GALAXIES AND CLUSTERS AS PROBES OF THE LARGE-SCALE STRUCTURE OF THE UNIVERSE

Robert Graham Mann

Presented for the Degree of Doctor of Philosophy at the

University of Edinburgh

1993

DECLARATION

This thesis is my own composition, except where specifically indicated in the text.

R.G. Mann

November 1993

CONTENTS

1 INTRODUCTION 1

1.1 THE BIG BANG MODEL 2

1.1.1 The Cosmological Principle 2

1.1.2 The Friedmann-Lemaître-Robertson-Walker Metric 2

1.1.3 The Recession of the Nebulæ 5

1.1.4 General Relativity 8

1.1.5 The Friedmann Equations 10

1.1.6 Primordial Nucleosynthesis in the standard Big Bang model 12

1.1.7 The Cosmic Background Radiation 18

1.1.8 Summary 22

1.2 INFLATIONARY COSMOLOGY 23

1.2.1 Inadequacies of the relativistic Big Bang model 23

1.2.2 An Inflationary Universe 25

1.2.3 New Inflation and the generation of density perturbations 27

1.2.4 Summary: Generic Inflation 29

1.3 Dark Matter in the Universe 30

1.3.1 Evidence for the existence of dark matter 30

1.3.2 The nature of dark matter 35

1.3.3 A challenge to orthodox cosmology 37

1.4 The growth of cosmological density perturbations 41

1.4.1 Linear perturbation growth in a pressureless Newtonian fluid . . . 42

1.4.2 The Jeans Length 46

1.4.3 Silk Damping 47

1.4.4 Transfer Functions 48

1.4.5 Normalisation of the power spectrum 49

1.4.6 Non-linear growth of density perturbations 51

1.5 Plan for the remainder of this thesis 53

2 CLUSTER CORRELATIONS AS PROBES OF LARGE-SCALE STRUCTURE **54**

2.1 INTRODUCTION 54

2.2 THE ABELL/ACO CATALOGUE 56

2.2.1 Construction of the Abell/ACO catalogue 56

2.2.2 Spatial correlations of Abell/ACO clusters 63

2.2.3 The Integrity of the Abell/ACO Catalogue 68

2.3 AUTOMATED CLUSTER CATALOGUES 77

2.3.1 The Edinburgh/Durham Cluster Catalogue 78

2.3.2 The APM Cluster Catalogue 82

2.4 OBSERVATIONAL BACKGROUND: SUMMARY 84

2.5 THEORETICAL BACKGROUND 90

2.5.1 The Kaiser Effect 91

2.5.2 Further approximations to the correlations of peaks in Gaussian noise 93

2.5.3 The dynamic evolution of peak-peak correlations 100

2.5.4 Numerical studies of cluster correlations 102

2.6 CLUSTERING IN THE ZELDOVICH APPROXIMATION 103

2.6.1 The Zeldovich Approximation 104

2.6.2 Mass correlations in the Zeldovich Approximation 107

2.6.3 Defining a population of objects 112

2.6.4 Cluster correlations in the Zeldovich Approximation 118

2.6.5 Models and Normalisation 122

2.6.6 Sensitivity to parameters 125

2.7 COMPARISON WITH RECENT NUMERICAL RESULTS 128

2.7.1 Comparsion with Bahcall & Cen (1992) 128

2.7.2 Comparison with Croft & Efstathiou (1993) 128

2.7.3 Comparison with Cole (1993) 130

2.7.4 Discussion and Summary 130

2.8 RESULTS: OPTICAL CLUSTER SAMPLES 134

2.9 RESULTS: A *ROSAT* CLUSTER SAMPLE 144

2.10 SUMMARY 150

3 THE CLUSTERING OF WARM AND COOL IRAS GALAXIES 153

3.1 INTRODUCTION 153

3.2 THE DATA SAMPLES 159

3.2.1 The QDOT Survey 159

3.2.2 The 2 Jy Survey 160

3.2.3 The Mask 161

3.2.4 Temperature assignment and subsample selection 162

3.2.5 Selection Functions 164

3.3 CORRELATION FUNCTION ANALYSIS 168

3.3.1 Method 168

3.3.2 Results 176

3.4 COUNTS-IN-CELLS ANALYSIS 187

3.4.1 Method 187

3.4.2 Results 189

3.5 SUMMARY 194

4 GALAXY CLUSTERING WITH EULERIAN BIAS 196

4.1 INTRODUCTION 196

4.2 SIMULATING GALAXY CLUSTERING 198

4.3 DETAILS OF THE NUMERICAL SIMULATIONS 208

4.4 THE POWER SPECTRUM OF OPTICAL GALAXY CLUSTERING . . 211

4.4.1 Observational Data 211

4.4.2 Method 213

4.4.3 Results 215

4.5 THE DISTRIBUTIONS OF OPTICAL AND IRAS GALAXIES 219

4.6 THE POWER SPECTRUM OF IRAS GALAXY CLUSTERING 223

4.6.1 Observational Data 224

4.6.2 Method 224

4.6.3 Results 224

4.7 THE PAIRWISE VELOCITY DISPERSION OF GALAXIES 226

4.7.1 Observational estimations of the pairwise velocity dispersion . . . 226

4.7.2 Theoretical determinations of the pairwise velocity dispersion . . . 229

4.7.3 Method 231

4.7.4 Results 231

4.8 SUMMARY 233

5 DISCUSSION, CONCLUSIONS AND FUTURE WORK **236**

5.1 THE SPATIAL CORRELATIONS OF CLUSTERS 237

5.2 THE CLUSTERING OF IRAS GALAXIES 241

5.3 GALAXY CLUSTERING WITH EULERIAN BIAS 244

5.4 CONCLUDING REMARKS 245

MATHEMATICAL APPENDICES **247**

A Derivation of $n_{\rm pk}(\nu_s)$ 248

B Derivation of $\langle\tilde{\nu}_s\rangle$ 249

C Derivation of $\xi_{\rho Z}(x)$ 250

D Derivation of $\xi_{\mathcal{CZ}}(x)$ 256

REFERENCES 267

PUBLISHED PAPER 285

ABSTRACT

The large-scale structure of the Universe is delineated by the spatial distributions of galaxies and clusters of galaxies. This thesis describes three projects concerned with the use of galaxies and clusters as cosmological probes, following the presentation of necessary background material in Chapter 1.

Chapter 2 is concerned with spatial correlations of clusters of galaxies. After comprehensively reviewing previous work addressing this topic from both observational and theoretical points of view, we present, test and apply an important new method for computing theoretical cluster correlations. Our method combines the theory of peaks in Gaussian random fields with the evolution of the cosmological density field by the Zeldovich Approximation: this is the first analytic calculation of the cluster correlation function to take account of the nonlinear evolution of the cosmological density field on cluster scales. We find good agreement between our results and those from recent numerical simulations, except for the richest cluster samples, for which our method yields stronger clustering. Comparison of our predicted correlations with those observed in recent optical cluster samples reveal that the once-popular Einstein – de Sitter Cold Dark Matter (CDM) model lacks the large-scale power required to match the observed clustering. We also apply our method in the first theoretical study of the spatial correlations of *ROSAT* clusters. Our results here favour cosmogonies with more large-scale power than CDM, in accordance with those we obtained from optical cluster samples.

The projects in Chapters 3 and 4 are concerned with galaxy clustering. In Chapter 3 we consider the redshift-space clustering of samples of *IRAS* galaxies selected on the basis of their dust emission temperature, having argued that there might be a relation

between the temperature of the galaxy and density of the environment in which the galaxy is located. We find, however, no conclusive evidence for a difference in the clustering strength of the "warm" and "cool" samples in redshift space. This validates the use of redshift samples of *IRAS* galaxies as tracers of large-scale structure, as well as constraining models of merger-induced star formation.

In Chapter 4 we show, through the novel analysis of high-resolution numerical simulation data, how the observed power spectra of optical and *IRAS* galaxy clustering constrain the initial power spectrum of density fluctuations and the relation between the galaxy distribution and the underlying density field. Motivated by recent N–body/hydrodynamic simulations, we employ a biasing prescription in which the local galaxy number density at redshift zero is determined by the present local mass density. We determine which combinations of initial power spectrum and biasing prescription are consistent with the observed clustering of optical galaxies and use the observed relation between the distributions of optical and *IRAS* galaxies to predict corresponding redshift-space *IRAS* power spectra. These are compared with observations, as are the pairwise velocity dispersions predicted by the models. In this way, building, in part, on our results from Chapter 3, we are able to construct a coherent picture of galaxy clustering which is in accord with our results on cluster correlations from Chapter 2, showing that galaxies and clusters are consistent probes of the large-scale structure of the Universe.

ACKNOWLEDGEMENTS

It has been my good fortune and my great pleasure to study for the past three years under the supervision of Alan Heavens and John Peacock. They have given me the freedom to make my own mistakes (and, hopefully, learn from them!), while ensuring that I did not stray too far. They have always taken time to answer my questions, no matter how dumb or oft-repeated, and have pored over more lines of my FORTRAN code than is good for anyone. A lot of their enthusiasm for cosmology has rubbed off on me during that time, as has, I hope, a little of their knowledge. The work contained in this thesis would not have been possible without them: that may sound trite and be something of a truism, but the gratitude it expresses is very sincerely felt – thanks a lot, chaps.

Draft chapters from this thesis benefitted greatly from being read by Alan, John, Phil James, Will Saunders and Andy Taylor. I am very grateful to all of them, but I would like to thank Phil in particular, as he has always been very generous with his time and very insightful with his criticisms whenever I have had work to be read.

A number of people have kindly provided data for my use and patiently answered my subsequent queries: my thanks to Chris Collins, Gavin Dalton, Stuart Lumsden, Bob Nichol and Kathy Romer for information regarding their various cluster redshift surveys; Shaun Cole and Rupert Croft for the cluster correlation data from their respective simulations; the QDOT and 2 Jy survey teams for access to their galaxy catalogues; Seb Oliver for the counts-in-cells software used in Chapter 3; and Ed Bertschinger and Jim Gelb for providing the numerical simulations which form the basis of Chapter 4. I am especially grateful to Will Saunders for a large number of very informative emails about

IRAS galaxies, which made possible the project outlined in Chapter 3 and to Andy Taylor, who has been a fine office-mate for the past year and a great sounding board for ideas.

I am very grateful to the SERC for the studentship that supported me for three years and helped finance my attendance of a number of conferences and summer schools, further support for which was gratefully received from the Department of Astronomy and from the organisers of the meetings. Thanks as well to Ian Roxburgh and the Astronomy Unit at QMW for offering me a postdoc: knowing I had a job to go to removed a lot of stress from the past few months and a deadline that proved to be flexible ensured that I wrote up in reasonably good time, while maintaining some semblance of sanity.

I have enjoyed the company of all the students in the Department that I have got to know during the last three years. While it is always dangerous to mention some, but not all, by name, special mentions must go to several groups: the other members of my year (Ruth Doherty, Isabelle Parkes and Kathy Romer); those (namely Keith Ballard, Dave Buckley, Antonio Chrysostomou, Pippa Goldschmidt, Bob Nichol, Mike Read, Suzie Ramsay, Steve Torchinsky and Brian Williams) who made the Observatory such an easy place to settle into; and, of course, the other members (Antonio Chrysostomou, Amadeu Fernandes, Alistair Glasse and Mike Read) of the magnificent five-a-side footie team which trounced the Bank of Scotland in a glorious match, the memory of which still brings the broadest of grins to my face a full year later and which I count as one of the finest moments of my life so far. Many thanks to Dave Buckley, Ruth Doherty and Mike Read, whose cassette players I have borrowed and, in one case, broken (sorry again Mike!) while writing up: it is left as an exercise for the reader to deduce which sections were written listening to Purcell and which to Dinosaur Jr.

In addition to those named above, I have enjoyed discussions of things astronomical and down-to-earth with many other members of the Department and the Observatory, especially Peter Brand, Andy Connolly, Liz Gibson, Mike Hawkins, Lance Miller, Phil Puxley, Adrian Webster, Martyn Wells and the guys in the Lodge (Drew Anderson, Bill MacLeod, Jack Paterson and Scott Watson) plus, of course, Mary Heaton: the stress of writing up is nothing in comparison to the onerous responsibility of being left to water

Mary's plants.

Outside the Observatory I have been very lucky with three great sets of flatmates: from my first year Bob Bliss and Eric Wolf deserve a special mention (Eric was also the source for the few pious quotes I have included and for which I make no apology – I can resist everything except quotation); Antonio Chrysostomou, Andy Connolly, Bob Nichol and Mike Read were "The Lads in The Flat" in my second year; while, for the past year or so, Bob Akroyd and Kay Tisdall have shared with me the cold of Findhorn Place, as well as many good times.

All the aforementioned people have contributed to three very enjoyable years here, as well as putting up with me during the suspension of joviality that is writing up. Of particular note on both counts is, of course, Kay Tisdall. My time in Edinburgh would have been far less memorable had not so much of it been spent with her and in the past few months she has had to put up with my coming and going at odd hours and in odder moods, while writing up her own thesis and starting a demanding job: the consideration she has shown has been remarkable, although, of course, nothing more than I would have expected from her.

Finally, I come to my family: Malcolm, Ann and David Mann. Their love and support – emotional, intellectual and financial – has enabled me to do pretty much all I have wished over the past twenty-five years. I doubt I can ever repay them for that but, in dedicating this thesis to them, I can at least acknowledge my great debt to them.

Chapter 1

INTRODUCTION

ταῦτα δεδοξάσθω μὲν 'εικὸτα τοῖς 'ετὺμοισι ...

Let these things be opined as resembling the truth ...

Xenophanes

Amongst the most fundamental questions addressed by the physical sciences are those of cosmology, concerning the formation and evolution of the large-scale structure of the Universe. Cosmology has matured greatly in recent decades, as increasingly sophisticated techniques – observational, theoretical and computational – have been brought to bear upon its subject matter. Rather than attempt to survey all of these advances, this Introduction provides the context in which to set the work described in subsequent chapters of this thesis. General references covering this background material are the books by Weinberg (1972), Kolb & Turner (1990) and Peebles (1993) and the proceedings of the Nobel Symposium 79 (Nilsson *et al.* 1991), to which the reader is referred for a fuller list of original references than is presented here.

1.1 THE BIG BANG MODEL

1.1.1 The Cosmological Principle

The basic tenet of the Big Bang cosmology is the Cosmological Principle. This is, basically, a statement of cosmic humility: it posits that we do not occupy a privileged position in the Universe or, more formally, that the Universe is isotropic about all points and, hence, homogeneous. The weight of astronomical evidence is against this assumption, as we see inhomogeneities on scales up to tens of megaparsecs, so it is conventional to say that the homogeneity only holds in some statistical sense, when the Universe is averaged on some sufficiently large scale. Formal problems arise when one takes averages in relativistic cosmologies (Ellis 1988), but we shall follow convention in neglecting these and press on with discussion of isotropic, homogeneous models.

1.1.2 The Friedmann-Lemaître-Robertson-Walker Metric

The Cosmological Principle motivates the study of homogeneous, isotropic spacetimes. The metric for the full set of such spacetimes is that studied variously by Friedmann, Lemaître, Robertson and Walker (FLRW). The FLRW metric may be derived by employing the following chain of reasoning (Islam 1992): for a full, rigorous derivation the

reader is referred to Weinberg (1972).

Consider a spacetime. The Cosmological Principle requires that for this spacetime to be used as a model for the Universe it must be homogeneous. It follows, therefore, that we can foliate the spacetime into a set of space-like hypersurfaces, labelled by some parameter t and that these hypersurfaces are homogeneous. If this foliation is performed such that the surface of simultaneity of the local Lorentz frame of any observer lying in a particular hypersurface coincides locally with that hypersurface then the parameter t may be identified with the proper time of the observer.

We may introduce coordinates (x^1, x^2, x^3) on the space-like hypersurfaces which are constant along each of the time-like geodesics orthogonal to the hypersurfaces. It follows that the line element in the spacetime is given by

$$\mathrm{d}s^2 \;=\; c^2\,\mathrm{d}t^2 \;-\; h_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j, \tag{1.1}$$

for $i, j = 1 \ldots 3$. In general $h_{ij} = h_{ij}(t, x^1, x^2, x^3)$ and the cross terms in the metric g_{0i} must vanish if the definition of simultaneity given by $t = \text{const.}$ is to agree with that of the local Lorentz frame. The Einstein summation convention is assumed throughout and the signature of the metric is taken to be $(+ - - -)$.

Consider a particular space-like hypersurface, given by $t = t_0$. The spatial separation between a close pair of observers with coordinates (x^1, x^2, x^3) and $(x^1 + \Delta x^1, x^2 + \Delta x^2, x^3 + \Delta x^3)$ is $h_{ij}\Delta x^i \Delta x^j$. A triangle is defined by this pair of observers and a third observer in their neighbourhood on the hypersurface $t = t_0$. Consider the triangle described by the same three observers (assumed to travel along geodesics) on the hypersurface corresponding to some later time t_1. The isotropy and homogeneity of the Cosmological Principle requires the three angles within the triangle to remain constant with time, so that the lengths of its sides must increase in proportion to each other. It follows that the 3-metric h_{ij} takes the form

$$h_{ij} \;=\; a^2(t)\,\gamma_{ij}, \tag{1.2}$$

where $a(t)$ is called the scale factor and $\gamma_{ij} = \gamma_{ij}(x^1, x^2, x^3)$. By assumption the 3-space described by the metric γ_{ij} is homogeneous and isotropic. It can be shown (see

Weinberg 1972) that such spaces (called maximally symmetric because they admit the maximum number of Killing vectors) are spaces of constant curvature, in the sense that their Riemann tensor is given by

$$^{(3)}R_{ijkl} = k(\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}), \tag{1.3}$$

where k is called the curvature constant.

It is simple to verify that one 3-metric that satisfies this equation is the one yielding the comoving line element

$$\gamma_{ij}\mathrm{d}x^i\,\mathrm{d}x^j = (1 + k\hat{r}^2/4)^{-2}\left[(\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2\right], \tag{1.4}$$

where $\hat{r}^2 \equiv (x^1)^2 + (x^2)^2 + (x^3)^2$. Thus, this metric satisfies our requirements of homogeneity and isotropy. It can be shown (Weinberg 1972) that any two maximally symmetric metrics with the same value of the curvature constant k and the same signature may be related by a coordinate transformation. It follows, therefore, that we have found the unique set of metrics satisfying the Cosmological Principle and that, by the coordinate transformations

$$\begin{cases} r &= \hat{r}/(1 + k\hat{r}^2/4) \\ x^1 &= \hat{r}\sin\theta\cos\phi \\ x^2 &= \hat{r}\sin\theta\sin\phi \\ x^3 &= \hat{r}\cos\theta \end{cases}$$

they take the form of the FLRW line element, which reads:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1 - kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right]. \tag{1.5}$$

It is clear from equation 1.5 that the value of k can be scaled by a suitable choice of units for the r coordinate. It follows that k can take values $-1, 0, 1$, yielding solutions in which the 3-space is described, respectively, as being hyperbolic, flat and spherical. In equation 1.5 θ, ϕ are the usual azimuthal, meridonal angles, respectively. The coordinate r runs from 0 to ∞ for $k = 0, -1$, but only from 0 to 2π if $k = 1$: in the former two cases the space is infinite, while it is finite, although unbounded, for $k = 1$.

1.1.3 The Recession of the Nebulæ

The isotropy of the FLRW model restricts its velocity field to be either pure expansion or pure contraction, since the presence of either shear or vorticity components would define preferred directions.

Consider light emitted by a source at some time t_e, which is detected by an observer at time t_o. Without loss of generality, we may place the observer at the origin of coordinates and consider a radial path for the light, since setting $d\theta = d\phi = 0$ is consistent with the vanishing shear and vorticity in the FLRW model. Photons follow null geodesics (*i.e.* $ds^2 = 0$), in which case equation 1.5 gives the equation of motion of a particular wave crest as

$$0 = c^2 dt^2 - a^2(t) \frac{dr^2}{1 - kr^2}. \tag{1.6}$$

Integrating this along the path of the light yields

$$c \int_{t_e}^{t_o} \frac{dt}{a(t)} = \int_{r_e}^{0} \frac{dr}{\sqrt{1 - kr^2}} \equiv f(r_e), \tag{1.7}$$

where r_e is the radial coordinate of the source. Consider now the next wave crest, which is emitted at time $t_e + \delta t_e$ and observed at time $t_o + \delta t_o$. It follows from the time-independence of $f(r_e)$ that

$$\frac{\delta t_o}{a(t_o)} = \frac{\delta t_e}{a(t_e)}, \tag{1.8}$$

since the change in $a(t)$ in times δt_e and δt_o will be negligible. From this it follows that the frequency, ν_e, of emission and of observation, ν_o, of light propagating in the FLRW model are related by

$$\frac{\nu_o}{\nu_e} = \frac{\delta t_e}{\delta t_o} = \frac{a(t_e)}{a(t_o)}. \tag{1.9}$$

Thus, if we define the quantity

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e}, \tag{1.10}$$

where λ_o, λ_e are the wavelengths corresponding to the frequencies ν_o, ν_e respectively, we see that

$$1 + z = \frac{a(t_o)}{a(t_e)}. \tag{1.11}$$

If $z > 0$ then light is redshifted by propagation through an FLRW model, while it is blueshifted if $z < 0$. It is convenient to normalise the scale factor to equal unity at the present, in which case equation 1.11 shows that the value of the scale factor, $a(z)$, at the time of emission of a photon which is observed now with a redshift z is given by

$$a(z) = \frac{1}{1+z}. \tag{1.12}$$

From 1910 onwards, Slipher undertook a programme of spectroscopic observations of spiral nebulæ. By 1922 he was able to report that out of a sample of 41 spiral nebulæ, 36 had redshifted lines and only five showed blue shifts. In 1918, Wirtz suggested that Slipher's observations should be interpreted as evidence for a general recession of the nebulæ and this view was strengthened as further work by Wirtz and Lundmark in the 1920s showed that Slipher's redshifts increased with the distance to the nebulæ, which is in accordance with their being cosmological redshifts in an FLRW universe, as we shall now show.

The argument advanced by Wirtz became widely accepted following Hubble's 1929 announcement of a "roughly linear relation" between redshift and distance. The distances to the nebulæ estimated by Slipher and Hubble were luminosity distances, determined using Cepheid variable stars as "standard candles" – objects whose luminosities could be deduced (from the period–luminosity relation in the case of Cepheids) and, hence, whose apparent magnitudes could be used as measures of distance. If the luminosity of a source is $\mathcal{L}$ and the measured flux of energy from it is $\mathcal{F}$, then its luminosity distance, d_L, is defined by the relation

$$d_L^2 = \frac{\mathcal{L}}{4\pi\mathcal{F}}. \tag{1.13}$$

The reason for this choice of definition is obvious in a static universe and to see its utility in a FLRW model, consider again the emission of light by a source at coordinate

r_e at time t_e, detected by an observer at the origin at time t_o. The fraction of the area of the sphere surrounding the source that is covered by the observer's detector (with a collecting area of dA) is $dA/4\pi a^2(t_o)r_e^2$. The energy flux reaching the detector is reduced by a factor $(1+z)^2$, where z is the redshift of the source: the first factor of $(1+z)$ comes from the reduction in the energy of each individual photon that results from equation 1.9, while the second comes from the decrease in the arrival rate of photons, which follows from equation 1.8. It follows that the luminosity distance is given by

$$d_L = a(t_o)\, r_e\, (1+z). \tag{1.14}$$

The coordinate distance r_e is unobservable, so we must eliminate it. This may be done as follows, for the case of a nearby source. We first make use of the fact that $f(r_e) \simeq r_e$ for small values of r_e, irrespective of the value of the curvature parameter k. If we make a Taylor expansion for the scale factor

$$a(t_e) = a(t_o) + (t_e - t_o)\dot{a}(t_o) + (t_e - t_o)^2 \frac{\ddot{a}(t_o)}{2} + \ldots, \tag{1.15}$$

where the dot denotes differentiation w.r.t. t and substitute this into equation 1.7 we obtain

$$r_e \simeq \frac{c}{a(t_o)} \left[(t_o - t_e) + \frac{1}{2}\frac{\dot{a}(t_o)}{a(t_e)} + (t_e - t_o)^2 \frac{\ddot{a}(t_o)}{2a(t_o)} + \ldots \right] \tag{1.16}$$

The same Taylor series may be used in equation 1.11 to obtain

$$1 + z = \left[1 + (t_e - t_o)\frac{\dot{a}(t_o)}{a(t_o)} + (t_e - t_o)^2 \frac{\ddot{a}(t_o)}{2a(t_o)} + \ldots \right], \tag{1.17}$$

which we may invert for small $(t_e - t_o)\,\dot{a}(t_o)/a(t_o)$ and solve for $(t_e - t_o)$, to yield

$$(t_o - t_e) = \frac{a(t_o)}{\dot{a}(t_o)} \left\{ z - \left[1 - \frac{\ddot{a}(t_o)a(t_o)}{\dot{a}^2(t_o)} \right] z^2 + \ldots \right\}. \tag{1.18}$$

Substituting this into equation 1.16 and then substituting the resulting expression for r_e into the definition for the luminosity distance we obtain

$$\frac{\dot{a}(t_o)}{a(t_o)}\, d_L = cz + \frac{c^2 z^2}{2} \left[1 + \frac{\ddot{a}(t_o)a(t_o)}{\dot{a}^2(t_o)} \right] + \ldots. \tag{1.19}$$

Comparing this with Hubble's observations in the low redshift limit that he studied shows that the quantity $\dot{a}(t_o)/a(t_o)$ is to be identified with Hubble's constant, H_0, if we are to interpret the redshifts observed by Slipher and Hubble as resulting from the expansion of a FLRW universe.

A comprehensive historical account of the observational estimation of H_0 is presented by Rowan-Robinson (1985), while the review articles by Jacoby *et al.* (1992) and van den Bergh (1992) summarise the current state of play in this controversial field. In this thesis we adopt the convention of writing Hubble's constant as $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, with the observational uncertainty as to its value translating to the fact that h can only be constrained to lie within the range $1 \gtrsim h \gtrsim 0.5$.

1.1.4 General Relativity

The determination of the dependence of the scale factor, $a(t)$, in equation 1.5 on time, t, requires the specification of the dynamics of the FLRW model. The standard Big Bang cosmology assumes that this is correctly given by the field equations of general relativity, which read

$$G_{\mu\nu} = 8\pi G\, T_{\mu\nu}, \tag{1.20}$$

where the Einstein tensor, $G_{\mu\nu}$, describes the curvature of the spacetime, $T_{\mu\nu}$ is the energy-momentum tensor, G is Newton's gravitational constant and we use units in which the speed of light is unity.

The symmetries of the FLRW model dictate that the energy-momentum tensor must take the form

$$T_{\mu\nu} = Au_\mu u_\nu + Bg_{\mu\nu} \tag{1.21}$$

where u_μ is the 4-velocity. By analogy with the corresponding case in special relativity, this form for $T_{\mu\nu}$ may be identified with that of the perfect fluid

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + pg_{\mu\nu}, \tag{1.22}$$

with the additional constraint that the pressure, p, and energy density, ρ, must depend only on time. With this form for $T_{\mu\nu}$ it follows that

$$T^{\mu\nu}_{;\nu} = 0, \tag{1.23}$$

where the semi-colon denotes covariant differentiation. The spatial components of this equation are satisfied identically, while the temporal component yields the First Law of Thermodynamics, in the form

$$\mathrm{d}(\rho a^3) = -p\,\mathrm{d}(a^3). \tag{1.24}$$

The Bianchi identities ensure that

$$\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)_{;\nu} = 0, \tag{1.25}$$

where $R^{\mu\nu}$ is the Ricci tensor, $g^{\mu\nu}$ is the metric tensor and R is the Ricci scalar. The term in brackets in equation 1.25 is, therefore, a candidate for being the Einstein tensor, but $g^{\mu\nu}_{;\nu} \equiv 0$, so that the Einstein tensor can only be defined up to a constant multiple of the metric tensor. It follows, therefore, that Einstein's equations take the form

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R - \Lambda g^{\mu\nu} = 8\pi G T^{\mu\nu}, \tag{1.26}$$

where Λ is called the cosmological constant. Einstein first used the freedom to introduce the Λ-term into equation 1.26 when he noticed that without it the field equations would admit no static solutions, which was abhorrent to his intuition. In the light of the discovery of the recession of the nebulæ by Slipher and Hubble, Einstein removed the Λ-term from the field equations, dubbing it "the biggest blunder of my life". Since then the cosmological constant has often been the last resort of a theorist unable to account for some observational data (*e.g.* the redshift distribution of quasars: Petrosian *et al.* 1967): cosmologies with non-zero values of Λ will be met from time to time in this thesis and the reader may wish to recall these historical precedents at each re-appearance. An excellent review of the chequered career of the cosmological constant and the controversies that still surround it is presented by Carroll *et al.* (1992).

1.1.5 The Friedmann Equations

Substituting the FLRW metric into the Einstein equations with the perfect fluid form for $T_{\mu\nu}$ yields the following pair of coupled equations

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G\rho}{3} + \frac{\Lambda}{3} \tag{1.27}$$

and

$$2\left(\frac{\ddot{a}}{a}\right) + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = -8\pi G p + \frac{\Lambda}{3}. \tag{1.28}$$

These are the Friedmann equations, which describe the dynamics of the FLRW model, when supplemented with an equation of state, $p = p(\rho)$. Three equations of state will be of particular interest to us: those for matter–, radiation– and vacuum-dominated universes. The Universe today is matter-dominated and is well approximated by dust (*i.e.* $p_{\rm m} = 0$), in which case equation 1.24 gives the variation of the density, $\rho_{\rm m}$, in non-relativistic matter as

$$\rho_{\rm m} \propto a^{-3}. \tag{1.29}$$

Another interesting equation of state is that for radiation, where $p_{\rm m} = \rho_{\rm r}/3$, in which case the energy density, $\rho_{\rm r}$, in radiation varies as

$$\rho_{\rm r} \propto a^{-4}. \tag{1.30}$$

Finally, the false vacuum state in the inflationary models discussed in Section 1.2 has the equation of state $p_{\rm v} = -\rho_{\rm v}$, in which case the vacuum energy density, $\rho_{\rm v}$, does not vary with scale factor.

The Friedmann equations may be combined to yield the Raychaudhuri equation for isotropic, homogeneous models, which reads

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}. \tag{1.31}$$

Consider the situation where $\Lambda = 0$ and $(\rho + 3p) \geq 0$: this latter is the so-called Strong Energy Condition (*e.g.* Hawking & Ellis 1973) and it holds true for all known kinds of classical matter. In that case, $\ddot{a} \leq 0$ which, coupled with the fact that $\dot{a} > 0$ today (by the observed recession of the nebulæ) implies (Hawking & Penrose 1970) that $a = 0$ at some finite time in the past, which it is clearly sensible to label as $t = 0$. This is the Big Bang. It can be shown that scalars defined by contraction from the curvature tensor diverge as $t \to 0$, so that the Big Bang is a curvature singularity. Such a point must be excluded from the spacetime manifold and, thus, the Big Bang marks the point at which classical physics breaks down. This immensely important result follows solely from the symmetries of the FLRW metric and a reasonable assumption about the equation of state.

A positive value of the cosmological constant can push the singularity further back in time and can prevent it happening for a sufficiently large, positive value of Λ. Another way to get around the singularity is to argue that quantum processes will be important at times before the Planck era ($t_{\mathrm{P}} \sim 10^{-43}$ s) and thus a general relativistic analysis should not be expected to hold earlier than that time.

We may define the (dimensionless) density parameter, Ω, as

$$\Omega \equiv \frac{\rho}{\rho_{\mathrm{c}}}, \tag{1.32}$$

where we have defined

$$\rho_{\mathrm{c}} \equiv \frac{3(\dot{a}/a)^2}{8\pi G}, \tag{1.33}$$

in which case equation 1.27 may be written as

$$\left(\frac{\dot{a}}{a}\right)^2 (1 - \Omega) = -\frac{k}{a^2} + \frac{\Lambda}{3}. \tag{1.34}$$

From equation 1.34 we can see that the density, ρ_{c}, is that which yields $\Omega = 1$ in a spatially flat (*i.e.* $k = 0$) universe with $\Lambda = 0$: as a consequence of this, ρ_{c} is called the critical density.

In general, the density parameter, Ω, will be the sum of contributions from a number of components, such as relativistic matter, non-relativistic matter, vacuum energy, etc. We have seen above that the densities in these components vary differently with the scale factor. Given the knowledge of the contributions from these components at a given time, say, the present, it is possible to integrate the Friedmann equations to determine the expansion age of the Universe — *i.e.* the age of the FLRW model which would have the expansion rate quantified by the value of H_0, given the assumed model for the matter constituents of the Universe. A particularly simple and important case is that of the Einstein – de Sitter universe, which is the FLRW model with $\Lambda = 0$ and $\Omega_0 = 1$ in a component whose equation of state is $p = \omega\rho$, in which case the Friedmann equations can be integrated to yield the following relation for the expansion age, t_0:

$$t_0 = \frac{2}{3(1+\omega)H_0}. \tag{1.35}$$

Evaluating this for the case of a matter-dominated model ($\omega = 0$), which seems most appropriate for the Universe, gives an expansion age of 7 Gyr $< t_0 <$ 14 Gyr, where the age range results from uncertainty in the value of H_0. The oldest globular clusters appear to be $\sim 12 - 18$ Gyr old, so that the larger possible values of H_0 are excluded by this argument if we live in an Einstein – de Sitter universe.

1.1.6 Primordial Nucleosynthesis in the standard Big Bang model

If we can neglect the cosmological constant term in the Friedmann equations and if, as we have seen, the scale factor, $a(t)$, is a monotonically increasing function of time, t, then it follows from equations 1.29 and 1.30 that some early phase of the universe must have been radiation-dominated.

We can estimate the redshift, $z_{\rm eq}$, of matter-radiation equality from a knowledge of the energy densities in radiation and non-relativistic matter today. The dominant contribution to the energy density in radiation today is the cosmic background radiation (CBR, see Section 1.1.7), which has the spectrum of a black body with a temperature $T_0 \simeq 2.7$K (Mather *et al.* 1990). From this it follows that the current radiation density

is $\rho_{r,0} \simeq 5 \times 10^{-31}$ kg m^{-3}. The current density in non-relativistic matter may be written as $\rho_{m,0} \equiv \Omega_{m,0}\, \rho_{c,0}$, which may be evaluated as $\rho_{m,0} = 1.9 \times 10^{-26}\, \Omega_{m,0}\, h^2$ kg m^{-3}, where the presence of the factor h^2 indicates that the uncertainty in the value of the Hubble constant translates into an uncertainty in the definition of the critical density, through equation 1.33. Combining these values for $\rho_{r,0}$ and $\rho_{m,0}$ with equations 1.12, 1.29 and 1.30 yields the following expression for the redshift of matter-radiation equality:

$$1 + z_{eq} = 4 \times 10^4\, \Omega_{m,0}\, h^2. \tag{1.36}$$

This will be of great importance when we come to discuss the evolution of cosmological density perturbations in Section 1.4, but for the time being we just note that equation 1.36 shows that at epochs earlier than $z \sim 10^4$ we may treat the Universe as being radiation-dominated. In this case, we may obtain the following scalings for the conditions of the early universe:

$$\frac{T}{10^{10}\mathrm{K}} \sim \frac{k_B T}{1\mathrm{MeV}} \sim \frac{1+z}{10^{10}} \sim \left[\frac{t}{1\mathrm{sec}}\right]^{-1/2}, \tag{1.37}$$

where k_B is the Boltzmann constant. In the 1940s, Gamow realised that the relations of equation 1.37 imply that the nuclei of light elements could have been produced in the hot, early phase of a FLRW universe, since their binding energies are typically $\sim 1 - 10$ MeV. Gamow erroneously believed that neutrons were the only matter constituents present initially but by 1950, when Hayashi recognised the role of neutron-proton equilibration, the framework for standard Big Bang nucleosynthesis (SBBN) was set. The classic paper by Wagoner, Fowler & Hoyle (1967) saw the first computation of the predicted abundances of light elements using a full network of SBBN reactions. The codes used today are essentially the same as that employed by Wagoner *et al.* , with only minor modifications resulting from improved experimental determinations of parameters of the model, such as reaction rates. This attests not only to the quality of the work of Fowler *et al.* , but also to the fact that the computation of light element abundances is a well posed problem in the hot Big Bang model, producing robust predictions for comparison with observations. An excellent recent review of all aspects of SBBN is presented by Smith *et al.* (1993).

The period of interest here is $10^{-2}\text{s} \lesssim t \lesssim 10^{2}\text{s}$, at the start of which the constituents of the universe are assumed to be a homogeneous mix of (black body) photons and leptons, with trace amounts of nucleons. The determination of the width of the Z° decays at CERN indicates that we should consider three generations of leptons (although, in fact, this result had been predicted from the SBBN limit on the ^{4}He abundance before the discovery of the Z°). This fixes the total density of non-relativistic matter and, hence, the expansion rate and temperature of the universe at a particular time, t, given the assumption of radiation-domination. The thermal history of the Universe through the period of primordial nucleosynthesis is profitably described in the following five logarithmic time-steps:

1. $t < 10^{-2}\text{s}$, $T > 10^{11}\text{K}$: The reaction rates of weak interactions are much higher than the expansion rate of the universe, so neutral current weak interactions

$$\text{e}^+ + \text{e}^- \longleftrightarrow \nu_\text{i} + \bar{\nu}_\text{i} \quad \text{i} = \text{e}, \mu, \tau \tag{1.38}$$

maintain the leptons in equilibrium, while charged current weak interactions

$$\text{p} + \text{e}^- \longleftrightarrow \text{n} + \nu_\text{e} \quad \text{and} \quad \text{n} + \text{e}^+ \longleftrightarrow \text{p} + \bar{\nu}_\text{e}, \tag{1.39}$$

plus beta decay and its inverse

$$\text{n} \longleftrightarrow \text{p} + \text{e}^- + \bar{\nu}_\text{e} \tag{1.40}$$

maintain the nucleon equilibrium ratio

$$\frac{n_\text{n}}{n_\text{p}} = \exp\left(\frac{-\Delta m}{T}\right). \tag{1.41}$$

Here, n_n and n_p are the neutron and proton number densities respectively, $\Delta m = m_\text{n} - m_\text{p}$ is the neutron-proton mass difference and we use units where the Boltzmann constant and the speed of light are unity.

2. $t \sim 10^{-1}\text{s}$, $T \sim 3 \times 10^{10}\text{K}$: The neutrinos freeze out (*i.e.* decouple from the other material components) when, first, the neutral current and then, slightly later, the charged current weak interactions become unimportant, as their reaction rates become comparable with the expansion rate and the expansion of the universe moves particles apart before they have time to react with each other. Beta decay, being a one-body process, persists and so $\bar{\nu}_\text{e}$ freeze out slightly later.

3. $\underline{t \sim 1\text{s}, T \sim 10^{10}\text{K}}$: Charged current weak interactions have ceased by now, so that the nucleons cannot stay in equilibrium and they freeze out at a temperature $T_{\rm f}$, when

$$\frac{n_{\rm n}}{n_{\rm p}} = \exp\left(\frac{-\Delta m}{T_{\rm f}}\right). \tag{1.42}$$

The n/p ratio falls below this slightly due to neutron decay, but soon all the neutrons are bound up in nuclei, where they are stable, and the neutron/proton ratio remains constant thereafter.

4. $\underline{t \sim 10\text{s}, T \sim 3 \times 10^{9}\text{K}}$: Before this time, electrons and positrons had been in equilibrium with the radiation, due to pair creation/annihilation

$$\mathrm{e}^{+} + \mathrm{e}^{-} \longleftrightarrow \gamma + \gamma \tag{1.43}$$

plus Compton scattering

$$\mathrm{e}^{\pm} + \gamma \longrightarrow \mathrm{e}^{\pm} + \gamma. \tag{1.44}$$

Now, pair production slows, but annihilation continues, increasing the number density of photons by a factor 11/4: the origin of this factor being in the requirement of a continuous change in the entropy density (which is proportional to the product of the photon density and the number of degrees of freedom in the system) across the annihilation transition, where the number of degrees of freedom falls by a factor 11/4, due to the loss of $\mathrm{e}^{\pm}$.

5. $\underline{t > 10^{2}\text{s}, T < 10^{9}\text{K}}$: As the radiation cools in the expanding universe, the photons become less capable of photodissociating those nuclei that form and by $t \sim 10^{2}$s element production starts in earnest. By $t \sim 10^{3}$s, the epoch of primordial nucleosynthesis draws to a close, as the Coulomb repulsion between nuclei prevents reactions from taking place at rates faster than the expansion of the universe can disperse the would-be reactants and the primordial abundances of the light elements are fixed.

The abundances of light elements left at the end of this period of activity depend on the conditions in the universe, which may be quantified in terms of

$$\eta_{10} = 10^{10}\,\eta = 10^{10}\,\frac{n_{\rm N}}{n_{\gamma}}, \tag{1.45}$$

where n_N, n_γ are the present number densities of nucleons and photons, respectively. The value of η_{10} is unknown and so a number of values of it are studied. Note that the following relation holds between the value of η_{10} and the fraction of the closure density present in baryons, Ω_B:

$$\Omega_B h^2 = 0.015 T_{2.75}^3 \eta_{10}, \tag{1.46}$$

where $T_{2.75}$ is the current black body temperature of the CBR in units of 2.75K ($T_{2.75} \simeq 1$, see Section 1.1.7).

The light element abundances which are studied are those of d (deuterium), ^{3}He, ^{4}He and ^{7}Li. Heavier elements than these are not produced in SBBN to any appreciable extent due to the absence of stable mass-5 and mass-8 nuclei which can be produced by two-body collisions: these are the only processes that can take place in the early universe, when the baryon density is low and the available time short — heavier nuclei can be formed only in stellar interiors, when the density is high enough and the timescales long enough for appreciable yields from much rarer three-body collisions.

The abundances of these four light nuclei (relative to hydrogen) are computed from the following network of twelve crucial reactions:

$n \leftrightarrow p$	$p(n,\gamma)d$	$d(p,\gamma)^3He$	$d(d,n)^3He$
$d(d,p)^3H$	$^3H(d,n)^4He$	$^3H(\alpha,\gamma)^7Li$	$^3He(n,p)^3H$
$^3He(d,p)^4He$	$^3He(\alpha,\gamma)^7Be$	$^7Li(p,\alpha)^4He$	$^7Be(n,p)^7Li$

The cross-sections for these reactions may be determined experimentally and applied, with appropriate temperature-dependent corrections, to SBBN. As would be expected from this complicated network of reactions, the final abundances of the nuclei of d,^{3}He and ^{7}Li reflect a balance between competing production and destruction processes, with the position of the balance depending sensitively on η_{10}. The abundances of d and ^{3}He are both decreasing functions of η_{10} (reflecting the fact that if the density is higher, they are more likely to be destroyed, to produce, eventually, ^{4}He) while ^{7}Li exhibits a bimodal behaviour, with a dip in its abundance at $\eta_{10} \sim 3$. The bimodality exhibited by ^{7}Li is readily explainable in terms of two competing methods of production: at low densities,

^{7}Li is primarily produced by ^{3}H$(\alpha,\gamma)^7$Li and both ^{3}H and ^{7}Li are readily destroyed, so the lithium abundance declines with increasing density; at higher densities, lithium production proceeds largely through the electron capture decay of ^{7}Be, which is more robust, so the lithium abundance increases with density.

By contrast, the primordial abundance of ^{4}He, which is denoted by $Y_{\rm P}$, is very insensitive to η_{10}. This is because virtually all the neutrons that remain after freeze out end up in ^{4}He because it is the most tightly bound nucleus in the network of accessible reactions, so $Y_{\rm P}$ is only sensitive to the freeze-out temperature. This depends on the strength of the weak interactions, which may be quantified by the half-life, $\tau_{1/2}$, of free neutrons: increasing $\tau_{1/2}$ weakens the weak interactions, which leads to decoupling at a higher temperature, hence a higher n/p ratio at freeze-out and a higher relative abundance of ^{4}He. The determination of the observed abundance of ^{4}He may be made from observations of emission lines in metal-poor extragalactic HII regions and dwarf galaxies, where helium is seen due to the recombination of He^+. The ^{4}He abundance is determined as a function of metallicity and the results extrapolated to zero metallicity to yield an estimate of $Y_{\rm P}$. Opinions differ as to the best metallicity tracer to use, but a primordial helium abundance of $0.21 \leq Y_{\rm P} \leq 0.24$ is advanced by Smith *et al.*, which provides an upper bound to η_{10}.

The observations of deuterium can only yield a lower limit to its primordial abundance, since deuterium is always destroyed in stars, largely to produce ^{3}He. The amount of deuterium that survives depends sensitively on the stellar models, limiting the utility of the deuterium abundance as a cosmological probe. Much of this model-dependence can be removed by considering, instead, the combination d + ^{3}He, which provides a lower bound to η_{10}.

Combining the upper bound from ^{4}He and lower bound from d + ^{3}He constrains η_{10} to lie in the interval $2.86 \leq \eta_{10} \leq 3.77$ for SBBN to be consistent. Observations of the ^{7}Li abundance are consistent with this interval, although their interpretation is complicated by the apparent sensitivity of the ^{7}Li destruction rate in stars to complexities of rotation, convective mixing and mass loss. This allowed range of η_{10} corresponds to $0.011 \leq \Omega_{\rm B}\, h^2 \leq 0.015$. Combining this with sensible limits on the value of h leads to

the conclusion that there cannot be a closure density in baryons. (It also leads one to conclude that there must be some baryonic dark matter, but discussion of this point is left until Section 1.3.)

There remain, however, ways to avoid the conclusion that $\Omega_B \ll 1$. The first is to doubt the reliability or universality of the observed abundances. A second, and more interesting, approach is to drop the assumption that nucleosynthesis took place in a homogeneous plasma (see Sato & Teresawa 1991 and references therein). If the quark-hadron phase transition is strongly first order, then large inhomogeneities could be produced at temperatures $T \geq 100$ MeV. Neutrons can pass through the primordial plasma much more easily than protons, so they will preferentially diffuse from these overdense regions, resulting in nucleosynthesis taking place in conditions of varying density and neutron/proton ratios. The earliest inhomogeneous Big Bang nucleosynthesis models appeared to be able to reconcile $\Omega_B = 1$ with the observed abundances of d,^{4}He and ^{3}He, but only at the price of overproduction of ^{7}Li. This lithium overproduction problem seemed to disappear in later work, following the inclusion of the diffusion of neutrons back to the overdense regions after nucleosynthesis had taken place in the low density regions, but this back diffusion also had the unfortunate consequence of over-producing ^{4}He for $\Omega_B = 1$. Fine-tuning of the parameters of the phase-transition model was found to be capable of reconciling the abundances of ^{4}He and ^{7}Li, but only in a very contrived way, since any realistic model would include a distribution of conditions and would then over-produce helium, lithium or both. Furthermore, it has been argued that that the phase boundaries at the quark-hadron transition are likely to be fractal, giving them large surface areas, increasing the interaction between the phases and reducing the magnitude of any of these exotic effects.

1.1.7 The Cosmic Background Radiation

Gamow realised that a corollary of the synthesis of light nuclei in the hot Big Bang model is the presence of a cosmic background of photons today, principally those generated in the $e^+ e^-$ annihilation at $t \sim 10 - 100$s. Two of Gamow's collaborators (Alpher & Herman 1948) calculated that this cosmic background radiation (CBR) should have the

spectrum of a black body at a temperature of $T_0 \sim 5$K. Although it was not realised until much later, direct evidence for the existence of the CBR had been observed several years earlier. Adams (1941) had detected absorption lines in interstellar clouds due to rotationally excited states of CN as well as its ground state and the relative strengths of these lines had been used (M^{c}Kellar 1941) to infer an effective excitation temperature of $T \sim 2.3$K for the CN molecules, in accordance with the predictions of Alpher & Herman, given the various uncertainties involved. The importance of these results was not, however, appreciated until twenty-five years later, after what is commonly called the discovery of the CBR.

The fascinating series of events leading up to this "discovery" are well described by Peebles (1993). In 1964, Dicke's group at Princeton and Zeldovich's in Moscow were interested in the CBR as a probe of the early universe. Two members of the Moscow School (Doroshkevich & Novikov 1964) calculated a strong bound on the present CBR temperature based on a report by Ohm (1961) of excess noise present in a radio telescope at Bell Telephone Laboratories in Holmdel, New Jersey: Ohm suspected that the source of the noise was ground radiation being picked up by the back lobes of the antenna, but it seemed stronger than expected from similar antennæ. The CBR temperature bound ($T_0 \lesssim 1$K) computed by Doroshkevich & Novikov was judged to be too low for compatibility, within the context of Gamow's model, with the observed ^{4}He abundance, thus leading Zeldovich (1965) to consider a cold Big Bang model.

Meanwhile, unaware of the excess noise problem in the radio telescope thirty miles away at Bell Labs, Dicke proposed the construction at Princeton of a radiometer capable of detecting the CBR. Dicke's goal, however, was not the CBR of the Gamow model, but the background of thermal photons generated in an oscillating universe. Dicke assigned the task of building the radiometer to Roll & Wilkinson and the task of interpreting its results to Peebles.

In February 1965, Peebles gave a seminar on the project at Johns Hopkins University. News of this talk reached Penzias & Wilson at Bell Labs, who had been planning to use the Holmdel radio telescope for radio astronomy, but whose efforts had been thwarted by the excess noise problem detailed by Ohm, which they had been unable to account for by

any instrumental effect. A meeting was arranged between Penzias & Wilson and Dicke's group, which reached the view that Ohm's excess noise, now known to be isotropic over the sky, could well be the CBR. Following this meeting there came the publication of the famous pair of papers: one by Penzias & Wilson (1965), announcing an "excess antenna temperature" of $T_0 = 3.5 \pm 1.0$ K at a wavelength of 7 cm and one by Dicke *et al.* (1965) interpreting this as "cosmic blackbody radiation". A few months later Roll & Wilkinson (1966) announced at a detection at 3.2 cm with a temperature of $T_0 = 3.0 \pm 0.5$ K, in excellent agreement with the interpretation of the Penzias & Wilson result as being the CBR.

The spectrum of the CBR has subsequently been shown to follow that of a black body over four orders of magnitude in frequency. The FIRAS (Far InfraRed Absolute Spectrophotometer) instrument on board *COBE* (The Cosmic Microwave Background Explorer) has shown that between 0.5 mm and 1 cm the CBR spectrum deviates from a black body with a temperature of $T_0 = 2.735 \pm 0.010$ K by less than 0.03% of the peak intensity: this result (Mather 1993) comes from just the first six weeks of data and a great improvement should be possible by the end of the *COBE* mission.

The CBR photons observed by *COBE* probe the conditions of the Universe at the epoch of last scattering at a redshift of $z \sim 1000$: this is the point at which the CBR photons decoupled from the primordial plasma, since Thomson scattering could no longer maintain thermal contact once the photon mean free path became comparable to, and then exceeded, the Hubble radius. Anisotropies in the sky distribution of CBR photons are produced by anisotropies on the surface of last scattering (primary anisotropies), as well as by the conditions along the line of sight (secondary anisotropies). By far the largest anisotropy is, however, the dipole component (with amplitude $\Delta T/T \sim 10^{-3}$), which is interpreted as being due to the motion of the satellite with respect to the cosmic rest frame defined by the CBR.

The relative importance of the various primary anisotropies depends on angular scale. At large angular separations, such as those studied with the DMR (Differential Microwave Radiometer) on *COBE*, the dominant term is the Sachs-Wolfe (Sachs & Wolfe 1967) effect, due to gravitational potential fluctuations on the surface of last scattering. In the

Newtonian gauge the Sachs-Wolfe effect can be understood readily (Peacock 1991) as the sum of gravitational redshift and time dilation terms of opposite sign resulting from photons climbing out of potential wells on the last scattering surface. The large angular scales probed by the *COBE* DMR correspond to super-horizon scales at decoupling. This greatly simplifies their analysis by removing any dependence on microphysics and, hence, sensitivity to the material constituents of the universe. The *COBE* DMR results reported by Smoot *et al.* (1992) indicate temperature anisotropies consistent with a scale-invariant (*i.e.* $|\delta_k|^2 \propto k$) primordial power spectrum with an rms–quadrupole–normalised amplitude of $Q_{\rm rms} = 16 \pm 4\,\mu$K: we shall use this amplitude extensively in what follows and postpone discussions of possible complications in the interpretation of the *COBE* DMR result due to a contribution from gravitational waves until later.

Anisotropies on scales $\lesssim 1^\circ$ are much more complicated to analyse, arising as they do from microphysical phenomena, making them much more sensitive to the assumed material constituents of the universe. Two effects contribute to the small-scale anisotropy: the first is a simple Doppler term, due to the CBR photons scattering off baryons falling into potential wells on the last scattering surface, while the second, the adiabatic term, is slightly more complicated in origin. The photon fluid will respond to variations in the density of the plasma, so that an overdense region at $z = 10^3$, say, will have a larger photon density too (excepting the case of isocurvature fluctuations, when the photon density will be lower, to compensate for the matter overdensity). This means that the photons will be hotter and so will recombine slightly later. If recombination takes place at a fixed temperature (ignoring density effects), then this means that the photons from this overdense region which are observed today will have suffered a smaller redshift due to their more recent last scattering event and, hence, will appear to be hotter. The adiabatic term is dominant on the smallest scales, with a region of Doppler dominance on intermediate scales. On the arcminute scales, anisotropies can be washed out as a result of the superposition of different temperature components, due to the finite thickness of the surface of last scattering.

Potential fluctuations in an Einstein - de Sitter universe are time-independent to first order in perturbation theory, so that photons do not undergo a change in energy at first order due to their passage through collapsing overdensities/expanding underdensities

along the line of sight from the surface of last scattering. This time-independence does not hold at all orders, however, and the resulting Rees-Sciama effect (Rees & Sciama 1968) is our first example of a secondary anisotropy. Another source of anisotropy generated along the line of sight is the Sunyaev-Zeldovich (Sunyaev & Zeldovich 1972) effect, which is caused by the inverse Compton scattering of CBR photons passing through a very hot plasma. The Rees-Sciama effect may be neglected for practical purposes, while the Sunyaev-Zeldovich effect is only important for those lines of sight which pass through rich clusters of galaxies, where there is hot intracluster gas: this will be seen to have an important consequence in Section 1.3.

