

ARA&A: Annual Review of Astronomy and Astrophysics

ASP Conf. Proc.: The Astronomical Society of the Pacific conference series

Helv. Phys.: Helvetica Physica

IAUS: International Astronomical Union Symposium

JCAP: Journal of Cosmology and Astroparticle Physics

JHEP: Journal of High Energy Physics

MNRAS: Monthly Notices of the Royal Astronomical Society

New AR: New Astronomy Reviews

Nat: Nature

Nucl. Phys. B: Nuclear Physics B

Phys. Lett. B: Physical Letters B

Phys. Rep.: Physics Reports

Phys. Rev.: Physical Review

Phys. Rev. D: Physical Review D

Phys. Rev. Lett.: Physical Review Letters

Proc. NAS: Proceedings of the National Academy of Science

PW: Physics World

SSR: Space Science Reviews

Abell, G.O., Corwin, H., Olowin, R., 1989, ApJS, 70, 1

Acquaviva, V., Bartolo, N., Matarrese, S. & Riotto, A. 2003, Nucl. Phys. B, 667, 119

Afshordi, N., Loh, Y-S., & Strauss, M.A. 2003, AAS, 203, 6907

Alpher, R.A., Bethe, H.A. & Gamow, G. 1948, Phys. Rev., 73, 803

Arkani-Hamed, N., Creminelli, P., Mukohyama, S. & Zaldarriaga, M. 2004, JCAP, 04, 001

Avelino, P.P., Shellard, E.P., Wu, J.H. & Allen, B. 1998, ApJ, 507, L101

Bahcall, N.A., & Soneira, R.M.1983, ApJ, 270, 20

Banday, A.J., Gorski, K.M., Bennett, C.L., Hinshaw, G., Kogut, A., & Smoot, G.F. 1996, ApJ, 468, L85

Baugh, C.M. & Efstathiou, G. 1993, MNRAS, 265, 145

Baugh, C.M., Gaztañaga, E. & Efstathiou, G. 1995, MNRAS, 274, 1049

Baugh, C.M. 1996, MNRAS, 280, 267

Baugh, C.M. et al. 2004, MNRAS, 351, L44

Bell, E.F., McIntosh, D.H., Katz, N. & Weinberg, M.D. 2003, ApJS, 149, 289

Bell, E.F. et al. 2004, ApJ, 608, 752

Bennett, C. L. et al. 2003, ApJS, 148, 97

Bernardeau, F. 1992, ApJ, 392, 1

Bernardeau, F. 1994a, A&A, 291, 697

Bernardeau, F. 1994b, ApJ, 433, 1

Bernardeau, F., Colombi, S., Gaztañaga, E. & Scoccimarro, R. 2002, Phys. Rep., 367, 1

Bernardeau, F. & Uzan, J.P. 2003, Phys. Rev. D., 67, 121301

Blake, C.A., Abdalla, F.B., Bridle, S.L. & Rawlings, S. 2004, New AR, 48, 11

Bond, J.R. & Szalay, A.S. 1983, ApJ, 274, 443

Bouchet, F.R., Juszkiewicz, R., Colombi, S. & Pellat, R. 1992, ApJ, 394, L15

Bouchet, F.R., Strauss, M., Davis, M., Fisher, K., Yahil, A. & Huchra, J. 1993, ApJ, 417, 36

Boughn, S. & Crittenden, R. 2004, Nat, 427, 45

Broadhurst, T.J. Ellis, R.S. & Shanks, T. 1988, MNRAS, 235, 827

Bruzual, A.G., & Charlot, S. 1993, ApJ, 405, 538

Busswell, G.S., Shanks, T., Outram, P.J., Frith, W.J., Metcalfe, N. & Fong, R. 2004, MNRAS, 354, 991

Cabella, P., Liguori, M., Hansen, F.K., Marinucci, D., Matarrese, S., Moscardini, L. & Vittorio, N. 2005, MNRAS, 358, 684