Even when secondary anisotropies may be neglected the analysis of primary anisotropies is very complicated, requiring detailed comparisons of observational results with what would be detected by a particular experiment, with a particular observational set-up, for a particular cosmological model. All that, furthermore, is possible only after the correction for anisotropy induced by other Galactic and extragalactic sources: *COBE*, for example, made observations at three different frequencies to try to overcome this problem. We shall say no more here, except to note that the Doppler and adiabatic terms ensure that the CBR anisotropy should be larger on $\sim 1^\circ$ scales than on the larger scales observed by *COBE*, provided that the primordial fluctuation spectrum was scale-invariant and that $\Omega_0 = 1$. Recent observations (*e.g.* Meinhold *et al.* 1993; Gundersen *et al.* 1993) which indicate roughly the same level of temperature fluctuations on degree scales as seen by *COBE* are, therefore, difficult to reconcile with a scale-invariant primordial spectrum in an $\Omega_0 = 1$ model: we shall have cause to reconsider this point later.

1.1.8 Summary

As discussed here and, for example, in the review by Peebles *et al.* (1991), the case for the relativistic hot Big Bang, resting on the recession of the nebulæ, the abundances of light elements and the isotropy of the CBR, is a strong one. It is made even stronger by the *COBE* DMR results which may be interpreted as indicating the presence of density fluctuations at the epoch of decoupling of the right amplitude to form the large-scale

structure seen today, as we shall discuss in later chapters. Despite the great successes of this model, there remain those (*e.g.* Arp *et al.* 1990) who reject the Big Bang model and favour some variant of the Steady State cosmology of Hoyle (1948) and Bondi & Gold (1948). Whilst one may admire the tenacity and ingenuity of those who manage to steer the Steady State model through the unfriendly seas of new observational data, there will be no cause to consider their views further in this thesis.

1.2 INFLATIONARY COSMOLOGY

1.2.1 Inadequacies of the relativistic Big Bang model

The successes of the Big Bang model are impressive, but there remain a number of enigmas, as Dicke & Peebles (1979) call them, which indicate that perhaps the relativistic Big Bang model is not the whole story. These are discussed in some detail by Blau & Guth (1987), upon whose account the following description is based.

The first of these is the "flatness problem". Consider equation 1.34 for the case $\Lambda = 0$. It can be shown that this may be rewritten in the form

$$\Omega(t) = (1-x)^{-1}, \text{ where } x(t) = \frac{3k}{8\pi G\rho a^2}. \tag{1.47}$$

It follows from equations 1.29 and 1.30 that $x(t) \propto a(t)^2$ for $t \lesssim t_{\rm eq}$ and $x(t) \propto a(t)$ for $t \gtrsim t_{\rm eq}$: thus $\Omega = 1$ is an unstable solution to the Friedmann equations for equations of state relevant to the Big Bang model. Given that $1 \gtrsim \Omega_0 \gtrsim 0.1$, equation 1.47 places a very tight constraint on the possible deviation of Ω from unity at earlier times in the history of the Universe: for example, at the Planck time, $t_{\rm P} \sim 10^{-34}$s, we obtain

$$|\Omega - 1| \lesssim \mathcal{O}(10^{-60}). \tag{1.48}$$

Given that Ω is, in some sense, a free parameter, the "flatness problem" is the statement of surprise that equation 1.48 constrains Ω to be close to unity at $t_{\rm P}$. It has been argued that the form of equation 1.47 ensures that $\Omega(t) \to 1$ at early times, so that the equation 1.48 should come as no surprise: if you look at early times then Ω should be

very close to unity. This is mistaken, however, since $t_{\rm P}$ is the natural unit of time and, thus, $t = t_{\rm P}$ is not an early time in the history of the Universe. This leads to another way of looking at the "flatness problem", recasting it as the question as to how the universe could reach such an age (in units of the Planck time) and still have $\Omega \sim \mathcal{O}(1)$.

The second enigma is the "horizon problem". The essence of this is that the observed level of isotropy in the CBR requires the universe to have been homogeneous on scales relating regions of space that have never been in causal contact in the standard Big Bang model: Blau & Guth (1987) show that antipodal regions of the sky were $\sim$75 horizon distances apart at the epoch of last scattering, so what process could cause them to be at temperatures differing by only one part in $\sim 10^5$?

A third problem comes from the consideration of the growth of structure in the Big Bang model by gravitational instability. Blau & Guth (1987) show that the density perturbations that would give rise to a galaxy-mass object in the standard model are, at the Planck time, some seventeen orders of magnitude smaller than the Poisson fluctuations expected on that scale.

A further problem arises if one includes the principle of grand unification into the standard model, as one may wish to do, since the baryon non-conserving reactions predicted to take place in GUTs by the Sakharov mechanism provide a way of explaining the small, non-zero value of the η parameter discussed in Section 1.1.6. GUT phase transitions to the $SU(3) \otimes SU(2) \otimes U(1)$ low temperature standard model of particle physics produce topological defects in the Higgs field, called magnetic monopoles, with a number density (Kibble 1976) of roughly one per horizon volume. Monopole-antimonopole annihilation is very inefficient at this number density, so these monopoles should survive to the present era. Estimates of the monopole mass from simple GUTs imply that the mass density of monopoles present now is $\Omega_{\rm mono} \gtrsim 10^{11}$, which is clearly not compatible with observations.

There is also a problem with the "bang" in Big Bang itself: the standard model suggests no origin for the expansion of the universe, it just posits an expanding early state. These enigmas are not inconsistencies in the Big Bang, they just indicate inadequacies in its predictive power, requiring the assumption of very special initial conditions: for exam-

ple, Collins & Hawking (1973) show that the set of spatially homogeneous cosmological models that approach spatial isotropy at late times is a set of measure zero in the space of homogeneous models. One way around this problem is to assert that the conditions of the early universe before $t_{\rm P}$ will be set by quantum gravitational processes, so that until we have a quantum theory of gravitation we are in no position to question the "naturalness" or otherwise of any possible set of conditions at $t_{\rm P}$. One may also appeal to anthropic arguments, suggesting that perhaps only a very special set of initial conditions could produce sentient beings capable of pondering these issues. Another explanation is furnished by the idea of an inflationary epoch in the early universe, as first proposed by Guth (1981) and it is to this that we now turn our attention.

1.2.2 An Inflationary Universe

The story of the genesis of inflationary cosmology is a fascinating one and like that relating the discovery of the CBR it shows how science can often progress along very tortuous paths. Many particle physicists had considered the cosmological effects of GUT phase transitions before Guth (see the review by Olive 1990), but his stroke of genius was to see how this work could account for the enigmas described in the article by Dicke & Peebles (1979).

The essence of the model advanced in Guth's seminal paper (Guth 1981, now referred to as old inflation) is that a phase transition in the early universe could produce a period of exponential growth of the scale factor and that, provided the exponential expansion lasted sufficiently long, it could resolve the horizon and flatness problems. Old inflation is based on the minimal $SU(5)$ GUT of Guth & Tye (1980), with a first-order phase transition to $SU(4) \otimes U(1)$ occurring at a critical temperature of $T_{\rm c} \sim 10^{14}$ GeV. As the universe cools to $T_{\rm c}$, the Higgs field, ϕ, gets caught in a false vacuum state – a local minimum of the potential, $V(\phi)$ – with a potential barrier preventing ϕ from falling to the true vacuum with $V(\phi_{\rm true}) < V(\phi_{\rm false})$. Symmetry considerations dictate that the energy-momentum tensor of the false vacuum be a multiple of the metric tensor, so that its equation of state is $\rm p_f = -\rho_f =$ constant. If we consider a region of space which is

homogeneous, isotropic and flat, then equation 1.27 has the solution

$$a(t) = \mathrm{e}^{\chi t}, \quad \text{where } \chi = \left(\frac{8\pi G}{3}\rho_{\mathrm{f}}\right)^{1/2}. \tag{1.49}$$

This solution to the Friedmann equations, with a constant value of χ and, hence, an exponentially increasing scale factor, is called de Sitter space and it is the hallmark of inflation. Equation 1.49 was derived using the assumption that the region of space under consideration was homogeneous and isotropic, so that we could use the Friedmann equations, but this is not necessary: it is believed (*e.g.* Blau & Guth 1987) that there exists a 'cosmological no-hair theorem', by which perturbations about de Sitter space are smoothed out by the expansion, so that homogeneous, isotropic de Sitter will result from generic initial conditions.

Suppose that the inflationary phase continues for some time Δt, in which case the universe will expand by a factor $Z \equiv \mathrm{e}^{\chi\Delta t}$. In a first-order phase transition like the one under consideration here, the decay of the false vacuum to the true vacuum happens by quantum tunneling (Coleman 1977), with the production of bubbles of the true vacuum, which expand at the speed of light in the surrounding sea of false vacuum. Resolution of the flatness problem requires $Z \gtrsim 10^{29}$, which may seem an extraordinarily large number, but it is found (Guth & Weinberg 1981) that the minimal $SU(5)$ model can yield bubble nucleation rates sufficiently low as to produce $Z \sim 10^{10^{10}}$. So, inflation replaces the flatness problem with a prediction that Ω_0 should equal unity to great precision. The horizon problem is solved equally trivially: the horizon distance is stretched by a factor Z during inflation, so that (with $Z \sim 10^{29}$) the entire observable universe was well within the horizon at the onset of inflation and the equality of the CBR temperature at antipodal points on the sky can be explained causally.

Old inflation does, however, suffer from a fatal flaw: the so-called "graceful exit" problem. If inflation is to last sufficiently long for the horizon and flatness problems to be solved, the bubble nucleation rate must be low. It can be proved that, with a sufficiently low nucleation rate, the bubbles of true vacuum form finite sized clusters only; they do not percolate and fill all space. The bubble size reflects that of the horizon at the time of their formation, so the clusters of bubbles are much smaller than the horizon size today.

This yields a picture of a much more inhomogeneous universe than the one we observe and so marks the failure of old inflation.

1.2.3 New Inflation and the generation of density perturbations

Guth's seminal paper was quickly followed by two papers (Linde 1982, Albrecht & Steinhardt 1982) independently advancing an inflationary scheme based on a second-order GUT phase transition. As the effective temperature falls through T_c, the shape of the potential $V(\phi)$ changes: ϕ is no longer constrained to remain in the false vacuum state at $\phi = 0$ and can roll down the potential curve to the true vacuum. The requirement for sufficient inflation translates here into a requirement of a potential which is very flat near $\phi = 0$, so that ϕ takes a long time to roll down the potential, during which time the energy density remains very nearly constant and the scale factor undergoes the exponential expansion characteristic of inflation: such potentials are found (Coleman & Weinberg 1973) in the minimal $SU(5)$ GUT. At the end of the period of inflation the Higgs field rolls quickly down the final, steep part of the Coleman-Weinberg potential and goes into oscillation about the true vacuum. These oscillations are equivalent to the production of zero momentum scalar particles, which will decay, releasing energy, which will soon thermalise and reheat the universe to a temperature $\sim T_c/3$.

The field is in unstable equilibrium at $\phi = 0$, so quantum and/or thermal fluctuations are required to start it rolling down the potential. These fluctuations are random in nature and so will vary throughout space. At different points in space, therefore, the Higgs field will experience a different initial jolt down the hill of the potential, which will result in the potential reaching the true vacuum at slightly different times at different points in space. Since energy deposition results from the arrival of the Higgs field at the true vacuum, this means that the universe will be inhomogeneous after reheating. The new inflationary model, therefore, not only solves the horizon and flatness problems, but it also provides a mechanism for the generation of density perturbations.

It transpires that the slope of the power spectrum of density fluctuations is related to the slope of the potential $V(\phi)$, with a flat potential producing a scale-invariant power spectrum. This may be understood quite readily, by the following chain of reasoning. If

the potential is flat, then $\phi \simeq \phi_{\rm false}$ and the energy density remains close to that of the false vacuum, so the inflationary spacetime remains close to de Sitter space. Each Fourier mode of the density perturbation field freezes out once the horizon size becomes equal to its wavelength. De Sitter space is time-invariant, so that the amplitudes of different modes at the time of freeze out are the same. A density power spectrum whose modes have the same amplitude at horizon crossing is, by definition, scale invariant (Harrison 1970, Zeldovich 1972). Consequently, an inflationary model with a potential that is very flat will produce expansion which differs only slightly from de Sitter space and, thus, a power spectrum of density fluctuations whose amplitudes differ only slightly at horizon crossing – *i.e.* a power spectrum very close to the Harrison-Zeldovich scale-invariant form.

The density fluctuations are produced by random, quantum processes, so the phases of different modes are unrelated, with the consequence that the density perturbations obey Gaussian statistics. The density fluctuations are also adiabatic, in the sense that they do not produce spatial variations in the ratio of baryon to photon number densities. This may be explained readily as well. The density perturbations are produced because of the time delay between the start of the slow roll of the potential at different points in space. Different regions of space therefore underwent the same evolution, but slightly out of synchronisation. If the regions underwent the same evolution, then they will go through the same baryogenesis process, resulting in the production of a universal baryon-to-photon ratio. We shall not consider isocurvature fluctuations (spatial variations in the equation of state) in this thesis, due to the lack of a plausible mechanism for their generation.

The new inflationary scenario seems a great success: it can solve the horizon and flatness problems in the same way as old inflation, plus it can also generate a scale-invariant spectrum of Gaussian density perturbations. It too has a fatal flaw, however. The amplitude of the density fluctuation spectrum is determined by the strength of the coupling of the Higgs field. It is found that, for the fluctuation amplitude to be low enough not to contradict the observed level of CBR isotropy, the coupling must be so weak that the Higgs field would not be in thermal contact with the other fields present in early universe and, thus, the finite-temperature quantum field theory methods used

above would not be valid. One way around this is to have the inflation produced by some other scalar field, not the Higgs field. This has the added advantage of being able to solve the monopole problem as well since, if the Higgs field can be arranged to acquire its non-zero expectation value before or during the inflationary epoch, then the number density of monopoles surviving today will be diluted by the huge factor of Z^3 by inflationary expansion. Introducing a scalar field, the inflaton, specifically to produce inflation, removes the GUT motivation for new inflation and leaves it looking rather contrived.

1.2.4 Summary: Generic Inflation

We have seen the successes and failures of both new and old inflation. We described these in some detail for their pedagogical value: our description of old inflation showed how exponential expansion can solve the horizon and flatness problems, while our outline of the new inflation model illustrated the principles behind inflation's mechanism for generating density fluctuations. A bewildering array of other inflationary models have been suggested over the past decade, but none of them has been shown to work: in the words of Kolb & Turner (1990), inflation is a paradigm in search of a model – its general principles are so appealing that one just wishes for a concrete way of realising them.

In this spirit, one may consider a generic inflationary model, which combines the successes of specific models with a deft avoidance of their individual problems. Generic inflation features a phase of exponential expansion, which is sufficiently long to solve the horizon, flatness and monopole problems. Quantum fluctuations are generated in generic inflation, which freeze out to become the classical density fluctuations which grow by gravitational instability to produce the large-scale structure we observe today. Note, however, that these perturbations need not be scale-invariant (power law inflation and extended inflation produce power spectra which differ from the Harrison-Zeldovich form), nor do they have to be Gaussian (some inflationary models featuring multiple, interacting fields produce non-Gaussian fluctuations). The latter complication will, however, be generally neglected in what follows and we shall usually assume that generic inflation produces Gaussian fluctuations. We shall, however, have cause to consider

power spectra which are not scale-invariant: such models tend to generate gravitational waves which contribute to the CBR anisotropy and, thus, we shall have to consider them when using the *COBE* DMR results to normalise power spectra. With a single CBR anisotropy measurement it is impossible to distinguish the contributions from density fluctuations and gravitational waves, but one may differentiate between them given data from different angular scales (Crittenden *et al.* 1993).

1.3 Dark Matter in the Universe

The possibility of the existence of non-luminous matter was first raised by Zwicky (1933), who noticed that the motions of the galaxies in the Coma cluster implied a larger dynamical mass than that associated with the galaxies themselves. Whilst it may now be widely accepted that dark matter exists, there is still no agreement as to its amount, distribution or nature. These are questions which have a strong bearing on the field of large-scale structure which is the subject of this thesis, so it is important to consider them here. We can give, however, only a brief summary of the salient points in this Introduction, referring the reader to the reviews by Ashman (1992) and Trimble (1987) for a more detailed discussion.

1.3.1 Evidence for the existence of dark matter

Over the past sixty years, evidence for the existence of dark matter has been inferred from observations on all scales from the solar neighbourhood to the large-scale distribution of galaxies. These observations are often quantified in terms of the mass-to-light ratio (M/L), which is the ratio (in solar units) of the mass of the system to its luminosity in a particular pass band: the mass-to-light ratio corresponding to the critical density, ρ_c, is $\sim 1500h$, although there is considerable uncertainty in this due to uncertainty in the luminosity density from which it is derived and there is some variation dependent on the pass band in which the luminosity is measured.

The study of the density of the disk of the Milky Way was pioneered by Oort (1932), who used the observed motions of stars perpendicular to the Galactic plane to show

(Oort 1960) that the mass density of the disk exceeds that of its luminous components. Bahcall (1984) used a development of the Oort method, plus a constraint of consistency with the Galactic rotation curve, to estimate that the density of dark matter in the solar neighbourhood is at least half that of the luminous matter and that the dark matter is concentrated in a disk with an exponential scale-height of 0.7 kpc. Contrasting results were obtained by Bienaymé *et al.* (1987), who found no evidence for a significant quantity of dark matter in the Galactic disk. A suggested resolution to this disagreement came from Crézé *et al.* (1989), who argued that the uncertainties in the mass determinations used in previous work had been seriously under-estimated and that the results of Bahcall and Bienaymé *et al.* are consistent once realistic errors are used. Kuijken & Gilmore (1989a-c) advanced a new method for determining the surface density of the disk and found no evidence for a disk-like component of dark matter. More recently, a comprehensive study by Bahcall *et al.* (1992) reached the following two conclusions: (*a*) that the hypothesis that there is no disk dark matter is consistent with data at the 14% level; and (*b*) that if the suggested dark matter component is distributed like the luminous matter in the disk, then there must be 53% more mass in dark matter than in luminous matter. The question of a disk-like dark component to the Milky Way is clearly a very controversial one but, equally, it is a very important one: for the hypothesised dark matter to have settled into a disk would imply that it is dissipative, which would place strong constraints on its nature (see Section 1.3.2).

Estimations of the total mass of the Milky Way have been made using three dynamical methods. The first of these is to study the tidal radii of the Milky Way's satellite galaxies and globular clusters: a number of uncertainties limit the practical utility of this method. A second method makes a statistical application of the virial theorem to the dynamics of the globular clusters and satellite galaxies. This method is worryingly sensitive to the velocity of one particular satellite, Leo I: for example, Little & Tremaine (1987) and Kulessa & Lynden-Bell (1992) obtain masses of $5 \times 10^{11}\ M_{\odot}$ and $1.3 \times 10^{12}\ M_{\odot}$ respectively by this method with the only significant difference between their analyses being the assumed velocity of Leo I. A mass similar to that determined by Kulessa & Lynden-Bell was estimated by Zaritsky *et al.* (1989) by timing arguments, through a consideration of the relative motions of the Milky Way and M31 and the Milky Way and

Leo I. There is, therefore, some degree of consensus for a Milky Way mass of $\sim 10^{12} M_{\odot}$, which implies about ten times more mass in the dark halo than in the luminous portion of the Galaxy.

Nearby dwarf spheroidal galaxies have been the subject of much study, with the advent of fibre systems facilitating the accurate determination of their velocity dispersions and, hence, their masses, given the assumption of virial equilibrium. These systems tend to show mass-to-visible light ratios of $M/L_{\rm V} = 50 - 100$. Care must be taken in making these measurements to ensure that the stars observed are not pulsating variables or members of binary systems, which would yield misleading velocities. A more serious objection to this method (Kuhn & Miller 1989) is that these galaxies may not be in virial equilibrium and that tidal effects in the gravitational field of the Milky Way could boost the velocity dispersions of the dwarf spheroidals and lead to an over-estimation of their M/L values.

Perhaps the most compelling piece of evidence for the existence of dark matter in galaxies comes from the observed rotation curves of spirals: the rotation curves of many spirals, determined either from observing stars in the optical (*e.g.* Rubin *et al.* 1980) or from HI studies of interstellar gas (Bosma 1978), stay roughly flat out to large radii, in contrast to the decline that would be expected if the luminous matter comprised the total gravitational mass of the galaxies. Estimations of the mass of dark matter in spirals from this effect require models decomposing the galaxy into disk and halo mass components. It is common to make the "maximal disk assumption" (*e.g.* Sancisi & van Albada 1987) in which one gives the disk component the maximum M/L value consistent with the inner portion of the rotation curve and deduces the properties of the dark halo from the discrepancy at outer radii between the rotation curve that is observed and that produced by the maximal disk: support for the maximal disk assumption comes (*e.g.* Casertano & van Albada 1990) from the observation that features in the inner rotation curve seem to correlate with features in the light distribution. Application of the maximal disk assumption to real spirals indicates that the disk and halo masses must be comparable within the optical radii of the galaxies to produce flat rotation curves. Many authors have questioned the fine-tuning apparent in this, dubbing it the "disk-halo conspiracy".

A possible resolution of the disk-halo conspiracy is suggested by the work of Persic & Salucci (1988), who suggest a variation in the relative importance of disk and halo components with galactic luminosity. They claim that the disk component dominates the dynamics of the inner portions of bright spirals, with the halo only making a significant contribution at larger radii, while the halo is more likely to dominate at all radii in spirals with lower luminosities: this is borne out by the observation by Casertano & van Gorkom (1991) that the HI rotation curves of bright spirals decline beyond their optical radii. Salucci *et al.* (1991) find that the M/L values they deduce for spirals in their sample decrease with increasing galactic luminosity and it appears that this relation may be extended to include dwarf irregular galaxies: no clear explanation of this phenomenon exists, but it might seem to indicate some variation of star formation efficiency with the depth of the galactic potential well.

Spiral galaxies furnish evidence for the existence of dark matter in two further ways. Ostriker & Peebles (1973) argue that the luminous portions of spiral galaxies must be embedded in dark halos to stabilise them against the formation of bar instabilities, although Kalnajs (1983) has suggested that this function may be performed by the luminous material in the central bulges which many spirals possess. Athanassoula *et al.* (1987) have used the swing-amplification theory of Toomre (1981) to relate the number of spirals arms present in a galaxy to the ratio of disk mass to total mass, arguing that the fact that most spirals have two arms implies that most spirals have comparable disk and halo masses.

Huchra & Brodie (1987) have studied the dynamics of the globular cluster system surrounding M87 and deduced a mass-to-blue light ratio of $M/L_{\rm B} \sim 150$, implying ~ 15 times as much dark mass as there is luminous mass: this conclusion is supported by a similar study by Mould *et al.* (1990). The mass determined by Huchra & Brodie is consistent (given an assumed $M(R) \propto R$ mass profile) with that obtained by Stewart *et al.* (1984) from observations of X-ray emission from M87 using the *Einstein* satellite: the poor spatial resolution of current X-ray instruments limits the utility of this method for determining the masses of elliptical galaxies, as it requires accurate knowledge of the temperature profile of the gas.

An obvious parallel with the field of stellar mass determinations suggests the study of dynamics of binary galaxies, but this has a number of practical difficulties. The first obvious problem concerns the selection of true bound systems, rather than pairs of galaxies coincidentally close together on the sky – clearly, the orbital periods are many orders of magnitude too long for the observer just to wait and see! The difficulty in translating the projected velocity difference and separation between the galaxies into an orbital velocity and physical separation means that one must analyse a large sample of binary galaxies and make some statistical allowance for effects of orbital inclination. White *et al.* (1983) have performed such an analysis and have deduced larger M/L values for the galaxies in binary systems than determined on the basis of their individual rotation curves: this is to be expected, since this technique probes the mass distribution in the galaxies to larger radii than is possible using HI rotation curve observations, although note that this would also be the result produced by contamination by unbound systems.

Puche & Carignan (1988, 1991) have performed a detailed study of the dynamics of the Sculptor group of galaxies and have derived $M/L_{\rm B} \sim 83 \pm 10$. This anomalously low value has been questioned by Ashman (1992), who believes that the Sculptor group is not dense enough to have virialised and doubts over group membership and virialisation plague the dynamical study of groups of galaxies.

We shall postpone discussion of the virial analysis of clusters of galaxies until Section 1.3.3, where we shall discuss the recent work on the Coma cluster by White *et al.* (1993) in some detail. The largest scale on which evidence for dark matter has been found is the scale probed by large galaxy surveys. These provide indirect evidence for dark matter by implying a value for the cosmological density parameter in excess of the value determined by summing the luminous matter in galaxies. These methods typically do not yield an estimation of Ω directly, but rather a combination of Ω and the linear bias parameter, b: we shall postpone detailed discussion of this work until later, but simply note that Dekel *et al.* (1993) determine a bound of $\Omega > 0.3$ from a comparison of the density field traced by *IRAS* galaxies with the velocity field in the same region of space sampled by ~ 500 galaxies.

In this section we have outlined the evidence for the existence of dark matter, which we may summarise as follows:

- the luminous matter in the observable universe appears to have a density of $\lesssim 1\%$ of the critical density
- dark matter outweighs luminous matter by at least an order of magnitude
- the observable universe contains at least $\Omega \sim 0.3$ in matter that clusters on scales of a few tens of megaparsecs: dynamical measurements do not preclude a component contributing the remainder of the critical density, provided it does not cluster on these scales.

1.3.2 The nature of dark matter

Dark matter candidates are many and varied: from axions to supermassive black holes they span roughly eighty orders of magnitude in mass. Some of the most mundane candidates can be ruled out quite easily. The Gunn-Peterson test (Gunn & Peterson 1965) implies that there can only be a tiny amount ($\Omega_{\rm HI} \lesssim 4 \times 10^{-7}\, h^{-1}$) of diffuse neutral hydrogen along lines of sight to quasars, while closure density in hot gas is ruled out by the absence of an X-ray background with the correct spectrum. Rees (1987) argues that constraints from background light, the observed heavy element abundances in Population II stars and the absence of dynamical effects disturbing the disks of spiral galaxies imply that baryonic dark matter candidates must have masses less than $0.1\, M_\odot$ or in the range $10^2 - 10^6\, M_\odot$. The low mass option here corresponds to brown dwarfs: these are self-gravitating bodies which are insufficiently massive to initiate core hydrogen fusion and which radiate the gravitational potential energy of their collapse. No convincing brown dwarf candidates have been observed so far, but they should be detectable through either their proper motion (which must be high, as they must be very close to be detected at all), through their very cool near-infrared colours or through microlensing (Canizares 1982) of light from quasars.

If $\Omega \gtrsim 0.06$ then the primordial nucleosynthesis constraint on $\Omega_{\rm B}$ dictates that some, at least, of the dark matter must be non-baryonic. Non-baryonic dark matter candidates

are classified as hot, warm or cold, depending on their thermal velocities at early times: the importance of this classification will be seen in Section 1.4, when we consider the growth of cosmological density perturbations – for now we shall just consider which particles fall into each category.

The principal candidate for hot dark matter is a hypothesised massive neutrino. Interest in neutrino-dominated cosmogonies has fluctuated since their introduction twenty years ago (Marx & Szalay 1972, Cowsik & M$^{\rm c}$Lelland 1973), peaking perhaps in the aftermath of the announcement of a detection of a non-zero mass for the electron neutrino of $m_\nu \sim 36$ eV by Lyubimov *et al.* (1980). Controversy still surrounds the question of whether the electron neutrino has a rest mass, but one attraction of the hot dark matter scenario is that neutrinos are known to exist, which is clearly desirable and which distinguishes them from almost all other non-baryonic dark matter candidates. Another attraction of the hot dark matter model is that it is well specified: the Big Bang model specifies the number density of neutrinos and so the present fraction, Ω_ν, of the closure density contributed by a single massive species of neutrino with rest mass m_ν is given (Weinberg 1972) by

$$\Omega_\nu h^2 \;=\; 0.32 \left(\frac{m_\nu}{30\,\mathrm{eV}} \right). \qquad (1.50)$$

The entire mass range $100 h^2 \mathrm{eV} - 3$ GeV is, therefore, ruled out: if $m_\nu > 3$ GeV, then the rest mass term in the Boltzmann factors maintaining equilibrium in equations 1.38,1.39 & 1.40 would ensure that few neutrinos would survive to decouple. Tremaine & Gunn (1979) show, by means of a very neat phase space argument, that this allowed mass range, coupled with the Pauli exclusion principle, precludes neutrinos from comprising the dark matter in dwarf galaxies. Observations of high redshift quasars are also difficult to explain within the context of a neutrino-dominated universe when structure will form in a top-down fashion (see Section 1.4) as is the absence of the huge X-ray emitting structures that the model predicts, so the hot dark matter model has fallen out of favour, although mixed dark matter models (see below) do contain a significant mass fraction in a heavy neutrino.

There are no well-motivated warm dark matter candidates, so we move on to consider cold dark matter (CDM), which is the most studied cosmogony. There exist a plethora of

possible CDM candidates, including the axion (the particle whose existence was invoked by Peccei & Quinn 1977 to explain why strong interactions do not exhibit large CP violations) and a number of supersymmetry particles. The difference between these is that the supersymmetry particles are cold due to their high rest mass ($\sim$ GeV), while the axion is cold because it is produced out of thermal equilibrium. An even more exotic possibility is that of macroscopic nuggets of quark matter: Witten (1984) has shown how stable clumps of such material, comprising roughly equal quantities of up, down and strange quarks, can be produced in GUT phase transitions in the early universe, while Alcock & Fahri (1985) conclude that the quark nuggets must be of at least planetary mass to avoid subsequent evaporation.

In Section 1.4 and in later chapters of this thesis we shall see that the observed large-scale structure of the universe can be accounted for by a model in which there are significant quantities of both hot and cold dark matter. Such mixed dark matter models have been dismissed as being contrived: they have been derided for "invoking the tooth fairy twice" to explain how the universe could contain roughly comparable mass fractions of baryons, neutrinos and some CDM particle. This criticism may now be out of place, following the work by Madsen (1992) and Kaiser, Malaney & Stockman (1993) on what Kaiser *et al.* call 'neutrino-lasing'. The basic principle is that a mixture of hot and cold dark matter can be the natural result of decays of heavy fermions in the early universe. The model advanced by Kaiser *et al.* produces a hot/cold ratio of $\sim 70\%/\sim 30\%$, which is the inverse of that required to account for the observed large-scale structure, but, nevertheless, this may prove to be a very important step to understanding the dark matter composition of the universe.

1.3.3 A challenge to orthodox cosmology

The title for this subsection comes from that of a recent paper (White *et al.* 1993) investigating the baryon fraction in the Coma cluster. The importance of the questions raised by this work merits its being described in a separate subsection and it provides a suitable coda to this section.

White *et al.* start their paper with an inventory of the Coma cluster, their aim being to

estimate the mass of the various components of the cluster contained within the central sphere of radius $1.5h^{-1}$ Mpc (this being $R_{\rm A}$, the Abell radius; Abell 1958). The first mass component is that in galaxies, which may be estimated as the product of the V-band luminosity density within the central Abell radius (determined from the data of The & White 1988 and Kent & Gunn 1982) and the mean $M/L_{\rm V}$ value ($M/L_{\rm V} = 14.7\,h$) quoted by Binney & Tremaine (1987) for a sample of bright elliptical galaxies, similar to those found in the core of the Coma cluster. This yields a mass in galaxies, $M_{\rm gal}$, within the central $R_{\rm A}$ of Coma of

$$M_{\rm gal}(R_{\rm A}) \;=\; (3.15 \pm 0.66) \times 10^{13}\, h^{-1}\, M_{\odot}, \tag{1.51}$$

where the uncertainty corresponds to an assumed 0.2 mag zero-point error in the photometry.

The mass in hot intracluster gas is determined from the recent observations of Coma by the *ROSAT All-Sky Survey* (Briel *et al.* 1992). Briel *et al.* fit the X-ray surface brightness profile to a standard form and, using a range of mass profiles, estimate the gas mass within the central $2.5\ h^{-1}$ Mpc. The value so obtained may be scaled using the surface brightness profile fit to yield a gas mass, $M_{\rm gas}$, within the central Abell radius of Coma of

$$M_{\rm gas}(R_{\rm A}) \;=\; (5.66 \pm 1.02) \times 10^{13}\, h^{-5/2}\, M_{\odot}. \tag{1.52}$$

By summing the quantities in equations 1.51 and 1.52 it is possible to obtain a lower limit to $M_{\rm bar}(R_{\rm A})$, which is the baryonic mass within the central Abell radius of Coma: it is a lower limit, since there may also be baryonic dark matter.

The total dynamical mass, $M_{\rm tot}$, within the central Abell radius is harder to estimate, since it requires assumptions. Assumptions of spherical symmetry and dynamical equilibrium are common to all the methods, each of which requires further assumptions: to perform a virial analysis using optical data requires an assumption about the relationship between the distributions of light and mass, while use of X-ray data requires an assumed density and temperature profiles. Virial analyses (The & White 1986, Merritt 1987) which assume that light traces mass yield $M_{\rm tot}(R_{\rm A}) = (6.7 \pm 1.0) \times 10^{14}\, h^{-1}\, M_{\odot}$.

Hughes (1989) assumes the density profile resulting from this analysis to determine $M_{\rm tot}(R_{\rm A}) = (6.82 \pm 0.34) \times 10^{14}\, h^{-1}\, M_{\odot}$ from X-ray data.

To test the sensitivity of the $M_{\rm tot}$ determinations on the assumption that light traces mass, White *et al.* study Coma-like clusters in numerical simulations of $\Omega = 1$ CDM universes. The dissipationless simulations of Frenk *et al.* (1990) produced 152 clusters with velocity dispersions in excess of 800 km s^{-1}. Scaling these to account for the mass profiles of the simulated clusters being shallower at $R_{\rm A}$ than the galaxy number density profile in Coma produces a set of 'galaxy' velocity dispersions for the simulated clusters. Viewing each of these along one of three orthogonal axes facilitated an estimation of the mean value and dispersion of the ratio of the quantity $R_{\rm A}\langle v_{\rm g}^2\rangle/G$ (the observational virial mass estimator) to the mass within $R_{\rm A}$. Applying these values to correct the raw virial mass obtained for Coma yielded $M_{\rm tot} = (8.5 \pm 2.6) \times 10^{14}\, h^{-1}\, M_{\odot}$, although the uncertainty would have been higher had cluster members been identified in projection, rather than in three dimensions (Frenk *et al.* 1990). Simulations employing Smooth Particle Hydrodynamics (SPH, see Section 1.4) were used to generate clusters with X-ray profiles similar to that of Coma. From these a total mass for Coma of $M_{\rm tot} = (1.0 \pm 0.18) \times 10^{15}\, h^{-1}\, M_{\odot}$ was deduced. So, from these two results it is seen that the assumption of light tracing mass produces a total mass about 50% smaller than that found by assuming cluster profiles similar to those in $\Omega = 1$ CDM simulations. Since the aim here is to determine the a lower limit to the baryon mass fraction in Coma, White *et al.* adopt the CDM-simulation mass

$$M_{\rm tot} \;=\; 1.1 \times 10^{15}\, h^{-1}\, M_{\odot} \tag{1.53}$$

to be conservative. From equations 1.51, 1.52 and 1.53, White *et al.* deduce the following limit on the baryon mass fraction in the central Abell radius of Coma:

$$\frac{M_{\rm bar}}{M_{\rm tot}}(R_{\rm A}) \;\geq\; 0.029 \;+\; 0.051\, h^{-3/2}. \tag{1.54}$$

Comparing this with the maximum baryon density ($\Omega_{\rm B} = 0.015\, h^{-2}$) allowed by nucleosynthesis shows that this figure appears to be inconsistent with the baryon fraction in an $\Omega = 1$ universe by a factor of about ~ 3 for $h = 0.5$ and a factor of ~ 5 for $h = 1$.

Now, dissipative processes are likely (Babul & Katz 1993) to produce in rich clusters like Coma an enhancement in the baryon fraction above its cosmological value, so White *et al.* use a combination of analytical and numerical analyses to question whether it is possible for such mechanisms to produce a baryon fraction as high as that observed in Coma. The first analysis used is that of Bertschinger (1985), studying the self-similar infall of spherical shells of matter onto a central point mass in an otherwise unperturbed Einstein - de Sitter universe. Bertschinger finds two solutions to this problem: shells of collisionless mass fall in and oscillate with a decaying amplitude about the central mass, while an extremely dissipative medium is modelled by a solution in which shells of matter fall onto the compact object at the centre and just stick. Modelling the formation of clusters as comprising a mixture of these two components, it is possible to calculate the baryon enhancement, Υ, as a function of the mean overdensity, Δ, enclosed within a certain radius. Through the use of numerical simulations White *et al.* show that Bertschinger's spherically symmetric infall model produces the maximum baryon enhancement at a given value of Δ: simulations with aspherical proto-clusters and a slightly more realistic treatment of dissipation using SPH produce lower values of Υ at a given value of Δ. White *et al.* find that the baryon enhancement, Υ, predicted for the value of Δ corresponding to the central Abell radius of Coma is well below that required to reconcile the observed baryon fraction in Coma with SBBN in an $\Omega = 1$ universe: a simple model for the uncertainties in the parameters of the models rules out consistency at the 99.99% level.

This leads White *et al.* to the following four possible resolutions:

- open universe: clearly, the results of White *et al.* are not in conflict with SBBN in a low density universe and, indeed, if enhancement effects are small, they imply $\Omega_0 = 0.15\, h^{-1/2}/(1 + 0.55\, h^{3/2})$
- non-standard nucleosynthesis: if the upper bound on Ω_B provided by SBBN is incorrect then baryon mass fractions as high as that deduced here for Coma need not be incompatible with $\Omega = 1$
- non-gravitational effects: it is possible that non-gravitational effects could be responsible for enhancing the baryon mass fraction in the cores of clusters, but no

plausible mechanism suggests itself

- mass uncertainties: the mass estimates used by White *et al.* could be seriously in error.

White *et al.* dismiss the latter two possibilities and leave the reader with a choice between an open universe and non-standard nucleosynthesis. The depth of this challenge to orthodox cosmology merits a few comments. Firstly, to follow up a point discussed by White *et al.* , there is the possibility that their value of $M_{\rm gas}$ is an over-estimate, due to inhomogeneity in the X-ray emitting gas: the bremsstrahlung emissivity varies as the square of the gas density, so appreciable clumpiness in the gas could produce a misleading estimate of the amount of gas required to produce the observed X-ray flux. White *et al.* counter this argument by saying that the observed Sunyaev-Zeldovich (S-Z) effect in Coma constrains the clumping of the gas to a low level but, until, a detailed account of the S-Z detection in Coma is published there must be some uncertainty on this point. Secondly, the results presented by White *et al.* are entirely consistent with a scenario in which there is $\Omega \sim 0.8$ in a component which is not clustered on the scale of the Coma cluster. This places a constraint on the mass of such a particle and it would appear that currently favoured mixed dark matter scenarios including a ~ 30 eV neutrino would not do, since they have only $\Omega \sim 0.3$ in the hot component. These caveats aside, the White *et al.* result provides a very important datum, which should be borne in mind by the reader of the remainder of this thesis.

1.4 The growth of cosmological density perturbations

In this section we consider the growth of cosmological density perturbations through gravitational instability. We shall present a purely Newtonian description of perturbation growth, which is adequate as long as our concern is with fluctuations on scales much smaller than the Hubble radius and we neglect issues relating to the gauge freedom in general relativity, since an unambiguous choice of hypersurfaces of simultaneity is possible in this regime. The paper by Bardeen (1980) gives a rigorous treatment of

gauge-invariant perturbation theory, while the excellent review article by Efstathiou (1990) is recommended as a general reference for the material covered in this section.

1.4.1 Linear perturbation growth in a pressureless Newtonian fluid

The Newtonian theory of perturbation growth in a pressureless perfect fluid is essentially the study of the following set of the three equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{1.55}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\nabla \phi \tag{1.56}$$

and

$$\nabla^2 \phi = 4\pi G \rho. \tag{1.57}$$

These are, respectively, the continuity, Euler and Poisson equations and they govern the evolution of a self-gravitating Newtonian fluid: ρ is the fluid density, $\boldsymbol{v}$ its velocity and ϕ the Newtonian gravitational potential. It is convenient to transform these to a comoving coordinate system $\boldsymbol{x}$ related to the proper coordinates $\boldsymbol{r}$ by

$$\boldsymbol{r} = a(t)\boldsymbol{x}. \tag{1.58}$$

The proper velocity, $\boldsymbol{v}$, may be written as

$$\boldsymbol{v} = \dot{a}\boldsymbol{x} + a\dot{\boldsymbol{x}}, \tag{1.59}$$

where $a\dot{\boldsymbol{x}}$ is the peculiar velocity and we may define a comoving peculiar velocity $\boldsymbol{u} \equiv \dot{\boldsymbol{x}}$. We may define a density fluctuation field $\delta(\boldsymbol{x}, t)$ about the background density $\bar{\rho}(t)$ by

$$\rho(\boldsymbol{x}, t) = \bar{\rho}(t)\,[1 + \delta(\boldsymbol{x}, t)], \tag{1.60}$$

and then transform equations 1.55, 1.56 and 1.57 to the comoving frame, to obtain

$$\frac{\partial \delta}{\partial t} + \nabla_x \cdot \boldsymbol{u} + \nabla_x \cdot (\boldsymbol{u}\delta) = 0 \tag{1.61}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + 2\frac{\dot{a}}{a}\boldsymbol{u} + (\boldsymbol{u} \cdot \nabla_x)\boldsymbol{u} = -\nabla_x(\hat{\phi}/a^2) \tag{1.62}$$

$$\nabla_x^2(\hat{\phi}/a^2) = 4\pi G\bar{\rho}\delta, \tag{1.63}$$

where $\hat{\phi}$ is the peculiar potential and ∇_x denotes gradients w.r.t. comoving coordinates.

We assume that our perturbed FLRW model is periodic over a large volume V, so that we may define Fourier components $\{\delta_k\}$ of the density field by

$$\delta(\boldsymbol{x}, t) = \sum_{\boldsymbol{k}} \delta_k \, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{x}} \tag{1.64}$$

$$\delta_k(t) = \frac{1}{V}\int \delta(\boldsymbol{x}, t)\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \mathrm{d}^3x \tag{1.65}$$

where $\boldsymbol{k}$ is the comoving wavevector.

The statistical properties of the density perturbation field $\delta(\boldsymbol{x}, t)$ are given by the infinite hierarchy of n-point correlation functions $\langle \delta(\boldsymbol{x}_1) \ldots \delta(\boldsymbol{x}_n) \rangle$, where the brackets denote ensemble averages. The notion of an ensemble average can be a somewhat tricky one when considering the universe but, in any case, this difficulty is avoided by assuming that the density field is ergodic, so that the ensemble averages in the definitions of the correlation functions may be replaced by spatial averages taken over a sufficiently large volume of space. The ergodic property of the density field may be proved (Adler 1981) for the case where it obeys Gaussian statistics. The statistical distribution of a Gaussian density field is straightforward too, with its statistical properties being fully specified by the two-point correlation function, ξ,

$$\xi(x) = \langle \delta(\boldsymbol{x} + \boldsymbol{x}')\delta(\boldsymbol{x}') \rangle, \tag{1.66}$$

or, equivalently, the power spectrum, $P(k)$,

$$P(k) = \langle \mid \delta_k \mid^2 \rangle, \tag{1.67}$$

which are related by a Fourier transform:

$$\xi(x) \;=\; \frac{V}{2\pi^2}\int_0^\infty P(k)\,\frac{\sin kx}{kx}\,k^2\mathrm{d}k. \tag{1.68}$$

In much of what follows it will be useful to use a dimensionless form of the power spectrum, $\Delta^2(k)$, which is defined by

$$\Delta^2(k) \;=\; \frac{V}{2\pi^2}\,k^3\,P(k), \tag{1.69}$$

so that $\Delta^2(k)$ is the contribution to the variance per unit interval in $\ln k$.

Physically, Gaussian density fields will be produced whenever the phases of different Fourier components are uncorrelated, as a result of the Central Limit Theorem: an example of this is the generation of density fluctuations in the new inflationary scenario discussed in Section 1.2.3. The probability distribution function $p(\delta)$ for a Gaussian density field is given by

$$p(\delta)\,\mathrm{d}\delta \;=\; \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\delta^2}{2\sigma^2}\right)\mathrm{d}\delta, \tag{1.70}$$

where σ^2 is the variance of the density field: equation 1.70 shows that $\sigma < 1$ necessarily for a Gaussian density field.

Following this digression, we may then obtain the Fourier space form for our set of coupled equations in the comoving frame:

$$\frac{\mathrm{d}\delta_{\boldsymbol{k}}}{\mathrm{d}t} \;+\; i\boldsymbol{k}\cdot\boldsymbol{u}_{\boldsymbol{k}} \;+\; \sum_{\boldsymbol{k}} i\,\delta_{\boldsymbol{k}'}\,(\boldsymbol{k}\cdot\boldsymbol{u}_{\boldsymbol{k}-\boldsymbol{k}'}) \;=\; 0 \tag{1.71}$$

$$\frac{\mathrm{d}\boldsymbol{u}_{\boldsymbol{k}}}{\mathrm{d}t} \;+\; 2\frac{\dot{a}}{a}\boldsymbol{u}_{\boldsymbol{k}} \;+\; \sum_{\boldsymbol{k}'} i\left[\boldsymbol{u}_{\boldsymbol{k}'}\cdot(\boldsymbol{k}-\boldsymbol{k}')\right]\,\boldsymbol{u}_{\boldsymbol{k}-\boldsymbol{k}'} \;=\; i\frac{\boldsymbol{k}}{a^2}\hat{\phi}_{\boldsymbol{k}} \tag{1.72}$$

$$\hat{\phi}_{\boldsymbol{k}}/a^2 \;=\; -4\pi G\bar{\rho}\,\frac{\delta_{\boldsymbol{k}}}{k^2}. \tag{1.73}$$

The presence of the summation signs in these equations indicates that the different Fourier modes of the density perturbation field do not evolve independently. Peebles (1980) provides a plausibility argument why these mode-coupling terms may be negligible

in comparison to the linear terms and, neglecting them, we may obtain the following second-order differential equation which gives the evolution of δ_k:

$$\frac{\mathrm{d}\delta_k}{\mathrm{d}t^2} + 2\frac{\dot{a}}{a}\frac{\mathrm{d}\delta_k}{\mathrm{d}t} - 4\pi G\bar{\rho}\,\delta_k = 0. \tag{1.74}$$

Different solutions to equation 1.74 clearly follow for different equations of state. We know that the early universe was radiation-dominated, so our first solution of equation 1.74 will be for matter fluctuations in a radiation-dominated universe. We may take the radiation to constitute a smooth background, as it will show negligible fluctuations on the scales of interest ($\lambda \ll ct$), in which case equation 1.74 becomes

$$\frac{\mathrm{d}^2\delta_k}{\mathrm{d}t^2} + 2\frac{\dot{a}}{a}\frac{\mathrm{d}\delta}{\mathrm{d}t} - 4\pi G\bar{\rho}_{\mathrm{m}}\delta_k = 0, \tag{1.75}$$

with the effect of the relativistic component just being felt through its contribution to the expansion rate

$$H^2 = \frac{8\pi G}{3}\left(\bar{\rho}_{\mathrm{m}} + \bar{\rho}_{\mathrm{r}}\right). \tag{1.76}$$

Recalling that $\bar{\rho}_{\mathrm{m}} \propto a^{-3}$ and $\bar{\rho}_{\mathrm{r}} \propto a^{-4}$, it is convenient to transform from t to

$$\eta \equiv \frac{\bar{\rho}_{\mathrm{m}}}{\bar{\rho}_{\mathrm{r}}}, \tag{1.77}$$

in which case equation 1.75 becomes

$$\frac{\mathrm{d}^2\delta_k}{\mathrm{d}\eta^2} + \frac{(2+3\eta)}{2\eta(1+\eta)}\frac{\mathrm{d}\delta_k}{\mathrm{d}\eta} - \frac{3\delta_k}{2\eta(1+\eta)} = 0. \tag{1.78}$$

This has (Mészáros 1974, Peebles 1980) two solutions: a 'growing mode'

$$\delta_k \propto 1 + \frac{3}{2}\eta \tag{1.79}$$

and a 'decaying mode'

$$\delta_k \propto \left(1 + \frac{3}{2}\eta\right)\ln\left[\frac{(1+\eta)^{1/2}+1}{(1+\eta)^{1/2}-1}\right] - 3(1+\eta)^{1/2}. \tag{1.80}$$

From these it is clear that perturbation growth is not possible until the universe becomes matter-dominated – *i.e.* $\eta > 1$.

Now, in a matter-dominated Einstein - de Sitter universe, $a(t) \propto t^{2/3}$, so equation 1.74 becomes

$$\ddot{\delta}_k + \frac{4}{3t}\dot{\delta}_k - \frac{2}{3t^2}\delta_k = 0, \tag{1.81}$$

where the dot denotes differentiation w.r.t. t. This has the general solution

$$\delta_k = A_k t^{2/3} + B_k t^{-1}, \tag{1.82}$$

where, clearly, A_k is the growing mode and B_k is the decaying mode: so, linear theory growth is

$$\delta(t) \propto a(t). \tag{1.83}$$

1.4.2 The Jeans Length

The analysis given in the previous subsection ignores the effects of pressure. If we include the effects of pressure, then the Euler equation becomes

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\nabla\phi - \frac{1}{\rho}\nabla p, \tag{1.84}$$

where p is the pressure. This, in turn, means that equation 1.74 becomes

$$\frac{\mathrm{d}^2\delta_k}{\mathrm{d}t^2} + 2H\frac{\mathrm{d}\delta_k}{\mathrm{d}t} = \left[4\pi G\bar{\rho} - (c_s k/a)^2\right]\delta_k, \tag{1.85}$$

where c_s is the adiabatic sound speed, defined by

$$c_s^2 = \left(\frac{\partial p}{\partial \rho}\right)_s, \tag{1.86}$$

where the subscript s denotes partial differentiation at constant entropy. Equation 1.85 is the Jeans-Bonnor (Jeans 1928, Bonnor 1956) equation and its form shows that the

evolution of a perturbation with wavenumber k depends on the sign of the quantity in square brackets. This quantity defines a characteristic scale, the Jeans length, $\lambda_{\rm J}$,

$$\lambda_{\rm J} \;=\; c_{\rm s}\left(\frac{\pi}{G\rho_{\rm m}}\right)^{1/2}, \tag{1.87}$$

which marks the balance between the effects of pressure gradients and of gravity: for perturbations with wavelengths $\lambda \gg \lambda_{\rm J}$, the effects of pressure are negligible and evolution proceeds as in the case of a pressureless fluid described in the previous subsection; if $\lambda \ll \lambda_{\rm J}$ then gravity may be neglected and the perturbations will oscillate like acoustic waves, without growing.

Before decoupling, Thomson scattering kept the matter and radiation tightly coupled, so that they act as a single fluid, with an adiabatic sound speed given by

$$c_{\rm s} \;=\; \frac{c}{\sqrt{3}}\left(\frac{3\rho_{\rm m}}{4\rho_{\rm r}}+1\right). \tag{1.88}$$

Now, since $\rho_{\rm m}/\rho_{\rm r} \propto a$, the adiabatic sound speed and, hence, the Jeans mass, $M_{\rm J}$ ($M_{\rm J} \equiv 4\pi\rho_{\rm m}\lambda_{\rm J}^3/3$) decrease with time, so that perturbations corresponding to increasingly small masses start growing: by decoupling it has fallen to $M_{\rm J} \simeq 9\times10^{16}\,(\Omega h^2)^{-2}\,M_\odot$. After decoupling the relevant sound speed is that of a monotonic gas:

$$c_{\rm s} \;=\; \left(\frac{5k_{\rm B}T}{3m}\right), \tag{1.89}$$

where m is the proton mass. So, for a temperature $T \sim 3000$ K at the epoch of decoupling, $M_{\rm J} \simeq 1\times10^{6}\,(\Omega h^2)^{-1/2}\,M_\odot$: this great drop in the Jeans mass at decoupling facilitates the growth of perturbations on scales of relevance to the study of large-scale structure.

1.4.3 Silk Damping

The coupling between photons and electrons due to Thomson scattering is tight, but not perfect, resulting in the phenomenon of Silk damping (Silk 1968), whereby small-scale perturbations in the baryonic component are damped before decoupling.

Silk damping may be explained quite simply, as follows. The imperfect coupling of the photons to the matter means that each photon performs a random walk between scatterings. If the mean time between collisions is t_c, then the number of steps in the random walk made in time t is $N = t/t_c$ and it follows that the photons drift a mean distance $\lambda_D \sim \sqrt{N} c t_c \sim c(t t_c)^{1/2}$. Photons can, therefore, diffuse out of overdense regions, or into underdense regions, of size $\lambda \lesssim \lambda_D$. They will drag baryonic matter with them since, even though the coupling is not perfect, it is tight. Thus it follows that Silk damping will erase perturbations on mass scales below $M_D = 4\pi\rho_m \lambda_D^3/3 \sim 10^{12} (\Omega h^2)^{-3/2} M_\odot$ by decoupling in a purely baryonic universe.

1.4.4 Transfer Functions

From the previous subsections it should be clear that even the linear theory of perturbation growth in a realistic, multi-component cosmogony is complicated and will require numerical, rather than analytical, study. As long as mode-coupling terms are negligible, the results of such a treatment may be expressed in terms of a transfer function, $T(k)$,

$$\delta_k(t_0) \propto T(k)\, \delta_k^P, \tag{1.90}$$

relating the density perturbations $\{\delta_k\}$ at the present, t_0, with the primordial fluctuations $\{\delta_k^P\}$: it is conventional to normalise the transfer function such that $T(k) \to 1$ as $k \to 0$.

The form of the transfer function for a particular cosmogony depends on the matter components present and their microphysical interactions. Bond & Szalay (1983) find that the transfer function for a neutrino dominated universe takes the form

$$T(k) \propto 10^{-2(k/k_\nu)^{3/2}}, \tag{1.91}$$

where

$$\lambda_\nu = \frac{2\pi}{k_\nu} \simeq 41 \left(\frac{m_\nu}{30\,\mathrm{eV}}\right)^{-1} \mathrm{Mpc} \tag{1.92}$$

is the neutrino damping length. This arises because neutrino free-streaming erases fluctuations that enter the horizon while the neutrons are relativistic: λ_ν is, therefore, roughly equal to the distance that a neutrino has travelled since the Big Bang.

Bond & Efstathiou (1984) give the following fit to the transfer function for CDM-dominated models with negligible baryon content:

$$T(k) = \left\{1 + \left[ak + (bk)^{3/2} + (ck)^2\right]^\nu\right\}^{-1/\nu}, \tag{1.93}$$

where $a = 6.4(\Omega h^2)^{-1}$ Mpc, $b = 3.0(\Omega h^2)^{-1}$ Mpc, $c = 1.7(\Omega h^2)^{-1}$ Mpc and $\nu = 1.13$. Equation 1.93 shows that the shape of the CDM transfer function depends only on the quantity Ωh, given wavenumbers expressed in observational units of h^{-1} Mpc. This is because the only length scale imposed on the density perturbation field of a dissipationless fluid is the Hubble radius at matter-radiation equality which, as we saw in equation 1.36, depends on Ωh^2.

Comparison of equations 1.91 and 1.93 shows that a neutrino-dominated universe has more power on large scales (low k) than a CDM-dominated universe. This is important, as it dictates the form of large-scale structure in these two cosmogonies. In a neutrino-dominated universe, large-scale structures are expected to form first, with small structures forming by fragmentation: this is called the 'top-down' scenario of structure formation, for obvious reasons. By contrast, smaller structures will form first in a CDM-dominated universe, with larger structures forming through their agglomeration: this is the 'hierarchical' scenario.

1.4.5 Normalisation of the power spectrum

We have seen how inflationary processes can generate the power spectrum of density perturbations and how microphysics modifies it: we now need to discuss its normalisation.

In this thesis we shall consider only power spectra that tend to the Harrison-Zeldovich ($n = 1$) slope at large scales. Since the CBR anisotropy data from the *COBE* DMR

experiment are consistent with this spectrum we may use them to normalise our model power spectra.

We can write the power spectrum in the form

$$\Delta^2(k) = \frac{4\epsilon^2}{9}\left(\frac{ck}{H_0}\right)^4 T^2(k), \tag{1.94}$$

where ϵ is a dimensionless normalisation constant, whose value we are to determine: Peacock (1991) shows that ϵ^2 is the gravitational potential variance per unit $\ln k$.

The background radiation temperature fluctuations on the sky may be expanded in spherical harmonics, $Y_l^m(\theta,\phi)$:

$$\frac{\Delta T}{T} = \sum_{l,m} a_l^m Y_l^m(\theta,\phi), \tag{1.95}$$

where the angular power spectrum, $C_l = \langle \mid a_l^m \mid^2 \rangle$, is related to the power spectrum of density perturbations at the present epoch. For a scale-invariant spectrum of the form

$$\Delta^2(k) = \frac{4\epsilon^2}{9}\left(\frac{ck}{H_0}\right)^4, \tag{1.96}$$

in a spatially flat universe, this relationship is given (Efstathiou, Bond & White 1992, EBW) as

$$C_l = \frac{2\pi}{9}\frac{\Omega_0^{1.54}}{l(l+1)}\epsilon^2. \tag{1.97}$$

Smoot *et al.* (1992) fit such a formula to the *COBE* DMR data for multipoles higher than the quadrupole and express the results in terms of the value of $Q_{\rm rms}$, which is related to C_2 by

$$Q_{\rm rms} = \left(\frac{5C_2}{4\pi}\right)^{1/2} T_0, \tag{1.98}$$

and so using equations 1.97 and 1.98 this gives, for a scale–invariant spectrum,

$$\epsilon^2 = \frac{108}{5}\left(\frac{Q_{\rm rms}}{T_0}\right)^2 \Omega_0^{-1.54}. \tag{1.99}$$

Fitting to the *COBE* data, subject to the constraint of a scale-invariant spectrum, Smoot et al. (1992) obtain a value of $Q_{\rm rms} = 16.7 \pm 4\ \mu$K. This does not include, however, the effect of the 'cosmic variance', an added uncertainty due to the power spectrum of the observed sky not being equal to that of the mean of the ensemble of all possible skies in the Universe. Inclusion of the cosmic variance requires another fit to the higher order multipoles, subject, for our purposes, to the constraint of a scale-invariant spectrum, which will modify the $Q_{\rm rms}$ value and increase its uncertainty slightly. We take the fluctuation amplitude indicated by *COBE* to be $Q_{\rm rms} = 17 \pm 5\ \mu$K, which, from equation 1.99, yields $\epsilon = 2.9 \pm 0.9 \times 10^{-5}\ \Omega_0^{-0.77}$.

Before the advent of *COBE*, the most common way to express the amplitude of a linear density power spectrum was in terms of its predicted value for σ_8, the variance in spheres of radius $8\ h^{-1}$ Mpc: the origin of this somewhat arcane measure being that the galaxy distribution is observed to have roughly unit variance when samples in $8\ h^{-1}$ Mpc spheres. This leads to another measure of the fluctuation amplitude, the bias factor, b_8, defined by $b_8 = 1/\sigma_8$. This measure is motivated by a particularly simple picture of the relationship between the distribution of matter and galaxies in the universe, in which it is assumed the fluctuations in the galaxy number density field, $n_{\rm gal}$, are related to the density fluctuations by

$$\left(\frac{\delta n_{\rm gal}}{n_{\rm gal}}\right) = b \left(\frac{\delta \rho}{\rho}\right), \tag{1.100}$$

where b is called the bias parameter. This association of a measure of the density fluctuation amplitude with a specific hypothesis for the relation between the distribution of mass and galaxies in the Universe is unfortunate and leads to 'bias' being the most confusingly-used word in cosmology: we shall avoid it wherever possible in what follows.

1.4.6 Non-linear growth of density perturbations

Our treatment of the growth of density fluctuations heretofore has been based on the assumption of their being small. Observations that $\sigma_8 \sim \mathcal{O}(1)$ indicate that this perturbative approach is inadequate, on small scales at least, and that we must seek a non-linear treatment instead.

Both analytic and numerical methods may be used to study perturbation growth into the non-linear regime, with the most appropriate method depending on the details of the particular problem and, especially, on how non-linear the density field is on the scales of interest. In the mildly non-linear regime, one may profitably use the Zeldovich Approximation (Zeldovich 1970, Shandarin & Zeldovich 1989), which we discuss at some length in Chapter 2. The density field on galactic scales is, however, sufficiently non-linear that one must resort to numerical simulations.

Numerical simulations may be separated into two categories, depending on whether or not they include prescriptions to model dissipative processes. Dissipationless simulations (commonly called N-body simulations) may themselves be classified, according to the method by which they compute the gravitational forces between their particles. The simplest of these are the particle-mesh (PM) codes, which, in their simplest form, calculate the net force acting on each particle at each time-step by moving it to the nearest node of a mesh laid out throughout the simulation volume and then using a fast Fourier transform (FFT) routine to solve Poisson's equation on this (conventionally cubic) grid. The accuracy with which this models the gravitational interaction between particles depends, of course, on the fineness of this grid, which is limited by the computer memory available. For many purposes, the force resolution obtainable by this method is inadequate, in which case a particle-particle/particle-mesh (P^3M) code may be used instead. This supplements the computation of the gravitational forces between widely-separated particles using the FFT prescription of a PM code with the direct computation of forces between nearby pairs of particles: the computational expense of this exercise limits the range over which this direct force computation can be performed.

N-body codes have been used extensively in cosmology for the past fifteen years, but it is only recently that advances in computer hardware have facilitated the creation of codes including dissipative processes. These codes supplement the evolution of dissipationless components by standard N-body techniques with that of a dissipative gaseous component, modelled using simple prescriptions for cooling and star formation. The development of these codes is still in its infancy, but we shall have cause to consider some of the results of one such simulation in Chapter 4.

1.5 Plan for the remainder of this thesis

As its title suggests, this thesis is concerned with the use of galaxies and clusters as tracers of the large-scale structure of the Universe.

In Chapter 2 we consider the use of rich clusters of galaxies as probes of large-scale structure. The observational study of the spatial distribution of clusters has advanced greatly in recent years and this chapter combines a detailed analysis of recent clustering data with the development of a new method for calculating predictions for the cluster correlation function in cosmological models.

Chapters 3 and 4 are principally concerned with the clustering of galaxies. A great deal of attention has been focussed in recent years on the study of galaxies observed in the infrared by *IRAS*. It has been argued that *IRAS* galaxies are unbiased tracers of the cosmological density field, that "*IRAS* galaxies trace the mass". Implicit in such a statement is a belief in the homogeneity of the class of *IRAS* galaxies. In fact, *IRAS* sees a broad range of galaxies: one may divide them into "cool" galaxies (isolated spirals) and "warm" galaxies (interacting galaxies, starbursts and the like). This classification is somewhat crude, but it suggests a difference in the clustering of the two populations, since interacting galaxies would be expected in more dense environments than isolated spirals. In Chapter 3 we investigate the clustering of "cool" and "warm" *IRAS* galaxies, to see whether any such difference is discernible.

In Chapter 4 we show how a synthesis of recent observational data may be used to determine the power spectrum of density fluctuations over a wide range of scales and produce a consistent picture of the large-scale structure. Chapter 5 presents a discussion of the results of the previous chapters, along with the conclusions that may be drawn from them. Finally, there are several mathematical appendices, containing derivations that it would be too tedious to include in full in the main text of this thesis, followed by a complete list of works cited and by a copy of the paper (Mann, Heavens & Peacock, 1993, MNRAS, 263, 798) describing some of the results of Chapter 2.

Chapter 2

CLUSTER CORRELATIONS AS PROBES OF LARGE-SCALE STRUCTURE

2.1 INTRODUCTION

The large-scale structure of the Universe is delineated by the spatial distributions of galaxies and clusters of galaxies. Later chapters in this thesis are concerned with the galaxy distribution; in this chapter we consider what one may learn from a study of the spatial distribution of rich clusters of galaxies.

There are both observational and theoretical motivations for studying the cluster distribution in addition to the galaxy distribution. On the observational side, clusters offer a much more efficient coverage of space; they trace the same large-scale structure as galaxies, but do so much more sparsely, marking, presumably, the highest points in the cosmological density field.

One of the attractions of the study of the cluster distribution for the theorist is its relative lack of dynamical evolution: the cosmological density field is only mildly non-linear on cluster scales today, so clusters have not moved far from the sites of their formation.

This is of particular importance to the present work, since it means that the dynamical component to the clustering of clusters is amenable to an analytic treatment, as pursued here.

In Sections 2.2 - 2.4 we present a comprehensive review of the observational study of cluster correlations: it is essential for the theorist to understand the complexities of, and controversies surrounding, these observational data before using them to constrain cosmological models and that is why a lengthy review is necessary, particularly since a great deal of work (both observational and theoretical) has been performed since the publication of the last substantial review of the subject (Bahcall 1988). The most widely used cluster catalogue in studies of large-scale structure is the Abell/ACO catalogue, which is the subject of Section 2.2. The advent of plate-measuring machines has facilitated the production of cluster catalogues with far more objective selection criteria than those of catalogues constructed from visual scans of photographic plates, like the Abell/ACO catalogue: two such catalogues, produced using the COSMOS and APM plate-measuring machines, are described in Section 2.3. In Section 2.4 we summarise the current state of the observational study of cluster correlations, before turning to the theoretical side of the question.

Section 2.5 reviews previous theoretical studies of cluster correlations, using both analytical and numerical methods. Analytical work has centred on the computation of the statistical component to the clustering, through the study of the spatial correlations of peaks in Gaussian random fields, while the success of numerical simulations of the full dynamical problem has been restricted by its inherent dynamical range.

In Section 2.6 we present a new method for computing cluster correlations. This is the first analytical method to take account of the nonlinear evolution of the cosmological density field. It employs the framework devised by Bond & Couchman (1987, 1988; hereafter, collectively, BC) to supplement the statistical peak-peak methods of Section 2.5 with dynamical evolution using the Zeldovich Approximation.

Section 2.7 compares the results of our analytical method with those from recent numerical simulations, while Section 2.8 compares them with observations of optical cluster samples. The application of our methods to the study a sample of *ROSAT* clusters is

outlined in Section 2.9. A detailed discussion of the results of this chapter will be given in Chapter 5.

2.2 THE ABELL/ACO CATALOGUE

The most widely used cluster catalogue in large-scale structure work has been the Abell/ACO catalogue. This comprises the original Abell (1958) catalogue of rich clusters, together with a revision and southern extension prepared by Abell, Corwin & Olowin (1989, ACO): the construction of the Abell/ACO catalogue is described in Section 2.2.1. (In what follows the word "Abell" will generally be used to denote northern clusters from Abell (1958), while "ACO clusters" will be those in the southern extension presented by ACO.) A number of redshift samples of Abell/ACO clusters have been produced and these are discussed in Section 2.2.2, along with the spatial correlations deduced from them. Many authors have questioned the integrity of the Abell/ACO catalogue and its suitability for use in studying large-scale structure: a review of the long and contentious debate on this subject that has raged in the literature for the past fifteen years is presented in Section 2.2.3.

2.2.1 Construction of the Abell/ACO catalogue

The scientific aim of George Abell's 1958 paper was two-fold: to produce a large finding list of rich clusters of galaxies and to draw from it a smaller sample, selected in a homogeneous fashion, that would be suitable for statistical studies investigating the clustering of clusters.

Abell's raw material was the National Geographic Society - Palomar Observatory Sky Survey. This comprises 879 pairs of photographic plates, (one blue- and one red-sensitive plate for each field), taken by the 48-inch Schmidt telescope of the Palomar Observatory and covering the sky from the northern celestial pole to a declination of $-27°$: Abell used only the red plates from each pair, as these were found to be more suitable for identifying clusters, particularly at higher redshifts.

Earlier analyses of the galaxy distribution on plates taken at the Lick Observatory per-

formed by Neyman, Scott and collaborators (Neyman & Scott 1952, Neyman, Scott & Shane 1953, 1954, Scott, Shane & Swanson 1954) had motivated Abell's work by revealing the presence of clusters, but also illustrated the difficulty of selecting true clusters from a distribution of galaxies seen in projection on the sky: Abell was well aware of projection effects and many subsequent authors have been uncharitable in failing to acknowledge his understanding of the difficulty, as well as the scale, of the task he set himself in producing a cluster catalogue from the Palomar sky survey.

Abell pictured the distribution of galaxies on the sky as a field population, whose surface density varies with position, upon which is superposed a number of very rich and conspicuous associations. He assumed that these are physical clusters, not chance projections: some of the galaxies in these associations will, of course, be members of the field and should be excluded by some suitable means but, he argued, this contamination should be relatively unimportant if only the richest and most compact clusters are considered.

In order to select associations that match this idealised picture, Abell imposed the following four selection criteria:

1. Richness Criterion: a cluster must contain, after background subtraction, at least 50 galaxies in the magnitude range $m_3 \rightarrow m_3 + 2$, where m_3 is the apparent magnitude of the third brightest cluster member. The validity of the galaxy count in this magnitude range as a richness estimator depends on the form of the bright end of the luminosity function for cluster galaxies, but Abell provided evidence in support of it.

2. Compactness Criterion: the fifty or more members must be contained within a physical distance, $r_{\rm A}$, of the cluster centre. This distance was given by Abell as a redshift-dependent distance on the plate, but it translates to a physical distance of $r_{\rm A} = 1.5\, h^{-1}$ Mpc: this is commonly called the Abell radius.

3. Distance Criterion: the cluster must be sufficiently distant that its members do not extend over more than one plate or, at most, part of an adjacent plate, but sufficiently near that the counting strip ($m_3 \rightarrow m_3 + 2$) lies well within the magnitude limit of the plate. These constraints correspond to a range of estimated

redshifts of $0.02 \leq z \leq 0.2$, where the upper limit results from considering clusters with $m_3 \leq 17.5$.

4. Galactic Position Criterion: the position of the cluster should be such as to ensure identification of clusters satisfying the other three criteria. This region of "complete identification" is given in Table 1 of Abell (1958) and excludes those areas of the sky where Abell felt that either a high surface density of stars or particularly high Galactic absorption would hamper cluster detection.

Abell found 1682 clusters that satisfied all four criteria and these constitute his *statistical sample*, which he supplemented with further objects to produce a catalogue containing a total of 2712 members.

These selection criteria should, in principle, select clusters of the desired sort in a homogeneous, redshift-independent manner, but there are a number of caveats relating to their performance in reality. The first caveat concerns the estimates made of the apparent magnitudes of cluster members. Magnitude information is required for two purposes: firstly to define the counting strip of $m_3 \rightarrow m_3 + 2$ and, secondly, to estimate the redshift of the clusters from the $m_{10} - z$ relation derived from the redshift sample of clusters studied by Humason, Mayall & Sandage 1956) to determine the angular distance on the plate corresponding to the Abell radius $r_{\rm A}$. Abell estimated the magnitudes of galaxies on a plate by comparing their images with a set of calibration images of galaxies on film. This method assumes that plate quality is consistent and that the conditions of observation and development are fixed, which is clearly inaccurate at some level. Zero-point errors in the derived magnitude scale will be unimportant, since the counting strip is defined by an interval in magnitude and the $m_{10} - z$ relation is normalised using observations of clusters with known redshift. Doubts must remain, however, about systematic errors that would not shift the zero-point, but could alter the width of the counting strip and wrongly determine the redshift used to set the counting radius, $r_{\rm A}$: one such systematic effect is the Scott effect (Scott 1957), whereby richer clusters have brighter M_{10} values (provided that the shape of the luminosity function of cluster galaxies does not vary much with cluster richness).

Internal checks of the parameters assigned to the selected clusters were made possible by

the overlap of survey plates: Abell found 120 clusters on more than one plate, 90 of which satisfied the conditions for inclusion into the statistical sample. From comparing these twice-selected clusters he estimated that his stated cluster positions were accurate to a couple of minutes of arc and that the error on the estimated magnitudes was ~ 0.2 mag. His estimated error on the counts of cluster members was 17%, but there may be a significant systematic error on that, due to his method of background subtraction: to obtain the background count Abell turned to an area of the plate that appeared to be free of clusters and counted galaxies down to the appropriate limiting magnitude. This procedure produced corrections of up to 30% of the raw count, but this could be a systematic underestimate, since it was made on an area with a low surface density of galaxies, by construction: this effect is illustrated by Lucey (1983) who finds that for his simulated plates (see Section 2.2.3) areas selected to be cluster-free have typically $\sim$70% of the mean surface density of galaxies.

These estimated errors for the counts and magnitudes were used by Abell to select appropriate richness and distance classification schemes for the clusters in the statistical sample: in each case the width of the class was set to be 3.5 times the estimated error at the lower limit of the class, so that, for normally-distributed errors, there would be a very small chance of a cluster being in a group more than one level away from its true one. The resultant classification of the statistical sample into richness and distance classes is shown in Tables 2.1 and 2.2, respectively. Clusters with 30–49 galaxies were assigned to richness class $R = 0$: Abell advised that this class not be used in statistical work, because of its incompleteness.

In 1975 Abell suggested that a southern rich cluster survey be performed, using the IIIa–J plates then being taken by the 1.2m UK Schmidt telescope for its Southern Sky Survey. The resultant catalogue, constructed in broadly the same manner as the original Abell (1958) northern catalogue, is presented by ACO, along with a revision and correction of the earlier work. Partly, no doubt, in response to some of the criticism levelled at the original catalogue since its publication, ACO were less confident in the claims they made for their catalogue: they did not consider it possible to construct a statistical sample, writing "since the present all-sky catalogue is based on purely visual surveys of apparent areal densities of galaxies, *it should not be taken as a definitive catalog of*

Richness Class	Galaxy Count	No. of clusters
1	50–79	1224
2	80–129	383
3	130–199	68
4	200–299	6
5	300 and over	1
Total		1682

Table 2.1: Richness Class classification for clusters in Abell's statistical sample (from Abell 1958)

Distance Class	m_{10} magnitude range	No. of clusters	Cumulative count
1	13.3–14.0	9	9
2	14.1–14.8	2	11
3	14.9–15.6	33	44
4	15.7–16.4	60	104
5	16.5–17.2	657	761
6	17.3–18.0	921	1682

Table 2.2: Distance Class classification for clusters in Abell's statistical sample (from Abell 1958)

clusters, but rather as a finding list of apparent clusters which need further investigation [ACO's italics]."

In addition to the problems associated with the Abell (1958) catalogue, some of which were mentioned above and more of which will be discussed in detail in Section 2.2.3, ACO faced further difficulties, one might say of their own making: their catalogue was produced by three observers, in two continents, using two different types of material (Abell & Corwin in Edinburgh using original plates and Olowin in Oklahoma using film copies), using calibration procedures that varied with time and between observers. Not surprisingly, a sizeable section of their paper is concerned with comparing clusters that more than one of them observed, due to plate overlaps, to find systematic differences in the cluster parameters assigned and to "correcting" their raw results accordingly.

One significant difference between the methods of ACO and those of Abell (1958) is the method of background subtraction: ACO correct cluster counts on the basis of a universal luminosity function for field galaxies (from Rainey 1977). This is clearly not in keeping with Abell's initial picture relating clusters to a position-dependent field, since it leads to their being selected as excursions above a global, rather than a local, surface density threshold: the unsuitability of this procedure is evinced by the negative "corrected" counts that it returns for a number of poorer ACO clusters in sparsely-populated regions of the sky.

ACO produced a catalogue containing a total of 4073 clusters, divided into the same richness and distance classes as Abell (1958). They do not define a statistical sample and write "we hope that this will be the last such catalog prepared by visual scans of photographic plates, and we urge future investigators to compile cluster catalogs using high-speed microphotometric scanning machines and objective selection criteria." To compare this quote and the one reproduced above with the more confident tone of Abell (1958) is to see how far observational cosmology had progressed in the intervening three decades: Abell's methods may have been "state-of-the-art" in 1958, but by 1989 not only was it realised that a more objective approach was required, but such an approach was becoming possible, due to advances in astronomical technology.

Comparisons of the Abell (1958) and ACO catalogues are presented by Batuski *et al.*

(1989, BBOB) and Scaramella *et al.* (1991, SZVC) and both groups find systematic differences between them. SZVC begin their comparison by looking at the clusters in the overlap region between the northern Abell catalogue and its southern extension presented by ACO. Two points are clear from this: (*i*) the number density of ACO clusters is larger than that of Abell clusters in the overlap region (the ratio is 1.47 $\pm$ 0.13 for $R \geq 0$ clusters and 1.39 $\pm$0.14 for $R \geq 1$); (*ii*) ACO tend to estimate higher richnesses for the clusters (the best fit linear correlation between the number of galaxies, N_g, counted by Abell and ACO for overlap clusters is $N_g(\mathrm{ACO}) \simeq -30 + 1.5\ N_g(\mathrm{Abell})$, with a quoted 1σ error of 0.07 on the slope of the fit: this fit comes after the counts for the four most discrepant clusters have been excluded and assumes equal errors on the two N_g values). The first observation is probably the result of the higher completeness in the ACO catalogue, while the second may well be the result of the difference between the background subtraction procedures adopted for the two catalogues: both these effects are discussed in more detail below.

BBOB and SZVC both compare the radial density distributions of the two catalogues, finding the distributions consistent with each other and with homogeneity out to an estimated distance of $\sim$300 h^{-1} Mpc. Differences between the distributions determined by the two groups at larger redshifts result from the different distance estimators they use. BBOB find that the difference between the radial distributions of Abell and ACO subsamples is greater when distance classes are used rather than estimated redshifts: they interpret this as indicating a systematic shift in the magnitude limit of distance classes between the two catalogues, with the redshift limit of the $D \leq 4$ class in the ACO catalogue being $\sim$30% higher than that for the Abell catalogue. SZVC find that there is an offset in Abell's magnitude calibration between the overlap region and the rest of the northern catalogue, with Abell's estimated distances in that region being systematically larger by $\sim (15 \pm 6)\%$. There are clearly serious problems with magnitude estimates and, consequently, distance estimation in the Abell/ACO catalogue.

Both groups find Galactic-latitude selection dependences for both catalogues, as well as a selection bias with declination for the ACO catalogue. SZVC find that the Galactic-latitude selection bias is much stronger in the full catalogues than in volume-limited samples (with depths of 300 and 350 h^{-1} Mpc for the Abell and ACO samples respectively;

these being the distances out to which the samples appear to be homogeneous). This is important: it shows that Galactic-latitude selection biases affect more distant clusters more seriously, (as would be expected if the dependence were due to Galactic extinction pushing the faint end of the counting strip beyond the magnitude limit of the plates), which means it is inappropriate to correct for this effect using a redshift-independent Galactic-latitude selection function, $P(b)$, as is conventional practice, followed even by SZVC themselves!

The complexity (especially the redshift-dependences) of the selection biases discussed by SZVC and the differences both they and BBOB find between their Abell and ACO subsamples, (as well as the biases introduced by using estimated distances that their analyses reveal), can only reinforce the belief, implicit in the quotes from ACO given above, that the Abell/ACO catalogue is far from ideal for use in statistical studies.

2.2.2 Spatial correlations of Abell/ACO clusters

Abell (1958, 1961) was the first author to study the distribution of clusters in his statistical sample. His work was soon followed by that of others (Bogart & Wagoner 1973, Hauser & Peebles 1973, Rood 1976) who used statistical tools such as the angular correlation function and the nearest-neighbour distribution to study the clustering of Abell clusters. This early work was principally concerned with the distribution of clusters seen in projection on the sky: studies of the three-dimensional distribution of clusters had to wait until the 1980s, when large and complete redshift samples of Abell clusters became available.

The first determination of the spatial correlation function of Abell clusters to be published was that of Klypin & Kopylov (1983). Their sample contained 158 $R \geq 0$ clusters with redshifts (largely from Fetisova 1982) of $z \lesssim 0.08$ and $\mid b \mid \geq 30°$: they estimated their completeness to be 85%. Inclusion of $R = 0$ clusters, which Abell (1958) deemed to be unsuitable for statistical work, was justified by noting that the class apparently only becomes seriously incomplete at redshifts higher than those considered by Klypin & Kopylov. Fitting the spatial correlation function of this sample to the canonical form of $\xi(r) = (r_0/r)^\gamma$ in the range $2.5 \leq r \leq 50\ h^{-1}$ Mpc yields $r_0 \simeq 25\ h^{-1}$ Mpc and $\gamma \simeq 1.6$.

Partly, no doubt, due to its appearance in a Soviet, rather than a Western, journal, the work of Klypin & Kopylov received much less attention than that enjoyed by a second study, published a few months later by Bahcall & Soneira (1983, BS83). The redshift sample used by BS83 is that of Hoessel, Gunn & Thuan (1980), comprising all 104 $R \geq 1$ Abell clusters in distance class $D \leq 4$ and with $\mid b \mid \geq 30°$. Before determining the spatial correlation function, BS83 make a correction for the variation of the cluster surface density with Galactic latitude: apparently the region of "complete identification" defined by Abell (1958) does not exclude all the areas of the sky where interstellar absorption or star/galaxy confusion hamper cluster detection. This effect is modelled with a latitude selection function $P(b) = \text{dex}\ [0.3(1 - \text{cosec} \mid b \mid)]$, which has become standard procedure in subsequent work, although, as discussed above, SZVC show that this is inappropriate, as the Galactic-latitude selection bias varies with redshift.

Bahcall & Soneira detect clustering out to $\sim 150\, h^{-1}$ Mpc, with a correlation length of $r_0 \simeq 25\, h^{-1}$ Mpc for a fit to a $\gamma = 1.7$ power law. They find that this spatial correlation function is consistent, through Limber's equation (Limber 1953, Peebles 1980), with the observed angular clustering of the $D \leq 4$ and $D = 5 + 6$ Abell subsamples. From this scaling test they conclude that the $D \leq 4$ redshift samples of Hoessel *et al.* is a fair sample of the larger $D = 5 + 6$ sample and that the observed spatial correlations of Abell clusters are genuine.

BS83 also study the angular correlations of $R = 1$ and $R \geq 2$ subsamples of $D = 5 + 6$ clusters. They find that both richness classes produce correlation functions well approximated by the same power law, $[w(\theta) \propto \theta^{-1}]$, but that that of the $R \geq 2$ sample has a higher amplitude. Even stronger small scale angular correlations are found for $R \geq 3$ clusters, but the small sample size precludes meaningful comparisons here: BS83 conclude that they have demonstrated the presence of a strong richness dependence for the correlations of Abell clusters.

The validity of these results is challenged by Postman, Geller & Huchra (1986, PGH). They question whether the Hoessel *et al.* survey is a "fair sample", due to its inclusion of the rich Corona Borealis supercluster. PGH note that "mean value of the spatial correlation function drops by a factor of ~4 in the 10-20 Mpc [with h=1] range when

Corona Borealis is removed from the sample. The drop is significant at the 3σ level. No such drop in the spatial correlation function occurs when six northern clusters are randomly removed from the sample". Enhancements due to Corona Borealis are also seen at larger separations, albeit at a lower significance level, so PGH are correct to conclude that the Hoessel *et al.* survey is not a fair sample for the purposes of clustering analysis.

To circumvent this problem PGH construct a larger cluster sample, whose clustering should be less sensitive to individual superclusters like Corona Borealis. Their sample is constructed by imposing on Abell's statistical sample the twin distance limits of $z \leq 0.1$ and $M_{10} \leq -19.38$; the latter corresponding to the catalogue limit of $m_{10} = 18$ at $z = 0.1$, beyond which PGH judge that foreground contamination affects estimation of m_{10} seriously. The sample so defined comprises 152 clusters, for which there are known redshifts for all but sixteen clusters: estimates for these are made using the methods of Leir & van den Bergh (1977), who advocate estimating redshifts from a number of cluster parameters, to reduce scatter due to inaccurate magnitude estimations for individual galaxies. PGH detect significant correlations in their sample out to $\sim 60h^{-1}$ Mpc. Fitting the correlation function to the canonical form yields $\gamma \simeq 1.8$ and $r_0 \simeq 20h^{-1}$ Mpc, but with large uncertainties: the uncertainty on the slope of the power law is $\sim$10%, while that on its amplitude is $\sim$50%. These results are consistent with those of BS83, given their uncertainties, and it would appear from their Figure 2 that PGH have not been altogether successful in overcoming the sensitivity of the correlation function to Corona Borealis.

PGH also study the richness dependence of clustering strength. They determine the spatial correlation function for the full $R = 1$ and $R = 2$ subsamples of the Abell statistical sample, using Leir & van den Bergh (1977) redshift estimates in the absence of a measured redshift. Constraining the correlation functions to the canonical $\gamma = 1.8$ power law they obtain correlation lengths of $r_0 = 24\ h^{-1}$ Mpc and $r_0 \simeq 42\ h^{-1}$ Mpc for the $R = 1$ and $R = 2$ samples respectively, with respective uncertainties of $\sim$20% and $\sim$35% in the correlation length. This richness dependence is consistent with that found by BS83, but a degree of scepticism must greet observational results so reliant on estimated redshifts: $\sim$80% of the $R = 1$ clusters and $\sim$75% of the $R = 2$ sample

have estimated redshifts. PGH raise another caveat: they show that the $R = 3$ clusters are under-represented at low galactic latitude and suggest that star/galaxy confusion might lead to a systematic underestimation of cluster richness in that region and, thus, bias the clustering of richer clusters and spuriously enhance the strength of any richness dependence of clustering.

A deeper complete sample of clusters, containing no supercluster comparable in richness to Corona Borealis, is presented by Huchra *et al.* (1990, HHPG). This sample contains 145 clusters (103 $R \geq 1$) from distance classes $D \leq 6$ (137 $D \geq 5$), all with measured redshifts. HHPG justify their inclusion of $R = 0$ clusters by noting that the redshift histograms for their $R = 0$ and $R \geq 1$ subsamples are consistent out to $z \simeq 0.2$. Fitting the correlation function they calculate to a power law with $\gamma = 1.8$ in the interval $10 \leq r \leq 70\ h^{-1}$ Mpc, HHPG obtain correlation lengths of $r_0 = 20.3$ and $20.9\ h^{-1}$ Mpc for their $R \geq 0$ and $R \geq 1$ subsamples, respectively: the respective 1σ bootstrap errors on these are (+4.79,–5.05) and (+6.69,–6.91) respectively, with all distances in h^{-1} Mpc. HHPG also consider a subsample of 132 clusters limited to $z \leq 0.24$, (beyond which they judge their deep sample to become seriously incomplete), for which they obtain $r_0 = 20.7\ h^{-1}$ Mpc, with 1σ uncertainties of (+6.32, –6.38) h^{-1} Mpc: these results are again consistent with those of BS83.

Following their analysis of a deep Abell sample, the same authors turn to a survey of nearby clusters in Postman *et al.* (1992, PHG). This sample contains all 350 Abell/ACO clusters (plus Virgo, which is too extended to satisfy the Abell/ACO selection criteria) north of $\delta = -27°30'$ with $m_{10} \leq 16.5$: this includes 15 southern ACO clusters. The richness distribution of this sample is: 195 $R = 0$ clusters, 126 $R = 1$, 28 $R = 2$ and 2 $R = 3$. The correlation function of this sample is well approximated by the canonical power law form with $r_0 = 20.0^{+4.6}_{-0.4}\ h^{-1}$ Mpc and $\gamma = 2.49 \pm 0.22$: these results are demonstrated to be very insensitive to the Corona Borealis supercluster. The sensitivity to Corona Borealis is more visible in the $R \geq 1$ subsample considered by PHG: they find $r_0 = 23.7^{+7.9}_{-9.0}\ h^{-1}$ Mpc when Corona Borealis is included, but $r_0 = 18.9^{+7.8}_{-8.5}\ h^{-1}$ Mpc when it is left out, although, of course, these results are consistent, given the large quoted uncertainties.

PHG define a "statistical sample", in which they include all clusters in their sample with $z \leq 0.08$ and $\mid b \mid \geq 30°$: the redshift limit marks the distance out to which the number density of PHG clusters remains constant. For this "statistical sample" a correlation length of $r_0 = 20.6^{+4.5}_{-4.8}\ h^{-1}$ Mpc and a power law slope of $\gamma = 1.86 \pm 0.20$ are computed.

Three further samples of Abell/ACO clusters are considered by Plionis, Valdarnini & Jing (1992, PVJ). The computation of the spatial correlation function is not the prime concern of these authors, but they do calculate it in passing: the same material is presented by Plionis, Valdarnini & Coles (1992). The first two PVJ samples are drawn from the Abell (1958) catalogue and contain clusters with measured redshifts of $z \leq 0.1$ in the northern and southern Galactic caps: 206 clusters in the NGC and 131 in the SGC. The correlation functions for both samples are fitted to the canonical power law with $r_0 = 18 \pm 4\ h^{-1}$ Mpc for $\gamma = 1.8$: significant correlations are observed in the NGC subsample out to $\sim 50\ h^{-1}$ Mpc, but the bootstrap errors are larger than the signal for almost all bins in the SGC subsample, so these correlations have a low formal significance. The third sample considered by PVJ comes from the ACO catalogue and contains 118 clusters with $b \leq -20°, \delta \leq -17°$ and $m_{10} \leq 16.4$: redshifts for 25 of these are derived using the $m_{10} - z$ relation of Plionis & Valdarnini (1991). Significant correlations in this sample are seen out to $\sim 30\ h^{-1}$ Mpc, with a correlation length of $r_0 = 22 \pm 10\ h^{-1}$ Mpc for a fit to a power law constrained to $\gamma = 1.8$.

The final redshift sample of Abell/ACO clusters we shall consider is that compiled by Peacock & West (1992, PW92). Their sample covers the whole sky, excluding regions where Galactic extinction (deduced from *IRAS* 100μm observations) exceeds $A_V = 0.5$ mag and is volume-limited to $z \leq 0.08$. PW92 consider the spatial correlations of three subsamples: $R \geq 1$ (for which they have 179 clusters with measured redshifts and 16 with estimated redshifts), $R = 0$ (198 measured and 34 estimated) and $R \geq 2$ (42 measured and 3 estimated). PW92 estimate z using a number of cluster parameters, not just m_{10}: in the north they fit $\log_{10}z$ to m_{10}, A_V and $\log_{10}N$ (where N is the galaxy count for the cluster) and in the south supplement these with m_1, m_3 and $(\log_{10}N)^2$, finding that this second order 'Scott effect' term is important. Fitting the resultant correlation functions to the canonical power law model they obtain the following results: $r_0 = 21.1 \pm 1.3\ h^{-1}$ Mpc, $\gamma = 2.0 \pm 0.2$ for $R \geq 1$; $r_0 = 20.6 \pm 1.5\ h^{-1}$ Mpc, $\gamma = 1.5 \pm 0.2$

or $r_0 = 21.1 \pm 1.5\ h^{-1}$ Mpc with γ=2.0 forced for $R = 0$; and $r_0 = 45 \pm 5\ h^{-1}$ Mpc with γ=2.0 forced for $R \geq 2$.

2.2.3 The Integrity of the Abell/ACO Catalogue

If one takes these observational results at face value, then a consistent picture is apparent: the spatial distribution of Abell/ACO clusters is correlated out to large separations ($r \sim 100\ h^{-1}$ Mpc); the correlation function of the distribution is well approximated by the same power law as the galaxy correlation function – *i.e.* $\xi(r) = (r_0/r)^\gamma$ with $\gamma \sim 2$; the cluster correlation length is richness dependent – $r_0 \simeq 20 - 25\ h^{-1}$ Mpc for richness $R \geq 1$ (and $R \geq 0$ if these are included), but increases to $r_0 \sim 40\ h^{-1}$ Mpc for $R \geq 2$ clusters. The details of this picture have, however, been challenged by a number of authors who question the suitability of the Abell/ACO catalogue as the basis of statistical clustering studies: we review the voluminous literature on this topic in this section.

Fesenko (1979a) considered $6^\circ \times 6^\circ$ fields on the sky, all centred at the same declination and with serial numbers i=1,2,3... running in the direction of increasing right ascension: each field corresponds to a Schmidt plate. Let n_i^W, n_i^E be the number of Abell clusters whose centre falls in the western, eastern half of the i^{th} field respectively. Fesenko defines the following two quantities, γ_1, γ_2 which measure the fluctuations between contiguous half-fields within a given field and in adjacent fields, respectively:

$$\gamma_1 = \frac{\langle (n_i^E - n_i^W)^2 \rangle}{\langle n_i^E + n_i^W \rangle} \quad \text{and} \quad \gamma_2 = \frac{\langle (n_{i+1}^W - n_i^E)^2 \rangle}{\langle n_{i+1}^W + n_i^E \rangle}, \tag{2.1}$$

where the averages are taken over all fields in the declination zone. The two quantities should be statistically equivalent, giving $\gamma_1 = \gamma_2$, while $\gamma_2 > \gamma_1$ would indicate inhomogeneities between plates, either inherent in the plates (due to variations in the conditions of exposure and development, or even in the quality of the plates themselves) or in Abell's identification of clusters on them. Fesenko studied seven declination zones, with centres from -18° to $+18^\circ$ and finds $\gamma_2 > \gamma_1$ for six of them: the mean value of the difference, $\Delta = \gamma_2 - \gamma_1$, is found to be $\langle \Delta \rangle = 0.317 \pm 0.142$. Plate-to-plate variations like

these, whatever their cause, clearly hamper the study of clustering in the Abell/ACO catalogue. In a second paper Fesenko (1979b) studied the effects of variations in the surface density of the field population of galaxies, given that Abell estimated his background count some distance away on the plate from the cluster candidate. This analysis is strongly dependent on the toy model used for the field population and on the cluster multiplicity function, but Fesenko concludes that "most Abell clusters are spurious."

A less extreme conclusion is reached by Lucey (1983), who modelled the effects of contamination by field galaxies through the use of artificial "plates", constructed so as to mimic Abell's analysis of real plates. Lucey scattered points into a cone to mark cluster centres and then projected them onto the plane normal to the axis of the cone to generate a two-dimensional distribution. Galaxies were then placed around each cluster centre on this artificial plate so as to reproduce the radial dependence observed by Seldner & Peebles (1977) for the surface density of galaxies around rich clusters, with the cluster multiplicity function of Bahcall (1979) being used to set the number of clusters required in each richness class: an objection to this method is that the random scattering of cluster centres means that the galaxy distribution generated on the artificial plates will not be a particularly good approximation to the true distribution on the sky. The resulting two-dimensional galaxy distribution was then analysed using an algorithm implementing Abell's selection criteria: tests were conducted to check that this algorithm selected a very similar set of clusters as a visual application of these criteria.

Contamination by field galaxies is found by Lucey to have two effects: firstly, some rich clusters are missed, as contamination by a foreground galaxy leads to m_3 being estimated at too bright a magnitude, with the result that too few galaxies are found in the counting strip; and, secondly, some spurious clusters are identified, due to the inclusion of field galaxies in the counting strip, boosting the cluster richness above the threshold of fifty counts. Quantitative estimates of the seriousness of these effects are model-dependent, but Lucey estimated that the 15–30% of rich clusters were missed by Abell on account of a contaminated m_3, while 15–25% of Abell clusters have a true galaxy population which is less than half their Abell count, due to the inclusion of field galaxies.

Attempts have been made, by Dekel *et al.* (1989, DBPO) and Olivier *et al.* (1990,

OBDPS; 1993, OPBD), to model these projection effects and so to construct "decontaminated" catalogues of Abell/ACO clusters. This work assumes a model in which the galaxy distribution is composed of two components, an unclustered field population superposed on which is a number of clusters, which have spherically-symmetric halos of galaxies, following a universal radial profile: this is clearly very similar to the picture Abell had in mind when he set out to identify rich clusters for his catalogue. This model is quantitatively specified by the observed mean number densities of galaxies and clusters and the observed galaxy-cluster and cluster-cluster correlation functions. Two forms for the galaxy-cluster cross-correlation function are considered, those determined by Seldner & Peebles (1977) and Lilje & Efstathiou (1988) from the Lick galaxy catalogue and the Abell catalogue: the use of observations of (possibly contaminated) Abell clusters in this decontamination procedure is likely to lead to an underestimation of the importance of projection effects. This toy model, together with a Schechter galaxy luminosity function from Efstathiou *et al.* (1988), is then used to calculate the contribution to the Abell count of each cluster from every other cluster in the sample under consideration. This is done iteratively and, upon convergence, those clusters with an Abell count of less than fifty are removed, to leave a "decontaminated" sample.

DBPO use this model to decontaminate the $R \geq 1, D \leq 4$ sample of 102 (*sic*) Abell clusters, considering only contamination from other members of this sample. The results of this procedure depend in detail on the form of galaxy-cluster correlation function used: 12 clusters are eliminated if the Lilje & Efstathiou (1988) form is used, against 17 for that of Seldner & Peebles (1977). The spatial correlation function can then be determined for the decontaminated sample: DBPO find (with either form of the galaxy-cluster correlation function) that decontamination reduces the amplitude of the cluster correlation function by a factor of $\sim$2 at separations $\leq 30\ h^{-1}$ Mpc, so that the resultant correlations are significant only at the $\sim 1\sigma$ level.

DBPO find that significant contamination in their model is only found for clusters which are at similar redshifts and are separated by less than $\sim 30\ h^{-1}$ Mpc: they explain this by saying that once the ratio of distances becomes greater than about two the magnitude ranges comprising the counting strips of the two clusters have little overlap, so there can be little projection contamination. This explains their observation for the case where a

distant cluster is contaminating a nearby one (*i.e.* background contamination), but it does not account for the converse case, of foreground contamination. This is because, so long as the contaminating foreground cluster has a luminosity function that rises towards fainter magnitudes, there will *always* be galaxies in the foreground cluster that will fall into the counting strip of the more distant cluster. The reason why foreground contamination is not observed for clusters with wide redshift separations must just be a geometrical effect: there will not be significant overlap on the sky between the halo of the nearby cluster and the circle of radius r_A surrounding the more distant cluster once the distance between them becomes too large. Despite the fact that DBPO do not explain it correctly, their observation that projection contamination is only important for clusters relatively close in space is an important one since, as they argue, it undermines the use of the scaling test in deciding whether or not clustering is genuine: as DBPO write, projection "should not lead to an increased number of pairs at all separations in redshift; it enhances the number of pairs preferentially at small angular separations and relatively small redshift separation. The effect should therefore be similar at all depths."

The claims of DBPO are challenged by Szalay, Hollósi & Tóth (1989, SHT) on the basis of their study of angular correlations of Abell clusters in different richness classes. They take the sample of 1418 $R \geq 1$ Abell clusters with $\delta > -27°$ and $\mid b \mid \geq 40°$ and divide it into three subsamples: 95 clusters with $D \leq 4$, 565 $D = 5$ clusters and 758 from the $D = 6$ distance class. They compute the angular autocorrelation functions for the three samples, [$w_{ii}(\theta)$ for i=4,5,6], finding them all consistent with the power law $w(\theta) \propto \theta^{-1}$, with amplitudes $w_{44} > w_{55} > w_{66}$ scaling broadly as expected from Limber's equation for real clustering. SHT seek to refute the projection contamination model of DBPO by considering the cross-correlations between clusters in different distance classes [w_{ij} for $i \neq j$]. If the distance classes were well separated then the cross-correlation between different classes would be zero, but one might expect a small positive correlation here, since the distance classes are known to overlap (Struble & Rood 1987), as a result of the inaccuracy of distance estimation using the $m_{10} - z$ relation. SHT find that the cross-correlation functions for the adjacent classes (w_{45} and w_{56}) are indeed weak and positive, as expected, and they find that w_{46} is negative, which they claim contradicts the DBPO contamination model. This conclusion is incorrect: as OBDPS are quick to point out,

the model they advance in DBPO predicts significant projection contamination only for clusters which are fairly close together in space, so this effect could only boost w_{46} if Abell's distance estimates were so wildly inaccurate as to produce a significant overlap in redshift between distance classes $D \leq 4$ and $D = 6$. In fact, what SHT have done is to quantify another selection bias in the Abell catalogue, (one mentioned in passing by PGH), that Abell systematically missed distant ($D = 6$) clusters in the proximity of nearby ($D \leq 4$) clusters on the sky. SHT suggest two possible explanations: (*i*) that distant clusters were thought by Abell to be subclusters within the nearer ones; (*ii*) that Abell's distance estimates incorrectly placed (presumably due to foreground contamination of m_{10}) some distant clusters in classes $D \leq 4$. Supporting evidence for the latter suggestion is provided by the excess of pairs [above the $w(\theta) \propto \theta^{-1}$ power law] counts seen in w_{44} for angles $1^\circ \leq \theta \leq 3^\circ$, but SHT argue that this effect alone is insufficient to account for the strength of anticorrelation seen in w_{46}.

An extension of the work of DBPO is presented by the same authors in OBDPS, where they consider angular correlations in decontaminated samples of distant ($D = 5 + 6$) clusters in the Abell and ACO catalogues. They consider two samples, constructed to be as similar as possible, given the differences between the two catalogues: (*i*) 1650 Abell $R \geq 1$ clusters from all distance classes; (*ii*) 546 ACO clusters with $R \geq 1$, $m_{10} \leq 19.3$ and $b \leq -45^\circ$. These two samples were decontaminated following the method outlined in DBPO, with distances estimated for all clusters using a fit to the $m_{10} - z$ relation including a term to account for the Scott effect. The angular correlations of the decontaminated $D = 5 + 6$ clusters were then determined, using the standard latitude and declination selection functions, and are discussed for three angular ranges: (*i*) small angles, $\theta < 1^\circ$; (*ii*) intermediate angles of $1^\circ \leq \theta \leq 3^\circ$; and (*iii*) large angles, $\theta > 3^\circ$. They find that decontamination reduces $w_{cc}(\theta)$ by about a factor of two on small angles, by about ~30% for intermediate angles, but has a negligible effect on large scales ($\theta > 3^\circ$): this is, of course, just as expected from DBPO, where it is emphasised that projection contamination only affects clusters separated by $\lesssim 30\ h^{-1}$ Mpc. Consistent results are obtained on small and intermediate scales (both before and after decontamination) for the Abell and ACO subsamples, but they are discrepant on large scales: significant correlations are observed in the Abell sample out to $\theta \sim 40^\circ$, as against $\sim 3^\circ$ for the ACO

subsample.

This difference on large scales cannot be due to projection contamination of the DBPO variety, since the superclusters within which their effect is supposed to take place are smaller than $\theta \sim 3°$. As a further test, OBDPS try accounting for extinction effects using the Burstein & Heiles (1982) *HI* column density data to introduce a longitude dependence not present in the standard $P(b)$ selection function, but this too has no effect. It seems, in fact, that there are large-scale inhomogeneities in the Abell/ACO catalogue. OBDPS split their combined Abell/ACO sample into two longitude cuts, the first containing the regions $b < 0°$, $-60° < l < 120°$ and $b > 0°$, $90° < l < 270°$ and the other the rest of the sky. They find that the first zone contains 68% of the clusters in 50.8% of the sky, but that the angular correlations of clusters in the second region are much stronger on large scales.

To illustrate further that the large-scale angular correlations of Abell/ACO clusters are not terribly robust, OBDPS re-calculate the correlation functions of the Abell and ACO subsamples, normalising the pair counts using a local background (smoothing within 30° of each cluster), rather than a global one. This has a minor effect on the large-angle correlations of the ACO sample, but has a large effect on the Abell sample, bringing the two samples into agreement: this is in the opposite sense to that expected from the different (global versus local) background subtraction procedures used by Abell and ACO, indicating that Abell's selection was not a simple, algorithmic implementation of his criteria and may be significantly influenced by subjective effects.

This decontamination routine is applied to the HHPG and PHG samples by OPBD. They find that the PHG sample has roughly the same degree of contamination as the samples studied by DBPO and OBDPS, but that the HHPG deep sample appears to be anomalously 'clean'.

The decontamination procedure used by DBPO, OBDPS & OPBD corrects solely the Abell counts for clusters. It does not treat the effects of m_3 contamination [which the work of Lucey (1983) indicates may well be equally serious as the introduction of interlopers by projection contamination] and it assumes spherically-symmetric clusters [which are not expected from theoretical considerations of the shapes of collapsing density

perturbations (Lin, Mestel & Shu 1965; Peacock & Heavens 1985) nor from observations of the shapes of clusters projected on the sky (Plionis, Barrow & Frenk 1991)]. A more general criticism of this work follows from its model-dependence: it is able to correct for one particular selection bias in the Abell/ACO catalogue, but there are several others and, more importantly, the subjective nature of the selection method means that no algorithmic decontamination procedure can correct for all the biases Abell and ACO introduced into their catalogues; the psychological and physiological factors are simply too complex and are inherently unquantifiable.

An alternative approach which may seem more fruitful is that pioneered by Sutherland (1988, S88) and later applied by Sutherland & Efstathiou (1991) and Efstathiou *et al.* (1992b). The Sutherland method adopts a different philosophy: instead of assuming a model for the selection bias it seeks to correct for the results of the selection bias, irrespective of its origin. The starting point for this work is the observed anisotropy of the cluster-cluster correlation function, $\xi(r_{\rm p}, r_z)$, calculated as a function of redshift separation (r_z) and projected separation ($r_{\rm p}$). This anisotropy was first noted by BS83 and is discussed in some detail by Bahcall, Soneira & Burgett (1986, BSB). BSB construct scatter plots of pair separations from two samples of Abell clusters (the Hoessel *et al.* sample studied by BS83 and a larger, slightly incomplete sample of $R \geq 0$ clusters from Bahcall & Soneira 1984) in the R_z–R_δ, R_z–R_α and R_δ–R_α planes, subject to the restriction that the physical separations of the pairs $[(R_\alpha^2 + R_\delta^2 + R_z^2)^{1/2}]$ is less than 100 h^{-1} Mpc. They report that a " strong and systematic elongation in the z–direction is seen in all the real samples studied" (*i.e.* for $R \geq 0$ and $R \geq 1$ subsamples of the two cluster samples studied), with the elongation stronger in the $R \geq 0$ subsamples: by contrast the R_δ–R_α plane is populated relatively uniformly, showing evidence for clustering, but no pronounced elongation in either the α– or δ–direction. BSB find that this elongation in the redshift direction is equivalent to convolving the pair frequency distribution along one of the projected axes with a pairwise Gaussian velocity dispersion of width $\sigma_v = 2000\ ^{+1000}_{-500}$ km s^{-1}. They suggest that this redshift-space anisotropy is a combination of two effects: "a geometrical elongation of the large-scale structure (in a form of broad filaments, partial cells or shells)", plus a residual elongation due to cluster-cluster peculiar velocities.

Both the method of analysis of the redshift anisotropy used by BSB and the conclusions they draw from it are criticised in S88, in which Sutherland addresses the anisotropy first in the complete $R \geq 1$, $D \leq 4$ sample and then in a sample of Abell clusters drawn from the Struble & Rood (1987, SR87) redshift catalogue. The variation of the redshift space correlation function, $\xi(\Delta z)$, as a function of Δz for the $R \geq 1$, $D \leq 4$ sample is seen to be inconsistent with the 2000 km s^{-1} Gaussian velocity broadening model advanced by BSB: $\xi(\Delta z)$ remains positive and roughly constant out to redshift separations of $\Delta z \sim 0.07$, by which point the velocity broadening model would predict it to have tended to zero. The peculiar velocity model of BSB is rejected as a possible explanation of the anisotropy in the SR87 sample too, because $\xi(r_{\rm p}, r_z)$ does not decline with r_z in the way predicted and because implausibly large velocities ($\sim$ 7000 km s^{-1}) would be required to reproduce the magnitude of the anisotropy: cluster peculiar velocities determined by Aaronson *et al.* (1986) are much lower, although these are for isolated clusters, rather than clusters in the virialised superclusters envisaged by BSB.