Carlberg, R.G., Yee, H.K., Morris, S.L., Lin, H., Hall, P.B., Patton, D., Sarwicki, M. & Shepherd, C.W. 2000, ApJ, 542, 57

Chen, H.-S. et al. 2002, ApJ, 570, 54

Cimatti, A. et al. 2002, A&A, 391, L68

Cole, S.M., Hatton, S.J., Weinberg, D.H. & Frenk, C.S. 1998, MNRAS, 300, 945

Cole, S.M. et al. 2001, MNRAS, 326, 555

Cole, S.M. et al. 2005, accepted by MNRAS, astro-ph/0501174

Colless, M. Ellis, R.S., Taylor, K. & Hook, R.N. 1990, MNRAS, 244, 408

Colless, M. et al. 2001, astro-ph/0306581

Croom, S.M. & Shanks, T., 1999, MNRAS, 307, L17

Croton, D.J. et al. 2004, MNRAS, 352, 1232

Croton, D.J. et al. 2005, MNRAS, 356, 1155

Diego, J.M., Silk, J. & Sliwa, W., 2003, MNRAS, 346, 940

Donahue, M., Voit, G.M., Giola, I., Luppino, G., Hughes, J.P. & Stocke, J.T., 1998, ApJ, 502, 550

Driver, S. 2003, IAUS, 216, 97

Ebeling, H., Voges, W., Bohringer, H., Edge, A.C., Huchra, J.P. & Briel, U.G. 1996 MNRAS, 281, 799

Efstathiou, G. 1990, in Shanks, T. et al., eds, Observational Tests of Inflation. Kluwer, p425

Efstathiou, G., Sutherland., W.J. & Maddox, S.J. 1990, Nat, 348, 705

Eisenstein, D.J. & Hu, W. 1998, ApJ, 496, 605

Faber, S.M. & Gallagher, J.S. 1979, ARA&A, 17, 135

Fabian, A.C., Crawford, C.S., Ettori, S. & Sanders, J.S. 2001, MNRAS, 322, 11

Falk, T., Rangarajan, R. & Srednicki, M. 1993, ApJ, 403, L1

Feldman, H.A., Kaiser, N. & Peacock, J.A. 1994, ApJ, 426, 23

Feldman, H.A., Frieman, J.A., Fry, J.N. & Scoccimarro, R. 2001, Phys. Rev. Lett., 86, 1434

Fixsen, D.J., Cheng, E.S., Gales, J.M., Mather, J.C., Shafer, R.A., & Wright, E.L., 1996, ApJ, 473, 576

Fixsen, D.J., Hinshaw, G., Gales, Bennett, C.L., & Mather, J.C., 1997, ApJ, 486, 623

Fosalba, P. & Gaztañaga,E. 2003, MNRAS, 350, L37

Freese, K. & Lewis, M. 2002, Phys. Lett. B, 540, 1

Frith, W.J., Busswell, G.S., Fong, R., Metcalfe, N. & Shanks, T. 2003, MNRAS, 345, 1049

Frith, W.J., Outram, P.J. & Shanks, T. 2004, ASP Conf. Proc., Vol. 329, 49, astro-ph/0408011

Frith, W.J. 2004, MNRAS, 348, 916

Frith, W.J., Shanks, T. & Outram, P.J. 2005a, MNRAS, 361, 701

Frith, W.J., Outram, P.J. & Shanks, T. 2005b, MNRAS, 364, 593

Frith, W.J., Outram, P.J. & Shanks, T. 2005c, submitted to MNRAS, astro-ph/0507704

Frith, W.J., Metcalfe, N. & Shanks, T. 2005d, submitted to MNRAS, astro-ph/0509875

Fry, J.N. 1984, ApJ, 279, 499

Fry, J.N. & Gaztañaga, E. 1993, ApJ, 413, 447

Gangui, A., Lucchin, F., Matarrese, S. & Mollerach, S. 1994, ApJ, 430, 447

Gaztañaga, E. 1994, MNRAS, 268, 913

Gaztañaga, E. & Frieman, J.A. 1994, ApJ, 437, L13

Gaztañaga, E. & Baugh, C.M. 1995, MNRAS, 273, L1

Gaztañaga, E. & Mahonen, P. 1996, ApJ, 462, L1

Gaztañaga, E. & Fosalba, P. 1998, MNRAS, 301, 524

Gaztañaga, E. & Bernardeau, F. 1998, A&A, 331, 829

Gaztañaga, E., Norberg, P., Baugh, C.M. & Croton, D.J. 2005, submitted to MNRAS, astro-ph/0506249

Groth, E.J. & Peebles, P.J.E. 1977, ApJ, 217, 385

Giommi, P. & Colafrancesco, S. 2003, A&A, 414, 7

Guth, A.H. 1982, Phys. Rev. D, 23, 347

Guzzo, L. et al. 2000, A&A, 355, 1

Hamilton, A.J., Kumar, P., Lu, E. & Matthews, A. 1991, ApJ, 374, 1

Hauser, M.G. & Peebles, P.J.E. 1973, ApJ, 185, 757

Hawkins, E. et al. 2003, MNRAS, 346, 78

Hernandez-Monteagudo, C. & Rubino-Martin, J.A. 2003, MNRAS, 347, 403

Hinshaw, G. et al. 2003, ApJS, 148, 63

Hoyle, F., Szapudi, I. & Baugh, C.M. 2000, MNRAS, 317, 51

Hu, W., Fukugita, M., Zaldarriaga, M. & Tegmark, M. 2001, ApJ, 549, 669

Huang, J.S., Cowie, L.L, Gardner, J.P., Hu, E.M., Songalia, A. & Wainscoat, R.J. 1997, ApJ, 476, 12

Huang, J.S. et al. 2001, A&A, 368, 787

Hubble, E. 1929, Proc. NAS, 15, 168

Huterer, D., Knox, L. & Nichol, R.C. 2001, ApJ, 555, 547

Jarrett, T.H., Chester, T., Cutri, R., Schneider, S., Skrutskie, M. & Huchra, J.P. 2000, AJ, 119, 2498

Jarrett, T.H. 2004, astro-ph/0405069

Jenkins, A. et al. 1998, ApJ, 499, 20

Jones et al. 2004, MNRAS, 355, 747

Juszkiewicz, R., Bouchet, F.R. & Colombi, S. 1993, ApJ, 412, 9

Kaiser et al. 2005, AAS, 206, 2402

Kofman, L. & Pogosyan, D.Y. 1988, Phys. Lett. B, 214, 508

Kogut, A. et al. 2003, ApJS, 148, 161

Komatsu, E. & Kitayama, T. 1999, ApJ, 526, L1

Kummel, M.W. & Wagner, S.J. 2000, A&A, 353, 867

Kurki-Suonio, H. 2002, SSR, 100, 249

Lamarre, J.M. et al. 2003, New AR, 47, 1017

Lea, S.M., Silk, J., Kellogg, E. & Murray, S., 1973, ApJ, 184, L105

Lesgourgues, J., Polarski, D. & Starobinsky, A.A. 1997, Nucl. Phys. B, 497, 479

Linder, E. et al. 2003, IAUS, 216, 155

Loveday, J., Peterson, B.A., Maddox, S.J. & Efstathiou, G. 1996, ApJS, 107, 201

Loveday, J., Peterson, B.A., Efstathiou, G. & Maddox, S.J. 1992, ApJ, 390, 338

Loveday, J. 2000, MNRAS, 312, 517

Loveday, J. 2004, MNRAS, 347, 601L

Maddox, S.J., Sutherland, W.J., Efstathiou, G. & Loveday, J. 1990a, MNRAS, 243, 692