Having identified projection effects in the SR87 catalogue, Sutherland then proceeds to correct for them. His method makes no assumptions about the source of the projection effects, but it does make one assumption regarding the true clustering: it assumes that the mean value of $\xi(r_{\rm p}, r_z)$ is zero for $r_z \geq 50\ h^{-1}$ Mpc. The Sutherland correction method proceeds as follows:

- produce histograms of cluster-cluster and cluster-random pairs in both $r_{\rm p}$ and r_z
- normalise the columns at constant $r_{\rm p}$ in the cluster-cluster counts, $N_{\rm cc}$, to have the same total, $N_{\rm cr}$, as the corresponding cluster-random column: *i.e.* multiply each column by a correction factor, $B(r_{\rm p})$, given by

$$B(r_{\rm p}) = \frac{\sum_{r_z>50} N_{\rm cr}(r_{\rm p}, r_z)}{\sum_{r_z>50} N_{\rm cc}(r_{\rm p}, r_z)}, \tag{2.2}$$

  where the sums are for all bins $r_z > 50\ h^{-1}$ Mpc.

The cluster correlation function may then be calculated, with the pairs weighted by the appropriate value of $B(r_{\rm p})$. The resulting correlation functions show no significant

correlations beyond $\sim 50\ h^{-1}$ Mpc (this is not an artifact of the cut-off at $50\ h^{-1}$ Mpc: a cut-off at $80\ h^{-1}$ Mpc produces the same result) and are well approximated by the canonical $\gamma = 1.8$ power law with $r_0 \simeq 14\ h^{-1}$ Mpc: no significant difference is observed between $R \geq 0$ and $R \geq 1$ subsamples. Taking this correlation function and convolving it with Gaussian velocity dispersions of different widths gives a best fit of $\sigma_v \simeq 750$ km s^{-1} for the corrected $\xi(r_p, r_z)$, strongly excluding pairwise velocities of ~ 2000 km s^{-1}, as advocated by BSB.

Sutherland & Efstathiou (1991, SE91) apply the same correction method to the deep Abell sample of HHPG and again obtain a corrected correlation length of $r_0 \simeq 14\ h^{-1}$ Mpc. A further application of the Sutherland correction method to a sample of Abell/ACO clusters is presented by Efstathiou *et al.* (1992b, EDSM), who consider projection effects in the PHG redshift survey of nearby clusters. The effect of correction is quoted as reducing the correlation length of the PHG sample from $17.4\ h^{-1}$ Mpc to $13.9\ h^{-1}$ Mpc, but note that these figures are consistent within their quoted errors, and that the uncorrected correlation length quoted by EDSM differs from that given by PHG and calculated independently from the PHG data by Nichol (1992).

Sutherland & Efstathiou also present further evidence attesting to the weakness of the scaling test as a discriminator between genuine and spurious clustering. They show that to apply it profitably requires an accurate knowledge of the redshift distribution of the cluster sample and that its results depend on the assumed evolution of the clustering with redshift. They find that while the angular correlations of $D = 5$ and $D = 6$ subsamples of Abell clusters scale qualitatively as predicted for the BS83 model of the spatial correlation function, their amplitudes are too high, even for the evolution model that gives the largest $w(\theta)$. They suggest that the closer agreement that BS83 obtained in their scaling tests resulting from a combination of inaccurate distance estimation (BS83 use a simple m_{10}–z relation derived from the original Abell m_{10} magnitudes, while SE91 use the redshift distribution of the SR87 catalogue, which they argue should be a fair sample of that of the full Abell catalogue within each distance class) and using the combined $D = 5 + 6$ sample rather than the two samples individually, (since the weak cross-correlation w_{56} will reduce the amplitude of the auto-correlation function of the combined sample).

Two additional sources of spurious clustering in the Abell/ACO catalogue are discussed by SE91. The first of the these is the presence of large scale gradients, as discussed by OBDPS, while the second is variation between Schmidt plates: SE91 calculate the angular correlation functions for intra- and inter-plate pairs and find that the difference between them is not that which would be expected for a simple variation of the limiting magnitudes of the plates.

The analysis of redshift anisotropies conducted by Sutherland and his collaborators has principally been applied to samples of $R \geq 0$ clusters, with an implicit assumption that projection effects corrupt all samples of Abell/ACO clusters. Several authors (*e.g.* PW92, Soltan 1988 and Jing *et al.* 1992) have, however, found that samples of $R \geq 1$ clusters may not be as prone to corruption as those containing $R = 0$ clusters. In particular, PW92 find that their $R \geq 0$ sample is strongly corrupted, while the redshift-space anisotropy seen in their $R \geq 1$ sample is consistent with velocity broadening of ~ 1000 km s^{-1}, which they judge is not implausible, given the inhomogeneous origins of the redshifts for their sample: we shall discuss this point in more detail later.

2.3 AUTOMATED CLUSTER CATALOGUES

The previous section has shown that the Abell/ACO catalogue is far from ideal for use as the basis for statistical studies of cluster correlations and that the subjective way its selection criteria were implemented precludes correction of the catalogue to produce a homogeneous sample. All selection criteria will introduce biases, but only if they are implemented objectively, using a reproducible algorithm, can these biases be understood, quantified and corrected for, if necessary. As ACO noted, homogeneous cluster catalogues with objective selection criteria can be produced using plate-measuring machines. Two such catalogues now exist, constructed from scans of UK Schmidt plates by the COSMOS and APM machines, with objective selection criteria whose design was informed by the critique of the Abell/ACO catalogue summarised above: these two catalogues and the spatial correlations in redshift samples derived from them are the subject of this section.

2.3.1 The Edinburgh/Durham Cluster Catalogue

The Edinburgh/Durham Cluster Catalogue (EDCC) is composed of clusters selected from the Edinburgh/Durham Southern Galaxy Catalogue (EDSGC, Heydon-Dumbleton *et al.* 1989). The EDSGC consists of COSMOS scans of 60 UK Schmidt plates covering 0.5 steradians, centred on the South Galactic Pole. The catalogue is asserted to be 95% complete to a magnitude of b_j=20.5, with stellar contamination at the $\leq$10% level and plate-to-plate magnitude offsets of $\Delta m \sim 0.05$ mag.

The construction of the EDCC from the EDSGC is detailed by Lumsden *et al.* (1992, LNCG). The method employed by LNCG is similar in approach to that used to construct the Abell/ACO catalogue, to facilitate a direct comparison of the two catalogues.

The first step is to bin the EDSGC data: LNCG use binned data, rather than employ percolation analysis, because this latter method can introduce undesirable biases into the cluster catalogue, by merging distinct objects into a single structure. The EDSGC data were binned into equal area bins, whose size was chosen to produce a mean occupancy of roughly one galaxy per bin. Three runs over the EDSGC were made, down to different limiting magnitudes and, consequently, with different bin sizes: this was done to detect clusters over as broad a redshift range as possible.

The binned data for each run were then lightly smoothed, using a Shectman (1985) filter, which is a 3 $\times$ 3 grid, with weights

$$\begin{array}{ccc} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16}, \end{array}$$

which roughly approximates a Gaussian in two dimensions: this light smoothing reduces binning noise. A 'sky frame' was then made for each run, by taking its 'Shectman frame' and smoothing it with a filter of about 1° to 2°: the details of the filter have a negligible effect on the resulting cluster catalogue. The 'sky frame' was then subtracted from the 'Shectman frame' and cluster candidates identified as excursion sets of pixels above a threshold; the threshold appropriate to each run being determined empirically, so as to

produce minimum contamination but not to miss any clusters that had been identified by eye in small test regions. A 'deblending' algorithm was then applied, to assign each pixel to only one cluster candidate. From each run this procedure yielded a set of contiguous pixels and the COSMOS information relating to each of the EDSGC galaxies contained within those pixels.

An iterative Abell-type analysis was then applied to the candidate list from each run. For each candidate cluster the magnitude of the tenth brightest galaxy, m_{10}, was estimated and an Abell radius, r_A, deduced. The m_{10} value was then corrected for contamination, using the number-magnitude relation for galaxies in a $4^\circ \times 4^\circ$ square region of the plate centred on the centroid of the candidate cluster. This corrected m_{10} was then used to start the next iteration. This procedure was followed until the difference between successive estimates of m_{10} differ by less than 1%. Cluster candidates are rejected if: (*i*) the final m_{10} estimate is fainter than the EDSGC catalogue limit; (*ii*) the centroid of the candidate moves by more than a quarter of the original Abell radius during the course of the iteration procedure; or (*iii*) no convergence was achieved in ten iterations. Once convergence had been attained, the final m_{10} value and cluster centroid were used to make the Abell count in the $m_3 \rightarrow m_3 + 2$ interval, subject to background correction by the same method as described above for m_{10}.

The candidates listed for the three runs were then combined and a two-sided Kolmogorov-Smirnov test was then applied to the background-subtracted magnitude distributions of candidates with centroids within the Abell radius of a candidate from another run, to reject clusters detected in more than one run. This procedure yielded a final catalogue of 769 clusters, deblended so that each galaxy was assigned to only one cluster, for which the Abell parameters were then determined. All the candidates were examined by eye and doubtful ones flagged in the catalogue.

As with plate overlaps in the Abell and ACO catalogues, the selection of candidate clusters in more than one run facilitated an internal error estimation for the cluster richnesses and m_{10} magnitudes in the final catalogue: the results of this error analysis are summarised in Table 2.3. The scatter in the magnitudes is reduced when only the richest ($c > 30$) clusters are considered, but the scatter on the richness is not.

Table 2.3: Internal error estimates for the EDCC (from LNCG)

Sample	σ_{m_1}	σ_{m_3}	$\sigma_{m_{10}}$	σ_{counts}
All clusters: run 1 v. 2	0.54	0.36	0.24	10.4
All clusters: run 1 v. 3	0.52	0.31	0.34	9.2
All clusters: run 2 v. 3	0.42	0.31	0.26	9.5
$c > 30$: run 1 v. 2	0.57	0.24	0.15	10.4
$c > 30$: run 1 v. 3	0.60	0.24	0.19	8.2
$c > 30$: run 2 v. 3	0.46	0.16	0.15	9.5
Final catalogue:	0.34	0.23	0.18	6.9
Final catalogue: $c > 30$	0.38	0.15	0.15	6.5
Final catalogue: offset $< 2'$	0.32	0.14	0.08	4.6

LNCG explain that this is because the richness error is roughly equal to the product of the cluster richness and the error on m_3, so that even though the m_3 error decreases roughly in proportion to richness, the two terms will always give approximately the same product. The scatter on both the richness and the magnitudes is reduced by considering only those clusters whose centroids are offset by less than $2'$ in the two runs, to produce what LNCG call the "irreducible errors" in their catalogue, which they say are inherent in the iterative Abell analysis procedure they employ.

LNCG follow their description of the construction of the EDCC with a comparison between it and the Abell/ACO catalogue: this comparison requires correction between the R magnitude system of Abell, the V magnitudes of ACO and the b_j magnitudes of the EDCC. LNCG find ~80% of the Abell/ACO clusters in the EDCC area down to its limiting magnitude $m_{10}(b_j) = 18.75$, but find that only one third of those clusters which are quoted as $R \geq 1$ have a count of $\geq$30 galaxies and only one in six have more than the required 50 galaxies: this is broadly in line with the estimates of Lucey (1983) on the basis of his simulations, but greatly exceed the internal estimates of the scatter in Abell/ACO richnesses made by Struble & Rood (1991). LNCG show that Abell and ACO assigned higher richnesses to the clusters they found in common with the EDCC. This is, in part at least, due to poor m_{10} estimates in the Abell/ACO catalogue: the m_{10} magnitude difference increases with EDCC m_{10}, so that fainter Abell/ACO clusters

Table 2.4: Estimates of internal errors in the Abell/ACO catalogue, from comparison with the EDCC (from LNCG)

Sample	σ_{m_1}	σ_{m_3}	$\sigma_{m_{10}}$	σ_{counts}
All ACO clusters	1.28	0.98	0.82	33.0

have m_{10} increasingly too faint, and, hence, Abell radii and galaxy counts increasingly too large.

This comparison with the EDCC cannot be regarded as an absolute test of the reliability of the Abell/ACO catalogue, since the EDCC will have its own biases, but the results shown in Table 2.4 do indicate that the errors in the Abell/ACO magnitudes and richnesses are larger than these authors estimated: it would certainly appear that some clusters will have been assigned to distance and/or richness classes more than one level away from their true designation.

Spatial correlations of a sample of EDCC clusters are considered by Nichol *et al.* (1992, NCGL). For this purpose the following selection criteria are used:

- clusters must have ≥ 22 galaxies in the magnitude range $m_3 \rightarrow m_3 + 2$ within a distance $1h^{-1}$ Mpc of the cluster centre: by choosing a distance smaller than the Abell radius NCGL hope to suffer from less projection contamination
- the clusters must be near enough that $m_{10}(b_j) \leq 18.75$
- the clusters must be located on the sky in the region $21^h53^m \leq \alpha \leq 03^h35^m$, $-42°12' \leq \delta \leq -22°53'$.

These criteria selected 97 clusters from the EDCC, for which 79 have redshifts [63 from the Edinburgh/Milano Redshift Survey (EMRS, Nichol 1992), plus 16 from the literature], giving a 90% completeness. Fitting the correlation function of this sample to the canonical form, $\xi(r) = (r_0/r)^\gamma$, in the range $3 \leq r \leq 35\ h^{-1}$ Mpc, yields $r_0 = 16.4 \pm 4\ h^{-1}$ Mpc, $\gamma = 2.1 \pm 0.3$ with 1σ bootstrap errors or $r_0 = 16.2 \pm 2.3\ h^{-1}$ Mpc, $\gamma = 2.0 \pm 0.2$ for 1σ Poisson errors.

NCGL show that their sample yields a contour plot of $\xi(r_z, r_p)$ with a very low level of anisotropy: the best fit Gaussian velocity broadening for convolution with a spatial correlation function model of $\xi(r) = (16\ h^{-1}\ \mathrm{Mpc}/r)^2$ has a width of $\sigma_v = 300$ km s^{-1} (Nichol 1992), with a velocity dispersion of 1000 km s^{-1} being excluded at the 2σ level.

The correlation function of the EMRS sample has a relatively large statistical error, due to its low number of clusters, but its systematic errors are probably very low: the 71 clusters taken from the EMRS have, typically, ~ 10 measured redshifts, which should be sufficient to eliminate distance errors due to the presence of interlopers, while the low level of anisotropy present in $\xi(r_z, r_p)$ would seem to indicate that LNCG have succeeded in their aim of creating a catalogue free from serious contamination by projection effects.

2.3.2 The APM Cluster Catalogue

Dalton (1992) discusses the creation of cluster catalogues from the APM Galaxy Survey of Maddox *et al.* (1990a-c) and the spatial correlation function of a redshift sample of these clusters is presented by Dalton *et al.* (1992, DEMS). The APM Galaxy Survey contains over two million galaxies, selected from scans of 185 UK Schmidt plates by the APM plate-measuring machine. It is attested to be 90–95% complete at a limiting magnitude of $b_j = 20.5$, with stellar contamination at about the 5% level. Like LNCG, the APM group consider cluster selection algorithms based on that of Abell, but designed to circumvent some of its problems. They find (on the basis of simulations) no single algorithm that satisfies their requirement of stability and ability to select clusters equally well over a broad redshift range, but find a general improvement over the Abell algorithm if they use: (*i*) a counting circle with a radius smaller than r_A; (*ii*) a richness-dependent distance indicator, m_X, in place of m_{10}; and (*iii*) a galaxy count which gives a higher weight to galaxies near the cluster centre.

The redshift sample of DEMS is drawn from a catalogue constructed using a counting radius of $r_C = 0.75\ h^{-1}$ Mpc. The selection procedure starts with a percolation analysis, linking APM galaxies to produce a list of candidate clusters: this causes DEMS to miss nearby ($z \leq 0.035$) clusters, which are too extended on the sky to be picked up by percolation analysis, but there are so few of them that this should have a negligible

effect on the correlation function of the sample.

The following iterative selection procedure is followed for each cluster on the percolation list, starting with an initial guess of m_X=18.4:

1. predict search radius from a redshift, z_{est}, given by $z_{est} = \mathrm{dex}\,(0.125\, m_X - 3.37)$

2. determine the centroid of the cluster candidate

3. obtain the weighted richness count, $\mathcal{R}$, of the cluster, which is defined by

$$\mathcal{R} = \sum_i C_i \; - \; N_b, \tag{2.3}$$

where the weighted counts C_i are defined by

$$C_i \; = \; \frac{1}{\langle w \rangle} \sum_{j \leq i} w_j \; - \; \pi r_C^2 \bar{n}(< m_i), \tag{2.4}$$

with $\bar{n}(< m_i)$ being the surface number density of galaxies brighter than m_i, which is the apparent magnitude of the ith galaxy (in order of increasing magnitude), whose weight, w_i, is defined by

$$w_i \; = \; (1 + 2r_i/r_C)^{-1}, \tag{2.5}$$

where r_i is the distance of the ith galaxy from the cluster centroid. The background count, N_b, is equal to the number of galaxies in the magnitude range $[m_X - 0.5, m_X + 1.5]$ expected to lie within the counting circle, on the basis of the number found in an annulus 5–6 times the counting radius of the cluster and centred on the cluster centroid.

4. set X equal to $\mathcal{R}/2$ (taking X proportional to $\mathcal{R}$ should take account of the Scott effect)

5. evaluate m_X, by counting down the magnitude-ranked galaxy list until $C_i \geq$ X and setting $m_X = m_i$.

This loop is taken to converge if successive values of m_X differ by less than 0.025 mag and if the cluster centroids are within 40″. Clusters are rejected if the background-subtracted galaxy count falls below X at any time, or if convergence is not attained within 20 iterations.

This procedure produces a catalogue of 240 clusters with $\mathcal{R} \geq 20$ and $z_X \leq 0.1$. DEMS have constructed a redshift survey containing 211 of these clusters, which is, therefore, ~90% complete. From these they consider a sample of 190 clusters in the redshift range $5000 \leq cz \leq 35000$ km s^{-1} for correlation analysis, finding a correlation length of 13.2 h^{-1} Mpc for a fit to the canonical power law. Constraining this fit to $\gamma = 2$ yields $r_0 = 12.9 \pm 1.2\ h^{-1}$ Mpc for this full sample and $r_0 = 14 \pm 4\ h^{-1}$ Mpc for the richest 93 clusters, which have $\mathcal{R} \geq 35$. The redshift space anisotropy is low, with a good fit to $\xi(r_z, r_p)$ coming from convolving $\xi(r) = (13\ h^{-1}\ \mathrm{Mpc}/r)^2$ with a Gaussian of width $\sigma_v = 700$ km s^{-1}, which is about $\sqrt{2}$ times DEMS' estimated redshift error. Efstathiou *et al.* (1992b) consider a sample of $\mathcal{R} \geq 20$ clusters from a slightly wider redshift range ($5000 \leq cz \leq 40000$ km s^{-1}) and obtain a correlation length of $r_0 = 13.2\ h^{-1}$ Mpc.

2.4 OBSERVATIONAL BACKGROUND: SUMMARY

In the previous two sections we have reviewed the observational study of cluster correlations, through a discussion of the Abell/ACO, Edinburgh/Durham and APM cluster catalogues: these are the sources of the observations with which we shall compare our theoretical predictions in Section 2.8 below.

As we have seen, samples of Abell/ACO clusters display stronger correlations than those drawn from the two automated catalogues and it has been suggested that part of the clustering strength observed in samples of Abell/ACO clusters is spurious, resulting from selection biases in that catalogue. On the basis of the evidence presented above, we believe this undoubtedly to be the case for Abell $R \geq 0$ cluster samples and we shall not use these in what follows. Those authors who have included $R = 0$ clusters in their studies have justified their inclusion on the basis of the redshift distribution of $R = 0$ clusters, arguing, for example, that the distributions of $R = 0$ and $R = 1$ clusters

vary with distance in a consistent fashion out to the limiting redshift of their particular sample. That is not a sufficient justification, since it fails to consider the distribution of the clusters on the sky: poorer clusters are more likely to suffer from spurious angular correlations due to projection effects. We note that, in excluding $R = 0$ clusters, we are acting in accordance with Abell's belief that these should not be used in statistical work.

The question of significant projection contamination is far less clear-cut in the case of samples of Abell/ACO $R \geq 1$ clusters. In what follows we shall assume the truth of the hypothesis of PW92 that the correlations in their $R \geq 1$ sample are genuine and use them to constrain theory. The redshift-space anisotropy that PW92 find for this sample is less severe than that in their $R = 0$ sample and is equivalent to convolution of an isotropic real-space correlation function [$\xi(r) = (r/21\ h^{-1}\ \mathrm{Mpc})^{-2}$] with a Gaussian pairwise velocity distribution with a dispersion of about 1000 km s^{-1}. They argue that this is not unreasonable, given that, in their desire to construct as large a redshift sample as possible, they have included cluster redshifts from a large number of sources, some estimated, some from a single galaxy in a cluster and, thus, they are likely to include a number of inaccurate redshifts. This may be true, but it is also possible that, by adopting this cautious attitude to their redshift database, PW92 are inadvertently 'explaining away' projection contamination: what is certainly true is that the redshift-space anisotropy test cannot be profitably used to detect projection contamination in a redshift sample with such a possibility of anisotropy from other sources. It is arguable, indeed, whether the anisotropy test can ever be an unambiguous diagnostic for projection contamination in a redshift sample of the size currently feasible. The number of expected cluster pairs varies with projected separation in proportion to $r_\mathrm{p}\ \mathrm{d}r_\mathrm{p}$ and so a very small number of close pairs are expected. This means that the chance alignment of genuine clusters close to the line of sight can produce a large anisotropic feature in $\xi(r_\mathrm{p}, r_z)$ at low r_p. PW92 cite the example from their sample of the cluster A3565, which lies at a redshift of z=0.01 along the line of sight to the Shapley concentration, which contains ten $R \geq 1$ clusters at z=0.05: this alignment produces a high value for $\xi(r_\mathrm{p}, r_z)$ at $(r_\mathrm{p} = 0, r_z = 120\ h^{-1}\ \mathrm{Mpc})$, which might be taken as evidence of contamination were its origin not known.

It should, however, be possible to distinguish between such isolated high values of

$\xi(r_p, r_z)$ and the extended feature along the r_z axis that is expected to follow from projection effects. These isolated features could, however, cause problems for the Sutherland correction procedure (see Section 2.2.3), which assumes that there is no genuine clustering at separations beyond $r_z = 50\ h^{-1}$ Mpc. Efstathiou *et al.* (1992b) estimate that over-correction by the Sutherland method due to this effect could introduce biases of order ~ 0.01 into the correlation function of an otherwise unbiased catalogue. Whilst this will have a negligible effect on the correlation length derived for the cluster sample, it could have a significant effect on scales where $\xi < 1$ where, as we shall see below, comparison between theory and observation is most fruitful. For this reason we shall not compare our theoretical results with observational data to which the Sutherland correction method has been applied.

The two automated catalogues were designed to circumvent the selection problems that compromise the Abell/ACO catalogue, but they still select clusters as over-densities in the distribution of galaxies seen in projection on the sky and it could be argued (*e.g.* Frenk *et al.* 1990) that this procedure must necessarily introduce projection contamination at some level. It is desirable, therefore, to look for different methods of cluster selection which might be expected to produce contamination-free catalogues. One way of doing this is to use radio galaxies (Peacock & Nicholson 1991) which are found, typically, in relatively poor clusters (Prestage & Peacock 1988). Peacock & Nicholson (1991) determine a correlation function for their sample of ~ 300 radio galaxies which they fit to canonical power law form (for $r < 30\ h^{-1}$ Mpc) to obtain $r_0 = 11.0 \pm 1.2\ h^{-1}$ Mpc, $\gamma = 1.8 \pm 0.3$.

Another method, advocated by West & van den Berg (1991), uses cD galaxies, which are readily identifiable, due to their size and luminosity, and which are known to be good tracers of rich clusters: indeed, it is generally assumed that a high-density environment is a prerequisite for their formation, perhaps through merging of galaxies. West & van den Bergh (1991) have drawn a sample of 64 clusters containing cD galaxies with measured redshifts of $z \leq 0.1$ from the SR87 compilation of Abell clusters: 44 of these are $R \geq 1$ clusters. They fit the correlation function of this sample to the canonical power law form and obtain $r_0 = 22.1\ h^{-1}$ Mpc and $\gamma = 1.7$, with 68% error ranges from bootstrap resampling of $1.22 \leq \gamma \leq 2.18$, $15.3 \leq r_0 \leq 28.9\ h^{-1}$ Mpc. This correlation function

is consistent (perhaps not too surprisingly, given the large statistical uncertainties that result from such a small sample size) with that of the complete sample of all (138) $R \geq 1$ clusters with measured $z \leq 0.1$. West & van den Bergh claim that this indicates that the strong clustering seen in samples of Abell/ACO clusters is not spurious, since cluster selection using cD galaxies should not be prone to projection contamination. This is, however, an invalid deduction to make, since the cD cluster sample used is a subset of the Abell catalogue: the presence of cD galaxies had no bearing on its selection, which was conducted using the same criteria as the rest of the Abell/ACO catalogue and it cannot, therefore, be considered *a priori* as being any more reliable than any other subset of that catalogue. It remains true, however, that cD galaxies may provide a good way of selecting clusters, but selection of cD cluster samples must be performed solely on the basis of the presence of the cD galaxy, without any other criteria that might be prone to projection contamination.

An alternative approach which, at first sight, appears very promising is X–ray selection. Intracluster gas is known to emit X–rays, but it is only in the cores of the clusters, deep in their potential wells, that the temperature and density of the gas are sufficiently high that substantial thermal emission is possible: this is important, since it reduces the possibility of projection effects contaminating X–ray selected cluster samples, as the X–ray emitting region of the cluster covers a much smaller area on the sky than the halo of galaxies.

The first study of the spatial correlations of a sample of X–ray clusters was performed by Lahav *et al.* (1989, LEFP). They compiled a sample of the 53 brightest clusters in the sky from a miscellany of observations by the *HEAO-I*, Ariel V, *EXOSAT* and Einstein X–ray satellites. The number density of this sample is $\sim 10^{-5}\ h^3$ Mpc^{-3}, which is comparable to that of Abell $R \geq 1$ clusters, but LEFP estimate that the completeness of their sample is only ~60% and that, in particular, they are missing a large number of fainter clusters at low Galactic latitude, due to confusion with Galactic sources.

Fitting the correlation function of this sample to a $\gamma = 1.8$ power law gives a correlation length of $r_0 \simeq 21\ h^{-1}$ Mpc, but this falls to $r_0 = 17\ h^{-1}$ Mpc when only the 45 clusters with $\mid b \mid \geq 20°$ are considered. The large statistical errors here are compounded by

selection effects resulting from the different resolutions of the various instruments used to detect the clusters.

Some of the problems that faced LEFP should be overcome by those compiling samples of clusters on the basis of X–ray emission detected by *ROSAT*, since they may select clusters using homogeneous selection criteria, based on the capabilities of a single instrument: we shall consider in detail a sample of *ROSAT* clusters in Section 2.9. There are, however, problems inherent in X–ray cluster selection too, some of which are as difficult to circumvent as those which plague optical selection. A particularly serious problem concerns X–ray emission from AGN: given the spatial and spectral resolution of an instrument like*ROSAT*, it is difficult to distinguish between X–ray emission from cluster gas and from AGN, especially for distant clusters which are not extended and faint clusters, from which few photons will have been detected. More fundamental problems results from the likely association of AGN with clusters: should candidates which are revealed to be AGN be removed from cluster catalogues or, conversely, should AGN be used as tracers of cluster environments? It is clear that X–ray cluster selection is far from straightforward and *ROSAT* samples are being constructed using optical follow-up programmes, indicating the practical difficulties associated with selection based solely on X–ray emission.

Despite all the difficulties with cluster selection discussed in this chapter, it has been claimed, most forcefully by Bahcall and collaborators (*e.g.* Bahcall 1988, Bahcall & West 1992), that the complete suite of observations of cluster correlations described above fit neatly into a consistent picture. Bahcall asserts that there exists a universal correlation function, of the form $\xi(r) = (0.4d/r)^{1.8}$, where d is the characteristic separation of the sample of objects under consideration (and is related to the number density, N, of the objects by $d \equiv N^{-1/3}$). This relation is shown in Figure 2.1, which we have adapted from Bahcall & West (1992): the absence of number densities (from which to deduce d) from many of the original papers used to produce this figure (most of which are not identified by Bahcall & West) results in our being unable to reproduce this diagram for ourselves, adding, as we should like, the error bars that Bahcall & West omit. The solid line in this figure marks the prediction of Bahcall's universal correlation function: $r_0 = 0.4d$. Figure 2.1 includes data points from the cluster catalogues of Shectman (1985) and

Zwicky (1961–1968). These catalogues are not discussed in detail here since, for various reasons, they are not particularly suitable for statistical studies of cluster correlations. Shectman (1985) selected clusters from the Shane & Wirtanen (1967) Lick galaxy counts, as corrected by Seldner *et al.* (1977): the resultant catalogue is compromised by many of the problems (large-scale gradients, etc.) that are known to corrupt the Lick catalogue. Zwicky and his collaborators compiled a catalogue of ~9700 clusters from Palomar sky survey plates, but their richness classification is made in a redshift-dependent fashion, which is one reason why this catalogue is unsuitable for clustering studies.

Implicit in the use of the $d - r_0$ relation, as plotted in Figure 2.1, is a very strong assumption about the consistency of the methods used by theorists and observers to identify and select clusters in their work. The assumption is that associations of galaxies may be ranked in order by some property which is a convenient measure of richness (mass, for example) and that all selection procedures involve stepping down this ranked order to a particular point, with the result that cluster samples of a given number density, however selected, may be taken as being equivalent. This cannot, of course, be absolutely true. There must be some morphological influence on cluster selection using an Abell-type count of galaxies within a counting circle on the sky. Bahcall & West (1992) discuss a specific example of this. They argue that since Dalton *et al.* (1992) use a small counting circle they will select groups of galaxies (such as Morgan groups), whose central density is high, but whose total population is insufficient to lead to inclusion in a catalogue like Abell's where selection is determined by the total galaxy count over a wider area of sky. This is an interesting point and one that deserves further consideration but, in what follows, we shall assume that it has little effect on the correlation length of samples of optical galaxies, so that we are able to use the $d - r_0$ relation to compare the correlations of APM, EMRS and Abell/ACO clusters. A stronger objection should be raised, however, at the inclusion by Bahcall & West of the data point for the Lahav *et al.* X-ray cluster sample in Figure 2.1. Leaving aside the possibility of systematic differences in relative cluster richnesses determined by studying the optical and X-ray properties of clusters, the optical samples discussed above are mainly volume-limited (they are flux-limited insofar as their selection depends on the visibility of galaxies above the magnitude limit of the plate, but we assume that the imposition of a redshift

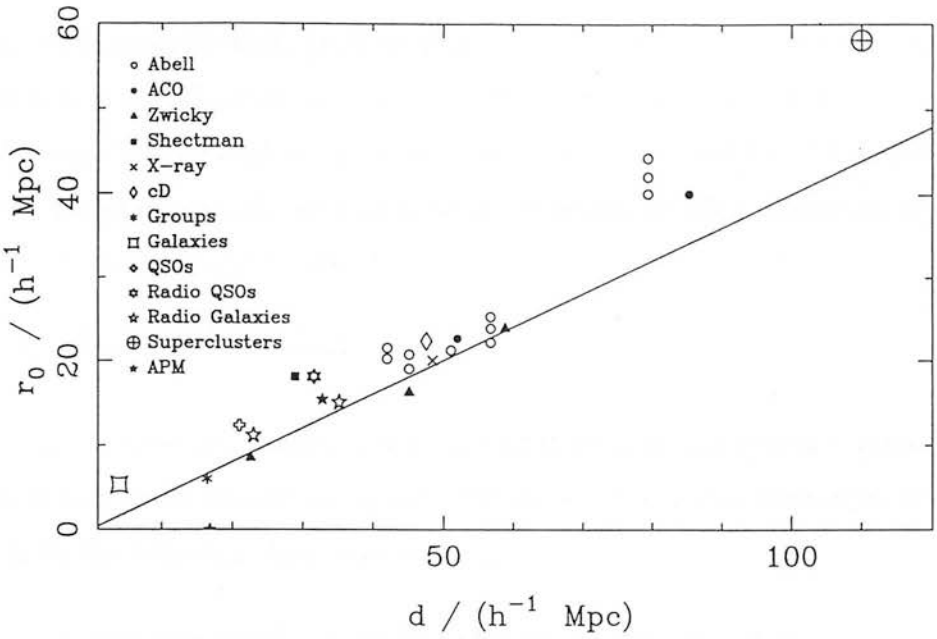


Figure 2.1: The universal correlation function of Bahcall & West

limit is the stronger bound) while the X-ray sample of Lahav *et al.* is flux-limited, as are those being constructed from the ROSAT All-Sky Survey. It is, therefore, incorrect to include X-ray cluster samples on the $d - r_0$ diagram for optically selected samples, since their correlation length will be determined, in essence, by integrating over exactly such a curve, weighted by the cluster luminosity function.

In the Bahcall picture, the observed correlations of Abell/ACO, EMRS and APM clusters are mutually consistent, with a richness dependence of clustering strength accounting for the differences in correlation length for samples drawn from the three catalogues. One of the main goals of this chapter is to see whether this supposed richness dependence can be reproduced by any theoretically plausible power spectrum of primordial density fluctuations.

2.5 THEORETICAL BACKGROUND

This section reviews previous theoretical studies of cluster correlations. We discuss in

some detail analytic work, predicated on a model in which sites of cluster formation are identified with high peaks in a Gaussian linear density field: cluster correlations in this model were first studied in the seminal paper of Kaiser (1984). We follow this with a brief description of early numerical work, reserving detailed discussion of more recent numerical studies until Section 2.7.

2.5.1 The Kaiser Effect

In a remarkable paper, Kaiser (1984) showed how the strong spatial correlations of Abell clusters may be accounted for by a simple model of structure formation, at the heart of which lie the following three assumptions:

1. the primordial density field obeyed Gaussian statistics
2. the primordial density field had fluctuations on the mass scale of clusters superposed on lower amplitude, longer wavelength modes
3. clusters formed at peaks in the primordial density field, smoothed on cluster mass scales, which exceeded some (moderately high) threshold.

These assumptions are all very plausible. The first assumption is particularly well motivated in inflationary models of the early universe, which predict Gaussian primordial fluctuations as a generic feature, but it is also the default hypothesis: the central limit theorem would lead one to expect Gaussian primordial fluctuations unless there is some very good reason why not in a particular case. The second assumption selects a generic hierarchical model for the power spectrum, and so excludes, for example, neutrino-dominated models, where small-scale fluctuations will have been damped at early times by free streaming. The final assumption is plausible too, if not beyond question: local density maxima seem eminently reasonable places for bound objects to form, as one would expect, *a priori*, that objects are more likely to form where there is more mass. There follows, therefore, a picture in which the highest density peaks collapse first, so that a threshold is naturally set by the requirement that a given density perturbation must have collapsed by the present to be observed today: this reasoning is supported by the work of White *et al.* (1992), who show that the linear rms density fluctuation today is roughly unity on the scales appropriate to the formation of clusters.

Calculation of the two-point correlation function of local maxima in a Gaussian random field that lie above a certain threshold is very complicated, as it involves correlations between the first and second spatial derivatives of the field at the two points, as well as correlations between the values of the field itself. To avoid this complexity, Kaiser chose to address a closely related, but much simpler, problem, that of computing the correlation function of points in the density field that lie above a threshold: by dropping the constraint that the points be local maxima one also drops the need to include correlations between the spatial derivatives of the field, which simplifies the analysis greatly.

If the threshold is set to be $\nu\sigma_0$, where σ_0 is the rms density fluctuation [$\sigma_0^2 = \xi(0)$, where $\xi(r)$ is the correlation function of the density field], then the desired correlation function (denoted by $\xi_{>\nu}$) is given by

$$1 + \xi_{>\nu}(r) = \frac{P_2}{P_1^2}, \tag{2.6}$$

where P_1 is the probability that a randomly selected point should lie above the threshold and P_2 is the joint probability that two such points should be found a distance r apart.

Now,

$$P_1 = \int_{\nu\sigma}^{\infty} P(y)\,\mathrm{d}y, \tag{2.7}$$

where $P(y) = (2\pi^2\sigma^2)^{-1/2}\exp(-y^2/2\sigma^2)$ and

$$P_2 = \int_{\nu\sigma}^{\infty}\int_{\nu\sigma}^{\infty} P(y_1, y_2)\mathrm{d}y_1\,\mathrm{d}y_2, \tag{2.8}$$

where

$$P(y_1, y_2) = (2\pi)^{-1}\left[\xi^2(0) - \xi^2(r)\right]^{-1/2}\exp\left\{\frac{-[\xi(0)y_1^2 + \xi(0)y_2^2 - 2\xi(r)y_1y_2]}{2[\xi^2(0) - \xi^2(r)]}\right\}. \tag{2.9}$$

It follows that $\xi_{>\nu}$ is given by

$$1 + \xi_{>\nu}(r) = (2/\pi)^{1/2}\left[\mathrm{erfc}(\nu/\sqrt{2})\right]^{-1/2}$$

$$\times \int_0^{\infty} \mathrm{e}^{-1/2y^2}\,\mathrm{erfc}\left\{\frac{\nu - y\xi(r)/\xi(0)}{\sqrt{2[1 - \xi^2(r)/\xi^2(0)]}}\right\}\mathrm{d}y. \tag{2.10}$$

In the limit that $\xi(r)/\xi(0) \ll 1$ and $\nu \gg 1$ this reduces to

$$\xi_{>\nu}(r) \simeq \left(\frac{\nu^2}{\sigma^2}\right) \xi(r), \tag{2.11}$$

i.e. the correlation function of points above a threshold is an amplified version of that of the underlying field, with the degree of amplification increasing with increasing threshold.

Kaiser applied this formalism to the correlations of Abell clusters, taking the density field to be an $n = -1$ power law (in accord with the observed slope of the galaxy correlation function on the appropriate scales), smoothed on the mass scale of a rich cluster. This smoothing length and the corresponding threshold, ν, were chosen such that they select objects with the correct abundance that would collapse (in the spherical collapse model) at time $t_{\rm coll}$ in the interval $1 \geq t_{\rm coll}/t_0 \geq 1/2$, where t_0 is the present age of the universe.

Thus specified, the Kaiser model estimates that the large-scale correlations of Abell $R \geq 1$ clusters are amplified by about an order of magnitude above those of the underlying density field. The richness dependence of clustering strength predicted by this model is seen to be in qualitative agreement with the observations of Shectman (1985), Abell $R \geq 1$ and $R \geq 2$ clusters.

The Kaiser model implies that the strength of the clustering of peaks of a given height on a given scale will depend on the amount of power in the density field at longer wavelengths and this is borne out by the work of Peacock & Heavens (1985), who made realisations of Gaussian noise with different power spectra and computed the correlation functions of the resulting peak distributions. The numerical work of Peacock & Heavens also reveals the importance of thresholding in producing the Kaiser effect: they find that the amplification increases dramatically with threshold, thus validating another prediction of the Kaiser model.

2.5.2 Further approximations to the correlations of peaks in Gaussian noise

Several authors have extended Kaiser's original study of the correlations of points in Gaussian random fields lying above a threshold. Politzer & Wise (1984) follow Kaiser's

treatment and obtain an expression for the n–point correlation function of thresholded points. In the limit that $(\nu/\sigma_0)^2 \gg 1$ and $\xi(\mathbf{r}_i - \mathbf{r}_j) \ll 1 \ \forall \ i,j$ they show that the n–point correlation function reduces to a combination of two–point functions; the two–point function taking the form:

$$1 + \xi_{>\nu}(r) = \exp\left[\nu^2 \xi(r)/\sigma_0^2\right], \tag{2.12}$$

which agrees with Kaiser's original result (equation 2.11) in the limit of weak clustering of thresholded regions. Silk & Vittorio (1987) use this approximation in their study of large-scale structure in low-Ω models, comparing its predictions with the BS83 correlation function.

Equations 2.11 and 2.12 are both approximations to equation 2.10 in different limits: Politzer & Wise require that the correlations in the underlying field be weak, while Kaiser needs the tighter constraint that the clustering of thresholded regions in that field is weak. Jensen & Szalay (1986, JS86) show how the n-point correlation functions for thresholded regions may be calculated to arbitrary accuracy, using a series solution. They find that the two-point correlation function may be written as

$$\xi_{>\nu}(r) = \sum_{m=1}^{\infty} \left[\frac{\xi(r)}{\xi(0)}\right]^m \frac{A_m^2}{m!}, \tag{2.13}$$

where $A_n = 2x H_{n-1}(x)\ 2^{-n/2}/[\sqrt{\pi} x e^{x^2} \mathrm{erfc}(x)]$, with $H_n(x)$ being the Hermite polynomials and $x = \nu/\sqrt{2}$: in the limit $\nu \to \infty$ this reduces to Kaiser's result (equation 2.11).

There are problems with using the correlations of thresholded points as an approximation to those of peaks. As noted by Coles (1986), the correlation function of thresholded points will overestimate the peak correlations at small scales (scales smaller than the typical size of the thresholded regions) because it will include counts of pairs which are within the same thresholded region: Kaiser (1984), in fact, acknowledged that the correlation function of thresholded regions will be more like a mass-weighted cluster correlation function than a simple cluster-cluster correlation function, anticipating Coles' objection. A second problem is that the thresholded points model predicts that the

zero-crossings of the peak-peak correlation function should coincide with those of the underlying field, but this is not seen in the numerical realisations of Gaussian random noise made by Otto, Politzer & Wise (1986) and Lumsden, Heavens & Peacock (1989, LHP). This may be readily explained as a result of the neglect of correlations of the first and second spatial derivatives of the field: if the correlation function of the field is steep when it crosses zero then its derivatives may not be neglected and their correlations may produce a non-vanishing peak-peak correlation function at that separation. Furthermore, Coles (1989) notes that the series solution presented by Jensen & Szalay (1986) is slow to converge, which reduces its utility.

Bardeen, Bond, Kaiser & Szalay (1986, BBKS) discuss the computation of the true n-point correlation functions of the peak distribution, defined by

$$1 + \xi^{(n)}_{\mathrm{pk-pk}} = \frac{\langle n_{\mathrm{pk-pk}}(\boldsymbol{r}_1) \ldots n_{\mathrm{pk-pk}}(\boldsymbol{r}_n) \rangle}{\langle n_{\mathrm{pk-pk}} \rangle^n}, \tag{2.14}$$

where $\langle n_{\mathrm{pk-pk}} \rangle \, \mathrm{d}V_i$ is the probability of their being a peak satisfying an appropriate threshold condition in the volume $\mathrm{d}V_i$ about the point $\boldsymbol{r}_i$ and $\langle n_{\mathrm{pk-pk}}(\boldsymbol{r}_1) \ldots n_{\mathrm{pk-pk}}(\boldsymbol{r}_n) \rangle \prod_i \mathrm{d}V_i$ is the joint probability of there being n such points in the volumes $\{\mathrm{d}V_i\}$ surrounding the points $\{\boldsymbol{r}_i\}$. This joint probability function is a function of $10n$ variables, being the values of the field, δ, its three first spatial derivatives and its six independent second spatial derivatives at the n points. The correlations between variables at a single point may be reduced to those between just two variables – *i.e.* $\nu = \delta/\sigma_0$ and $x = -\nabla^2\delta/\sigma_2$, where σ_i are the moments of the field. Further, BBKS assume that the cross-correlations between the derivatives of the field at different points may be neglected: the validity of this approximation in a particular case will depend on the field involved – for example, in the specific example BBKS consider (the CDM power spectrum smoothed on galactic scales), it is valid for separations greater than about four times the smoothing length used to define the peaks.

BBKS define n-component column vectors $\tilde{\nu}$ and $\boldsymbol{x}$, the components of which are

$$\tilde{\nu}(i) = \frac{[\nu(i) - \gamma x(i)]}{(1-\gamma^2)} \quad \text{and} \quad x(i) \quad \text{for} \quad i = 1 \ldots n, \tag{2.15}$$

where the spectral parameter, γ, defined by

$$\gamma \equiv \frac{\langle k^2 \rangle}{\langle k^4 \rangle^{1/2}} = \frac{\sigma_1^2}{\sigma_2 \sigma_0} \tag{2.16}$$

expresses the range of wavenumbers over which the field has significant power: $\gamma = 1$ if the power spectrum is a δ-function at a single wavenumber and is much less than unity for power spectra with power over a broad range of scales.

BBKS then define an $n \times n$ matrix, Ψ, whose diagonal elements are zero and whose off-diagonal elements are the normalised density cross-correlations, ψ_{ij}, $[\psi_{ij} = \xi(r_{ij})/\xi(0)]$ and a further matrix, $\mathbf{C}$, defined as

$$\mathbf{C} \equiv \{[\mathbf{\Psi}[\mathbf{I} + \mathbf{\Psi}/(1-\gamma^2)]\}^{-1}, \tag{2.17}$$

where $\mathbf{I}$ is the n-dimensional identity matrix. BBKS show that the joint probability distribution in the numerator of equation 2.14 then takes the form

$$\langle n_{\rm pk-pk}(\boldsymbol{r}_1) \ldots n_{\rm pk-pk}(\boldsymbol{r}_n) \rangle = (4\pi^2 R_*^3)^{-n} \{\det[\mathbf{I} + \mathbf{\Psi}/(1-\gamma^2)]\}^{-1/2}$$

$$\times \left\{ \prod_i \int \mathrm{d}\nu(i) \int_0^\infty \mathrm{d}x(i)\, t\left[\frac{\nu(i)}{\nu_t}\right] g\left[x(i), \gamma, \gamma\nu(i)\right] \mathrm{e}^{-\nu(i)^2/2} \exp\left(\frac{1}{2}\tilde{\nu}^{\rm T} \mathbf{C} \tilde{\nu}\right) \right\}, \tag{2.18}$$

where $R_* = \sqrt{3}\sigma_1/\sigma_2$, $t(\nu/\nu_t)$ is the peak selection threshold function $[t(\nu/\nu_t) = \theta(\nu - \nu_t)$, where θ is the unit Heaviside function, for a sharp threshold at $\nu = \nu_t]$. The quantity $g(x, \gamma, \gamma\nu)$ is given by

$$g(x, \gamma, \gamma\nu) \equiv \frac{f(x)}{[2\pi(1-\gamma^2)]^{1/2}} \exp\left[\frac{-(x-\gamma\nu)^2}{2(1-\gamma^2)}\right], \tag{2.19}$$

where a closed form for $f(x)$ is given in Appendix A. This $2n$-dimensional integration cannot, in general, be reduced to a product of simpler integrals, since $\mathbf{C}$ is non-diagonal by construction, so that evaluation of equation 2.14 is a formidable task, even for the case of n=2. Two further approximations are suggested by BBKS to enable progress beyond

this point. The first is to consider the weak clustering limit, in which $\nu_i^2\ \psi_{ij} \ll 1\ \forall\ i,j$. In this case

$$\exp\left(\frac{1}{2}\tilde{\nu}^{\mathrm{T}}\mathbf{C}\tilde{\nu}\right) \simeq 1 + \frac{1}{2}\tilde{\nu}^{\mathrm{T}}\mathbf{C}\tilde{\nu}, \tag{2.20}$$

and the integrals over x and ν may be evaluated separately at each point. In this limit equation 2.14 reduces to

$$\xi^n_{\mathrm{pk-pk}}(\boldsymbol{r}_1 \ldots \boldsymbol{r}_n) = \sum_{i<j} \langle\tilde{\nu}\rangle^2\ \psi(r_{ij}), \tag{2.21}$$

where $\langle\tilde{\nu}\rangle$ is the mean value of $\tilde{\nu}$ taken over all peaks selected by the threshold criterion, so that the two-point function takes the form

$$\xi_{\mathrm{pk-pk}}(r) = \frac{\langle\tilde{\nu}\rangle^2}{\sigma_0^2}\,\xi(r). \tag{2.22}$$

This is the same form as equation 2.11 derived by Kaiser (1984), but is not identical, because $\langle\tilde{\nu}\rangle < \nu$ for appropriate values of γ and ν. The second suggestion made by BBKS is to model $g(x,\gamma,\gamma\nu)$ as a Gaussian in x, with the result that the integrals over x in equation 2.14 may be performed analytically, reducing equation 2.14 to an n-dimensional integral, with a very complicated integrand.

BBKS present a second approach to the approximation of n-point peak correlations, using the so-called 'peak-background split' method. In this approach, the smoothed density field, δ_{s}, used to select the peaks is considered to be the sum of a 'peak' field, δ_{p}, and a 'background' field, δ_{b}, which is obtained by smoothing the full density field on a scale R_{b} larger than that (R_{s}) used to define δ_{s}: in a loose sense δ_{p} describes the local properties of the peaks of the field δ_{s}, while δ_{b} gives its large-scale correlations. This picture motivates the expression of $\langle n_{\mathrm{pk}}(\boldsymbol{r}_1)\ldots n_{\mathrm{pk}}(\boldsymbol{r}_n)\rangle$ in the form

$$\langle n_{\mathrm{pk}}(\boldsymbol{r}_1)\ldots n_{\mathrm{pk}}(\boldsymbol{r}_n)\rangle = \int \prod_i n_{\mathrm{pk}}[\nu_{\mathrm{b}}(i)]\,\mathrm{d}\nu_{\mathrm{b}}\ P[\nu_{\mathrm{b}}(1)\ldots\nu_{\mathrm{b}}(n)], \tag{2.23}$$

where $\nu_{\rm b} = \delta_{\rm b}/\sigma_{\rm b}$ and $\sigma_{\rm b}^2 = \xi_{\rm b}(0)$. The quantity $n_{\rm pk}(\nu_{\rm b})$ is the local density of peaks in the field $\delta_{\rm p}$ as a function of $\nu_{\rm b}$ and $P[\nu_{\rm b}(1)\ldots\nu_{\rm b}(n)]$ has a multivariate Gaussian distribution:

$$P[\nu_{\rm b}(1)\ldots\nu_{\rm b}(n)] \;=\; (2\pi)^{-n/2}\,[\det(\mathbf{I}+\mathbf{\Psi})]^{-1/2}\exp\left[-\frac{1}{2}\tilde{\nu}_{\rm b}^{\rm T}(\mathbf{I}+\mathbf{\Psi})^{-1}\tilde{\nu}_{\rm b}\right], \qquad (2.24)$$

where, as before, $\mathbf{\Psi}$ has vanishing diagonal elements and off-diagonal elements equal to the normalised cross-correlations, ψ_{ij} ; in this case the normalised cross-correlations of the background field, $\psi_{ij} = \langle\nu_{\rm b}(i)\nu_{\rm b}(j)\rangle$. Computation of $n_{\rm pk}(\nu_{\rm b})$ is not trivial, since $\delta_{\rm b}$ and $\delta_{\rm p}$ are correlated, by construction, as a result of their sum being equal to $\delta_{\rm s}$ at all points. In the 'peak-background split' approximation, BBKS ignore these correlations and *define* $\delta_{\rm p}$ to be a Gaussian field whose power spectrum is the difference between the power spectra of $\delta_{\rm s}$ and $\delta_{\rm b}$. With this simplification, it can be shown that $n_{\rm pk}(\nu_{\rm b})$ takes the form

$$n_{\rm pk}(\nu_{\rm b}) \;=\; n_0\exp\left(\alpha\nu_{\rm b}-\frac{1}{2}\beta\nu_{\rm b}^2\right), \qquad (2.25)$$

which we shall use later, in Section 2.6. This greatly simplifies the expression for $\langle n_{\rm pk}(\boldsymbol{r}_1)\ldots n_{\rm pk}(\boldsymbol{r}_n)\rangle$ and BBKS show how, in the limit that $\beta \ll 1$ (which holds when $R_{\rm b} \gg R_{\rm s}$), the peak-background split yields the following n-point correlation function

$$1 \;+\; \xi_{\rm pb}(r) \;\simeq\; \exp\left[\alpha^2\sum_{i<j}\psi_{ij}(r)\right], \qquad (2.26)$$

which is the same form as the correlation function of thresholded regions derived by Politzer & Wise (1986).

BBKS demonstrate that the peak-background split provides an excellent approximation to the true peak-peak correlation function for separations $r \gg R_{\rm b}$, where the correlations in $\delta_{\rm s}$ are well approximated by those of $\delta_{\rm b}$.

Lumsden, Heavens & Peacock (1989, LHP) argue that the approximations of BBKS outlined above may be less satisfactory for modelling the cluster distribution in a universe dominated by CDM than they are for the galaxy distribution. This is because

the larger smoothing scale appropriate to clusters pushes attention to scales where the power spectrum is flatter and the derivatives of ψ_{ij} may not be neglected so readily. To overcome this problem they consider a modification of the JS86 method, made in the light of the work of BBKS. They note that in the limit $\psi \to 0$, the JS86 formula (equation 2.13) reduces to $\xi_{>\nu} = A_1^2(\nu)\psi(r)$, while the asymptotic peak–peak correlation function (equation 2.22) is $\xi_{\rm pk-pk}(r) = \langle\tilde{\nu}\rangle^2\psi(r)$. Thus, by solving the implicit equation $A_1(\nu_{\rm eff}) = \langle\tilde{\nu}\rangle(\nu, \gamma\nu)$ they find an effective threshold, $\nu_{\rm eff}$. Using $\nu_{\rm eff}$ as the threshold in equation 2.13 they find that the JS86 formalism gives a much better fit at all separations to the true peak–peak correlation function (determined from a numerical realisation of the field $\delta_{\rm s}$) than is obtained using ν and they apply this modified JS86 approximation to flat baryon- and CDM-dominated models, as well as an open CDM model.