Maddox, S.J., Efstathiou, G., Sutherland, W.J. & Loveday, J. 1990b, MNRAS, 242, 43P

Maddox, S.J., Sutherland, W.J., Efstathiou, G., Loveday, J. & Peterson 1990c, MNRAS, 247, 1

Maldacena, J. 2002, JHEP, 0305, 013

Maller, A.H., McIntosh, D.H., Katz,N. & Weinberg, M.D. 2003, ApJ, 598, 1

Maller, A.H., McIntosh, D.H., Katz, N. & Weinberg, M.D. 2005, ApJ, 619, 147

Myers, A.D., Outram, P.J., Shanks, T., Boyle, B.J., Croom, S.M., Loaring, N.S, Miller, L., Smith, R.J., 2003, MNRAS, 342, 467

Myers, A.D., Shanks, T., Outram, P.J., Frith, W.J. & Wolfendale, A.W. 2004, MNRAS, 347, L67

Martini, P. 2000, AAS, 197, 6503

Martini, P. 2001, AJ, 121, 598

McCracken, H.J., Metcalfe, N., Shanks, T., Campos, A., Gardner, J.P. & Fong, R. 2000, MNRAS, 311, 707

McLeod, B.A., Bernstein, G.M., Rieke, M.J., Tollestrup, E.V, & Fazio, G.G. 1995, ApJS, 96, 117

Metcalfe, N., Fong, R. & Shanks, T. 1995, MNRAS, 274, 769

Metcalfe, N., Shanks, T., Campos, A., McCracken, H.J. & Fong, R. 2001, MNRAS, 323, 795

Metcalfe, N., Shanks, T., Weilbacher, P.M., McCracken, H.J., Campos, A., Fong, R. & Thompson, D. 2005, in prep.

Mo, H.J., Jing, Y.P. & White, S.D. 1997, MNRAS, 284, 189

Moy, E., Barmby, P., Rigopoulou, D., Huang, J.-S., Willner, S.P. & Fazio, G.G. 2003, A&A, 403, 493

Nolta, M. R. et al. 2003, ApJ, 608, 10

Norberg, P. et al. 2002, MNRAS, 336, 907

Oliver, S.J. et al. 1996, MNRAS, 280, 673

Outram, P.J., Hoyle, F., Shanks, T., Croom, S.M., Boyle, B.J., Miller, L., Smith. R.J. & Myers, A.D. 2003, MNRAS, 342, 483

Page, L. et al. 2003, ApJS, 148, 39

Pan, J. & Szapudi, I. 2005, submitted to MNRAS, astro-ph/0505422

Peacock, J.A. & Dodds, S.J. 1994, MNRAS, 267, 1020

Pearson, T. J. et al. 2003, ApJ, 591, 556

Peebles, P.J.E.. 1973, ApJ, 185, 413

Peebles, P.J.E. & Hauser, M.G. 1973, ApJ, 185, 757

Peebles, P.J.E. & Hauser, M.G. 1974, ApJS, 28, 19

Peebles, P.J.E. 1980, Principles of Physical Cosmology, Princeton University Press

Peebles, P.J.E. 1982, ApJ, 263, L1

Penzias, A.A. & Wilson, R.W. 1965, ApJ, 142, 419

Percival et al. 2001, MNRAS, 327, 1297

Perlmutter, et al. 1997, ApJ, 483, 565

Persi, F.M., Spergel, D.N., Cen, R. & Ostriker, J. 1995, ApJ, 442, 1

Press, W.H. & Schechter, P. 1974, ApJ, 187, 425

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. 1988, Numerical Recipes in Fortran, Cambridge University Press

Ratcliffe, A., Shanks, T., Parker, Q.A. & Fong, R. 1998, MNRAS, 296, 173

Ratcliffe, A. et al. 1998, MNRAS, 300, 417

Refregier, A., Spergel, D.N. & Herbig, T., 2000a, ApJ, 531, 31

Refregier, A., Komatsu, E., Spergel, D.N. & Pen U., 2000b, Phys. Rev. D, 61, 123001