This modification does not, however, circumvent one of the failings in the JS86 approach, since it scales the amplitude of the quantities A_n, but will keep the zero-crossings of the correlation function for thresholded regions coincident with those of the underlying density field which, as we have seen, compromises its use as a model for the peak-peak correlation function. In view of this, LHP and Coles (1989) consider the approximation of the true peak-peak correlation function for a three-dimensional field by that of a one-dimensional skewer through it: the restriction to one dimension simplifies the analysis to the point that the peak-peak correlation function may be computed quite readily. Although the peaks in the one-dimensional slice will not, of course, generally coincide with peaks in the three-dimensional field, LHP argue that since the contours around peaks in the three-dimensional field are simply-connected (for even moderately high peaks), a peak in a one-dimensional slice is likely to lie close to a true, three-dimensional peak: Coles (1989) gives the same argument in more mathematical terms. LHP find that the correlation function of peaks in a one-dimensional slice is, indeed, a very good approximation to the true peak-peak correlation of peaks in the three-dimensional field, much better than approximations based on thresholded regions.

2.5.3 The dynamic evolution of peak-peak correlations

The methods described above may be used to determine the spatial correlations of the distribution of nascent clusters at an early epoch (provided that we are justified in identifying sites of cluster formation with peaks in a Gaussian random field), but the cluster distribution has evolved over time and account must be taken of the resultant dynamical component to its observed clustering today. The coupling between the statistical and dynamical components is complicated and so BBKS restrict their treatment of this topic to the limit of weak correlations in the peak-background split.

The two-point correlation function in the peak-background split may be written as

$$\xi_{\rm pb}(x_{12},t) \;=\; \langle N_{\rm pk}(x_1,t)N_{\rm pk}(x_2,t)\rangle/\langle N_{\rm pk}\rangle^2, \tag{2.27}$$

for fixed Eulerian separation $x_{12} =\mid \boldsymbol{x}_1 - \boldsymbol{x}_2 \mid$, where

$$N_{\rm pk} \;=\; n_{\rm pk}[\nu_{\rm b}(\boldsymbol{r},0)] \times [1 + \nu_{\rm b}(\boldsymbol{r},t)\sigma_{\rm b}(t)], \tag{2.28}$$

where the factor $[1+\nu_{\rm b}\sigma_b(t)]$ comes from the transformation of volume elements from Lagrangian to Eulerian coordinates. Approximating equation 2.25 by $n_{\rm pk} \simeq n_0[1 + \alpha\nu_b(\boldsymbol{r},0)]$ BBKS show that equation 2.27 reduces to

$$\xi_{\rm pb}(t) \;=\; (\alpha + \sigma_b)^2\psi_{12} \;=\; \left\{[\xi_{\rm pb}(0)]^{1/2} + [\xi(t)]^{1/2}\right\}^2, \tag{2.29}$$

and they say that comparison with the numerical simulations of Davis *et al.* (1985) show that the second form given in equation 2.29 gives a good approximation to the evolution of clustering beyond the weak clustering limit assumed in its derivation, although it clearly breaks down when $\xi_{\rm pb}$ is negative. A modification of this treatment is used by Bardeen, Bond & Efstathiou (1987, BBE) to consider predicted cluster correlations in an array of fifteen models. BBE replace $\xi_{\rm pb}$ in equation 2.29 by a statistical peak-peak component, $\xi^{\rm stat}_{\rm pk-pk}$, given by the Fourier transform of the power spectrum $P^{\rm stat}_{\rm pk-pk}$, which is defined to be

$$P^{\rm stat}_{\rm pk-pk}(k) = \left\{[\langle\tilde{\nu}\rangle + (\langle\nu\rangle - \langle\tilde{\nu}\rangle)R_*^2k^2/3\gamma^2]^2\right.$$

$$+ \left.(15\langle xe\rangle^2 + 5\langle xp\rangle^2)R_*^4k^4/9\gamma^4\right\} P_\rho(k)/\sigma_0^2. \qquad (2.30)$$

In this equation P_ρ is the power spectrum of the density field and $\langle\nu\rangle$, $\langle\tilde{\nu}\rangle$, $\langle xe\rangle$ and $\langle xp\rangle$ are averages over the peaks selected by the threshold criterion of quantities describing their heights and shapes: e and p are the ellipticity and prolaticity of the peak respectively. BBE identify clusters as peaks which have collapsed by the present in the spherical collapse model – *i.e.* $\nu\sigma_0(R_s, \text{now}) > \delta_c$, where δ_c is in the range from 1.06 (corresponding to the turnaround of a spherical perturbation in an Ω=1 universe) and 1.69 (which corresponds to its collapse to a singularity): a smoothing length of $R_s = 5\ h^{-1}$ Mpc is used to select peaks in all models. This method has been used more recently by Holtzman & Primack (1993) to study cluster correlations in several CDM variants. It is unfortunate than Holtzman & Primack choose to compare their results with only one set of observations (the PHG sample, decontaminated by Olivier *et al.* 1992).

A very different approach to the study of the dynamical evolution of clustering is adopted by Coles (1989). The starting point for his work is the observation that the probability distribution function of a density perturbation field that obeys Gaussian statistics initially must develop a skew towards positive densities as it evolves, since it is bounded below at zero density. Following Coles (1989) we define a measure of skewness, $\aleph$, by $\aleph \equiv (\sigma_3^2/\sigma_2^3)^{1/2}$ and second-order perturbation theory (Peebles 1980, Coles 1989) enables one to estimate $\aleph \simeq 5$ for an initially Gaussian density field evolved to $\xi(0) \simeq 1$. Coles then seeks to model the correlations of the evolved peak distribution by the correlations of peaks in two non-Gaussian density fields bounded at zero and with $\aleph \simeq 5$: the two non-Gaussian models considered are the lognormal and χ^2 distributions. Coles concludes that the development of this (moderate) degree of skewness has little effect on the peak correlation function and suggests that the dynamical component of the cluster correlation function will be small compared to the statistical component. This investigation is, however, only made for two specific non-Gaussian distribution functions and a single

toy model for the correlation function of the underlying density field, so it is not clear how general conclusions may be drawn from it.

2.5.4 Numerical studies of cluster correlations

The past fifteen years have seen the widespread application of the methods of numerical simulation to the problems of cosmology. Their impact on the study of cluster correlations has, like that in most areas of cosmology, been somewhat mixed. The first serious N-body study was that reported by Barnes *et al.* (1983). This was somewhat inconclusive, since the number of particles used was so small (32^3 in the largest simulation and only 8000 in the smallest) that the size of the simulation volume (determined by identifying the correlation length of the particle distribution with that observed for galaxies - *i.e.* $\sim 5\,h^{-1}$ Mpc) was insufficient to produce many clusters as rich as those in the Abell/ACO catalogue, so that a quantitative comparison of theory with observation was impossible.

Clearly, the very rareness of rich clusters poses a problem for numerical simulators. Batuski *et al.* (1987, 1991) tried to overcome this by performing very large volume simulations (cubes of side $\sim 600 - 700\,h^{-1}$ Mpc), but their emphasis on simulation size left them with woefully inadequate resolution: the cell size in the simulations of Batuski *et al.* (1987) is $6\,h^{-2}$ Mpc and each particle has a mass of $\sim 10^{15} M_{\odot}$, which is the mass of a rich cluster.

A balance of some sort between the conflicting requirements of size and resolution is found in the work of White *et al.* (1987) on clusters in the standard CDM model but, like other authors, they have to employ a BBKS-inspired analytic prescription to produce a cluster distribution from their density field. More satisfactory simulations are becoming possible with improvements in computer capabilities and we shall discuss the current generation of simulations in Section 2.7.

2.6 CLUSTERING IN THE ZELDOVICH APPROXIMATION

In previous sections we have seen how to approximate the spatial correlation function of the distribution of nascent clusters using the peaks formalism and how to extend this formalism to include the dynamical component to the clustering *in linear theory*. This linear theory treatment is only strictly valid in the limit of weak statistical correlations and will not be accurate in situations where the density perturbations are sufficiently strong for linear theory to be inadequate. Studies of cluster correlations through N-body simulation face, as we have seen, difficulties resulting from the dynamic range inherent in the problem: the numerical simulations should ideally possess sufficient spatial resolution to produce plausibly galaxy-like objects, while the rarity of rich clusters means that a very large simulation volume is required to facilitate the computation of the correlation function with reasonable accuracy. Numerical modellers (*e.g.* Dalton *et al.* 1992) have been forced to make assumptions about the relation between the cluster distribution and the Mpc-scale mass distribution. Given this necessity, it is clearly preferable to implement such assumptions analytically, provided that one can employ an adequate analytical method for determining the Mpc-scale mass distribution. We do just that in the present work, using the Zeldovich Approximation to perform the evolution of the density field, within the framework devised by Bond & Couchman (1987, 1988; hereafter, collectively, BC). Their paper is an analytic *tour de force* but, unfortunately, it is plagued by typographical errors, as well as being concise to the point of terseness. In this section we describe the work of BC and employ the framework it presents in the calculation of the cluster correlation function. For a general discussion of the origin and properties of the Zeldovich Approximation the reader is referred to the marvellous review by Shandarin & Zeldovich (1989), while more recent applications in reconstruction analysis are discussed, for example, by Nusser & Dekel (1993) and Lachieze-Rey (1993a,b): in what follows we consider the Zeldovich Approximation only insofar as it is used in the method of Bond & Couchman.

2.6.1 The Zeldovich Approximation

Consider a three-dimensional space populated by a dissipationless fluid. In proper coordinates, the (Eulerian) position, $\boldsymbol{y}(\boldsymbol{r},t)$, of a particle at a time t is related to its initial (Lagrangian) position, $\boldsymbol{r}$, by

$$\boldsymbol{y}(\boldsymbol{r},t) \;=\; a(t)\boldsymbol{r} \;-\; \boldsymbol{D}(\boldsymbol{r},t), \tag{2.31}$$

where $a(t)$ is the cosmological scale factor, $\boldsymbol{D}(\boldsymbol{r},t)$ is the displacement field and the minus sign is a matter of convention. In linear theory (*e.g.* Peebles 1980; see also Section 1.4), $\boldsymbol{D}(\boldsymbol{r},t)$ can be written as the product of spatial and temporal parts, $\boldsymbol{D}(\boldsymbol{r},t) \equiv b(t)\boldsymbol{d}(\boldsymbol{r})$, so that equation 2.31 reads

$$\boldsymbol{y}(\boldsymbol{r},t) \;=\; a(t)\boldsymbol{r} \;-\; b(t)\boldsymbol{d}(\boldsymbol{r}), \tag{2.32}$$

where the function $b(t)$ depends on the value of the cosmological density parameter, Ω, and determines the evolution of density perturbations: an expression for the growing-mode solution for $b(t)$ is given in the appendix to Zeldovich (1970). The Zeldovich approximation consists of the extrapolation of equation 2.32 into the regime where density perturbations are no longer small. From equation 2.32 we can see that the comoving peculiar velocity and acceleration vectors are parallel in the Zeldovich Approximation: particles move along straight trajectories in the Zeldovich Approximation and its utility will depend on whether or not this simplification compromises its ability to describe the evolution of cosmological density perturbations.

We may define comoving coordinates, $\boldsymbol{x}(\boldsymbol{r},t) = \boldsymbol{y}(\boldsymbol{r},t)/a(t)$, in which case equation 2.32 becomes

$$\boldsymbol{x}(\boldsymbol{r},t) \;=\; \boldsymbol{r} \;-\; \frac{b(t)}{a(t)}\,\boldsymbol{d}(\boldsymbol{r}). \tag{2.33}$$

Our concern here is with the distribution of matter at a specific epoch, the present, rather than with following the evolution of the density field with time, so it is convenient to

define the variable

$$\boldsymbol{s}(\boldsymbol{r}) \equiv \frac{b(t_0)}{a(t_0)}\,\boldsymbol{d}(\boldsymbol{r}), \tag{2.34}$$

where t_0 denotes the present epoch, from which it follows that

$$\boldsymbol{x} = \boldsymbol{r} - \boldsymbol{s}. \tag{2.35}$$

Given knowledge of the density perturbation field, equation 2.35 is all that is required to determine the distribution of particles, since the displacement field, $\boldsymbol{s}$, is related to the density perturbation field in a way we shall now show.

Consider a small region of space. The tensor of deformation, $\mathcal{D}_{ij}$, is defined by

$$\mathcal{D}_{ij} \equiv \frac{\partial x_i}{\partial r_j} = \delta_{ij} - \frac{\partial s_i}{\partial r_j}, \tag{2.36}$$

using equation 2.35 and where δ_{ij} denotes the Kronecker delta. The derivatives $\{\partial s_i/\partial r_j\}$ define a set of coordinate axes, such that the unit cube is transformed into a parallelepiped. If we select this particular coordinate system then the deformation tensor becomes

$$\mathcal{D} = \begin{pmatrix} 1-\alpha & 0 & 0 \\ 0 & 1-\beta & 0 \\ 0 & 0 & 1-\gamma \end{pmatrix}$$

where α, β, γ are the eigenvalues of the matrix $\{\partial s_i/\partial r_j\}$. The deformation tensor gives the transformation of the unit cube, so that the volume of space occupied by the parallelepiped into which the unit cube is deformed is simply the determinant of $\mathcal{D}$ – *i.e.* $||\,\mathcal{D}\,|| = (1-\alpha)(1-\beta)(1-\gamma)$. Conservation of mass then requires that

$$\rho(1-\alpha)(1-\beta)(1-\gamma) = \bar{\rho}, \tag{2.37}$$

where ρ and $\bar{\rho}$ are the local and mean densities, respectively. For small perturbations this yields the following relationship between the displacement field, $\boldsymbol{s}$, and the density perturbation field, δ,

$$\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}} = \nabla \cdot \boldsymbol{s}, \tag{2.38}$$

where the divergence of $\boldsymbol{s}$ is taken with respect to the Lagrangian coordinates $\boldsymbol{r}$. This may be expressed in terms of the Fourier components, $\{\boldsymbol{s}_k\}, \{\delta_k\}$, of the two fields as

$$(\boldsymbol{s}_k)_j = -i \frac{\delta_k}{k} (\hat{\boldsymbol{k}})_j, \tag{2.39}$$

where $\hat{\boldsymbol{k}} = \boldsymbol{k} / \mid \boldsymbol{k} \mid$ and $(\boldsymbol{k})_j$ is the jth component of the vector $\boldsymbol{k}$, etc.

The Zeldovich Approximation, as outlined above, breaks down when caustics form – *i.e.* $\rho \rightarrow \infty$ locally. After this time it gives an increasingly poor description of the density field, as it leads to 'shell-crossing', with particles passing through each other, rather than coalescing, resulting in an underestimation of the density of collapsed objects. This may be remedied by introducing a viscosity term into the dynamical equations, to produce 'sticky' particles: this is the adhesion model, for a discussion of which the reader is referred to Williams *et al.* (1991) (and references therein, particularly Gurbatov *et al.* 1989). The advantage afforded by the introduction of 'sticky' particles is outweighed, in the present context, by the impossibility of incorporating the adhesion model into a straightforward analytic framework like that outlined here for the Zeldovich Approximation.

An alternative approach to solving the problem of 'shell-crossing' in the Zeldovich Approximation is to apply a filter to the density field, to remove small-scale power and to prevent 'shell-crossing'. It has been shown (Efstathiou & Silk 1983; Efstathiou 1990; Coles, Melott & Shandarin 1993) that, with the inclusion of such a filter mechanism, the Zeldovich Approximation can provide a remarkably accurate picture of the evolution of the density field well into the nonlinear regime, as judged by comparison with the results of N-body simulations using the same initial conditions. Coles *et al.* (1993) favour the simple truncation of the power spectrum at a particular wavenumber to remove small-scale power, while BC prefer smoothing with a Gaussian filter. We adopt the latter

approach here, since a sharp truncation in k-space can introduce undesirable ringing in the correlation function: we discuss the choice of smoothing scale for the Gaussian filter in Section 2.6.2.

In addition to this direct illustration of the efficacy of the Zeldovich Approximation in this application, through comparison with N-body simulations, BC and Coles (1990) have shown that the Zeldovich Approximation reproduces the quadratic nonlinear terms in the full dynamical power spectrum in the mildly nonlinear regime appropriate to analysis through second order perturbation theory.

Given that the Zeldovich Approximation is rooted firmly in linear theory, one might be surprised at how well it can describe the mildly nonlinear evolution of the cosmological density field, but this can be readily explained through a consideration of the properties of the collapsing density perturbations involved. It can be shown that proto-objects in Gaussian density fields are generically triaxial (Peacock & Heavens 1985; BBKS) and that gravitation acts to accentuate that deviation from spherical symmetry (Lin, Mestel & Shu 1965), leading to collapse along one axis first (Zeldovich 1970), with subsequent particle motion being preferentially directed parallel to this axis. While the details of this procedure depend on the nature of the cosmological density field, especially the slope of its power spectrum, it will be generally true that gravitational collapse in structure formation is well-approximated by a one-dimensional process and, hence, that the Zeldovich Approximation will give a good description of the evolution of the density field well into the nonlinear regime, for a wide range of cosmological models (Coles *et al.* 1993).

2.6.2 Mass correlations in the Zeldovich Approximation

The first application of the Zeldovich Approximation to be considered is in the determination of the correlation function of a density field with known linear power spectrum. We require knowledge of the mass correlation function as part of the procedure we adopt for determining the smoothing length of the filter that we must apply to the density field to prevent shell crossing. It is also instructive to consider this first part of the BC formalism (the evolution of the density field) before the addition of the further complication of

the prescription to relate the cluster distribution to the density field. The mass correlation function, $\xi_\rho(x)$, is related to the probability, $P(x\,|1,2)$, of finding masses at points 1,2 a distance x apart by

$$1 + \xi_\rho(x) = P(x\,|1,2). \tag{2.40}$$

This may be written as

$$1 + \xi_\rho(x) = \int \mathrm{d}^3\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{s}\,P(x\,|\boldsymbol{r},\boldsymbol{s})\,P(\boldsymbol{s}\,|\boldsymbol{r})\,P(\boldsymbol{r}), \tag{2.41}$$

where $P(x|\,\boldsymbol{r},\boldsymbol{s})$ is the conditional probability that two particles have an Eulerian separation x, given that their Lagrangian separation is $\boldsymbol{r}$ and that their relative displacement through subsequent motion is $\boldsymbol{s}$; $P(\boldsymbol{s}|\boldsymbol{r})$ is the conditional probability that two particles have a relative displacement $\boldsymbol{s}$ given a Lagrangian separation $\boldsymbol{r}$; $P(\boldsymbol{r})$ is the probability of there being two particles with Lagrangian separation $\boldsymbol{r}$; and the integral is taken over all $\boldsymbol{r}$ and $\boldsymbol{s}$.

This method would calculate the correlation function for a distribution of particles that had been perturbed in any way that produces a Gaussian displacement field: the Zeldovich Approximation is required only for relation of the displacement field to the initial density field. From equation 2.33 we see that the conditional probability $P(x\,|\boldsymbol{r},\boldsymbol{s})$ taking the simple form $P(x|\boldsymbol{r},\boldsymbol{s}) = \delta_{\mathrm{D}}(x-|\boldsymbol{r}-\boldsymbol{s}|)/(4\pi x^2)$, where δ_{D} denotes the Dirac δ-function.

We give here only a sketch of how the derivation of the final form of $\xi_\rho(x)$ proceeds from this point, reserving the full derivation for Appendix C.

The probability $P(\boldsymbol{s}|\boldsymbol{r})$ has a Gaussian distribution:

$$P(\boldsymbol{s}\,|\,\boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}\,||\,\mathbf{M}\,||^{1/2}} \exp\left(\frac{-s_i M_{ij}^{-1} s_j}{2}\right), \tag{2.42}$$

where the covariance matrix of the distribution is $M_{ij} = \langle s_i(\boldsymbol{y};\boldsymbol{r}) s_j(\boldsymbol{y};\boldsymbol{r})\rangle$, where $s_i(\boldsymbol{y};\boldsymbol{r}) \equiv s_i(\boldsymbol{y}+\boldsymbol{r}) - s_i(\boldsymbol{y})$ and $\{s_i\}$ are the components of the displacement field.

This may be readily evaluated in a Cartesian coordinate system whose 3-axis lies along the direction of the vector $\hat{\boldsymbol{r}} = \boldsymbol{r}/\,|\,\boldsymbol{r}\,|$, using the relationship between the Fourier

components of the displacement and density fields (equation 2.39) to obtain

$$M_{ij} = 2\sigma_s^2 \left[(1-\psi_\perp)\delta_{ij} + (\psi_\perp - \psi_{||})\hat{r}_i\hat{r}_j\right]. \tag{2.43}$$

The quantity σ_s is the one-dimensional rms particle displacement and is given in terms of the power spectrum, $\Delta^2(k)$, of the density field by

$$\sigma_s^2 = \frac{1}{3}\int_0^\infty \Delta^2(k)\, e^{-k^2R_f^2}\, \frac{dk}{k^3}, \tag{2.44}$$

and $\psi_{||}$ and $\psi_\perp$ are defined by

$$\sigma_s^2\psi_{||}(r) = \int_0^\infty \Delta^2(k)\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2\, e^{-k^2R_f^2}\, \frac{dk}{k^3}, \tag{2.45}$$

and

$$\sigma_s^2\psi_\perp(r) = \int_0^\infty \Delta^2(k)\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\left[\frac{1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2}{2}\right] e^{-k^2R_f^2}\frac{dk}{k^3}. \tag{2.46}$$

It follows that equation 2.42 may be written as

$$P(\boldsymbol{s}\mid\boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}(2\sigma_s)^{3/2}(1-\psi_\perp)(1-\psi_{||})^{1/2}} \cdot \exp\left\{-\left[\frac{(1-\psi_{||})s^2 - (\psi_\perp-\psi_{||})(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{4\sigma_s^2(1-\psi_\perp)(1-\psi_{||})}\right]\right\}. \tag{2.47}$$

Since $P(x|\boldsymbol{r},\boldsymbol{s})$ takes the form of a δ-function, the integration over $\boldsymbol{s}$ is restricted to the sphere at constant $|\,\boldsymbol{x}\,|$. In our coordinate system this means that $s^2 = r^2 + x^2 - 2xr\cos\theta$ and $\boldsymbol{s}\cdot\hat{\boldsymbol{r}} = r - x\cos\theta$. These may be substituted into the expression for $P(\boldsymbol{s}|\boldsymbol{r})$ and the resultant integrations over θ, ϕ performed, the latter one trivially. This leaves a one-dimensional integration (over r), which is straightforward to convert into the following form:

$$1+\xi_{\rho Z}(x) = \frac{1}{2\sqrt{\pi}\sigma_s^2 x}\int_0^\infty \frac{r^2dr}{(1-\psi_\perp)^{1/2}(\psi_\perp-\psi_{||})^{1/2}} \cdot \left\{F(u_+)\exp\left[-\frac{(r-x)^2}{4\sigma_s^2(1-\psi_{||})}\right] - F(u_-)\exp\left[-\frac{(r+x)^2}{4\sigma_s^2(1-\psi_{||})}\right]\right\}, \tag{2.48}$$

which is the final form for $\xi_{\rho Z}(x)$. In equation 2.48, $F(u)$ is Dawson's integral,

$$F(u) \equiv \int_0^u \mathrm{e}^{y^2-u^2}\,\mathrm{d}y, \tag{2.49}$$

and the arguments $u_\pm$ are given by

$$u_\pm \equiv \frac{(\psi_\perp - \psi_\parallel)^{1/2}}{2\sigma_s(1-\psi_\parallel)^{1/2}(1-\psi_\perp)^{1/2}}\left[r\,\frac{(1-\psi_\perp)}{(\psi_\perp-\psi_\parallel)} \pm x\right], \tag{2.50}$$

which differs slightly from the equivalent equation in BC, which contains typographical errors.

Equations 2.44 – 2.46 contain a Gaussian smoothing term ($\mathrm{e}^{-k^2R_\mathrm{f}^2}$) to filter small-scale power from the power spectrum $\Delta^2(k)$. The choice of the filter radius, R_f, for use with a particular power spectrum is motivated by several considerations: if R_f is too small, then shell crossing will not be suppressed; if R_f is too large, it will filter out real power at short wavelengths and produce a two-point function that underestimates the strength of correlations on small scales. This behaviour can clearly be seen in Figure 2.2, which shows the mass correlation function for a CDM model in the Zeldovich Approximation, with the linear power spectrum smoothed with Gaussian filters with different smoothing lengths, R_f. The largest smoothing length of $R_\mathrm{f} = 5\ h^{-1}$ Mpc filters out real small-scale power: as the smoothing scale is reduced, the small-scale correlations increase, until a maximum is produced by a filter radius of $3\ h^{-1}$ Mpc. As R_f is reduced further, it is clear that shell crossing is occurring to a large extent since, with filter radii of $R_\mathrm{f} \leq 1\ h^{-1}$ Mpc, the correlations are reduced at small separations, due to the smearing out of structure as shells of particles pass through each other. Motivated by these considerations, we adopt the filter radius which maximises $\xi(0)$ for each of the models we consider.

When calculating the cluster correlation function, however, consideration of the motion of the peaks in the density field leads us to modify this procedure and to impose the constraint that R_f cannot be smaller than the filter length, R_s, to be used to define the peaks. This follows from seeing that if the peaks are defined by smoothing the density field on some scale R_s then the motion of those peaks can only be due to modes with wavelengths longer than R_s and so, whenever the procedure outlined in the preceding paragraph gives $R_\mathrm{f} < R_\mathrm{s}$, we set R_f equal to R_s.

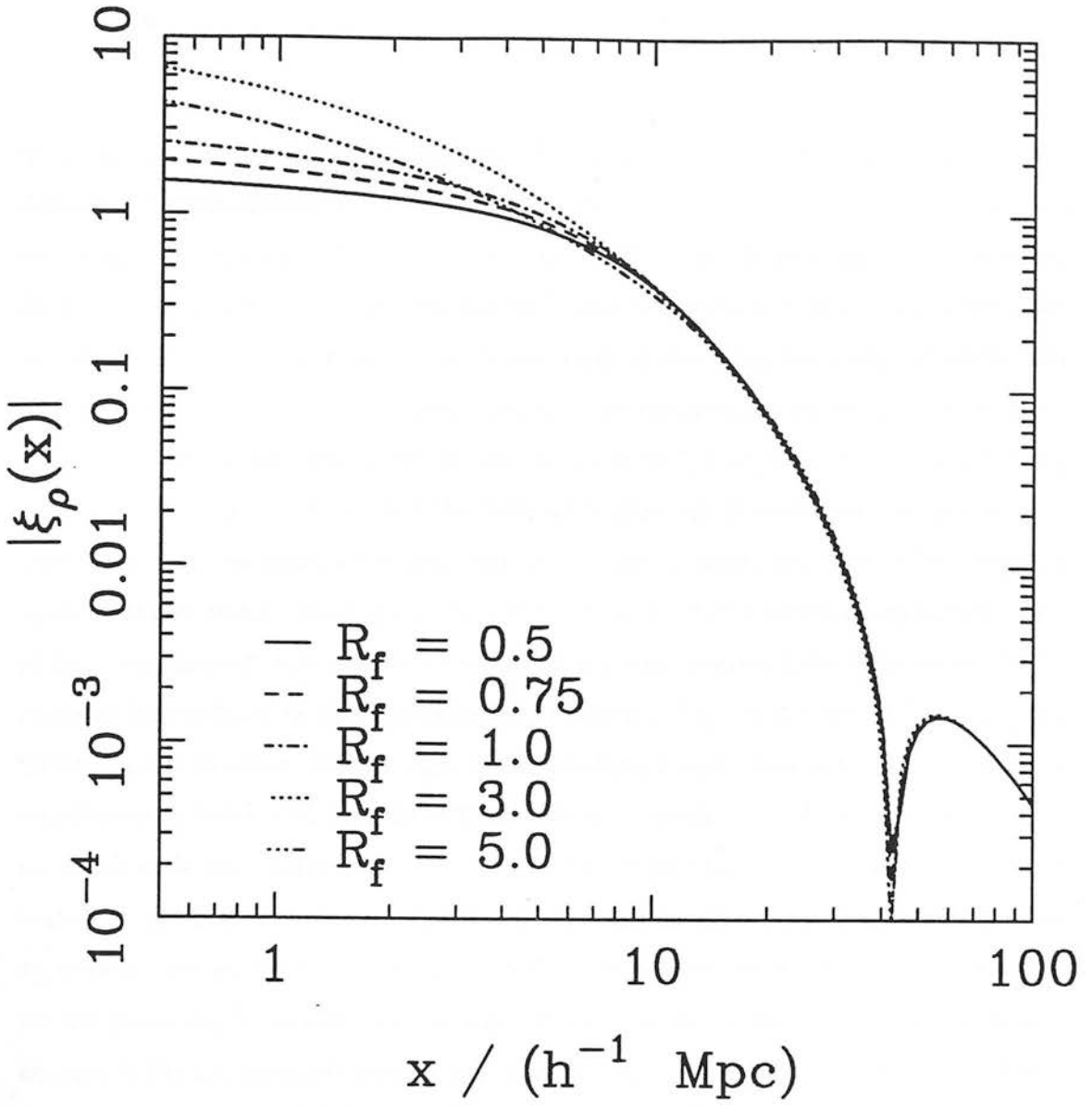


Figure 2.2: The mass correlation function for an $\Omega_0 = 1$, Γ=0.5 model with *COBE* amplitude, calculated for a number of values of the filter length, R_f, as indicated. (R_f in units of h^{-1} Mpc).

In Figure 2.3 we compare the mass correlation function in linear theory and in the Zeldovich Approximation for a CDM model, normalised to *COBE* and smoothed with Gaussians of filter length R_f=0.5, 1.0, 3.0 and 5.0 h^{-1} Mpc: from Figure 2.2 we see that Figure 2.3 (c) shows the correct smoothing length (R_f = 3.0 h^{-1} Mpc) to use with this power spectrum, as determined by our requirement of producing the maximal small-scale correlations in the Zeldovich Approximation. The relationship between the linear and Zeldovich mass correlation functions shown in Figure 2.3 (c) is generic: whenever the correct smoothing length is used the Zeldovich mass correlation function exceeds the linear prediction on small scales, dips below it on intermediate scales and then crosses it again on larger scales. The two correlation functions tend to a common asymptotic limit at large separations: this may be shown by expanding equation 2.48. It is instructive to compare Figure 2.3 with those for other power spectra. Figures 2.4 and 2.5 use the same CDM transfer function and the same four smoothing lengths, but with higher and lower amplitudes (ϵ twice and half the *COBE* values, respectively). From the variation of the small-scale mass correlations in these models we see that the appropriate smoothing scales for the higher, lower amplitude models are 5.0, 1.0 h^{-1} Mpc respectively. The agreement between linear theory and the Zeldovich Approximation is good over all scales for the lower amplitude, but for the larger dynamical evolution produced by the higher amplitude the agreement is poor on all scales. Figure 2.6 shows the corresponding figure for a power spectrum (Model 2 of Section 2.6.5, normalised to *COBE*) which has less small-scale power than CDM. This model has relatively more power at large scales, so that its correlation function has its first zero-crossing point at a significantly larger separation than the CDM models. One feature common to all four of these figures is that the first zero-crossing point of the mass correlation function is at a larger separation in the Zeldovich Approximation than in linear theory: this observation is important, since it shows that dynamical evolution can affect correlations out to surprising large separations.

2.6.3 Defining a population of objects

In Section 2.6.4 we shall use the asymptotic form of the peak–peak correlation function to normalize the prescription used by BC to relate the cluster distribution to the underlying

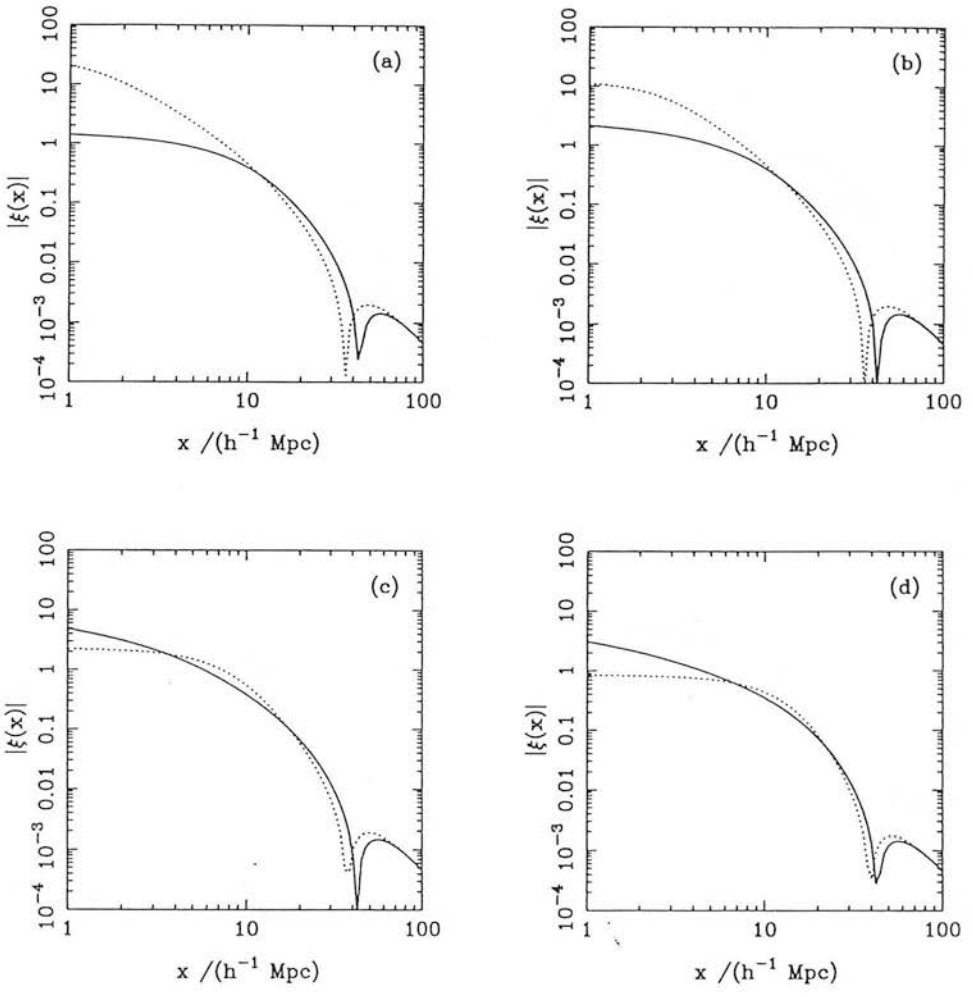


Figure 2.3: Comparison of mass correlation function in the Zeldovich Approximation and in linear theory, for a CDM model normalised to *COBE*. The smoothing lengths used are: (a) $R_{\rm f} = 0.5\ h^{-1}$ Mpc; (b) $R_{\rm f} = 1.0\ h^{-1}$ Mpc; (c) $R_{\rm f} = 3.0\ h^{-1}$ Mpc; and (d) $R_{\rm f} = 5.0\ h^{-1}$. In all four cases the solid line shows the mass correlation function calculated in the Zeldovich Approximation and the dotted line gives the linear theory prediction.

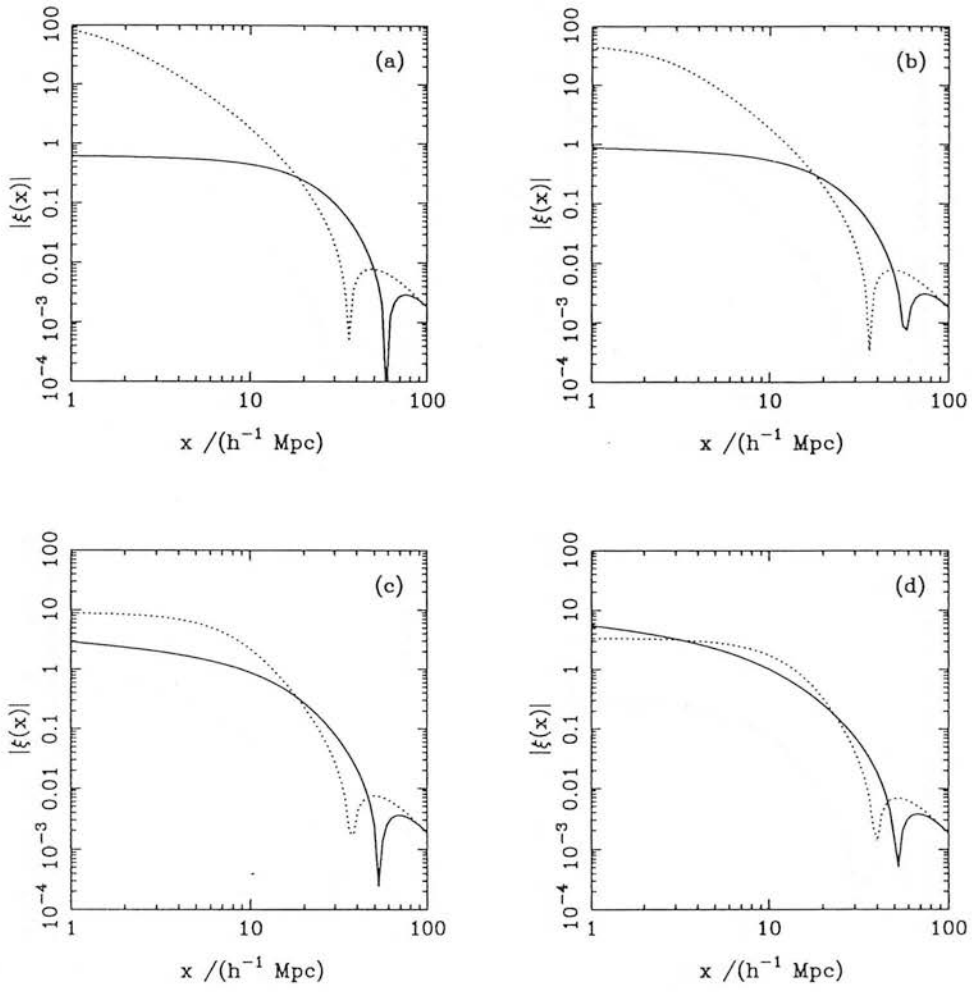


Figure 2.4: Comparison of mass correlation function in the Zeldovich Approximation and in linear theory, for a CDM model with a value of the normalisation parameter ϵ twice that for *COBE*. The smoothing lengths used are: (a) $R_{\rm f} = 0.5\ h^{-1}$ Mpc; (b) $R_{\rm f} = 1.0\ h^{-1}$ Mpc; (c) $R_{\rm f} = 3.0\ h^{-1}$ Mpc; and (d) $R_{\rm f} = 5.0\ h^{-1}$. In all four cases the solid line shows the mass correlation function calculated in the Zeldovich Approximation and the dotted line gives the linear theory prediction.

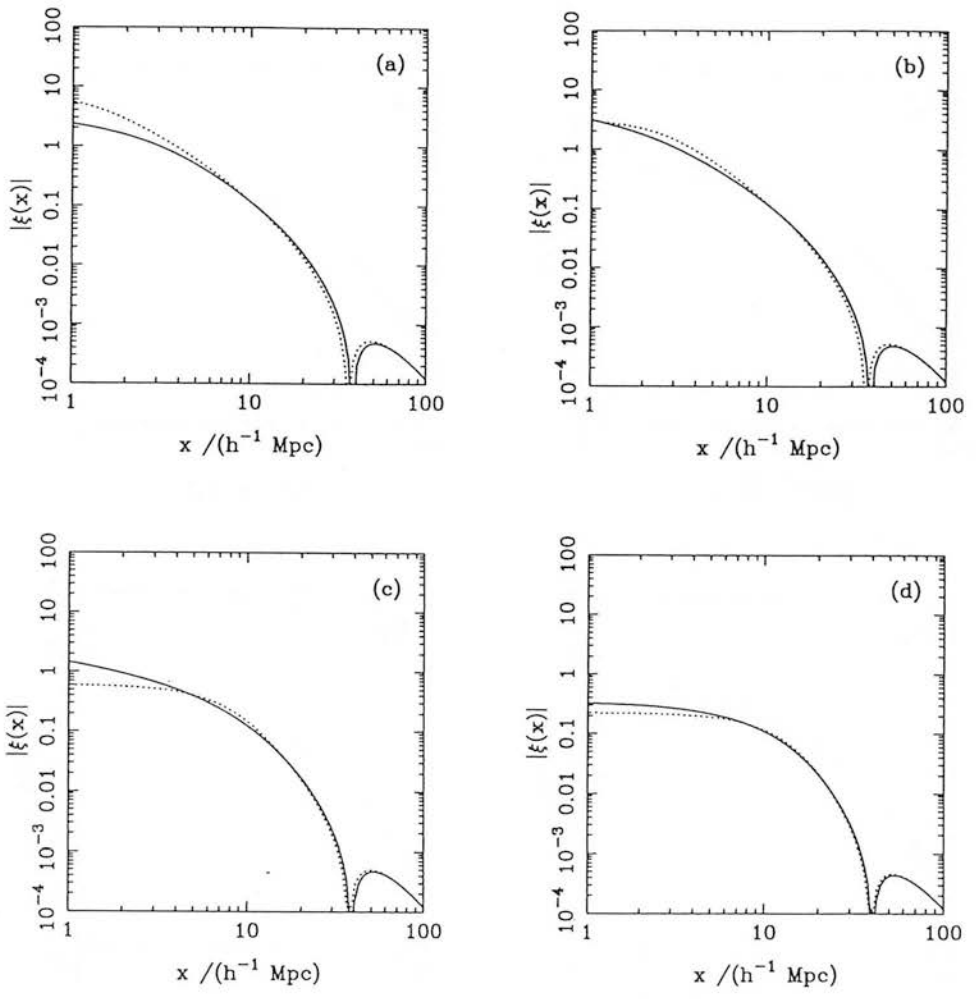


Figure 2.5: Comparison of mass correlation function in the Zeldovich Approximation and in linear theory, for a CDM model with a value of the normalisation parameter half that for *COBE*. The smoothing lengths used are: (a) $R_{\rm f} = 0.5\ h^{-1}$ Mpc; (b) $R_{\rm f} = 1.0\ h^{-1}$ Mpc; (c) $R_{\rm f} = 3.0\ h^{-1}$ Mpc; and (d) $R_{\rm f} = 5.0\ h^{-1}$. In all four cases the solid line shows the mass correlation function calculated in the Zeldovich Approximation and the dotted line gives the linear theory prediction.

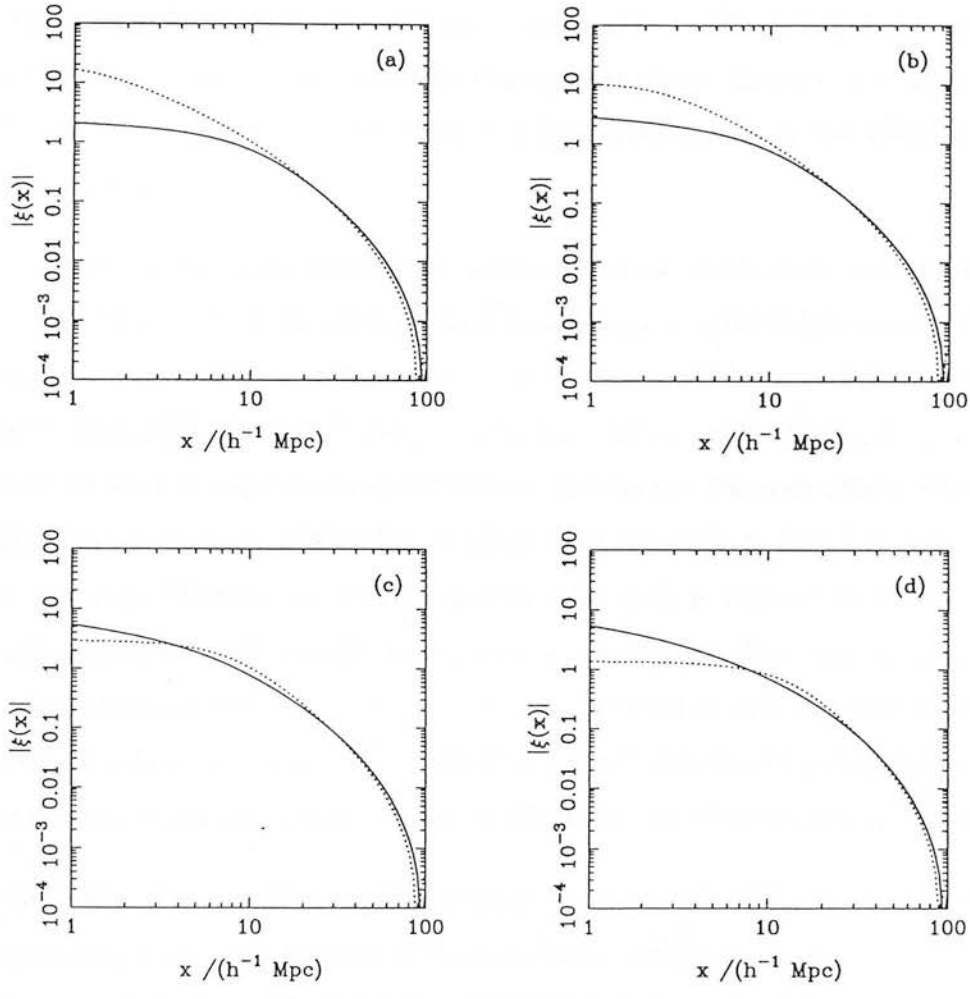


Figure 2.6: Comparison of mass correlation function in the Zeldovich Approximation and in linear theory, for Model 2 of Section 2.8, normalised to *COBE*. The smoothing lengths used are: (a) $R_{\rm f} = 0.5\ h^{-1}$ Mpc; (b) $R_{\rm f} = 1.0\ h^{-1}$ Mpc; (c) $R_{\rm f} = 3.0\ h^{-1}$ Mpc; and (d) $R_{\rm f} = 5.0\ h^{-1}$. In all four cases the solid line shows the mass correlation function calculated in the Zeldovich Approximation and the dotted line gives the linear theory prediction.

density field. To do this we require a procedure for calculation of the values of R_s and ν_s appropriate to the set of peaks which we identify with a population of clusters with number density N_c : we assume that the number density of clusters is conserved even though the individual cluster positions change over time. There is no unique way of making this identification. We employ a prescription based on the following simple collapse model.

Collapse models commonly identify the formation of an object with the time when the rms density contrast of the extrapolated local *linear* density field, smoothed on the appropriate mass scale, exceeds some critical value δ_c, which, in an Einstein–de Sitter universe, equals 1.7 for a spherical perturbation. Lower values ($1.7 \gtrsim \delta_c \gtrsim 1$) will be appropriate for a triaxial proto-object (More, Heavens & Peacock 1986), which is the generic case for peaks in a Gaussian random field (Peacock & Heavens 1985; BBKS). It is found (*e.g.* White *et al.* 1992) that the value of δ_c is very insensitive to Ω_0 for a flat universe, so we shall consider the range $1.7 \geq \delta_c \geq 1$ for all models in what follows. An independent constraint on the value of δ_c to be used is that this procedure should produce sensible cluster mass estimates: this point is discussed further in Section 2.9, with reference to the comparison of our results with those from numerical simulations.

It is clear that the first objects of a given mass to form in this picture will be those corresponding to the highest peaks in the initial density field and that, as time progresses, the objects of that mass which collapse will correspond to increasingly low peaks in the initial density field. A lower bound to the threshold ν_s may be obtained, therefore, by equating it with the peak height corresponding to peaks collapsing today. This would clearly be an underestimate of the threshold for objects like galaxies which formed some time ago, but for clusters, which are the most massive objects to have collapsed by the present and which are observed to be dynamically young, it is reasonable to equate the threshold with the height of peaks collapsing today, and that is what we shall do here. This provides the first constraint on the values of R_s and ν_s, which may be written as

$$\nu_s \, \sigma_0(R_s) \; = \; \delta_c. \tag{2.51}$$

A second constraint comes from matching the number density of clusters, N_c, to $n_{pk}(\nu_s)$,

which is the number density of peaks in the initial density field, smoothed on scale R_s, which lie above the threshold ν_s, and which is given by

$$n_{\rm pk}(\nu_s) = \frac{1}{8\pi^2 R_*^3} \int_0^\infty {\rm d}x f(x) {\rm e}^{-x^2/2} {\rm erfc}\left[\frac{(\nu_s - \gamma x)}{\sqrt{2(1-\gamma^2)}}\right], \tag{2.52}$$

where $R_* = \sqrt{3}\,\sigma_1/\sigma_2$: a derivation of this formula is presented in Appendix B.

From equations 2.51 and 2.52 we obtain a unique solution for R_s and ν_s for a particular cluster number density and a given power spectrum: this prescription is very similar to that used by BBKS and BBE. We discuss the application of this procedure to power spectra of practical interest in Section 2.8.

2.6.4 Cluster correlations in the Zeldovich Approximation

In order to calculate the cluster correlation function, it is necessary to supplement our mechanism for evolution of the density field with a prescription for relation of the cluster distribution to the underlying density field. This involves, of course, a rather subjective choice, given the current level of understanding as to how galaxies and clusters form. Given such a prescription, the cluster correlation function may be written in terms of conditional probabilities in a form analogous to, and derived from, equation 2.41 :

$$\begin{aligned} 1 + \xi_c(x) &= \int {\rm d}^3\boldsymbol{r}\, {\rm d}^3\boldsymbol{s}\, {\rm d}\nu_1\, {\rm d}\nu_2\; P(x\,|\,\boldsymbol{r},\boldsymbol{s})\, P(\boldsymbol{r})\, P(1c\,|\,\nu_1) P(2c\,|\,\nu_2) \\ &\cdot\; P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)\, P(\nu_1,\nu_2\,|\,\boldsymbol{r}). \end{aligned} \tag{2.53}$$

In this equation, $P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)$ is the conditional probability that two points, where the smoothed density field takes values ν_1,ν_2 (in units of the rms perturbation, σ_0), should have a relative displacement $\boldsymbol{s}$, given that their initial separation was $\boldsymbol{r}$. The term $P(\nu_1,\nu_2\,|\,\boldsymbol{r})$ is the conditional probability that the smoothed density field should take values ν_1,ν_2 at points with initial separation $\boldsymbol{r}$. From the definition of conditional probabilities it follows that

$$P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)\, P(\nu_1,\nu_2\,|\,\boldsymbol{r}) \;=\; P(\boldsymbol{s},\nu_1,\nu_2\,|\,\boldsymbol{r}), \tag{2.54}$$

which has a Gaussian distribution. The term $P(c \mid \nu)$, the probability of there being a cluster at a point where the density field (smoothed on a scale larger than that used to define the peaks) takes a value ν, encapsulates our statistical prescription for cluster formation. As discussed in Section 2.5 above, BBKS and others have studied models where objects form in thresholded regions or at peaks in the linear density field. Some authors have argued, however, that the physics of structure formation is more likely to produce a smooth non-linear relation between the mass density and the number density of objects, rather than a sharp threshold (*e.g.* Kaiser & Davis 1985; BBKS; Szalay 1988; Coles & Jones 1991). In particular, Kaiser & Davis (1985) suggest an exponential relationship between the number density of objects and the density field. This is an approximation to the peak-background split formula of equation 2.25, which tends to it in the limit that $R_{\rm f} \gg R_{\rm s}$. It gives the correct form for the statistical contribution to the correlation function at large separations (see below) and is very convenient computationally, so we follow BC in employing it here, giving the form of $P(c \mid \nu)$ as

$$P(c \mid \nu) \;=\; \exp[(E-1)\sigma_0\nu]\exp[-(E-1)^2\sigma_0^2/2], \tag{2.55}$$

where the second exponential term normalizes $P(c \mid \nu)$, so that it conserves mean density, and where the choice of the value of the enhancement factor, E, is considered below. The background smoothing scale used here is arbitrary, except for the constraint that it cannot be smaller than $R_{\rm s}$, so we take it to be equal to $R_{\rm f}$, the filter scale used to suppress shell crossing in the Zeldovich Approximation, which satisfies this constraint by definition.

The derivation of the final expression for the cluster correlation function from this point proceeds similarly to that for the mass correlation function and again we provide just a sketch of it here, reserving the full derivation for Appendix D.

The term $P(\boldsymbol{s}, \nu_1, \nu_2 | \boldsymbol{r})$ in equation 2.54 is given by

$$P(\boldsymbol{s}, \nu_1, \nu_2 | \boldsymbol{r}) \;=\; \frac{1}{(2\pi)^{5/2} \mid\mid \tilde{\mathbf{M}} \mid\mid^{1/2}} \exp\left(-\frac{1}{2}\tilde{y}_i^T \tilde{M}_{ij}^{-1} \tilde{y}_j\right), \tag{2.56}$$

where $\tilde{\boldsymbol{y}}^{\rm T} = (\boldsymbol{s}, \nu_1, \nu_2)$ and $\tilde{M}_{ij} = \langle \tilde{y}_i(\boldsymbol{z};\boldsymbol{r})\tilde{y}_j(\boldsymbol{z};\boldsymbol{r})\rangle$.

Using the same coordinate system as used in Section 2.6.2, the only non-vanishing matrix elements of the form $\langle \nu s_i \rangle$ are those involving the 3-component of the displacement field, which read

$$\langle \nu s_3 \rangle \;=\; -\frac{3\sigma_s^2}{\sigma_0}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right), \tag{2.57}$$

where $\psi_v \equiv (2\psi_\perp + \psi_{||})/3$ and σ_0 is the rms perturbation in the background density field. The matrix elements $\langle s_i s_j \rangle$ are as given above in equation 2.43, while the remaining elements of $\tilde{\mathbf{M}}$ are given by

$$\langle \nu_1 \nu_2 \rangle \;=\; \langle \nu_2 \nu_1 \rangle \;=\; \frac{\xi(r)}{\sigma_0^2} \;=\; \psi(r) \tag{2.58}$$

and

$$\langle \nu_1 \nu_1 \rangle \;=\; \langle \nu_2 \nu_2 \rangle \;=\; \psi(0) \;=\; 1, \tag{2.59}$$

where $\psi(r)$ is the normalised correlation function of the background density field.