Renzini, A. et al. 1996, ApJL, 465, L23

Riess, A.G. et al. 1998, AJ, 116, 1009

Salopek, D.S. & Bond, J.R. 1991, Phys. Rev. D, 43, 1005

Saunders, W. et al. 1991, Nat, 349, 32

Scaramella, R., Cen, R., & Ostriker, J. 1993, ApJ 416, 399

Scharf, C., Hoffman, Y., Lahav, O. & Lynden-Bell, D. 1992, MNRAS, 256, 229

Schlegel, D.J., Finkbeiner, D.P. & Davis, M. 1998, ApJ, 500, 525

Scoccimarro, R., Sefusatti, E. & Zaldarriaga, M. 2003, Phys. Rev. D, 69, 103513

Seldner, M., Siebers, B., Groth, E.J. & Peebles, P.J.E. 1977, AJ, 82, 249

Silk, J. & Juszkiewicz, R. 1991, Nat, 353, 386

Silk, J. 2002, PW, v. 15, No. 8, 21

Shanks, T. 1990, IAUS, 139, 269

Shectman, S.A., Landy, S.D., Oemler, A., Tucker, D.L., Lin, H., Kirshner, R.P. & Schechter, P.L. 1996, ApJ, 470, 172

Smoot, G.F. et al. 1992, ApJL, 396, L1

Soltan, A.M., Freyberg, M.J. & Hasinger, G. 2002, A&A. 395, 475

Spergel, D. N. et al. 2003, ApJS, 148, 175

Stubbs, C.W., Sweeney, D. & Tyson, J.A. 2004, AAS, 205, 10802

Sugiyama, N. 1995, ApJS, 100, 281

Szapudi, I. & Gaztañaga, E. 1998, MNRAS, 300, 493

Szapudi, I., Colombi, S., Jenkins, A. & Colberg, J. 2000, MNRAS, 313, 725

Szokoly, G.P. et al. 1998, ApJ, 492, 452

Tegmark, M. et al. 2002, ApJ, 571, 191

Tegmark, M., de Oliveira-Costa, A., & Hamilton, A.J.S. 2003, Phys. Rev. D, 68, 123523

Tegmark, M. et al. 2002, ApJ, 606, 702

Teplitz, H.I., Malkan, M. & McLean, I.S. 1998, ApJ, 506, 519

Thompson, R., Storrie-Lombardi, L. & Weymann, R. 1999, AJ, 117, 17

Tonry, J.L. et al. 2003, ApJ, 594, 1

Tucker, D.L. et al. 1997, MNRAS, 285, L5

Turok, N. & Spergel, D.N. 1991, Phys. Rev. Lett., 66, 3093

Vaisanen, P., Tollestrup, E.V., Willner, S.P. & Cohen, M. 2000, ApJ, 540, 593

Verde, L. et al. 2002, MNRAS, 335, 432

Vettolani, G. et al. 1997, A&A, 325, 954

Wang, L. & Kamionkowski, M. 2000, Phys. Rev. D, 61, 063504

Weinberg, D.H. & Cole, S.M. 1992, MNRAS, 259, 652

Wilson, G. 2003, ApJ, 585, 191

White, S.D., Navarro, J.F., Evrard, A.E. & Frenk, C.S. 1993, Nat, 366, 429

White, M. 1999, MNRAS, 310, 511

Yan, L., McCarthy, P., Storrie-Lombardi, L. & Weymann, R. 1998, ApJ, 503, L19

Yasuda, N. et al. 2001, AJ, 122, 1104

Zaldarriaga, M. & Seljak, U. 2000, ApJS, 129, 431

Zehavi, I. et al. 2002, ApJ, 571, 172

Zehavi, I. et al. 2004, ApJ, 608, 16

Zwicky, F. et al. 1933, Helv. Phys., 6, 110