The covariance matrix, $\tilde{\mathbf{M}}$, is of block-diagonal form, so there are no terms linking the sets $\{s_1, s_2\}$ and $\{s_3, \nu_1, \nu_2\}$ in the exponentiated quadratic form in the probability distribution $P\,(\boldsymbol{s}, \nu_1, \nu_2\,|\boldsymbol{r})$. So, to perform the integrations over ν_1, ν_2 in equation 2.56, we need only consider the reduced covariance matrix $\hat{\mathbf{M}}$ for the vector $\hat{\boldsymbol{y}}^{\mathrm{T}} = (s_3, \nu_1, \nu_2)$. It follows that the integral over ν_1, ν_2 takes the form

$$I_\nu \;=\; \frac{1}{(2\pi)^{3/2}\;||\,\tilde{\mathbf{M}}\,||^{-1/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{(E-1)\sigma_0(\nu_1+\nu_2)} \exp\left(-\frac{1}{2}\tilde{y}_i^T \tilde{M}_{ij}^{-1} \tilde{y}_j\right) \mathrm{d}\nu_1\, \mathrm{d}\nu_2. \tag{2.60}$$

Both integrations may be performed using the standard result (Gradshteyn & Ryzhik 1980, section 3.323)

$$\int_{-\infty}^{\infty} \exp(-p^2x^2 \pm qx)\,\mathrm{d}x \;=\; \frac{\sqrt{\pi}}{p}\exp\left(\frac{q^2}{4p^2}\right), \tag{2.61}$$

to obtain

$$\begin{aligned} I_\nu \;=\; & \frac{1}{\sqrt{4\pi\sigma_s(1-\psi_{||})}} \exp\left[\frac{-s_3^2}{4\sigma_s^2(1-\psi_{||})}\right] \exp\left[\frac{-3(E-1)s_3}{(1-\psi_{||})}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right)\right] \\ & \times \;\exp\left[(E-1)^2\sigma_0(1+\psi) \;-\; \frac{9\sigma_s^2(E-1)^2}{(1-\psi_{||})}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right)^2\right]. \end{aligned} \tag{2.62}$$

From this point, the derivation proceeds much as before to yield, after much tedious rearrangement, the following expression for ξ_{cZ}, the cluster correlation function in the Zeldovich Approximation:

$$1+\xi_{cZ}(x) = \frac{1}{2\sqrt{\pi}x\sigma_s^2}\int_o^\infty \frac{r^2\,dr\,\exp[(E-1)^2\xi(r)]}{(1-\psi_\perp)^{1/2}(\psi_\perp-\psi_{||})^{1/2}} \cdot \left\{F(\tilde{u}_+)\exp\left[\frac{-(x-\tilde{r})^2}{4\sigma_s^2(1-\psi_{||})}\right] - F(\tilde{u}_-)\exp\left[\frac{-(x+\tilde{r})^2}{4\sigma_s^2(1-\psi_{||})}\right]\right\}, \quad (2.63)$$

where

$$\begin{cases} \tilde{r} = r + 6(E-1)\sigma_s^2\,(d\psi_v/dr) \\ \tilde{u}_\pm = \frac{(\psi_\perp-\psi_{||})^{1/2}}{2\sigma_s(1-\psi_{||})^{1/2}(1-\psi_\perp)^{1/2}}\left[\tilde{r}\frac{(1-\psi_\perp)}{(\psi_\perp-\psi_{||})} \pm x\right] \end{cases}$$

and where $\psi_v(r) \equiv (2\psi_\perp + \psi_{||})/3$ and $\xi(r)$ is the correlation function of the smoothed linear density field. This also differs from the equivalent expression given by BC, due to the presence of typographical errors in that work.

The appropriate value of the enhancement factor, E, may be determined readily by matching the large-separation asymptotic limit of the statistical contribution to ξ_{cZ} to that of the statistical peak-peak correlation function. The one-dimensional rms particle displacement, σ_s, is a measure of the degree of dynamical evolution of the density field, so that the statistical contribution to ξ_{cZ} comes from considering the limit $\sigma_s \to 0$. It may readily be shown that

$$1 + \xi_{cZ}(x) \to \exp\left[(E-1)^2\xi(x)\right] \quad \text{as } \sigma_s \to 0, \quad (2.64)$$

in which case the large-separation asymptotic form of the statistical contribution to $\xi_{cZ}(x)$ is given by

$$\xi_{cZ}(x) \to (E-1)^2\xi(x) \text{ as } x \to \infty,\ \sigma_s \to 0. \quad (2.65)$$

Matching this to the large-separation asymptotic limit of the peak-peak correlation function derived by LHP

$$\xi_{\rm pk-pk}(r) \rightarrow \langle \tilde{\nu_{\rm s}} \rangle^2 \, \psi(r) \text{ as } r \rightarrow \infty, \tag{2.66}$$

yields the correct value of E to be used. From the results of LHP it is also possible to verify that the form of $P(c \mid \nu)$ given in equation 2.55 provides a good approximation to $\xi_{\rm pk-pk}(r)$ down to $r \sim R_{\rm s}$.

2.6.5 Models and Normalisation

We consider models in which the universe is spatially flat and in which the initial density fluctuations are Gaussian and have scale-invariant power spectra. Such models are consistent with the predictions of inflation and with the results of the *COBE* DMR experiment (Smoot et al. 1992), which we use to constrain the amplitude of the fluctuations at the present epoch.

The power spectrum of density fluctuations is given by

$$\Delta^2(k) = \frac{4\epsilon^2}{9} \left(\frac{ck}{H_0} \right)^4 T^2(k), \tag{2.67}$$

where $T(k)$ is the transfer function and the dimensionless normalization parameter, ϵ, is that defined by Peacock (1991). In Section 1.4.5 we showed how the value of ϵ could be determined from the *COBE* DMR results, upon the assumption that the contribution to these temperature anisotropies from gravitational radiation is negligible. We shall follow that procedure here and so consider fluctuation amplitudes in the range $\epsilon = (2.9 \pm 0.9) \times 10^{-5} \, \Omega^{-0.77}$ (Mann *et al.* 1993).

We follow Efstathiou, Bond & White (1992a, EBW) in adopting the following parametric form for the transfer function:

$$T(k) = \left\{ 1 + \left[a_1 k + (a_2 k)^{3/2} + (a_3 k)^2 \right]^{a_4} \right\}^{-1/a_4}, \tag{2.68}$$

where $a_1 = (6.4/\Gamma)h^{-1}$ Mpc, $a_2 = (3.0/\Gamma)h^{-1}$ Mpc, $a_3 = (1.7/\Gamma)h^{-1}$ Mpc and $a_4 = 1.13$. The origin of this parametric form is in fitting the linear power spectrum of CDM models (Bond & Efstathiou 1984) but, with suitable choices of the parameter Γ, it can fit the power spectra of a wide range of cosmological models to a reasonable accuracy. In retrospect, the definition of Γ chosen by EBW is slightly unfortunate. In pure CDM models, where Ω_B (the fraction of the critical density in baryons) is zero, $T(k)$ is a function of $k/\Omega h^2$ only, so one might interpret Γ as Ωh. For models where $\Omega_B \neq 0$, the appropriate scaling is found empirically to be roughly with $k/\Omega h^2 e^{-2\Omega_B}$ (Mann *et al.* 1993). The EBW transfer function for CDM is a fit to a model with $\Omega_B = 0.03$, which means that Γ differs from Ωh by about 6 per cent, which can lead to significant differences between a transfer function of the EBW form with a given value of Γ and, say, a BBKS CDM transfer function (for negligible baryonic content) which has Ωh equal to the same numerical value.

We consider the following four spatially flat models:

CDM: $\Omega_0 = 1$, $\lambda_0 = 0$, $\Gamma = 0.5$, (Model 1)

low density CDM: $\Omega_0 = 0.2$, $\lambda_0 = 0.8$, $\Gamma = 0.2$, (Model 2)

CDM + extra large-scale power: $\Omega_0 = 1$, $\lambda_0 = 0$, $\Gamma = 0.2$, (Model 3)

Bahcall & Cen model: $\Omega_0 = 0.2$, $\lambda_0 = 0.8$, $\Gamma = 0.1$, (Model 4)

where $\lambda_0 = \Lambda/3H_0^2$ and Λ is the cosmological constant. Model 2 is that advanced by Efstathiou, Sutherland & Maddox (1990) to account for the strong angular correlations seen in the APM galaxy catalogue: cluster-cluster correlations in this model were considered in linear theory by BBE (see Section 2.5). From EBW we see that one of the possible cosmogonies that could correspond to Model 3 (on these scales) is an $\Omega_0 = 1$, h=0.5, CDM-dominated universe in which one species of massive neutrino contributes $\Omega_\nu \simeq 0.3$ (van Dalen & Schaefer 1992): such a model is elaborated by Davis, Summers & Schlegel (1992) and Taylor & Rowan-Robinson (1992). Models 2 and 3 have the same transfer function, so any differences between them can only result from the difference in the fluctuation amplitude that they require to match the *COBE* observations: one

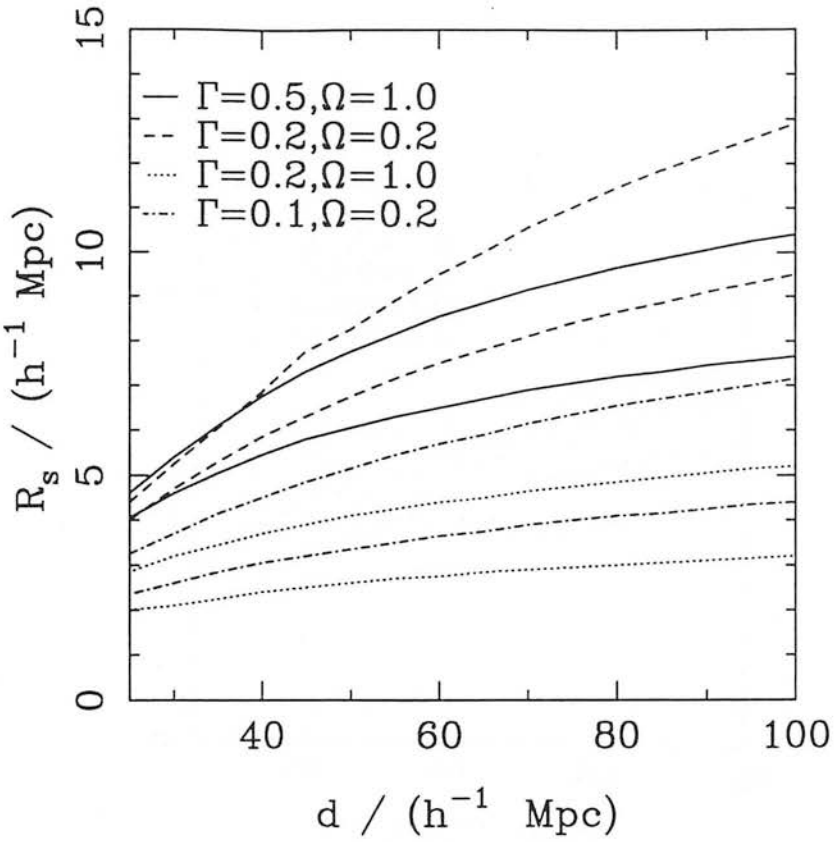


Figure 2.7: The variation of R_s with characteristic separation, d, for the four models we consider, normalized to *COBE*. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively. In all cases the upper line of the pair for each model is for δ_c=1.0 and the lower is for δ_c=1.7.

result of the higher amplitude of Model 2 is that the importance of the dynamical component of the clustering, relative to the statistical component, is greater than for Model 3. Model 4 is the transfer function which Bahcall & Cen (1992) claim is consistent with all cluster observations. The amplitude used by Bahcall & Cen is excluded by *COBE* , and we consider only the range of amplitudes that *COBE* allows. Bahcall & Cen (1992) claim that their cluster correlation function results are unchanged if this model is open or has a non-zero cosmological constant to make it spatially flat: we consider only the spatially flat model here.

In Figure 2.7 we show the variation of R_s with characteristic separation, d, that our procedure from Section 2.7.3 produces for the four models listed above. The differences in the value of R_s for a given separation, d, between the different models may be readily understood from a consideration of the relative amounts of small-scale power in their

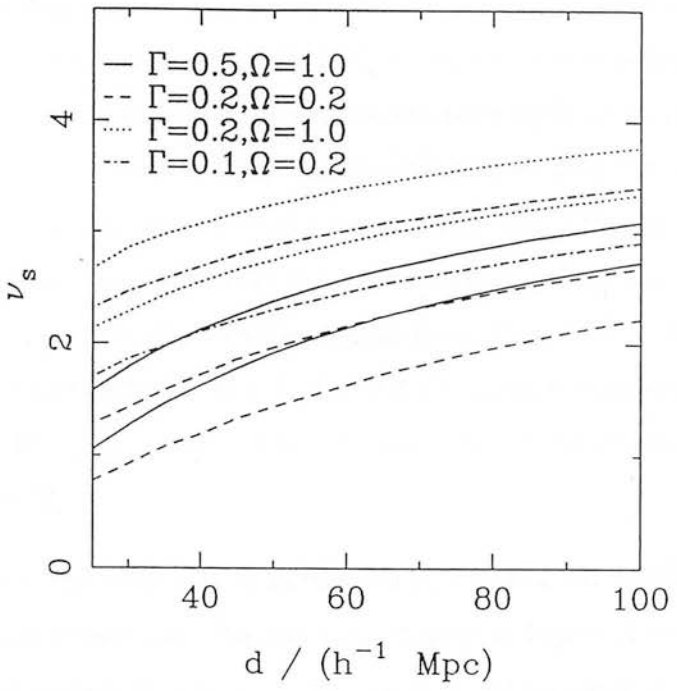


Figure 2.8: The variation of ν_s with characteristic separation, d, for the four models we consider, normalized to *COBE*. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively. In all cases the upper line of the pair for each model is for δ_c=1.7 and the lower is for δ_c=1.0.

respective power spectra. In all cases the filter radii correspond to reasonable mass estimates for, say, Abell $R \geq 1$ clusters, although we stress that the 'cloud-in-cloud' problem precludes the results shown in Figure 2.7 from being expressed as a cluster mass function. In Figure 2.8 we show the corresponding variation of ν_s versus d.

2.6.6 Sensitivity to parameters

In our formalism, the cluster correlation function for a given transfer function and cluster number density is a function of three parameters: the background smoothing scale, R_f; the linear density contrast at collapse, δ_c; and the amplitude of primordial fluctuations, ϵ. In the preceding sections we have discussed constraints on all these quantities, and now we consider the sensitivity of our results to variations in their values.

In Section 2.6.2 we discussed a procedure for determining the correct value of the smoothing scale, R_f, through the maximisation of the small-scale correlations of the density field. This is subject to a constraint that $R_f \geq R_s$, to ensure that excessive small-scale cluster correlations are not produced by power spectra with little small-scale power, resulting from our biasing prescription, $P(c \mid \nu)$. The effect of varying the value of R_f is, as one would naively expect, largely confined to separations of less than a few filter radii, where one must always be suspicious that the filter function will be dominant. For example, upon changing the value of smoothing length from $R_f = 0.5\ h^{-1}$ Mpc to $2.6\ h^{-1}$ Mpc, the correlation function for an $\Omega = 1$, $\Gamma = 0.2$ model only changes by more than 10 per cent for $r \leq 6\ h^{-1}$ Mpc. Our results are, therefore, quite robust to variations in the selected value of R_f.

The sensitivity to the value of δ_c is illustrated in Figure 2.10, which is discussed in more detail in the next subsection. We find that there is a degree of sensitivity to the value of δ_c for all four models that we consider, but it is not too serious: the difference in the correlation function for $\delta_c = 1.0$ and $\delta_c = 1.7$, for a given power spectrum and cluster number density is generally significant only on small scales, but is smaller than the error bars on the observational data. This sensitivity is a regrettable, but inevitable, feature of our formalism.

The sensitivity of our results to the value of the fluctuation amplitude, ϵ, is shown in Figure 2.9. This figure shows the correlation function for clusters of EMRS number density, calculated for the four models with the extreme allowed values of ϵ. It is clear that Model 3 is less sensitive to the fluctuation amplitude than are Models 1, 2 and 4. This is because Model 3 has less small-scale power and so its cluster correlations are dominated by the statistical contribution, rather than by the dynamical contribution, which means that its cluster correlation function will be less sensitive to the value of ϵ, which has a strong influence on the dynamical contribution to clustering, but not on the statistical one. The results in Figure 2.9 should be borne in mind when looking at the plots shown in Figures 2.14 – 2.17, which, for the sake of clarity, show the correlation function only for the *COBE* normalization and do not indicate the effects of uncertainty in the value of ϵ.

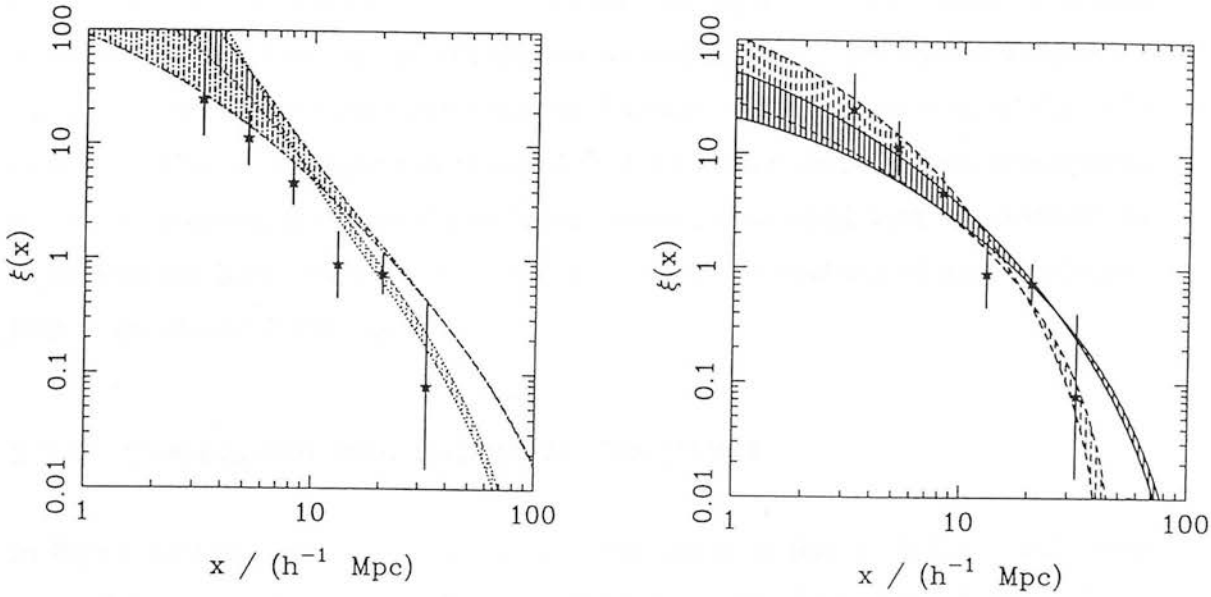


Figure 2.9: The effect on the correlation function of varying the amplitude of primordial fluctuations. The number density considered is that of the EMRS clusters of Nichol *et al.* (1992), from where the data points are taken. In (a) the dashed and solid lines give results for Models 1 and 2 respectively, while in (b) the dotted and dash-dotted lines are for Models 3 and 4 respectively: all models are for $\delta_c = 1.0$. In each case the hatched area indicates the region between the curves for the correlation function for the maximum and minimum value of the fluctuation amplitude ϵ.

2.7 COMPARISON WITH RECENT NUMERICAL RESULTS

Before comparing with observations the results of our computation of the cluster correlation function in Models $1-4$, we compare the results of three recent numerical simulations of the clustering of rich clusters with those determined by our method for the same power spectra and cluster number densities. It must be understood that this cannot be taken as a straight test of our method, since there are, of course, assumptions and approximations in numerical simulations just as in analytical work: approximations to the true non-linear evolution of the density field and assumptions relating the density field to the cluster distribution.

2.7.1 Comparsion with Bahcall & Cen (1992)

In Figure 2.10 we compare our method with the results of Bahcall & Cen (1992), who use a PM code with 250^3 particles in a simulation cube of side $400\,h^{-1}$ Mpc. They employ the BBKS CDM transfer function for zero baryons, which has the form

$$T(k) = \frac{\ln(1+2.34q)}{2.34q}\left[1+3.89q+(14.1q)^2+(5.46q)^3+(6.71q)^4\right]^{-1/4}, \qquad (2.69)$$

where $q \equiv k/(\Omega_0 h^2\ \mathrm{Mpc}^{-1})$. We perform a comparison with their results for $\Omega_0 = 1, h = 0.5$ and $\Omega_0 = 0.2, h = 0.5$ and we present results for the number densities of APM and Abell $R \geq 1$ clusters. Bahcall & Cen select clusters using an adaptive friends-of-friends algorithm which links pairs of particles using a density-dependent linking length. They appear to have run only one realisation of each power spectrum which, even for such a large simulation volume, must make one question the significance of their results.

2.7.2 Comparison with Croft & Efstathiou (1993)

Simulations with much better spatial resolution are described by Croft & Efstathiou (1993). They have used the $\mathrm{P^3M}$ code of Efstathiou *et al.* (1985) to follow the evolution of 10^6 particles in a simulation cube of side $300\,h^{-1}$ Mpc. They consider power spectra

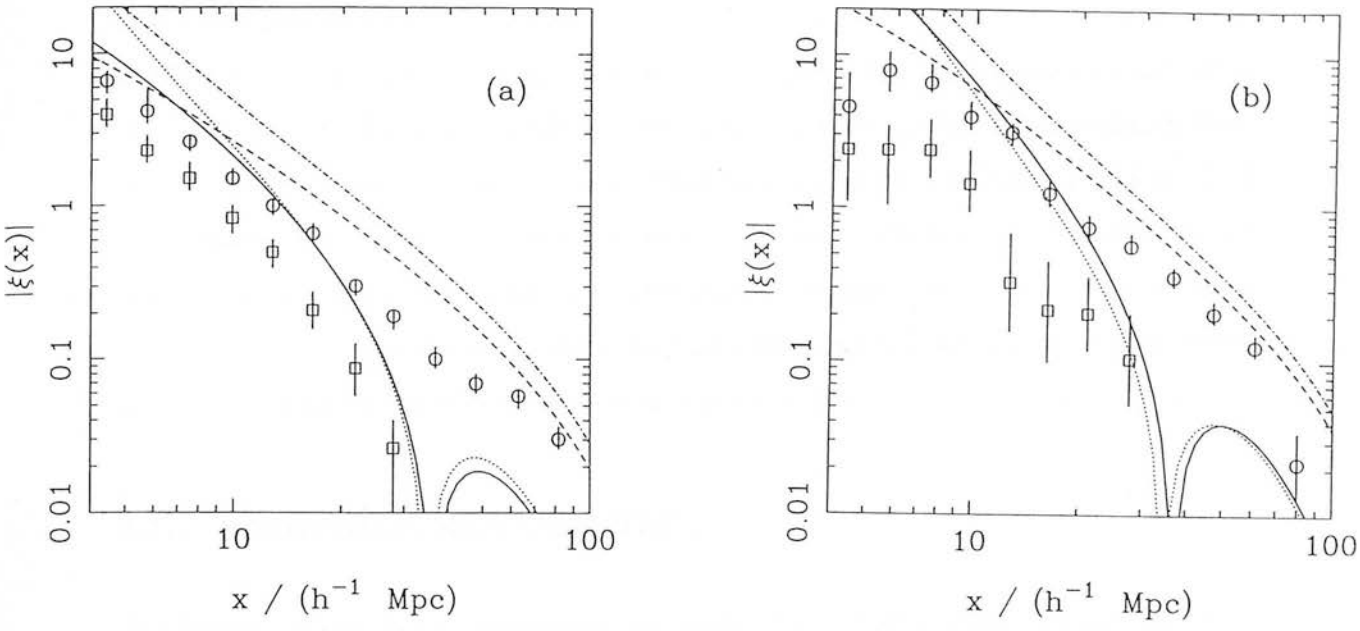


Figure 2.10: Comparison with numerical results from Bahcall & Cen (1992), for clusters with the number density of (a) APM clusters and (b) Abell $R \geq 1$ clusters. The solid (dotted) line shows the correlation function for an $\Omega_0 = 1, h = 0.5$ model with $\delta_c = 1.0$ ($\delta_c = 1.7$), while the dashed (dash-dotted) line shows that for an $\Omega_0 = 0.2, h = 0.1$ model with $\delta_c = 1.0$ ($\delta_c = 1.7$). The squares and circles are the numerical results of Bahcall & Cen for the $\Omega_0 h = 0.5$ and $\Omega_0 h = 0.1$ models respectively.

with $\Gamma = 0.2$ and $\Gamma = 0.5$, with amplitudes expressed in terms of σ_8 and run an ensemble of ten realisations for each model. The cluster selection algorithm employed by Croft & Efstathiou is designed to mimic that used by Dalton *et al.* in constructing the APM cluster survey (see Section 2.4), with the exception that it is applied in three dimensions, rather than in projection, as for the observational algorithm. In Figures 2.11 & 2.12 we compare their results (kindly provided by R. Croft) with those determined for the same power spectra and a variety of cluster number densities by our methods: we show results for $\sigma_8 = 1.0$, but a very similar level of agreement is found for $\sigma_8 = 0.59$, which was the other epoch for which the comparison was made.

2.7.3 Comparison with Cole (1993)

The final set of simulations we consider are those of Cole (1993, private communication). These were not performed for the purposes of studying the cluster-cluster correlation function and only two realisations of a CDM model with the BBKS zero-baryon transfer function have been run, evolved using a P^3M code, in a box of side 256 h^{-1} Mpc. Cole used a friends-of-friends algorithm with a fixed linking length equal to 20% of the mean interparticle separation to combine particles into clusters. The velocity dispersions of the clusters were then calculated and the top 100 and 200 clusters in rank order by velocity dispersion constitute the two samples shown in Figure 2.13: they have characteristic separations of $d = 44\,h^{-1}$ Mpc and $d = 55\,h^{-1}$ Mpc.

2.7.4 Discussion and Summary

Our comments here are principally concerned with the Cole and Croft & Efstathiou simulations, since they have much better resolution than those of Bahcall & Cen. From Figures 2.11 and 2.12 we can see that there is excellent agreement between the numerical and analytical results for low richness clusters: to $d = 40\,h^{-1}$ Mpc for the $\Gamma = 0.5$ model and up to $d = 50\,h^{-1}$ Mpc for $\Gamma = 0.2$. The differences between the two methods are more apparent for the richest clusters, where our analytic theory predicts stronger clustering than that found by Croft & Efstathiou. It is interesting to note, however, that the clustering strength found for the richest systems by Croft & Efstathiou is also

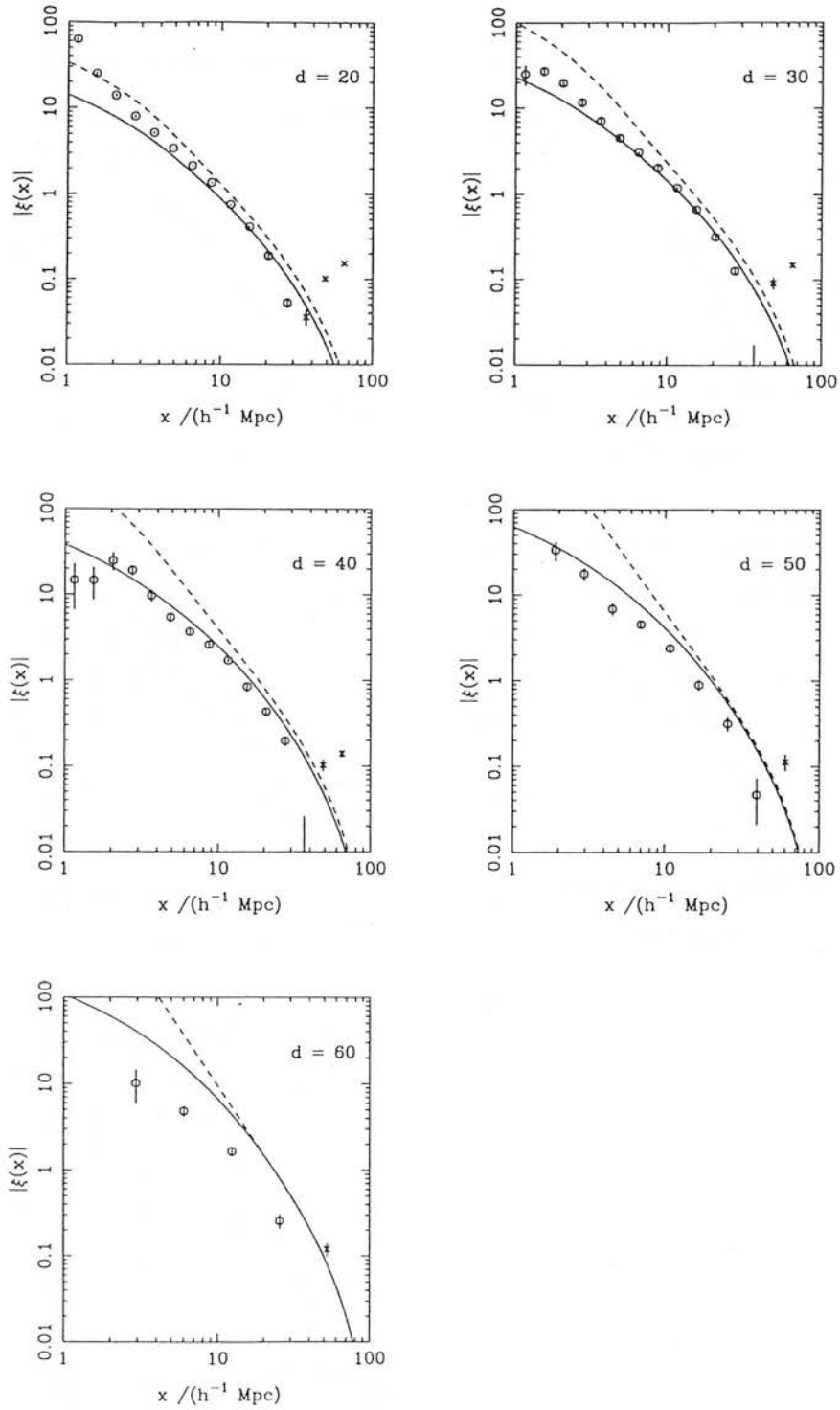


Figure 2.11: Comparsion with numerical results from Croft & Efstathiou (1993) for a $\Gamma = 0.2$ model with amplitude $\sigma_8 = 1.0$, for a number of different cluster number densities (d in units of h^{-1} Mpc). Circles denote positive values of the correlation function and crosses denote negative values. The error bars on the points are Poissonian. Solid and dashed lines are our predictions for $\delta_c = 1.0$ and $\delta_c = 1.7$ respectively.

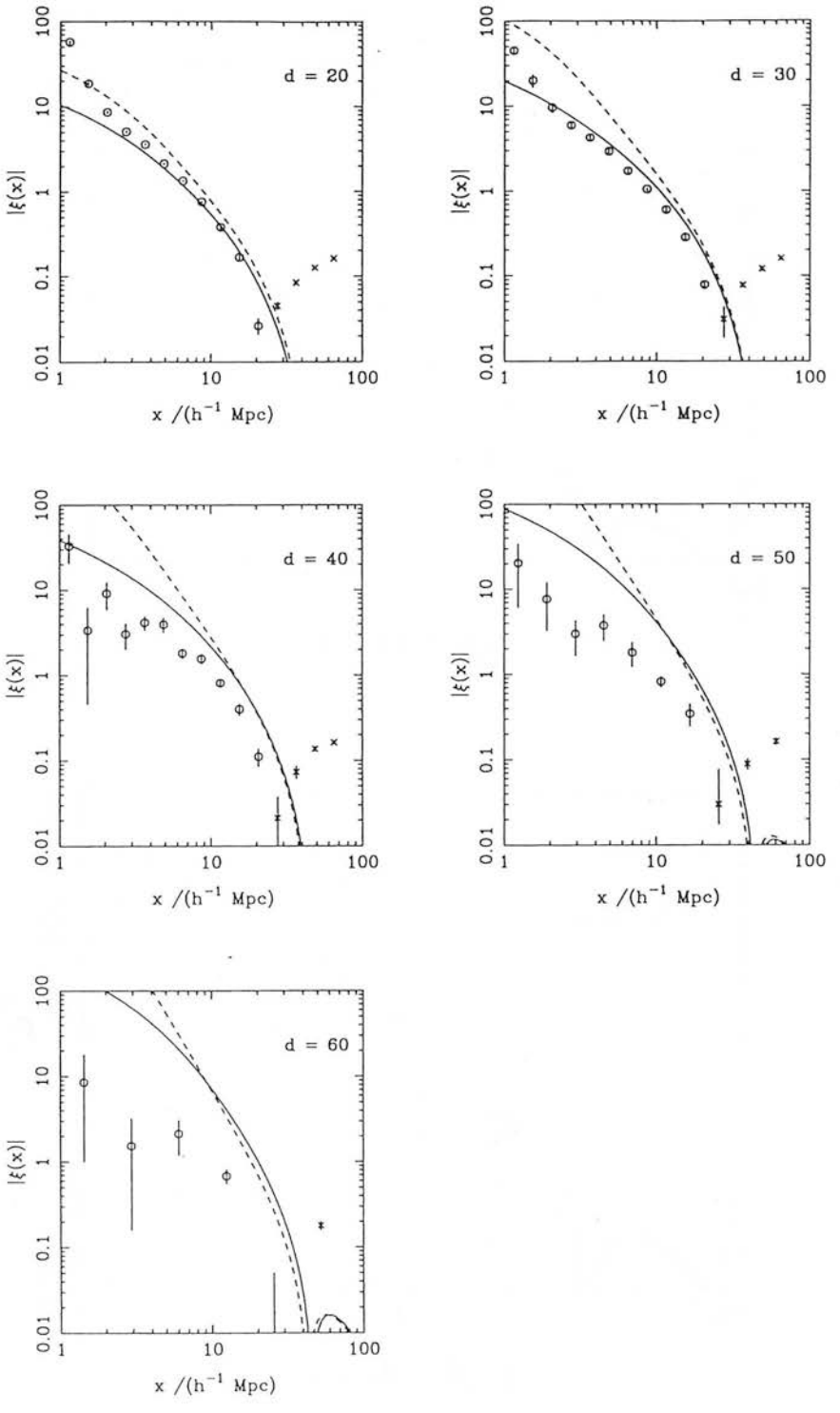


Figure 2.12: Comparsion with numerical results from Croft & Efstathiou (1993) for a $\Gamma = 0.5$ model with amplitude $\sigma_8 = 1.0$, for a number of different cluster number densities (d in units of h^{-1} Mpc). Circles denote positive values of the correlation function and crosses denote negative values. The error bars on the points are Poissonian. Solid and dashed lines are our predictions for $\delta_c = 1.0$ and $\delta_c = 1.7$ respectively.

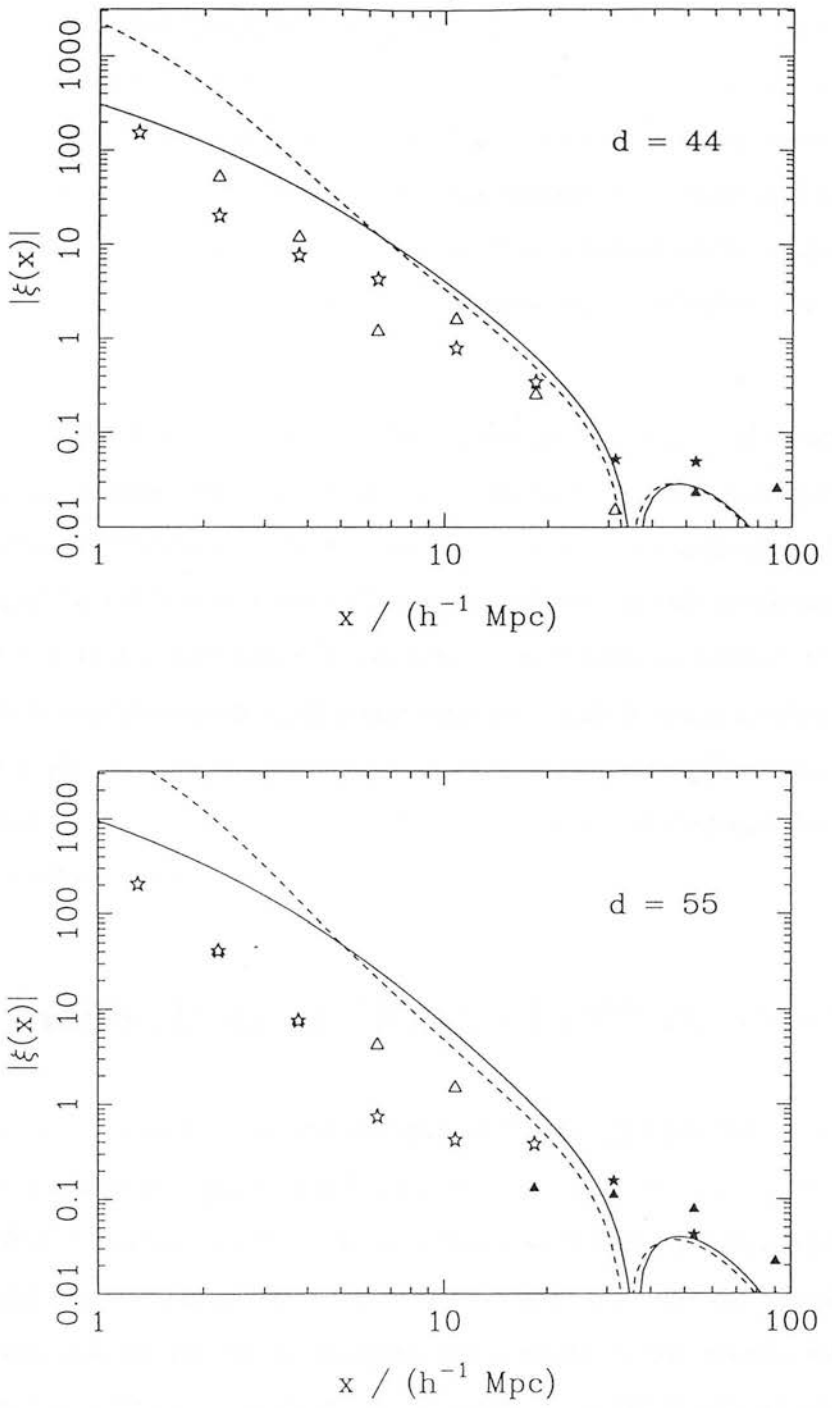


Figure 2.13: Comparison with numerical results from Cole (1993) for a $\Omega h = 0.5$ CDM model with amplitude $\sigma_8 = 0.63$, for cluster number densities $d = 44\,h^{-1}$ Mpc and $d = 55\,h^{-1}$ Mpc. Results from two realisations are shown: filled (empty) triangles show positive (negative) points from one realisation and filled (empty) squares show positive (negative) results from the other. Solid and dashed lines are our predictions for $\delta_c = 1.0$ and $\delta_c = 1.7$ respectively.

weaker than that found by Cole, although the scatter between the results from Cole's two realisations suggests that one must make such comparisons with caution. These results seem to indicate that differences between the selection criteria used to define the cluster samples significantly influence the spatial correlations deduced from them once the number density of the sample falls below a certain level. We eagerly await the results of other N-body studies of cluster correlations to see whether or not they confirm this assertion.

One striking difference between the numerical and analytical results is the respective strength of their large separation anticorrelations. Correlation functions derived from numerical simulations in finite sized boxes are self-normalised, which means that the integral $\int \xi(r) \mathrm{d}^3 r$ must vanish when evaluated over the full simulation volume. It follows from this that a correlation function from an N-body simulation which exhibits strong positive correlations on small scales *must* also exhibit compensating anticorrelations on large scales in order to satisfy the normalisation criterion. With this integral constraint in mind, it is to be expected that there is a significant discrepancy between the numerical and analytical results on large scales.

2.8 RESULTS: OPTICAL CLUSTER SAMPLES

We have computed the cluster correlation function for the four models for four cluster number densities, those for the EMRS and APM automated surveys and the Abell $R \geq 1$ and $R \geq 2$ samples of PW92: we note that PW92 urge caution in the use of their $R \geq 2$ sample, due to its small size. We do not consider Abell cluster samples including $R = 0$ clusters, nor do we choose to make comparisons of our results with observations to which the Sutherland method for correction for redshift anisotropy has been applied: the reasons behind these decisions are discussed in Section 2.4.

We show our results for the cluster correlation function in Figures 2.14 – 2.17. The observational points have 1σ Poisson error bars. Where no correlation is detected with significance at the 1σ level we plot a 90 per cent confidence upper limit, marked with an arrow. We compare theoretical predictions of configuration-space correlation functions

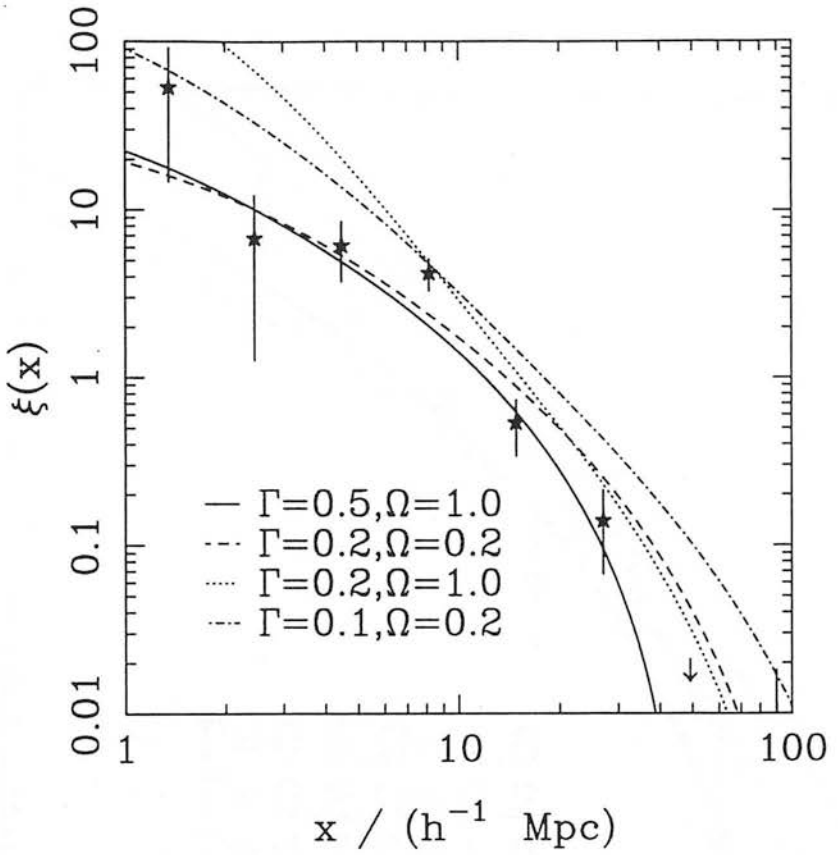


Figure 2.14: The cluster correlation function for APM $\mathcal{R} \geq 20$ clusters.The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively with *COBE* normalization and δ_c=1.0.

with observed results measured in redshift-space. Redshift-space corrections are considered by Kaiser (1987), who finds that the (isotropic) real-space correlation function, $\xi_r(x)$, is related to the direction-averaged redshift-space correlation function, $\xi_{\langle s \rangle}(x)$, by $\xi_{\langle s \rangle}(x) = [1+(2\Omega_0^{0.6}/3E)+(\Omega_0^{1.2}/5E^2)]\xi_r(x)$. For the models that we consider, this factor is a small fraction of the uncertainty in the *COBE* amplitude and much less than the errors in the observations with which we shall compare our results, so we neglect this correction. Figure 2.14 shows results for APM $\mathcal{R} \geq 20$ clusters, with number density $N_c = 2.4 \times 10^{-5}\ h^3\ \mathrm{Mpc}^{-3}$. The observational data are taken from Efstathiou *et al.* (1992b) and include clusters from a slightly wider redshift range than that considered by Dalton *et al.* (1992). In Figure 2.15 we give results for EMRS clusters, from Nichol *et al.* (1992), which have a number density of $N_c = 1.0 \times 10^{-5}\ h^3\ \mathrm{Mpc}^{-3}$. Figure 2.16 and 2.17 show correlation functions for the Abell $R \geq 1$ and $R \geq 2$ samples of PW92, which have number densities of 7.2×10^{-6} and $1.7 \times 10^{-6}\ h^3\ \mathrm{Mpc}^{-3}$ respectively.

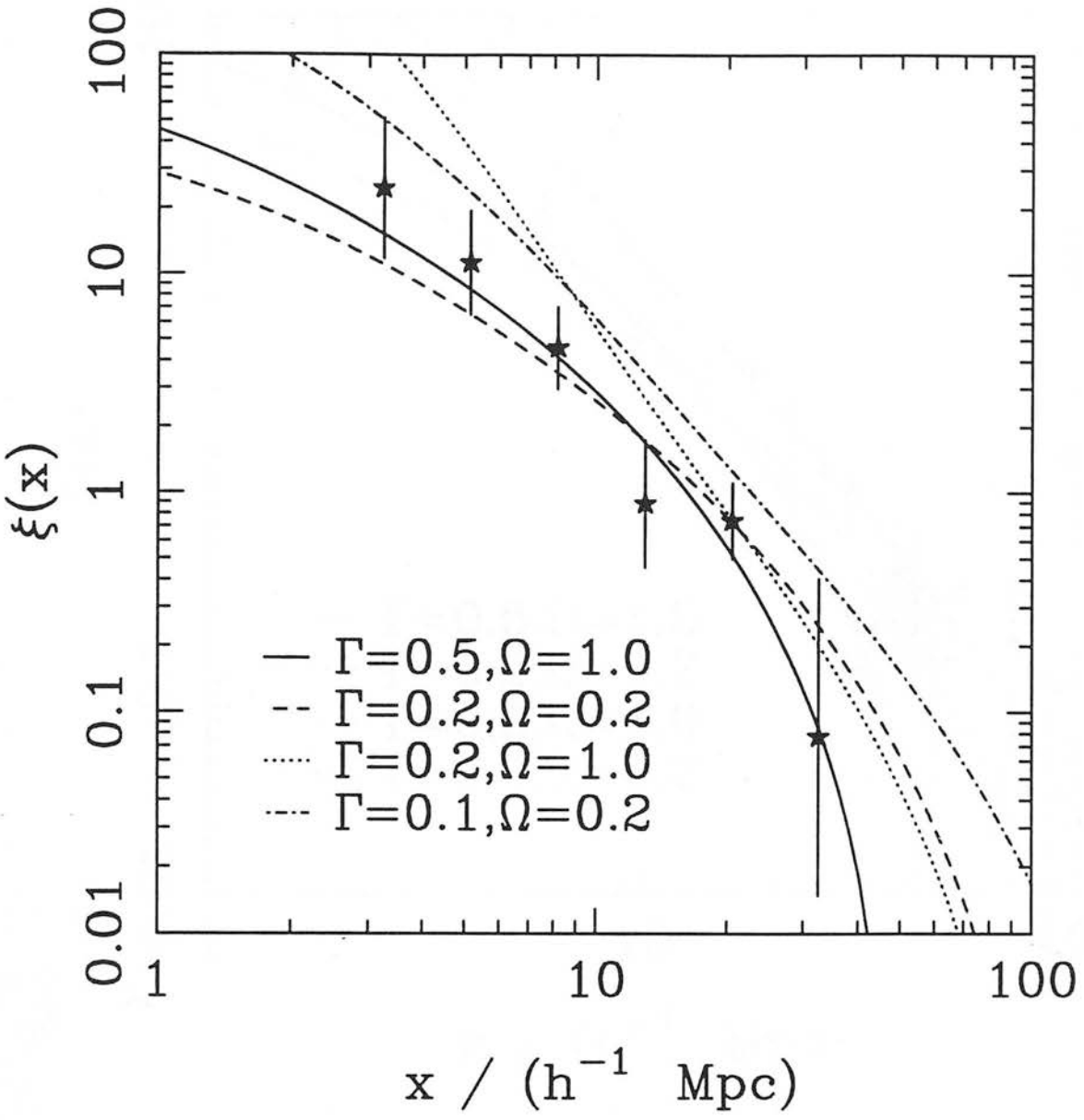


Figure 2.15: The cluster correlation function for EMRS clusters. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively with *COBE* normalization and δ_c=1.0.

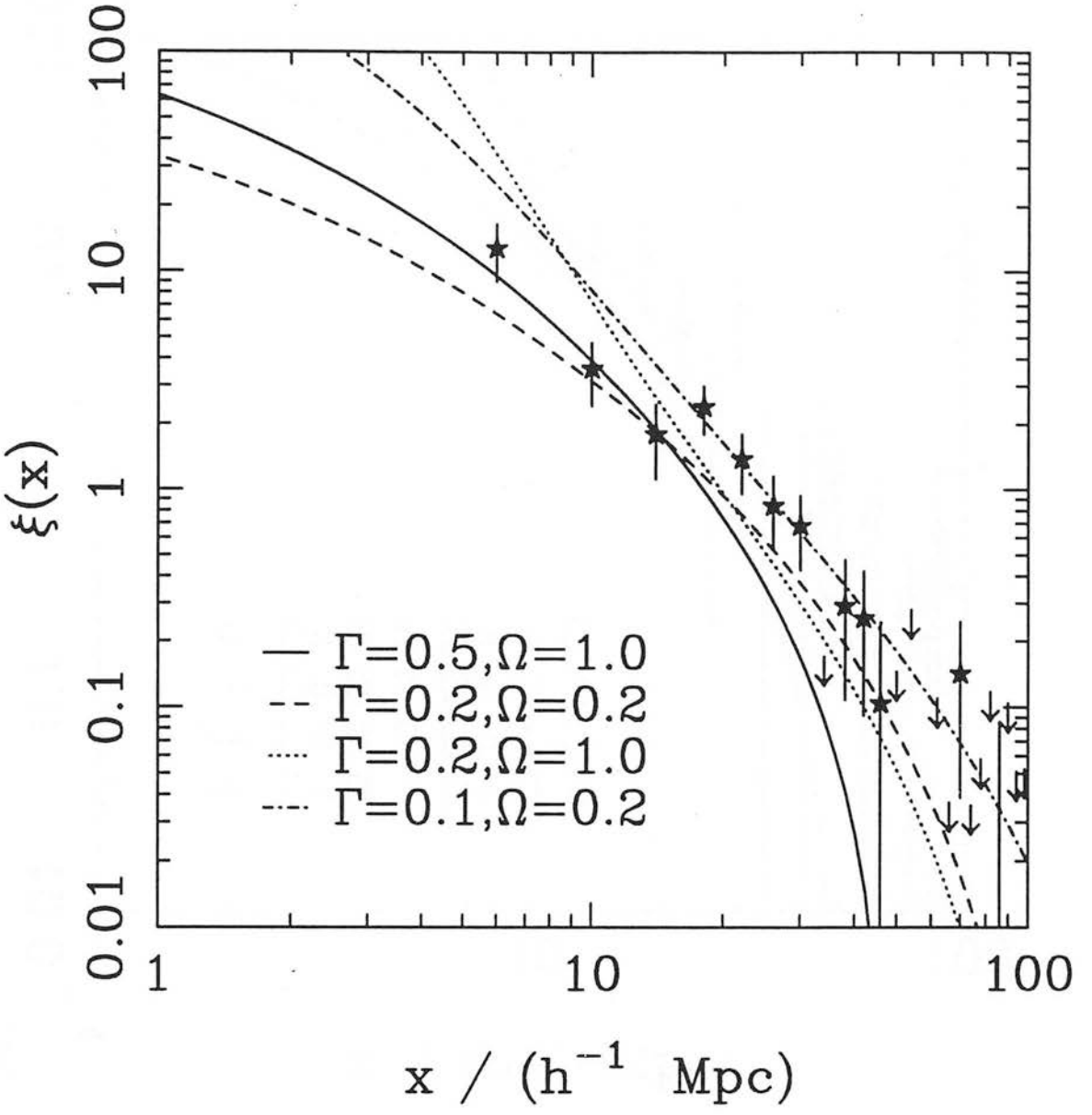


Figure 2.16: The cluster correlation function for Abell $R \geq 1$ clusters from PW92. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively with *COBE* normalization and δ_c=1.0.

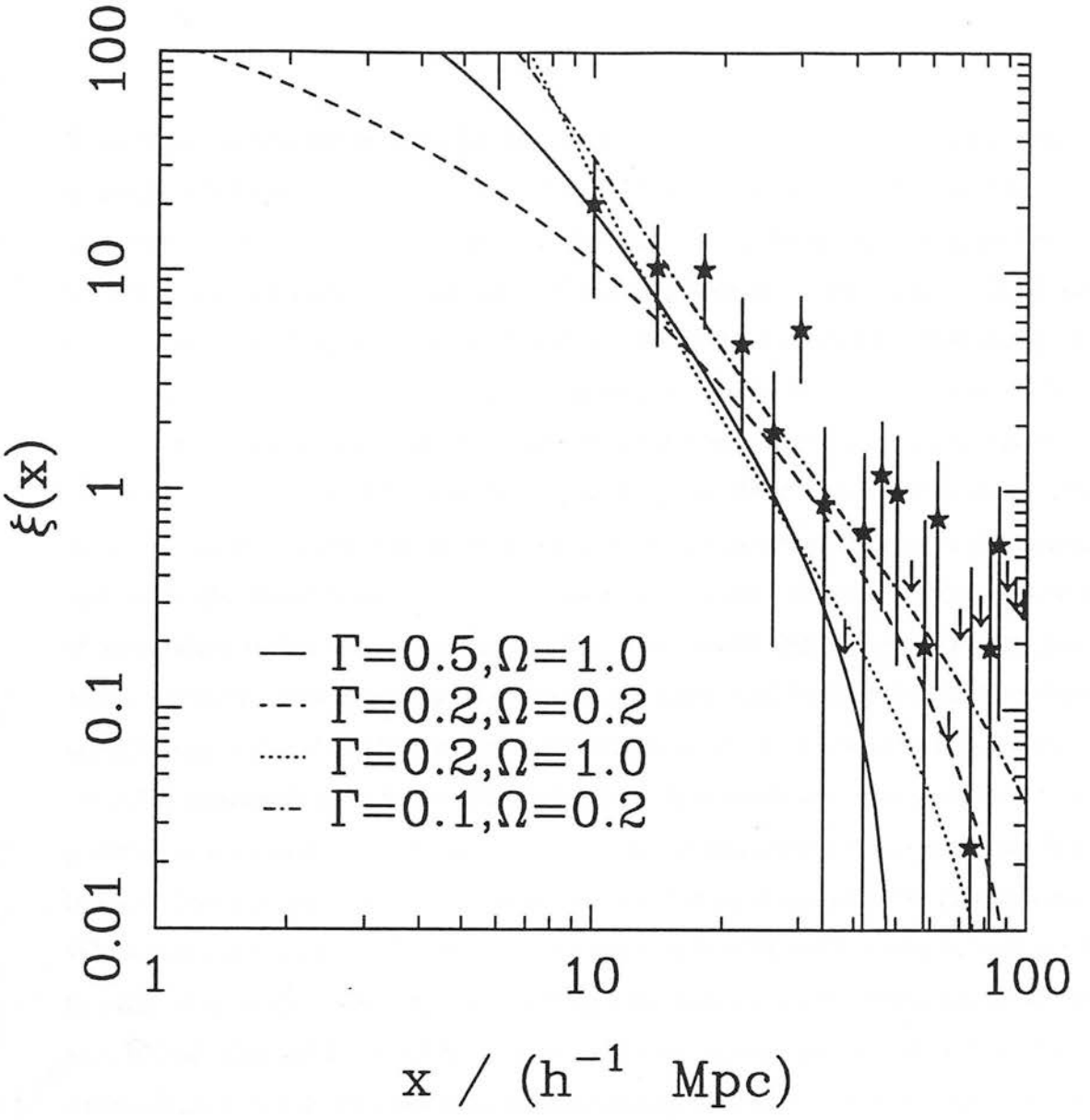


Figure 2.17: The cluster correlation function for Abell $R \geq 2$ clusters from PW92. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively with *COBE* normalization and δ_c=1.0.

None of the models can account for the correlation functions of all four cluster samples, although, conversely, the uncertainty in the *COBE* amplitude and the size of the error bars on the observations mean that it is difficult to exclude any of the models with a high degree of confidence on the basis of the observations as they stand. CDM with $\Gamma = 0.5$ appears incapable of reproducing the strength of clustering observed for rich Abell clusters, although it gives the best fit to the data for APM $\mathcal{R} \geq 20$ clusters and is the only model in which the first zero-crossing point of the correlation function is at a separation of $\sim 40\ h^{-1}$ Mpc, as indicated by the APM and EMRS (Nichol 1992) data. One must be careful, however, in using the first zero-crossing point as a diagnostic tool, since the observational data are prone to significant systematic errors as a result of uncertainty in the cluster number density, which could shift the zero-crossing point. Model 2 gives the best overall agreement with the correlation functions of the four cluster samples and, indeed, its only real inconsistency with the data comes at large scales for the APM number density, where the reliability of the observations is in some doubt. The presence of a cosmological constant affects the epoch and rate at which structure forms, but has little effect on the final phase-space distribution of the particles (Kauffmann & White 1992, and references therein), so that the results of Model 2 should be very similar to those of an open model ($\Omega_0 = \Gamma = 0.2, \lambda_0 = 0$) with the same fluctuation amplitude, and, indeed, Bahcall & Cen (1992) report such a finding on the basis of their numerical simulations of models with and without a cosmological constant. The problem, however, in making this identification between open and flat models with the same Γ lies with the amplitude: the standard analyses of the Sachs-Wolfe effect are for a spatially flat universe, so it is not clear how to determine the fluctuation amplitude for open models from *COBE* .

In Figures 2.18 – 2.21 we show results for the $d - r_0$ relationship. In addition to the four cluster samples discussed above, we also include observational data from the APM $\mathcal{R} \geq 35$ sample of Efstathiou et al. (1992b) and the Abell $R \geq 1$ samples of BS83 and Postman, Huchra & Geller (1992, PHG). We determine the value of r_0 by interpolation in log$r$ space from a correlation function evaluated at a hundred separations in the range $0.5 \leq r/h^{-1}\mathrm{Mpc} \leq 100$. If, instead, a power-law fit is made to the (positive portion of the) correlation function in this interval, the resultant values of r_0 differ, typically,

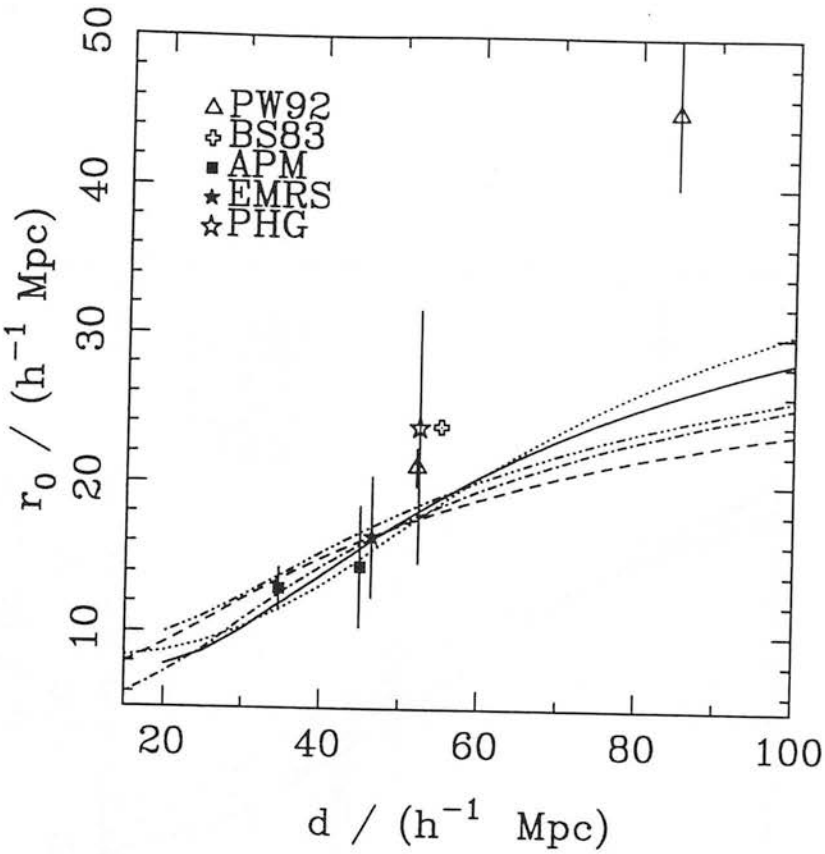


Figure 2.18: The $d - r_0$ relation for Model 1. The observational points come from the sources indicated in the figure and the theoretical curves are for (i) $\epsilon = 3.8 \times 10^{-5}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 2.9 \times 10^{-5}$, $\delta_c = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 2.0 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line).

by $\sim 1\ h^{-1}$ Mpc or less from those obtained by interpolation, and the slope of the best-fitting power law is seen to increase with d.

Two points are clear from these results. First, no model is able to produce the strength of the trend of correlation length increasing with richness that is required to account for all of the observations from APM $\mathcal{R} \geq 20$ clusters to Abell $R \geq 2$. Secondly, the $d - r_0$ relation, taken on its own, may be a misleading statistic: from Figures 2.18 – 2.21 one would conclude that all of the models are wildly inconsistent with the Abell $R \geq 2$ data, although the noise on the data given in Figure 2.17 shows that the situation is not so clear and, in addition, the Poisson error quoted for the correlation length may be something of an underestimate in this case.

The variation between models in the sensitivity of the $d - r_0$ relation to the value of

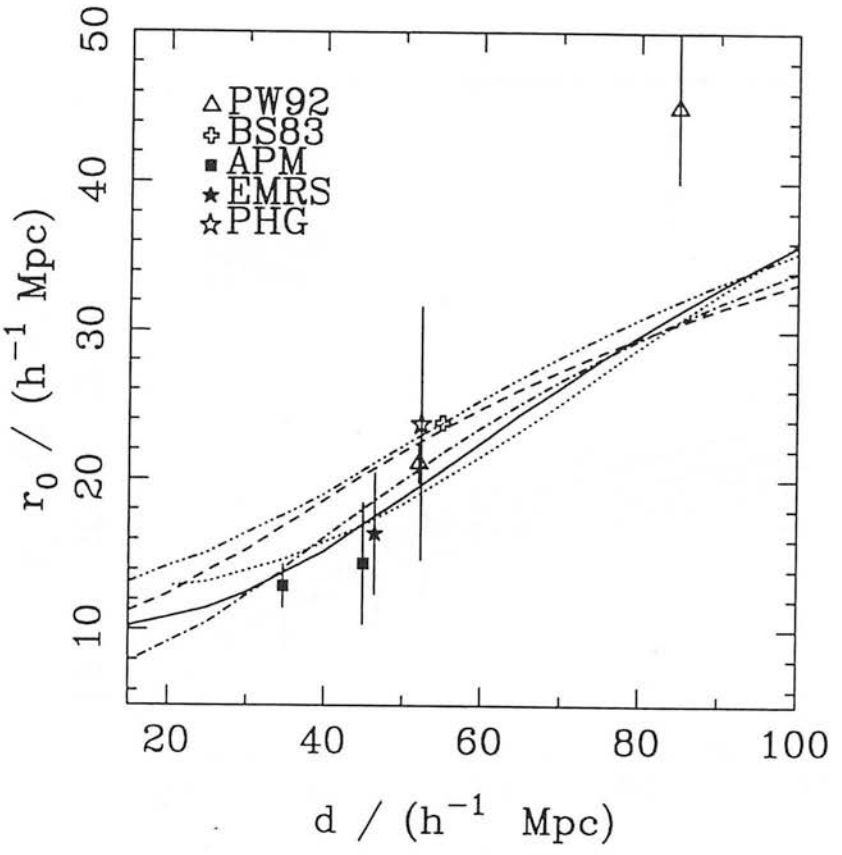


Figure 2.19: The $d - r_0$ relation for Model 2. The observational points come from the sources indicated in the figure and the theoretical curves are for (i) $\epsilon = 1.3 \times 10^{-4}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 1.0 \times 10^{-4}$, $\delta = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 6.9 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line).

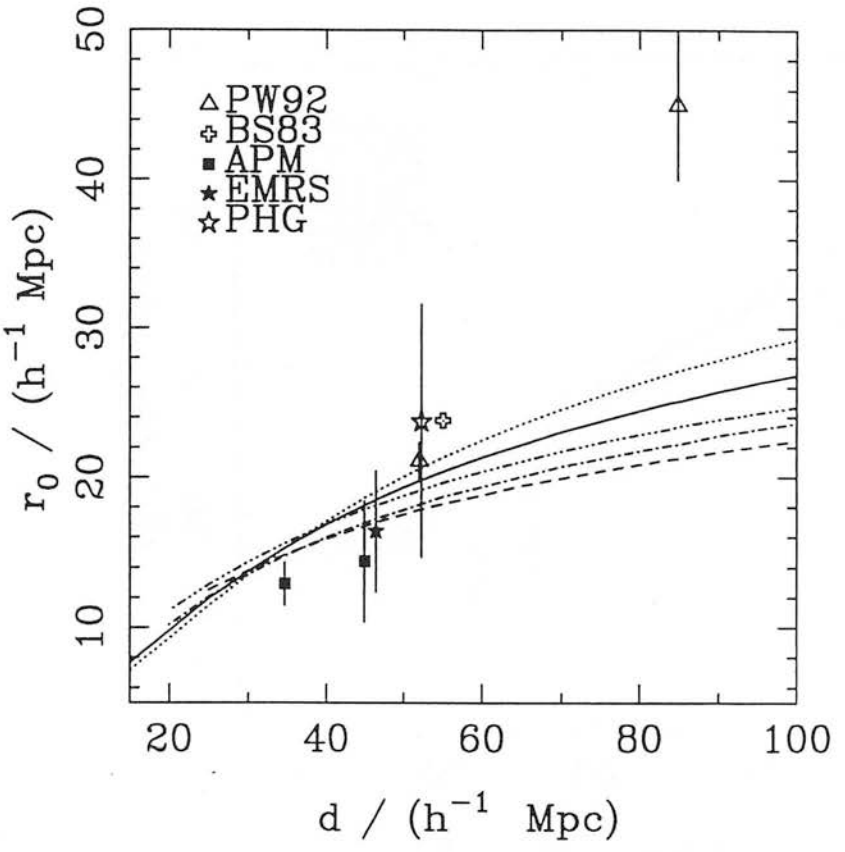


Figure 2.20: The $d - r_0$ relation for Model 3. The observational points come from the sources indicated in the figure and the theoretical curves are for (i) $\epsilon = 3.8 \times 10^{-5}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 2.9 \times 10^{-5}$, $\delta_c = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 2.0 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line).

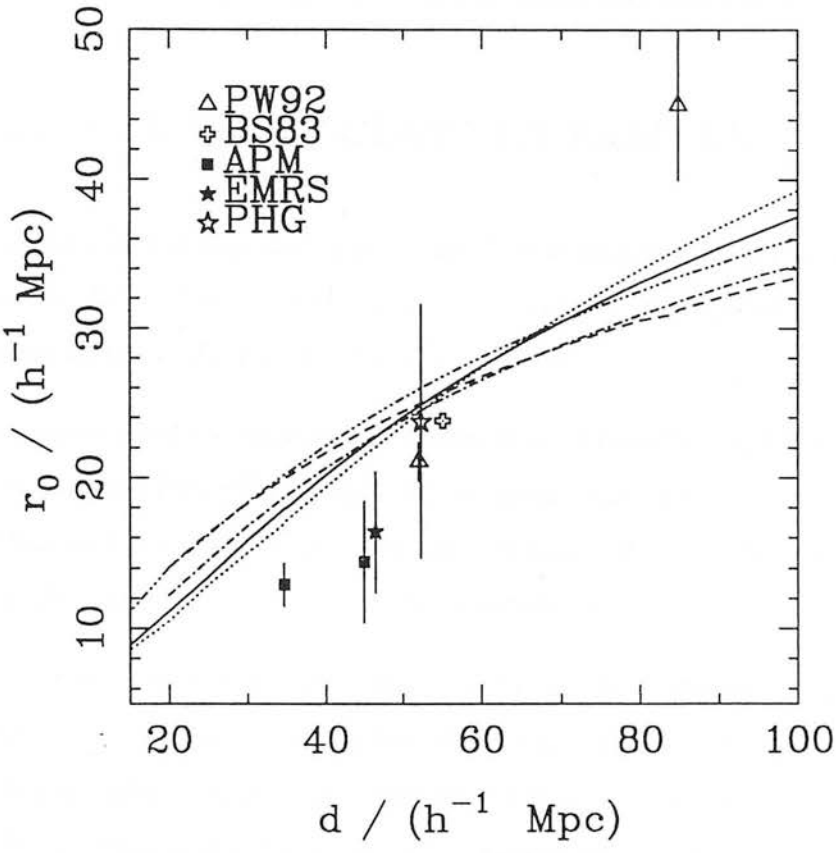


Figure 2.21: The $d - r_0$ relation for Model 4. The observational points come from the sources indicated in the figure and the theoretical curves are for (i) $\epsilon = 1.3 \times 10^{-4}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 1.0 \times 10^{-4}$, $\delta = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 6.9 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line).

ϵ can be readily explained by the same argument as given in Section 2.6.6 above: the sensitivity to ϵ is greater for models with more small-scale power, since these are the ones where the amplitude-sensitive dynamical contribution to the clustering is enhanced relative to the statistical component produced by the initial peak distribution.

2.9 RESULTS: A *ROSAT* CLUSTER SAMPLE

In this section we consider the application of our method to perform the first theoretical study of a sample of *ROSAT*-selected clusters, that of Romer *et al.* (1993), whom we thank for providing access to their data prior to publication.

Romer *et al.* have constructed a sample of 160 clusters in a 1.1 steradian region of the sky around the South Galactic Pole with *ROSAT* X-ray fluxes above a limit of 1.1×10^{-12} erg s^{-1} cm^{-2}. They have obtained redshifts for 128 of these clusters and have computed the spatial correlation function for this (80% complete) sample.

We may compute cluster-cluster correlation functions for a flux-limited sample like this by combining results for volume-limited samples calculated using the methods of Section 2.6. To see how this is done, consider clusters between distances R and $R + \mathrm{d}R$ from us. The clusters in this interval that are included in the survey are those with luminosities in excess of that corresponding to the survey flux limit at distance R. Now, the number of such clusters is defined to be $N(R)\,\mathrm{d}R$, where

$$N(R) \;=\; \Omega\, R^2\, \bar{n}(R), \tag{2.70}$$

where Ω is the solid angle subtended by the survey and $\bar{n}(R)$ is the mean number density of clusters above the appropriate luminosity threshold. The number of pairs, $N_p(r;R)$, of such clusters with separations in the interval r to $r + \mathrm{d}r$ is then given by

$$N_p(r;R)\mathrm{d}r\,\mathrm{d}R \;=\; N(R)\,\mathrm{d}R\,4\pi\, r^2\,\mathrm{d}r\,\bar{n}(R)\,[1 + \xi(r;R)], \tag{2.71}$$

where $\xi(r;R)$ is the correlation function of clusters detectable at distance R. Since the expected number of pairs is the same with $\xi(r;R)$ set to zero, it follows that the

correlation function of the flux-limited sample, $\xi(r)$, is given by

$$\xi(r) \;=\; \frac{\int \bar{n}^2(R)\, R^2\, \xi(r;R)\, \mathrm{d}R}{\int\, \bar{n}^2(R)\, R^2\, \mathrm{d}R}, \tag{2.72}$$

where the integrals are evaluated over the distance interval corresponding to the redshift limits of the sample. The only assumption made here is that pairs at a separation r lie above the luminosity threshold appropriate to the distance R. The validity of this assumption clearly depends on the orientation of the pair's separation vector relative to the line of sight and to the ratio of r to the scale over which the luminosity threshold varies but, in the case that we are concerned with here, of clustering on scales of tens of Mpc in a cluster sample that extends to $z \sim 0.25$, equation 2.72 should be an excellent approximation.

Evaluation of equation 2.72 is very straightforward. To obtain $\bar{n}(R)$ as a function of R, one simply bins the sample's redshift histogram and works out the number density of clusters in each bin. One may then interpolate between these values to obtain $\bar{n}(R)$ at the values of R required for the integration in equation 2.72. We compute the distance, R, to each cluster from its redshift, z, using the angular diameter distance – redshift relation appropriate to each of Models 1-4, using the following formula (Carroll *et al.* 1992):

$$d_{\mathrm{A}}(z) \;=\; \frac{c}{H_0(1+z)} \int_0^z \left[(1+x)^2(1+\Omega_{\mathrm{m}}x) \;-\; x(2+x)\Omega_{\mathrm{v}}\right]^{-1/2} \mathrm{d}x, \tag{2.73}$$

where $\Omega_{\mathrm{m}}, \Omega_{\mathrm{v}}$ are, respectively, the fractions of the critical density in (non-relativistic) matter and vacuum energy. Given $\bar{n}(R)$, the appropriate value of $\xi(r;R)$ may then be found by interpolation between the values of the correlation function at separation r calculated for volume-limited samples of clusters with different number densities.

We consider the same four spatially flat models as in Section 2.6, again normalised by *COBE* to $\epsilon = 2.9 \times 10^{-5}\, \Omega_0^{-0.77}$. The resultant cluster correlation functions are shown in Figure 2.22, together with the data points from Romer *et al.* .

There are two caveats, one theoretical and one observational, concerning the interpretation of Figure 2.22. The theoretical complication results form the fact that our method

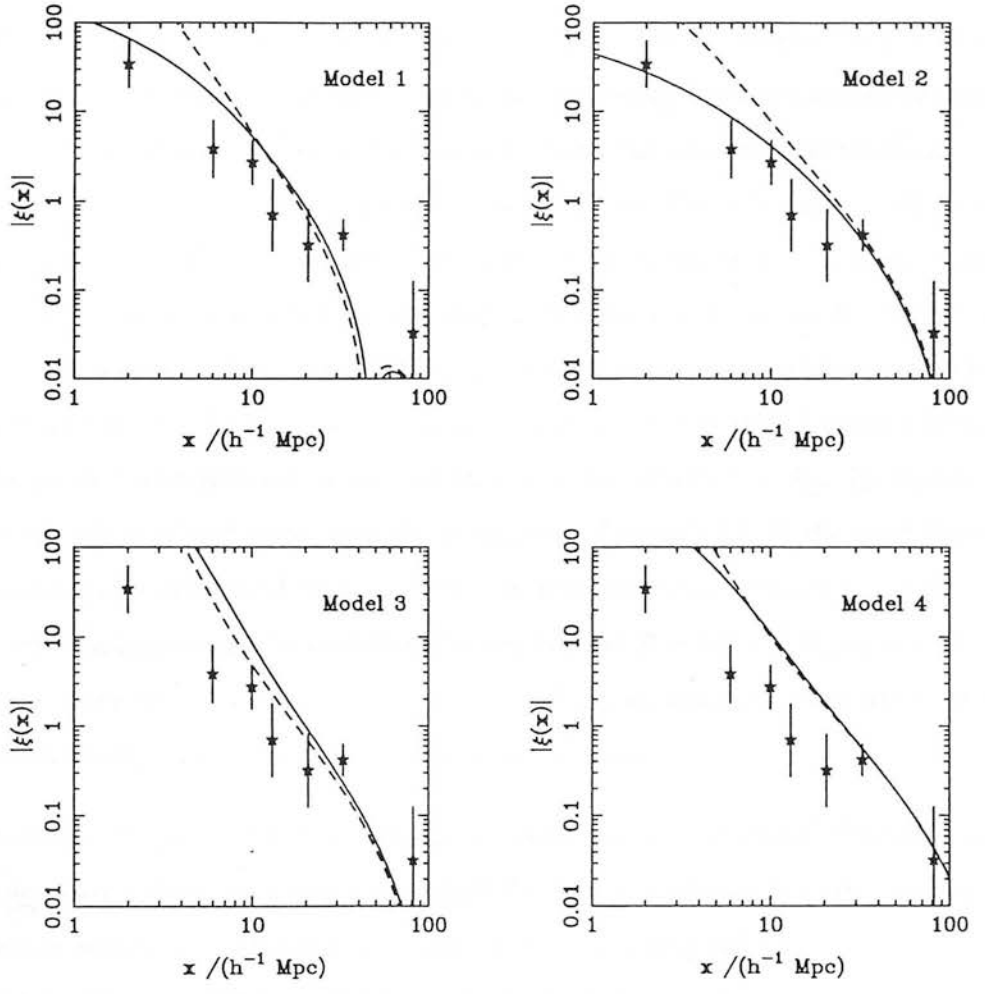


Figure 2.22: Predictions for the correlation function of the Romer *et al.* sample of *ROSAT* cluster sample for Models 1 – 4. In all cases the solid line corresponds to $\delta_c = 1.0$ and the dashed line to $\delta_c = 1.7$. This plot omits a negative value of the correlation function at $r \simeq 52\,h^{-1}$ Mpc.

yields accurate correlations only at separations larger than the smoothing length, R_s, used to define the clusters. For clusters of, say, Abell $R \geq 1$ number density this is only $\sim 5h^{-1}$ Mpc, but for the richest clusters in the Romer *et al.* sample this can be $\sim 15h^{-1}$ Mpc. This can introduce spurious small-scale clustering when evaluating equation 2.72. This effect is shown in Figure 2.23, which shows the relative contribution to $\xi(r)$ in Model 1 from cluster pairs at different distances: we plot $R^3\bar{n}^2(R)\xi(r;R)$ (normalised to unity at its maximum) against R, so that the contribution to $\xi(r)$ from cluster pairs out to a distance R is given by the area under the curve out to R. It is clear that cluster pairs beyond $R \sim 500\,h^{-1}$ Mpc are making a spurious contribution to the value of $\xi(r)$ at $r = 10\,h^{-1}$ Mpc (and, much less seriously, at $r = 20\,h^{-1}$ Mpc) where $r < R_s$, while no such contribution is seen on larger scales, where $r > R_s$. In Figure 2.24 we show the effect of excluding from the evaluation of equation 2.72 the contribution from clusters beyond different distances. It is clear from this that minimal changes to $\xi(r)$ are produced on large scales by omitting clusters beyond $R = 600\,h^{-1}$ Mpc, so we have done that to generate the results shown in Figure 2.22, to remove the spurious small-scale clustering component produced by the richest clusters.

The results of Figure 2.24 also make us confident that our predicted *ROSAT* correlation functions are robust, given the uncertainty in the correlations of the richest clusters discussed in Section 2.7: Figure 2.24 clearly shows that a very small change in the predicted *ROSAT* correlation function would result from a weaker richness dependence of clustering, like that found by Croft & Efstathiou. This insensitivity to the clustering strength of the richer clusters results from the presence of the $\bar{n}^2(R)$ factor in equation 2.72, which is a sufficiently steeply decreasing function of R to outweigh the increase in $\xi(r;R)$ with richness and, hence, R.

The observational caveat concerning the interpretation of Figure 2.22 results from the presence of a negative value for the cluster correlation function at $r \sim 50\,h^{-1}$ Mpc: is this point real or not? Its reality would lend support to CDM (Model 1), since that is the only model which produces a first zero-crossing point at about that scale. If this negative point is not real, then we can see from Figure 2.22 that Model 2 gives the best agreement with the correlations of the Romer *et al.* sample, as it did for the optical data of Section 2.8. Optical and X-ray observations of clusters appear, therefore, to give a

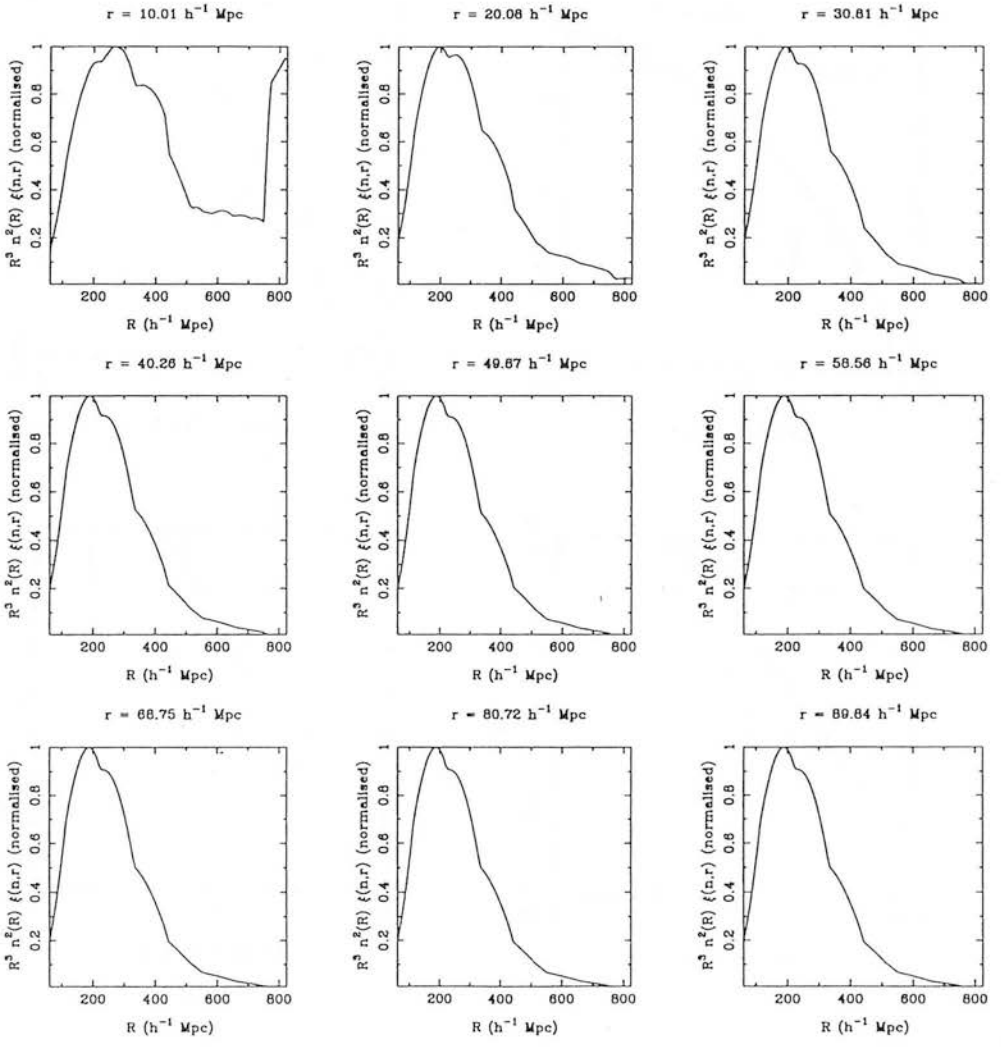


Figure 2.23: The relative contributions from cluster pairs at different separations to the predicted correlation function of the *ROSAT*-selected sample for Model 1, showing the spurious contribution from the richest clusters to the small-scale correlation function

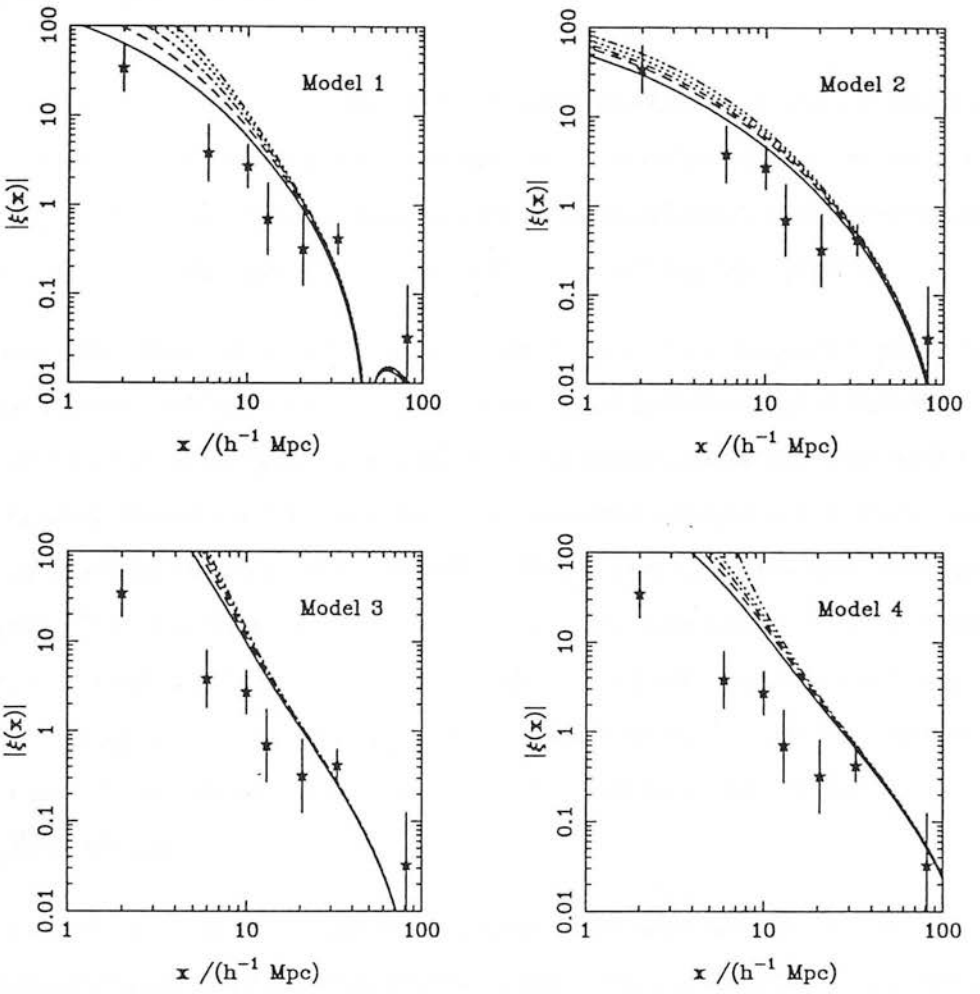


Figure 2.24: The effect on the predicted correlation function of the *ROSAT*-selected sample of excluding the contribution from cluster pairs beyond different distances: R = 400 (full line), 500 (dashed), 600 (dash-dot-dash-dot), 700 (dotted) and 800 (dash-dot-dot-dot) h^{-1} Mpc.

consistent picture of large-scale structure.

2.10 SUMMARY

In this Chapter we have considered the spatial correlations of cluster samples selected in both the optical and X-ray wavebands. We have reviewed previous work addressing this topic from both observational and theoretical standpoints and presented, tested and applied an important new theoretical method for tackling the problem.

Doubts concerning the reliability of the Abell catalogue have dogged theorists attempting to use cluster correlations to constrain cosmological models and it is, therefore, essential for the theorist to study the observational data before comparing them with theoretical predictions. Reliable determinations of the clustering strength of relatively poor clusters are provided by samples drawn from the EDSGC and APM machine-based galaxy catalogues. The clustering strength of richer clusters, investigated only by using samples from the Abell catalogue, remains less well constrained: we must await the results of work in progress on a richer sample of APM clusters (Dalton, private communication) for a reliable answer to whether there is a strong richness dependence to the clustering of galaxy clusters.

We have reviewed previous theoretical work which have used both analytical and numerical methods. Previous analytical work has generally considered only the statistical component to the clustering, neglecting the dynamical component which can be important in many popular models, such as *COBE*-normalised CDM. Those attempts which have been made to incorporate dynamics have employed linear theory, which will not be adequate if $\sigma_8 \sim 1$, as indicated by a wide range of recent work.

The central element to this Chapter has been our presentation of the first analytic calculation of the cluster correlation function to take account of the nonlinear evolution of the cosmological density field. Our method employs the sophisticated framework devised by Bond & Couchman. This combines the theory of the statistics of peaks in Gaussian random fields with the evolution of the cosmological density field by the Zeldovich Approximation, facilitating accurate computation of both the statistical and

dynamical components to the clustering.

We have extended and improved the Bond & Couchman method in a number of regards: we have shown how an understanding of the dynamical evolution described by the Zeldovich Approximation motivates the choice of the smoothing radius of the filter to be applied to the density field; we have constructed a self-consistent mechanism for defining the set of peaks in a Gaussian random field to be identified with a given cluster sample; and we have used the asymptotic large separation peak-peak correlation function to normalise the prescription used to construct a cluster distribution from the underlying density field.

Using this method we have computed the predicted correlation functions for cluster samples of different richness in four popular cosmogonies. We have performed a detailed comparison of these predictions with those determined from recent numerical simulations. Excellent agreement is found between our results and those of the best numerical simulations (those of Croft & Efstathiou) with the exception of samples of the richest clusters, where our method predicts stronger clustering: a detailed discussion of the importance of this point will be given in Chapter 5.

Comparison of our predictions with the observational results from the optical samples that we believe may be reliable shows that the once-favoured Einstein - de Sitter CDM model is incapable of reproducing the strength of clustering exhibited by rich clusters of galaxies. Much better agreement with the observational data is obtained for models with relatively more large-scale power than CDM has in an Einstein - de Sitter universe, in accordance with a number of recent studies of *galaxy* clustering. This consistency between galaxy and cluster clustering studies may be taken as support for the standard picture, due to Kaiser, in which objects form at the sites of peaks in an initial density field that obeys Gaussian statistics: we shall consider this point further in Chapter 5.

We have also computed predictions for comparison with the observed spatial correlations of a sample of clusters selected on the basis of their X-ray emission by the *ROSAT* satellite. This is the first study of the theoretical implications of the spatial correlations of *ROSAT* clusters. The results of this comparison are broadly in agreement with those we obtained from optical cluster samples. This is encouraging, suggesting as it does that

the selection criteria used in the optical and the X-ray bands are selecting similar sorts of object.

We shall discuss the issues raised by this project in more detail in Chapter 5.

Chapter 3

THE CLUSTERING OF WARM AND COOL IRAS GALAXIES

3.1 INTRODUCTION

Cosmology has benefitted greatly from the *IRAS* mission. The *IRAS* view of the extragalactic sky in the infrared (reviewed by Soifer, Houck & Neugebauer 1987) contains ~25,000 galaxies detected at 60 μm, roughly half of which were not previously listed in optical catalogues. *IRAS* offers the cosmologist excellent sky coverage (*IRAS* surveyed ~96% of the sky, but further areas at low Galactic latitude must be excluded, due to their high source density), together with uniform flux calibration (estimated to be better than 10% over nearly the entire sky by Soifer *et al.* 1987), good positional accuracy and insignificant Galactic absorption (there is no broad 'Zone of Avoidance' in the infrared). Of the four *IRAS* passbands (which are centred on 12, 25, 60 & 100 μm), it is the 60 μm band which provides the most successful means of selecting galaxies: more than three quarters of the sources in the *IRAS Point Source Catalog* (1985; hereafter PSC) detected at $\mid b \mid \geq 30^{\circ}$ in the 60 μm band are extragalactic while, conversely, ~75% of previously catalogued extragalactic objects to be found in the PSC were detected at 60 μm.

Cosmological studies have centred on two large redshift surveys of *IRAS* galaxies: the QDOT survey (Lawrence *et al.*, in preparation) and the Berkeley survey (Strauss *et al.*

1990, Fisher *et al.* 1992). The compilers of these surveys have adopted slightly different criteria (expressed in terms of constraints imposed on the fluxes of the sources in the four *IRAS* passbands) in attempting to meet their common goal of selecting uncontaminated galaxy catalogues. They have also used contrasting procedures to construct redshift samples from these parent catalogues of positions of *IRAS* sources on the sky. The QDOT survey comprises redshifts for a one-in-six random sample from the QMW *IRAS* Galaxy Catalogue (Rowan-Robinson *et al.* 1991, hereafter QIGC) of galaxies with 60 μm fluxes above 0.6 Jy: adoption of this sparse-sampling strategy seeks to optimise the information on large-scale galaxy clustering obtainable from a given amount of observing time (Kaiser 1986). The Berkeley group, by contrast, chose to construct a fully-sampled redshift catalogue of galaxies above a limit of 2 Jy (1.936 Jy, in fact), reduced to 1.2 Jy in the extension prepared by Fisher *et al.* (1992). These procedures are complementary: the sparse-sampling strategy of QDOT assays a larger volume of space, while the full sampling of the Berkeley survey provides better resolution of smaller scale features in the galaxy distribution.

The overwhelming majority of galaxies observed by *IRAS* are spirals, which are seen in preference to early-type galaxies by reason of their containing more dust which can re-radiate energy from galactic sources into the infrared. There are several mechanisms through which this can take place and these have been studied by many authors: the details of such work are well beyond both the scope of this thesis, so we shall only summarise them here. Rowan-Robinson & Crawford (1989) have modelled the *IRAS* spectra of galaxies as being composed of three components: a cool 'disc' component, a warm 'starburst' component and a hot 'Seyfert' component. The 'disc' component is due to the re-emission by interstellar dust grains of energy absorbed from the ambient galactic radiation field: this component peaks in the 100 μm band and is the extragalactic analogue of the infrared 'cirrus' (Low *et al.* 1984) observed in the Milky Way. The 'starburst' component peaks in the 60 μm band and is modelled as being produced by optically-thick dust clouds surrounding regions where massive stars are currently being formed. Finally, the 'Seyfert' component is to be identified with emission from the dust cloud surrounding a compact power-law continuum source and peaks in the 12 and 25 μm *IRAS* bands. Rowan-Robinson & Crawford argue that the *IRAS* spectra

of their sample of 227 galaxies are well explained in terms of varying fractions of these three components, with the great majority of galaxies being dominated by the 'disc' and 'starburst' components.

Reality is, no doubt, far more complicated than the simple picture presented by Rowan-Robinson & Crawford but the basic notion of the far-infrared emission from *IRAS* galaxies being a combination of ambient cirrus emission and localised emission from active star-forming regions appears to be correct in essence (Knapp, Helou & Stark 1987; Bothun, Lonsdale & Rice 1989). The relative contributions may be quantified by a dust emission temperature deduced (see Section 3.2.4 for details) from the 60 and 100 μm fluxes of the galaxies and Bothun, Lonsdale & Rice (1989) attest to its efficacy in their comprehensive study of 'normal' and 'active' samples of UGC (Nilson 1973) galaxies observed by *IRAS*. Saunders *et al.* (1990) have determined the 60 μm luminosity functions for warm and cool *IRAS* galaxies, defining their subsamples on the basis of the Rowan-Robinson & Crawford models, which predict that the emission from galaxies with emission temperatures above 36 K (classified as 'warm') is dominated by the 'starburst' component, while the 'disc' component dominates in ('cool') galaxies with emission temperatures below 36 K. They find that the overall 60 μm luminosity function of *IRAS* galaxies is dominated by cool galaxies for luminosities below $\sim 5 \times 10^6 \, h^{-2} \, L_\odot$ ($L_\odot = 3.826 \times 10^{33}$ erg s^{-1}) and by warm galaxies above that level. They find that there are no high luminosity cool galaxies, but that warm galaxies are found down to low luminosities.

The implications of this for a flux-limited sample are clear: as redshift increases, the fraction of galaxies included in the sample that are classified as warm will increase. In other words, there is a radial gradient in the local mean temperature of a flux-limited sample of *IRAS* galaxies: this trend is shown in Figure 3.1. The implication of this, in turn, for clustering analyses of flux-limited samples of *IRAS* galaxies is equally clear when one considers the relationship between the relative importance of the 'disc', 'starburst' and 'Seyfert' emission components for a particular galaxy and the environment within which the galaxy finds itself.

There is a growing body of evidence linking dynamical interactions between galaxies

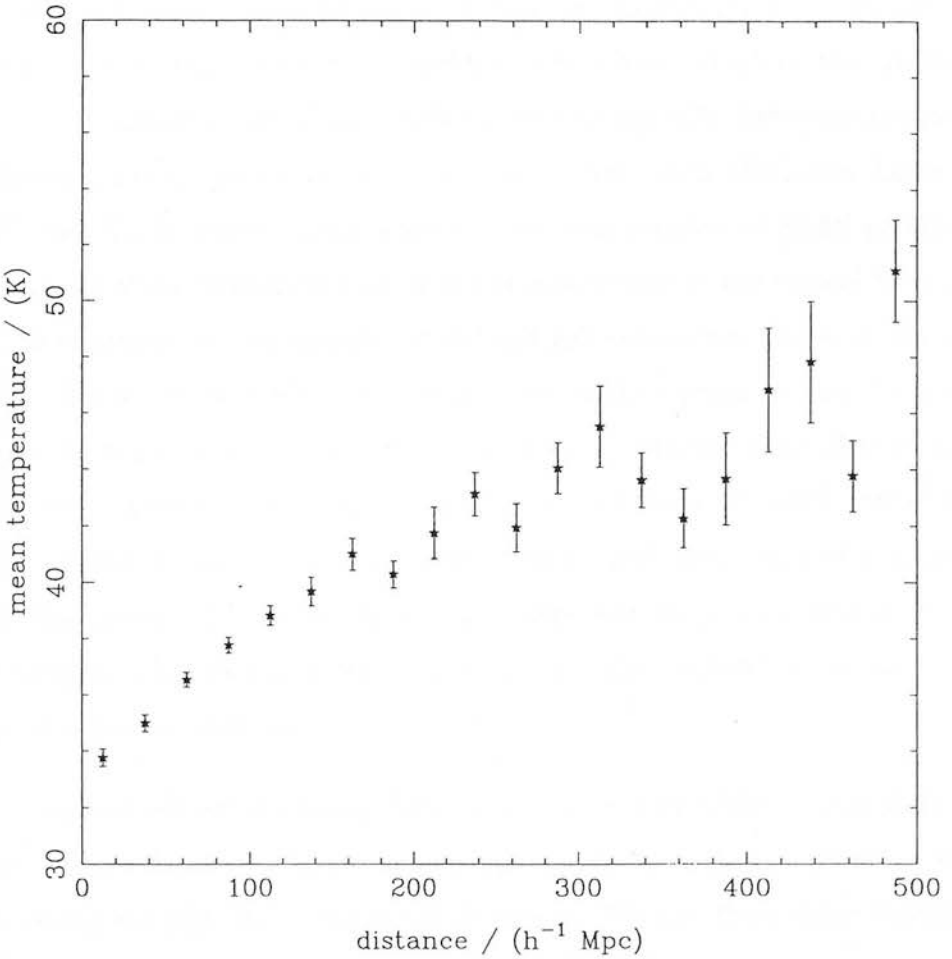


Figure 3.1: The radial gradient in galactic temperature in the QDOT survey.

with the starburst phenomenon of enhanced star fromation. Theoretical simulations of such events (*e.g.* Barnes & Hernquist 1991 and references therein) have reproduced morphologies like those seen in observed interacting pairs and, more importantly, shown how interactions can result in the gas in the interstellar media of the galaxies losing angular momentum, with a resulting infall of large quantities of gas, which could provide the necessary fuel to trigger a burst of star formation. Evidence on the observational side is equally forthcoming (see, for example, Sulentic, Keel & Telesco 1990 and references therein). For example, Larson & Tinsley (1978) have shown how the dispersion of *U-B* and *B-V* colours observed in samples of morphologically disturbed galaxies may be interpreted as resulting from recent bursts of star formation. Bushouse, Lamb & Werner (1988) and Xu & Sulentic (1991) have shown that samples of *IRAS* galaxies selected on the basis of morphological indications of interactions in the optical have higher far-infrared luminosities than samples of isolated galaxies showing no such features in the optical. Surace *et al.* (1993) have refined this study a stage further, by constructing samples of interacting and isolated *IRAS* galaxies with the same distribution of blue luminosities. They find that while infrared properties alone are insufficient to distinguish clearly an *individual* interacting galaxy from an *individual* isolated galaxy with the same blue luminosity, due to the intrinsic dispersion in galactic properties, it is seen that *samples* of interacting galaxies have substantially warmer far-infrared colours than *samples* of isolated galaxies.

One would naively expect galaxy interactions to be more likely to take place where the galaxy number density is higher than average and, by the Kaiser mechanism (see Chapter 2) these regions will exhibit enhanced clustering. We may draw these threads together to deduce a qualitative picture of the effect of the radial gradient in the mean emission temperature of a flux-limited sample of *IRAS* galaxies on studies of its clustering. Put rather crudely, the argument runs as follows: the 'starburst' component will tend to dominate in galaxies with higher emission temperatures; starbursts are expected to be found preferentially in interacting galaxies which, in turn, are more likely to be found in regions with above-average galaxy number density; these regions will exhibit enhanced clustering, due to the Kaiser mechanism, and, thus, we expect warm *IRAS* galaxies to be more strongly clustered than cool ones.

So, a radial mean temperature gradient in a galaxy sample could produce a radial gradient in clustering strength. Such an effect may have already been observed. Mo, Peacock & Xia (1993) report a significant increase in the strength of the redshift-space cross-correlations between Abell clusters and QDOT galaxies if they exclude QDOT galaxies within a redshift of $z = 0.03$, while Feldman, Kaiser & Peacock (1993) find that the amplitude of the redshift-space power spectrum of the QDOT survey is lower if they adopt a radial weighting scheme that gives more weight to nearby galaxies than one giving more weight to more distant galaxies.

Both of these observations would be explained by a picture in which there are two populations of *IRAS* galaxy: a cool, faint, weakly-clustered population of isolated spirals and a more strongly clustered population of hotter, more luminous starbursting galaxies. This scenario is undoubtedly a somewhat naive one, but the importance of any such effect merits its consideration. Cosmological studies using redshift samples of *IRAS* galaxies make implicit assumptions about the homogeneity of the properties of the galaxies. If, in fact, such samples comprise a mixture of these hypothesised two populations, then they will yield biased measures of clustering – in particular, a clustering strength with a non-trivial dependence on the effective depth of the sample.

In this Chapter we seek to address this very important issue, by comparing the clustering strengths of warm and cool subsamples drawn from the QDOT and 2 Jy surveys. In Section 3.2 we present details of the selection criteria used to calculate these subsamples and of the methods used to calculate their selection functions. Sections 3.3 and 3.4 describe our clustering analysis, using two complementary statistics: the two-point correlation function in Section 3.3 and a counts-in-cells analysis in Section 3.4. A brief discussion of the results of these analyses is presented in Section 3.5, while more general discussions are postponed until Chapter 5.

3.2 THE DATA SAMPLES

3.2.1 The QDOT Survey

The QDOT survey (Lawrence *et al.* , in preparation) is a redshift survey of a one-in-six random sample of galaxies drawn from the QMW *IRAS* Galaxy Catalogue (QIGC; Rowan-Robinson *et al.* 1991), limited to $\mid b \mid \geq 10^{\circ}$ and with a 60 μm flux limit of $S_{60} \geq 0.6$ Jy. The selection criteria for the QIGC are as follows, where $S_{\rm n}$ denotes the flux in the *IRAS* passband centred on n μm:

- Sources in the PSC are included if they satisfy

  1. $S_{60} > 0.5$ Jy.
  2. $\mathrm{Log}(S_{60}/S_{25}) > -0.3$. This condition is satisfied by all previously-catalogued galaxies, but not by the great majority of Galactic sources: if the source is not detected in the 25 μm band then the appropriate upper limit is used.
  3. $\mathrm{Log}(S_{25}/S_{12}) < 1.0$, if the source is detected at 25 μm, unless the source is a previously-catalogued galaxy: the 12 μm upper limit is used if there is no detection. Only two galaxies (Arp 220 and NGC 4418) had to be reinstated after failing this criterion, which is designed to exclude planetary nebulæ at low Galactic latitude.
  4. $\mathrm{Log}(S_{100}/S_{25}) > -0.3$, if the source is detected at 25 μm: this excludes further planetary nebulæ.
  5. $\mathrm{Log}(S_{60}/S_{12}) > 0$, if the source is detected at 12 μm: this removes further stars.
  6. $\mathrm{Log}(S_{100}/S_{60}) < 0.6$, unless the source is a previously-catalogued galaxy: the 100 μm upper limit is used in the absence of a detection. This condition is very effective at removing cirrus sources, but, not surprisingly, it also excludes nearby galaxies similar to the Milky Way. Rowan-Robinson, Helou & Walker (1987) have studied the known galaxies which fail this criterion and it is assumed on the basis of this work that virtually no galaxies cool enough to

fail this test, yet with a 60 μm flux in excess of 0.5 Jy, would be faint enough in the optical to have evaded detection by previous catalogues.

7. No identification with a catalogued star, *HII* region, star cluster or Local Group galaxy.

- Sources in the *IRAS Small Scale Structure Catalog* (SSS; Helou & Walker 1988) are included if they have an association with a previously-catalogued galaxy, excepting Local Group galaxies, and parts thereof, and if their 60 μm fluxes are in excess of 0.5 Jy.

- Sources from the *IRAS Large Galaxy Catalog* (LGC; Rice *et al.* 1988) are included if their coadded 60 μm fluxes exceed 0.5 Jy: galaxies must have optical diameters in excess of 8 arcmin for inclusion in the LGC.

The QDOT survey, constructed from the QIGC limited to $\mid b \mid > 10^{\circ}$ and $S_{60} \geq 0.6$ Jy is attested to be 98% complete (Efstathiou *et al.* 1990). We use the 1993 revision of the QDOT catalogue in which the redshifts for ~200 southern galaxies afflicted by a wavelength calibration error in earlier versions of the catalogue have been corrected.

3.2.2 The 2 Jy Survey

The 2 Jy survey (Strauss *et al.* 1990) was selected using a slightly different philosophy from the QDOT survey. Its compilers wished to use infrared properties alone, without reference to optical catalogues to reinstate known galaxies failing particular criteria, as was the case with QDOT. This, of course, necessitates the use of very cautious criteria and inevitably results in the inclusion of many Galactic sources, but, since Strauss *et al.* set out to obtain redshifts for all their objects, not just a sparse sample, this latter should not introduce bias into their catalogue. The criteria used to construct the 2 Jy are as follows:

- $S_{60} > 1.936$ Jy, after correction for various effects discussed by Strauss *et al.*

- Moderate or high flux quality at 60 μm.

- $S_{60}^2 > S_{12}S_{25}$, using upper limits in the absence of detections at 12 or 25 μm: this excludes most stars.

Using these criteria, Strauss *et al.* created a catalogue of 5014 objects, of which 2649 are believed to be galaxies.

3.2.3 The Mask

The standard *IRAS* data analysis procedure, which we follow here, involves the division of the sky into 41167 'lune bins', which measure $1^\circ \times 1^\circ$ defined in ecliptic coordinates (see Appendix X.1 of the *IRAS Explanatory Supplement* by Beichman *et al.* 1988). It is then necessary to construct a mask, which comprises a list of the lune bins to be excluded from the analysis. The standard QIGC/QDOT mask excludes the $\sim 4\%$ of the sky missed by *IRAS*, plus the region $\mid b \mid \leq 10^\circ$ and various other lune bins where high source density confuses discrimination between Galactic and extragalactic sources.

Our purposes, however, necessitate the exclusion of further lune bins, as we require accurate temperature determination on the basis of 60 and 100 μm fluxes. Not all QDOT galaxies were detected at 100 μm and the variation of the 100 μm Galactic cirrus emission across the sky results in a corresponding variation in the upper limits assigned to sources in the absence of confirmed detections at 100 μm. In the computation of the 60 μm luminosity functions of their warm and cool subsamples, Saunders *et al.* (1990) overcame this problem by assigning estimated 100 μm fluxes to the galaxies without confirmed detections in that passband on the basis of an empirical correlation between the 60 to 100 μm flux ratio and the 60 μm luminosity: a temperature–luminosity relation, in essence. While perfectly adequate for their purposes, such a procedure could introduce serious biases into our clustering analysis, by classifying all galaxies with a particular 60 μm flux in one region of the sky as, say, warm, whereas some should be classified as cool, with the obvious introduction of a bias into our estimates for the spatial correlations of the two subsamples. This situation may be greatly ameliorated by the exclusion of lune bins with low 100 μm upper limits and we implement that here, extending our mask to include lune bins where the 100 μm background flux due to Galactic cirrus exceeds

15 MJy per steradian. Thus extended, our mask (created by W. Saunders) leaves an unmasked region covering a solid angle of 8.84 steradian, about 70% of the sky. We use the same mask for the 2 Jy survey as for QDOT: this is being more conservative than we need to be, given the higher flux limit of the 2 Jy survey, but it has the advantage that the subsamples we draw from the two surveys cover the same area of sky and so are readily comparable. A number of galaxies with 100 μm upper limits remain in the unmasked regions of both catalogues even after this extension of the mask and we discuss what should be done with them in the next subsection.

3.2.4 Temperature assignment and subsample selection

Our warm and cool subsamples were constructed using the methods described by Saunders *et al.* (1990). We assume that the far-infrared spectra of the galaxies are well described by a single-temperature Planck function multiplied by an emissivity inversely proportional to wavelength, as often considered appropriate for dust: the modelling by Rowan-Robinson & Crawford (1989) shows that this is a reasonable approximation for both the 'disc' and 'starburst' components to the emission. For each source we fit a single-temperature $S_\nu \propto \nu B_\nu(T_{\rm obs})$ curve [where $B_\nu(T)$ is the Planck function for temperature T] to the spectrum, such that its convolution with the response curves for the 60 and 100 μm *IRAS* detectors (Beichman *et al.* 1984) gives the observed broadband fluxes. The emission temperature, $T_{\rm em}$, is then given by $T_{\rm em} = T_{\rm obs}(1+z)$ for a galaxy with redshift z. Saunders *et al.* use a temperature of 36 K to mark the division between the warm and cool subsamples, on the basis of the Rowan-Robinson & Crawford models. Bothun *et al.* (1989) show that the mean temperatures of their 'normal' and 'active' ('disc' and 'starburst/Seyfert' respectively in the parlance of Rowan-Robinson & Crawford) samples of UGC galaxies observed by *IRAS* are 35 and 38 K respectively, so a figure of 36 K is perfectly reasonable, although the temperature histograms of the 'normal' and 'active' samples overlap, so it cannot be regarded as a sharp cut-off.

As mentioned before, we still have to consider what to do with those galaxies with 100 μm upper limits that remain after the extension of our mask: 275 such galaxies remain in our QDOT sample and 61 in that from the 2 Jy survey. If the 100 μm

upper limit for a particular galaxy is sufficiently low compared with its observed 60 μm flux ($S_{60}/S_{100} > 0.6$), then the galaxy may safely be classified as warm, according to our criterion, but 161 galaxies remain with uncertain temperature classifications in the unmasked region of our QDOT sample after accounting for this. One possible course of action here would be to use 100 μm fluxes from the *IRAS Faint Source Survey* (FSS; Moshir *et al.* 1989) rather than the PSC, but that introduces further problems, as the FSS suffers from more serious cirrus contamination and confusion, while the scatter between the FSS and PSC indicates that the accuracy of low 100 μm fluxes is inherently poor (W. Saunders, private communication).

Instead, we decided to take the 100 μm upper limits for these problem sources as detections. This is not so questionable an action as at first it might seem. The majority of these sources (141 out of 161 in our QDOT sample) were detected at 100 μm in at least one *IRAS* scan, so that while an upper limit is quoted, to indicate the absence of a confirmed detection, its value is that obtained in the unconfirmed detection. We may be sure that the 94 galaxies in our QDOT sample classified as warm by this procedure truly belong to that category, as their 100 μm fluxes can only be lower than the assumed value and, thus, they can only be hotter than we have imagined. That leaves 57 galaxies with doubtful cool classifications in our QDOT sample: since 10 of the 20 galaxies not detected once at 100 μm were classified as cool on the basis of their 100 μm upper limits. Some of these will, no doubt, really belong in the warm sample, but we assume that this has a negligible bearing on our clustering analysis below. We may hope to assess the truth of this assumption by also performing our clustering analysis on subsamples selected above a higher 60 μm flux limit: the number of problem cool sources in our QDOT sample falls to 13 (out of 130 sources with upper limits) if the 60 μm flux limit is raised to 0.7 Jy, while it is only 7 (out of 71) for a cut at 0.8 Jy. We are unable to perform a similar analysis of the number of problem cool sources in our samples from the 2 Jy survey, since the catalogue does not include correlation coefficient data for the 100 μm fluxes, so we cannot determine how many of the galaxies were not detected even once at 100 μm. In general, however, we believe that the trend of decreasing percentage of problem cool sources with increasing 60 μm flux limit should continue and, thus, we expect our 2 Jy sample to be even less contaminated with doubtful temperature assignments

Sample	Survey	60 μm cut (Jy)	Warm/Cool	No. of galaxies
w36.6	QDOT	0.6	warm	1230
c36.6	QDOT	0.6	cool	817
w36.7	QDOT	0.7	warm	997
c36.7	QDOT	0.7	cool	635
w36.8	QDOT	0.8	warm	813
c36.8	QDOT	0.8	cool	530
w2j36	2 Jy	2.0	warm	1330
c2j36	2 Jy	2.0	cool	828

Table 3.1: Properties of the four sets of warm and cool *IRAS* galaxy samples to be considered in this Chapter.

than our QDOT samples.

In Table 3.1 we summarise the properties of the subsamples so produced and in Figure 3.2 and Figure 3.3 we plot the radial velocity histograms for the warm and cool samples derived from QDOT (with a 60 μm flux cut of 0.6 Jy) and the 2 Jy survey, respectively: note that throughout this work we correct radial velocities to the centroid of the Local Group by adding to the observed heliocentric radial velocity of a galaxy with Galactic longitude l and latitude b the correction term 300 $\sin(l)$ $\cos(b)$ km s^{-1}.

3.2.5 Selection Functions

The radial decline in the number density of galaxies in a flux-limited sample is quantified by the selection function, $\psi(r)$, which we define to be the expected number density of galaxies at distance r lying above the sample's flux limit in the absence of clustering: with this definition

$$\int_0^\infty \psi(x)x^2\mathrm{d}x = N/\Omega, \tag{3.1}$$

where N is the number of galaxies in a survey covering a solid angle Ω.

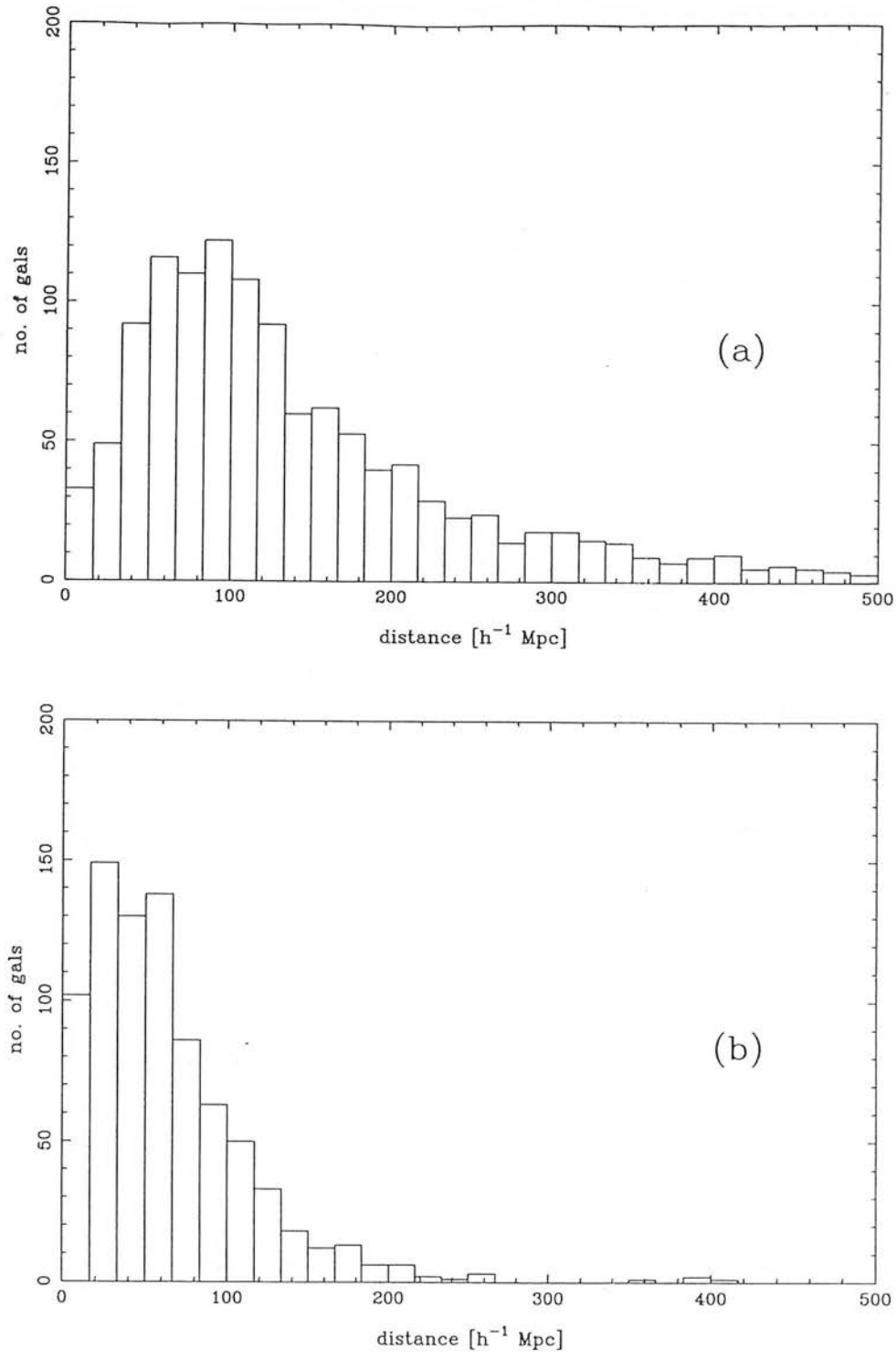
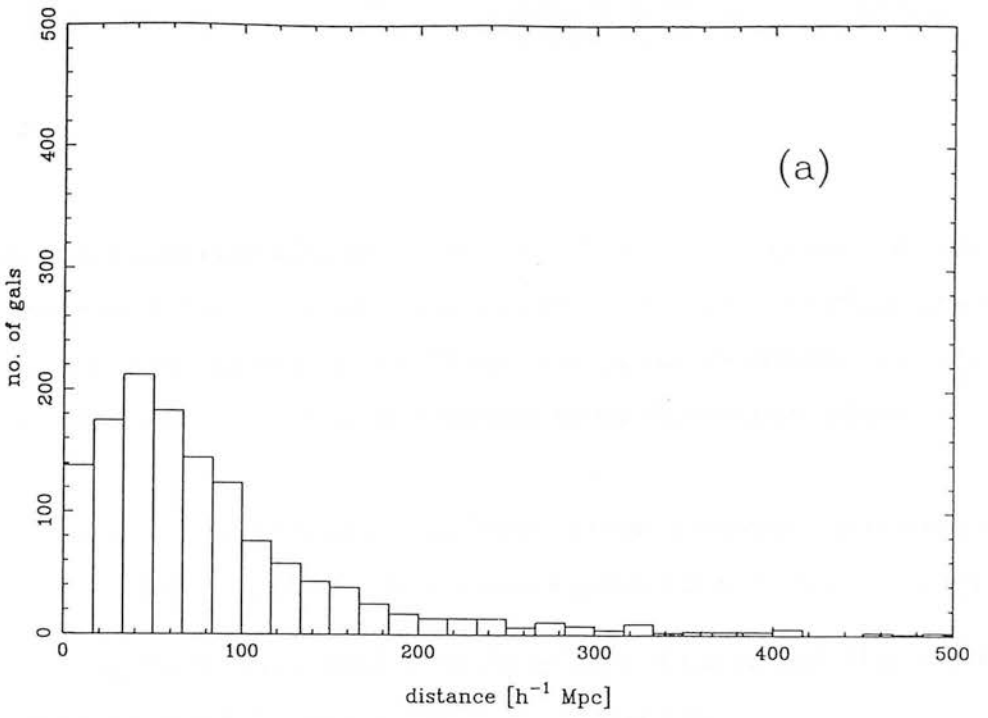


Figure 3.2: The radial distance distributions of (a) the warm and (b) the cool samples derived from QDOT with a 60 μm cut at 0.6 Jy – *i.e.* samples w36.6 and c36.6 respectively. The distance have been computed from radial velocities corrected to the rest frame of the Local Group.



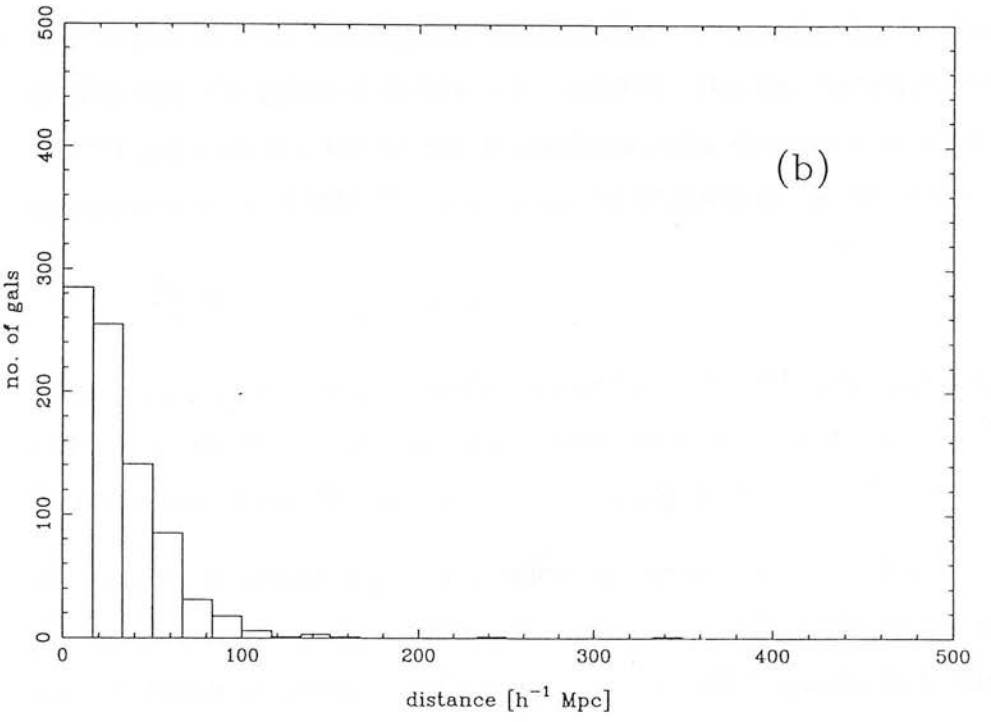


Figure 3.3: The radial distance distributions of (a) the warm and (b) the cool samples derived from the 2 Jy survey – *i.e.* samples w2j36 and c2j36 respectively. The distances have been computed from radial velocities corrected to the rest frame of the Local Group.

Details of the methods used in the computation of the selection functions for our sample (performed by W. Saunders) will be presented elsewhere (Saunders 1993, in preparation) and we provide only a summary here. We use both parametric and non-parametric forms for the selection function, which we determine by the following procedure:

- the galaxies in a given sample are binned by radial velocity to generate a redshift distribution, $n(z)$, which is the number of galaxies in a bin centred on redshift z.
- dividing this by $\mathcal{V}(z)$, which gives the volumes of the redshift bins, yields $\psi_1(z)$, which is our initial estimate of the selection function.
- $\psi_1(z)$ is not, however, an adequate estimate, since no correction has been made for the fact that the galaxy distribution is clustered. This may be done by dividing $\psi_1(z)$ by $\rho(z)$ which is the maximum-likelihood radial density estimator described by Saunders *et al.* (1990). This is obtained by maximising the likelihood

$$\mathcal{L} = \prod_i \frac{\rho(z_i)}{\int_0^{z_{\max,i}} \rho(z_i)(\mathrm{d}V/\mathrm{d}z)\mathrm{d}z}, \tag{3.2}$$

  where $z_{\max,i}$ is the maximum redshift which the source labelled i could have and still remain above the flux limit and V is the comoving volume out to redshift z: details of how this is done in practice are given by Saunders *et al.* (1990).
- the quantity so obtained [call it $\psi_2(z)$] is, however, still not a good estimate of the selection function, since dividing through by $\rho(z)$ will divide out the effects of number density evolution, which Saunders *et al.* (1990) showed to be important for QDOT. The final estimator for the selection function is obtained, therefore, by multiplying $\psi_2(z)$ by the evolution factor $g(z)$ which Saunders *et al.* (1990) have shown to be well approximated by the power law form $g(z) = (1+z)^{6.7}$. Combining these factors gives our final non-parametric estimate of the selection function as

$$\psi(z) = \frac{n(z)\, g(z)}{\mathcal{V}(z)\, \rho(z)}, \tag{3.3}$$

  which may be normalised using equation 3.1 and converted to $\psi(r)$ under the assumption that $\Omega_0 = 1, \Lambda = 0$.

It is found that the selection function so computed closely approximates a double power law, so a convenient parametric form for the selection function is

$$\psi(\Delta) \propto \frac{10^{(1-\alpha)\Delta}}{(1+10^{\gamma\Delta})^{\beta/\gamma}}, \tag{3.4}$$

where $\Delta = \log_{10}(r) - d_*$ and α, β, γ and d_* are parameters whose values are to be found by standard maximum likelihood methods (Saunders 1993).

In Figures 3.4 and 3.5 we show the selection functions derived by these methods for the warm and cool subsamples derived from QDOT with a 0.6 Jy cut at 60 μm (w36.6 & c36.6) and from the 2 Jy (w2j36 & c2j36): the selection functions of the QDOT subsamples selected above higher 60 μm flux cuts do not differ greatly from those for w36.6 and c36.6 shown here. Note that the results from the 2 Jy samples are not so well approximated by our parametric selection function as those from the QDOT subsamples, so we must be careful about using the parametric selection function in analyses involving 2 Jy samples.

3.3 CORRELATION FUNCTION ANALYSIS

3.3.1 Method

In this section we describe our comparison of the redshift-space clustering strengths of warm and cool *IRAS* galaxies using the two-point correlation function.

Our estimator for the redshift-space autocorrelation function, $\xi_{\rm AA}(s)$, of galaxies of species A is

$$1 + \xi_{\rm AA}(s) = \frac{N_{\rm AA}(s)}{N_{\rm AR}(s)} \cdot \frac{n_{\rm R}}{n_{\rm A}}, \tag{3.5}$$

where $N_{\rm AA}(s)$ is twice the number of distinct pairs in the sample of A galaxies (which have a number density $n_{\rm A}$) whose separation places them in the bin centred on separation s and $N_{\rm AR}$ is the corresponding number of cross-pairs between the A sample and a random catalogue with number density $n_{\rm R}$.

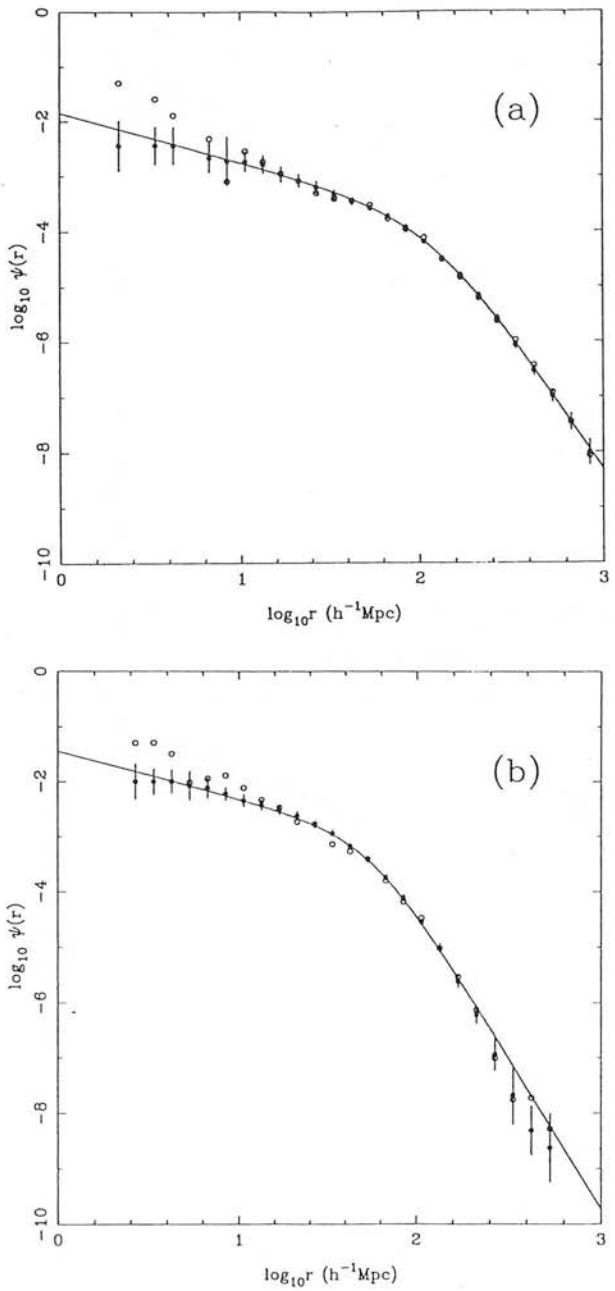


Figure 3.4: Selection functions for subsamples drawn from QDOT with a 0.6 Jy 60 μm flux limit: (a) w36.6 and (b) c36.6. The solid line shows the parametric fit of equation 3.4, while the solid and empty circles show, respectively, the parametric selection function $\psi(r)$ and the initial estimate $\psi_1(r)$.

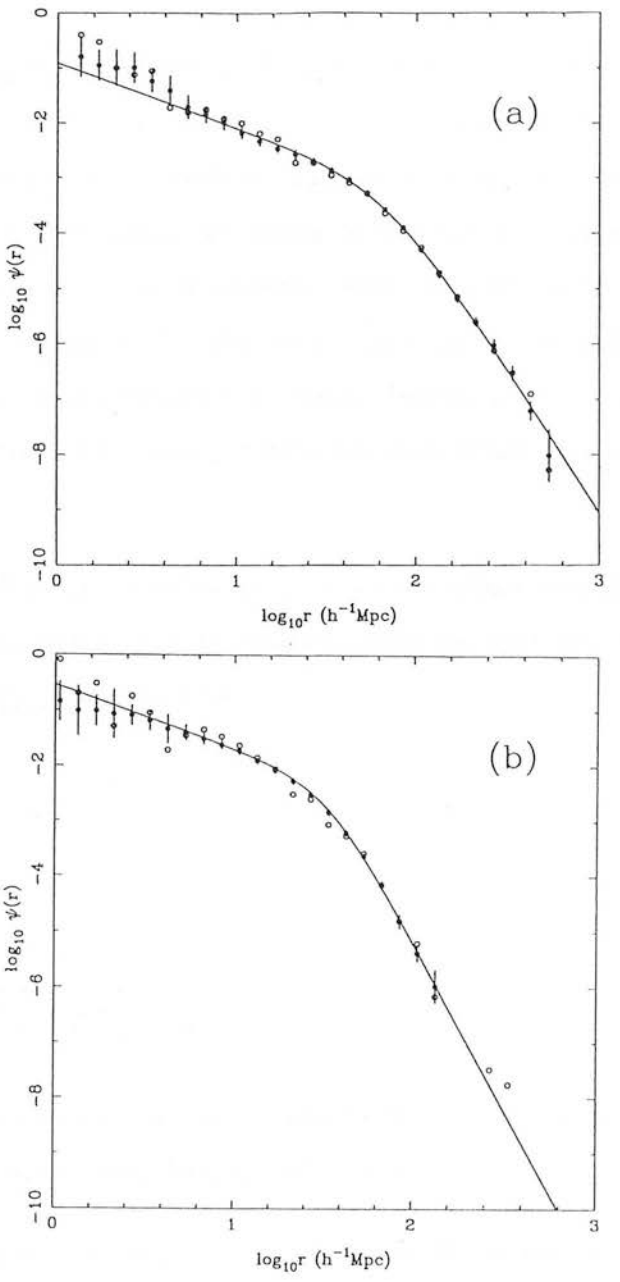


Figure 3.5: Selection functions for subsamples drawn from the 2 Jy survey: (a) w2j36 and (b) c2j36. The solid line shows the parametric fit of equation 3.4, while the solid and empty circles show, respectively, the parametric selection function $\psi(r)$ and the initial estimate $\psi_1(r)$.

The term $N_{\rm AR}(s)/(n_{\rm R}/n_{\rm A})$ represents a Monte Carlo estimation of the expected number of pairs of A galaxies in the bin centred on s in the absence of clustering. In order to evaluate this term accurately, we create the random catalogue by randomly distributing 100,000 particles across the unmasked region of the sky and make an assignment of radial velocities to them in a manner consistent with the selection function of the galaxy sample. By using an estimator with the term $N_{\rm AR}(s)$ in the denominator, rather than $N_{\rm RR}(s)$ we hope better to account for edge effects (Hewett 1982), although these should be relatively unimportant for a survey region like ours which covers almost the whole sky.

Where relatively small galaxy samples are used, a more robust description (Mo, Peacock & Xia 1993) of the clustering may be obtained using the spatially averaged two point correlation function, $\bar{\xi}_{\rm AA}(s)$, defined by

$$\bar{\xi}_{\rm AA}(s) \;=\; \frac{3}{s^3}\int_0^s \xi_{\rm AA}(x)\, x^2\, {\rm d}x, \tag{3.6}$$

and the estimator for which is

$$1 \,+\, \bar{\xi}_{\rm AA}(\leq s) \;=\; \frac{N_{\rm AA}(\leq s)}{N_{\rm AR}(\leq s)} \,\cdot\, \frac{n_{\rm R}}{n_{\rm A}}, \tag{3.7}$$

where $N_{\rm AA}(\leq s)$ is the sum of the pair counts in bins centred on separations $\leq s$ and the definitions of the other terms follows analogously.

By analogy with equation 3.5 we define our estimator for the cross-correlation function, $\xi_{\rm AB}(s)$, of galaxies of species A and B to be

$$1 \,+\, \xi^{(1)}_{\rm AB}(s) \;=\; \frac{N_{\rm AB}(s)}{N_{\rm AR}(s)} \,\cdot\, \frac{n_{\rm R}}{n_{\rm B}}, \tag{3.8}$$

where the meanings of the various terms are clear from above. This estimator is asymmetric under exchange of A and B, so we could clearly equally well define a second estimator of the same quantity by

$$1 \,+\, \xi^{(2)}_{\rm AB}(s) \;=\; \frac{N_{\rm AB}(s)}{N_{\rm BR}(s)} \,\cdot\, \frac{n_{\rm R}}{n_{\rm A}}. \tag{3.9}$$

The difference between the meanings of these two estimators is clear if one considers (Peebles 1980) the definition of the two point correlation function, as we now show. Consider choosing a galaxy of species X from a sample of them. Then the probability, δP, of finding a galaxy of species Y (with number density $n_{\rm Y}$) in a volume δV a distance s away is given by

$$\delta P \;=\; n_{\rm Y}\,\delta V\,[1+\xi_{\rm XY}(s)]. \tag{3.10}$$

It is clear, therefore, that the estimator of equation 3.8 is estimating $\xi_{\rm AB}(s)$ by comparing the probability of finding a neighbour of species B in volume δV of a clustered distribution of such galaxies at separation s from a galaxy of species A against that for an unclustered distribution of B galaxies, while the estimator of equation 3.9 compares the probability of finding a neighbour in the same volume of a clustered distribution of galaxies of species A at separation s from a galaxy of species B with that for an unclustered distribution. The symmetry of the situation dictates that these two estimators should yield identical results in the limit of arbitrarily large data samples, but there may be some difference between the answers they give for finite samples and one can see that the difference between the results so obtained is a measure of the uncertainty in the determinations that they give for the cross-correlation function.

Clearly, by analogy with equations 3.5, 3.7, 3.8 and 3.9 we may define two estimators for the spatially-averaged cross-correlation function by

$$1 \;+\; \bar{\xi}^{(1)}_{\rm AB}(s) \;=\; \frac{N_{\rm AB}(\leq s)}{N_{\rm AR}(\leq s)} \;\cdot\; \frac{n_{\rm R}}{n_{\rm B}}, \tag{3.11}$$

and

$$1 \;+\; \bar{\xi}^{(2)}_{\rm AB}(s) \;=\; \frac{N_{\rm AB}(\leq s)}{N_{\rm BR}(\leq s)} \;\cdot\; \frac{n_{\rm R}}{n_{\rm A}}. \tag{3.12}$$

An interesting quantitative measure of any difference between the clustering strengths of the two samples may be obtained (Alimi, Valls-Gabaud & Blanchard 1988; Strauss *et al.* 1992; Mo, Peacock & Xia 1993) by taking ratios of cross-correlation functions to autocorrelation functions. From the discussion above, picturing the meaning of the

two cross-correlation function estimators, it is clear that the correct ratios to consider are $\xi_{AB}^{(1)}/\xi_{AA}$ and $\xi_{AB}^{(2)}/\xi_{BB}$, since these represent comparisons between the numbers of neighbours of species A and B possessed by galaxies of species A and B respectively.

To optimise signal-to-noise, it is necessary to weight the galaxies in some way that reflects the variation in the mean galaxy number density with radial distance. The simplest way (Davis & Peebles 1983, hereafter DP83) is to weight each galaxy by the reciprocal of the selection function at its radial distance. This is motivated by the assumption that the galaxy distributions are homogenous and isotropic random processes, as it assigns to a given volume of space a weight in proportion to the number of galaxies expected in that region in a volume-limited sample. This prescription gives very high weights to the most distant galaxies in any sample, so a distance limit must be imposed if the estimated data-data pair count is not to be compromised by having a large uncertainty due to these few, highly-weighted distant galaxies: this, of course, means that some information is lost, as the most distant galaxies are excluded from the analysis.

An alternative scheme is to assign (Efstathiou 1988; Loveday *et al.* 1992; Saunders, Rowan-Robinson & Lawrence 1992) to each galaxy in a pair with separation s the weight

$$w(r) \;=\; 1/[1 + 4\pi f \psi(r) J_3(s)] \tag{3.13}$$

where f is the sparse sampling factor (*i.e.* $f = 1$ for a fully-sampled survey and $f = 1/6$ for QDOT) and

$$J_3(s) \;=\; \int_0^s \xi(x)\, x^2 \, \mathrm{d}x. \tag{3.14}$$

It can be shown (Efstathiou 1988) that this weighting minimises the variance in the data-data pair count if $\xi(x) = 0$ for $x > s$. Saunders *et al.* (1992) have shown that the real space $J_3(r)$ continues to rise out to at least $r = 100\,h^{-1}$ Mpc, so that equation 3.13 does not necessarily give the minimum variance estimate of the data-data pair count on scales smaller than that. It remains, however, a sensible weighting scheme: for small separations, where $4\pi f \psi J_3 \gg 1$ and where the contributions to the pair count variance are dominated by clustering rather than the discreteness of galaxies, the galaxies are weighted by the reciprocal of the selection function and, hence, equal volumes are

weighted equally; on large scales, galaxies are given equal weight, thus reducing the problem of shot noise from a few highly-weighted distant galaxies.

This has the advantage of using all the information in the galaxy sample, as none of its members are excluded, but it has the disadvantage that a model for J_3 must be assumed. Saunders *et al.* (1992) have shown that the *real space* J_3 for QDOT is well approximated as being the integral over a power law correlation function $\xi = (3.79\,h^{-1}\ \mathrm{Mpc}/r)^{1.57}$ out to $r = 30\ h^{-1}$ Mpc and, with rather less certainty, out to $r \sim 100\ h^{-1}$ Mpc. We may, therefore, model the *redshift space* J_3 as being this multiplied by a redshift space correction factor calculated according to the method of Kaiser (1987), whence, in linear theory, the (isotropic) real-space correlation function, $\xi_{\mathrm{r}}(r)$, is related to the direction-averaged redshift-space correlation function, $\xi_{\langle \mathrm{s} \rangle}(s)$, by $\xi_{\langle \mathrm{s} \rangle}(s) = F\xi_{\mathrm{r}}(r)$, where $F = [1 + (2\Omega_0^{0.6}/3b) + (\Omega_0^{1.2}/5b^2)]$ in a linear biasing model with bias factor b. Saunders *et al.* (1992) discuss the value of F appropriate to QDOT: from a comparison of their determination of the variance of counts-in-cells in real space to that of Efstathiou *et al.* (1990) in redshift space they conclude that $F = 1.23 \pm 0.13$, but they argue that this is an underestimate due to its omission of non-linear effects (Suto & Suginohara 1991) and, restricting themselves to cells of size $\geq 40\ h^{-1}$ Mpc, where such effects are small, they obtain $F = 1.57 \pm 0.32$. We may use these two values of F to model $J_3(s)$, requiring consistency between this model and the J_3 computed from our redshift space correlation function.

In Figure 3.6 we show a comparison between the model J_3 functions and those deduced observationally using this model as an input to the mininum variance weighting scheme. The model for $J_3(s)$ is seen to give consistent results for separations above a few Mpc: on smaller scales the Kaiser formula (whose basis is in linear theory) is invalid and the redshift space correlation function is not simply a constant multiple of that in real space. This does not constitute a serious problem for us, as we should not wish to compare our results for warm and cool samples on these scales anyway, simply because of the ambiguity in their interpretation introduced by redshift space effects: in our naive model, the warm sample is more strongly clustered in real space, but the higher pairwise velocities that the warm galaxies would possess as a result of their being in regions of higher density would have the effect of diluting the small-separation clustering in redshift

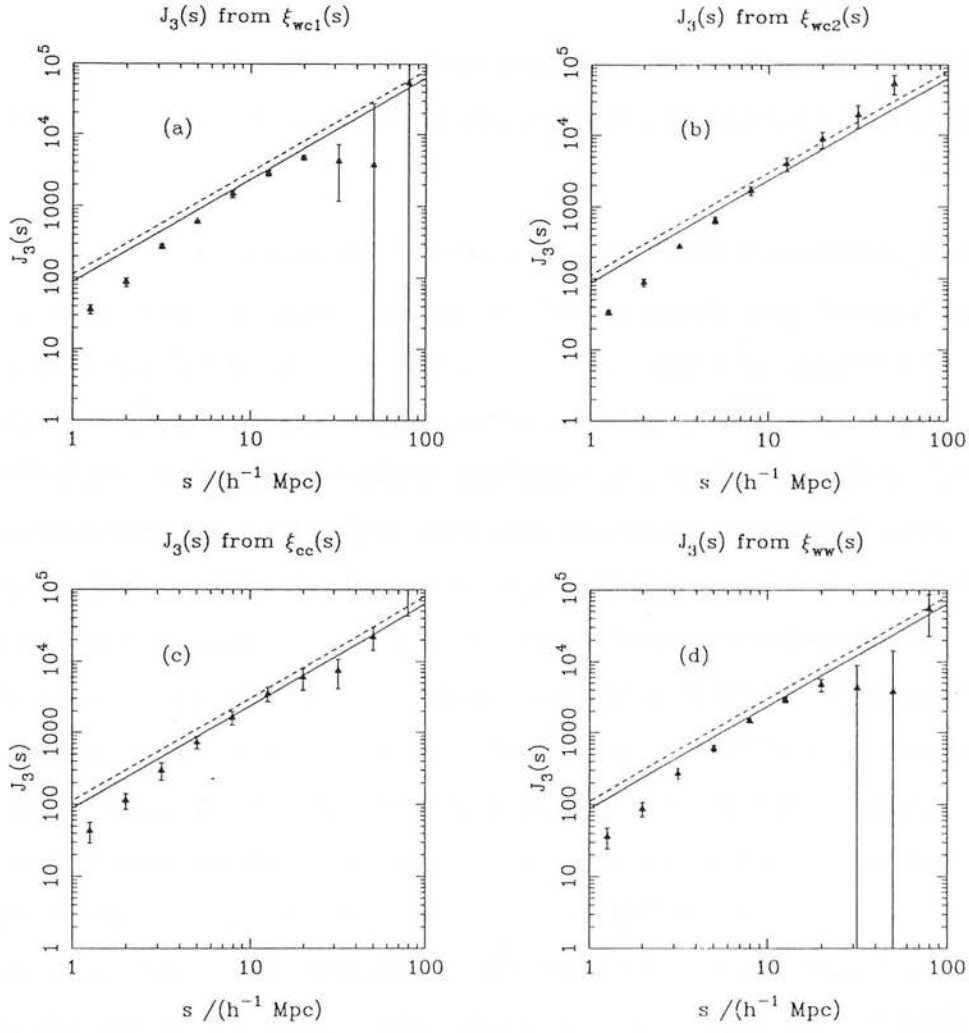


Figure 3.6: Comparison of the measured redshift space $J_3(s)$ functions for the QDOT samples with a 0.6 Jy flux cut at 60 μm with the model used in the minimum variance weighting prescription employed in their determination for: (a) $\xi_{\rm wc}^{(1)}$; (b) $\xi_{\rm wc}^{(2)}$; (c) $\xi_{\rm cc}$; and (d) $\xi_{\rm ww}$. The error bars on the points come from the scatter between the results computed from the four quadrants of the sky. The solid and dashed lines show the model $J_3(s)$ functions for $F = 1.23$ and $F = 1.57$ respectively.

space and so, any difference that we observed at small scales between the clustering strengths of the warm and cool samples in redshift space would be a lower limit on the difference in real space.

In what follows we shall employ both of the weighting schemes discussed above: by using both we may hope to test the robustness of our results and their sensitivity to weighting schemes.

The question of how to estimate error bars on observational correlation functions is one which still awaits a satisfactory answer. The commonly used 'Poisson' error (*i.e.* $\delta\xi/(1+\xi) = 1/\sqrt{DD}$, where DD is the data-data pair count, is incorrect as it assumes an unclustered sample. It may be improved by replacing $1/\sqrt{DD}$ by $(1+4\pi nJ_3)/\sqrt{DD}$ which (Peebles 1980) models the galaxy distribution as a number of randomly distributed clusters each with $N = 1 + 4\pi nJ_3$ members, but this is still conceptually unsatisfactory. Mo, Jing & Borner (1992) have used numerical simulations to validate analytic approximations to the true ensemble error on the correlation function of volume-limited samples, but their procedure is not readily applicable to flux-limited samples such as ours, nor is the boostrap resampling method of Ling, Frenk & Barrow (1986). We adopt, therefore, the robust method of estimating the uncertainty on the correlation function as being the standard error computed from independent estimates of it in the four quadrants of the sky defined by $0 \leq l < 180^\circ$ and $180 \leq l < 360^\circ$ in the northern and southern Galactic hemispheres. *N.B.*. Saunders *et al.* (1992) find good agreement between errors computed using the "$1 + 4\pi nJ_3$" cluster model and computed from independent regions on the sky.

3.3.2 Results

Our desire in this chapter is to obtain a robust comparison of the clustering strengths of cool and warm subsamples of *IRAS* galaxies. In this subsection we present, therefore, correlation function results for a number of samples with different 60 μm flux limits, extending to different redshifts, calculated using different weighting schemes. Note that the spatially-averaged correlation functions are denoted by $< \xi >$ rather than $\bar{\xi}$ in the figures, due to the limited labelling possibilities presented by the PGPLOT graphics

package.

We consider the clustering of our QDOT samples first. Our desire for robust conclusions from this study leads us to ensure that the results obtained through using the minimum variance weighting scheme should be consistent with those yielded by weighting the galaxies by the reciprocal of the selection function in the regime where the latter weighting scheme is applicable, so we must determine the maximum distance out to which this is so. Davis & Peebles (1983) argue that this weighting scheme be used only out to the point where the selection function has fallen by a factor of ten from its zero-redshift value. This corresponds to $r \sim 35\,h^{-1}$ Mpc for the c36.6 sample, but there are only 87 w36.6 galaxies within that distance, limiting the utility of any comparison made to that depth.

In Figure 3.7 we compare results by three methods for samples w36.6 and c36.6 limited to $40\,h^{-1}$ Mpc, within which there are 114 warm galaxies and 290 cool ones. We see that weighting by the minimum variance method gives very similar results to weighting by the reciprocal of the selection function, using either its parametric or non-parametric forms: weighting by the non-parametric selection function produces the most discrepant set of results, as a result of the effect of clustering on such a small sample. The results for the three methods are consistent within their estimated errors. These error bars (which are omitted for the sake of clarity) are quite large, of course, due to the small samples, so we compare the same three methods in Figure 3.8 for samples limited to $100\,h^{-1}$ Mpc – *i.e.* including 522 warm galaxies and 668 cool galaxies.

These three sets of results are again entirely consistent within their errors. Note that the larger numbers of galaxies in radial shells beyond $40\,h^{-1}$ Mpc has brought the results obtained by weighting by the reciprocal of the parametric and non-parametric selection functions into close agreement. There is a slight difference between those results and the ones yielded by the minimum variance weighting scheme, indicating that the inclusion of galaxies out to $100\,h^{-1}$ Mpc is introducing very highly-weighted galaxies which bias the correlation function: this interpretation of the discrepancy is supported by the fact that it is greater for $\xi_{cc}(s)$ than for $\xi_{ww}(s)$, as would be expected from the steeper decline of the selection function of the cool sample.

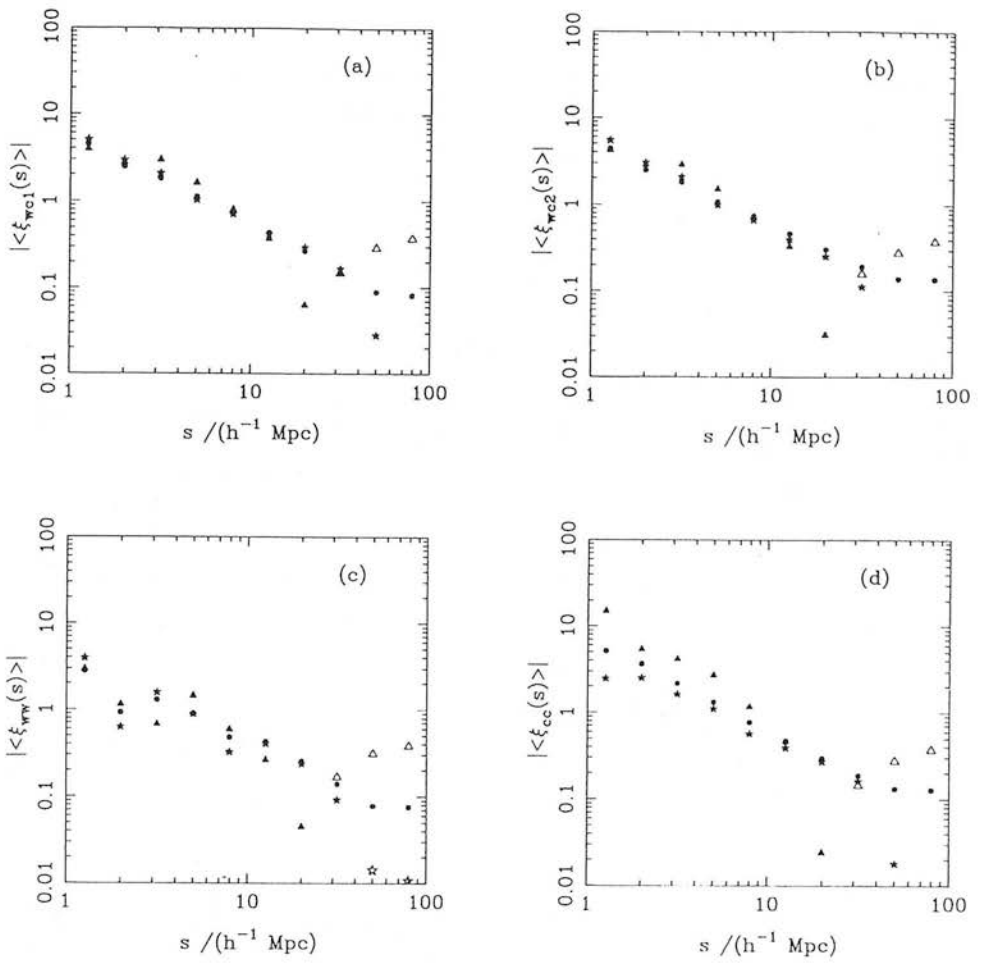


Figure 3.7: Results for w36.6 and c36.6 limited to 40 h^{-1} Mpc, using three weighting schemes: the stars and triangles are results determined by weighting with the reciprocal of the parametric and nonparametric selection functions respectively; the circles denote results obtained using the minimum variance weighting scheme with the parametric selection function. Filled symbols denote positive values of the correlation function and open symbols negative ones.

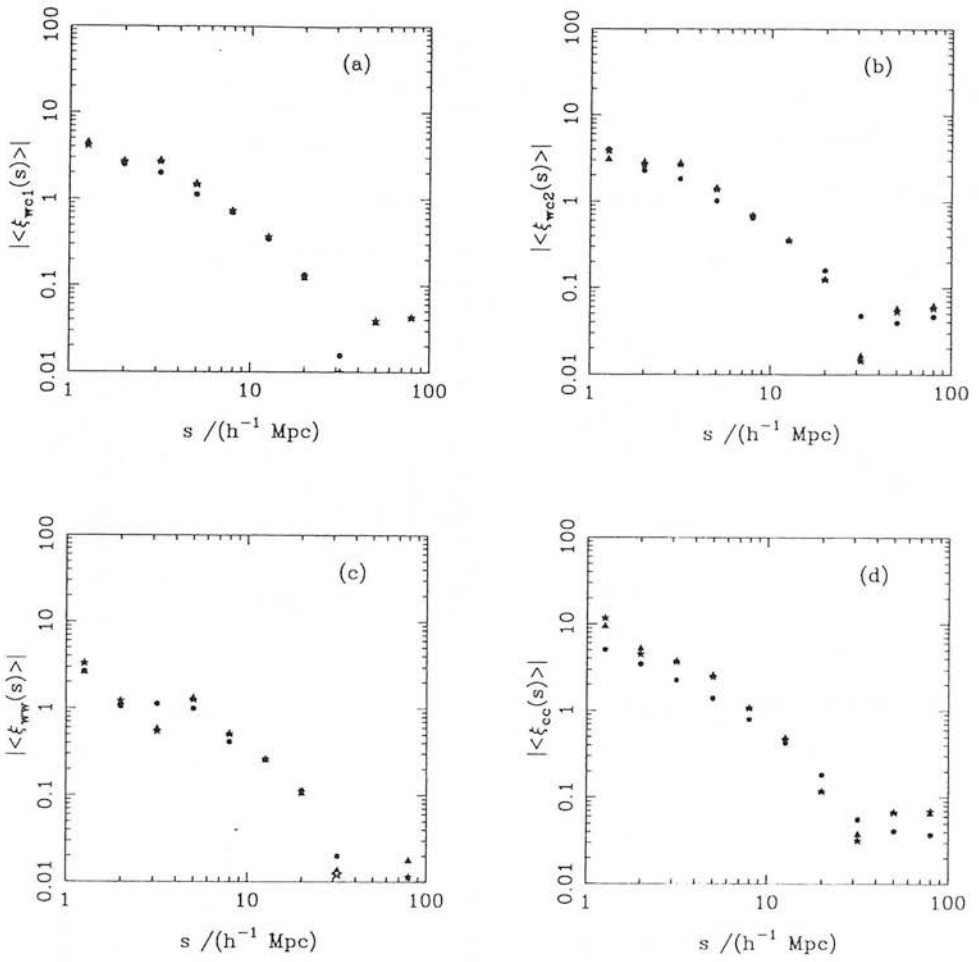


Figure 3.8: Results for w36.6 and c36.6 limited to 100 h^{-1} Mpc, using three weighting schemes: the stars and triangles are results determined by weighting with the reciprocal of the parametric and nonparametric selection functions respectively; the circles denote results obtained using the minimum variance weighting scheme with the parametric selection function. Filled symbols denote positive values of the correlation function and open symbols negative ones.

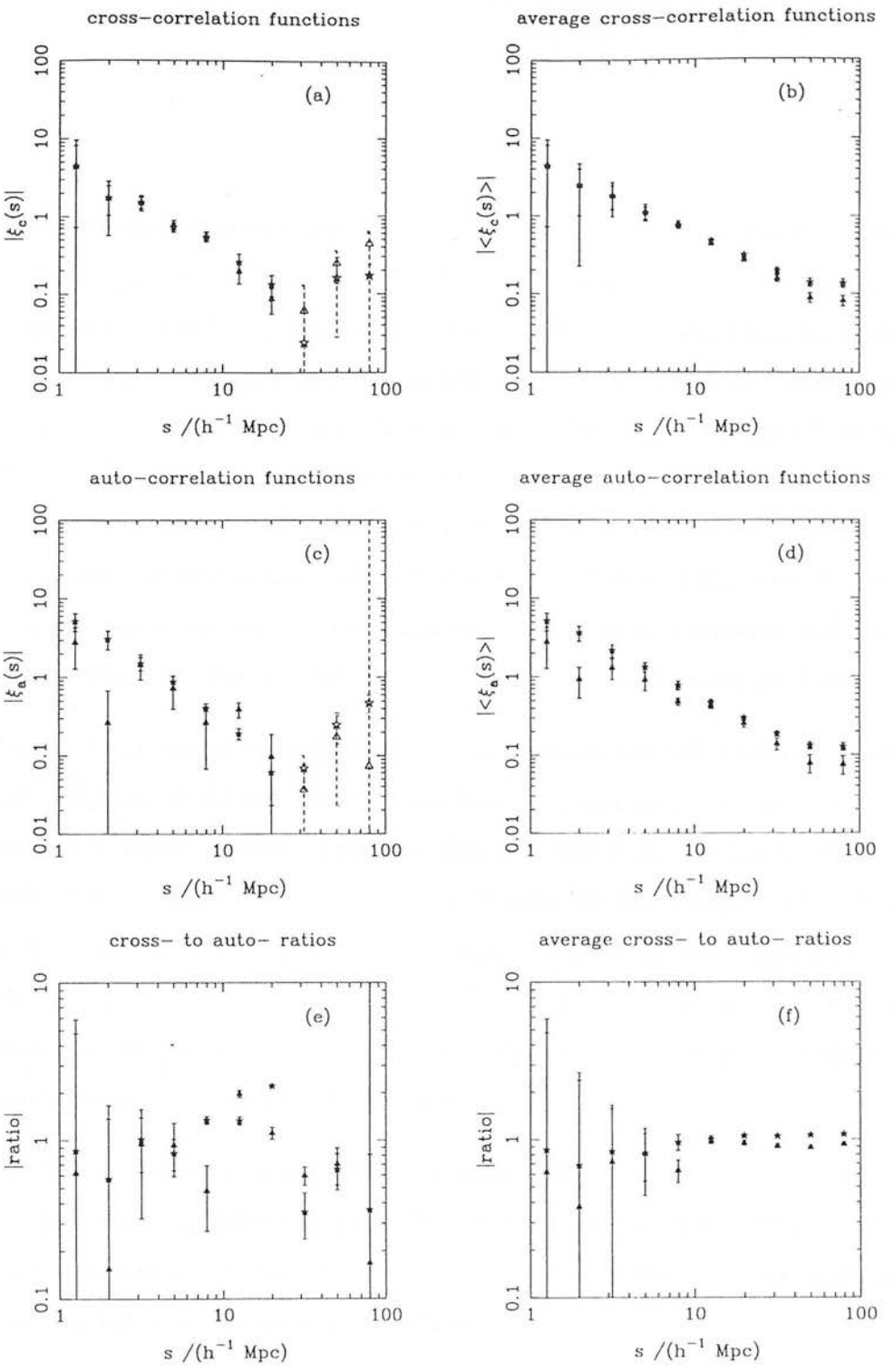


Figure 3.9: Results for w36.6 and c36.6 limited to 100 h^{-1} Mpc, using the minimum variance weighting scheme: in (a) and (b) the triangles and stars denote $\xi^{(1)}_{\rm wc}(s)$ and $\xi^{(2)}_{\rm wc}(s)$; in (c) and (d) the triangles [stars] denote $\xi_{\rm ww}(s)$ $[\xi_{\rm cc}(s)]$; in (e) and (f) the triangles and stars denote the ratios $\xi_{\rm ww}/\xi^{(1)}_{\rm wc}$ and $\xi^{(2)}_{\rm wc}/\xi_{\rm cc}$ respectively. Filled symbols denote positive values of the correlation function and open symbols negative ones.

We believe, therefore, that the minimum variance weighting scheme provides a robust measure of the clustering of the w36.6 and c36.6 galaxies within $100h^{-1}$ Mpc and we show those results in more detail in Figure 3.9. We show the correlation functions as well as their spatial averages, to illustrate the desirability of using the latter for obtaining much less noisy results. Figure 3.9 (f) shows that there is no discernible difference between the clustering strengths of the warm and cool QDOT samples selected above a 60 μm flux limit of 0.6 Jy and within a distance limit of 100 h^{-1} Mpc. Consistent results to these are obtained when QDOT samples selected with $S_{60} > 0.7$ Jy and $S_{60} > 0.8$ Jy are used instead, so we believe this result to be insensitive to the small number of dubious cool classifications resulting from the 100 μm upper limit problem discussed in Section 3.2.4.

Beyond $100h^{-1}$ Mpc we can use only the minimum variance weighting scheme, as weighting by the reciprocal of the selection function will be inappropriate for galaxies beyond this point, as it would produce extremely highly weighted cool galaxies towards the outer limit of any sample. In Figure 3.10 we compare results obtained for w36.6 and c36.6 samples limited to 40, 100 & 500 h^{-1} Mpc, weighted by the minimum variance prescription. The first two samples show no difference in the clustering of the warm and cool species, but the third does: in the 500 h^{-1} Mpc sample the warm subsample does appear to be more clustered than the cool one.

In Figure 3.11 we show the results from the 500 h^{-1} Mpc samples in more detail and see the evidence for any difference is marginal and that the difference between the two estimates for the cross-correlation function indicates the limited accuracy of the auto- to cross-correlation function ratios as quantifiers of it.

In Figure 3.12 we show clustering results for the QDOT subsamples selected above $S_{60} = 0.8$ Jy, including all galaxies within 500 h^{-1} Mpc. Comparison of Figures 3.11 and 3.12 suggests that the hypothesised difference between the clustering of the two samples is more pronounced when the higher flux limit is imposed. On the other hand, the difference between the two estimates of the cross-correlation function increases too, indicating, perhaps, that it is the smaller sample size that is responsible for the change. It is worth noting from Figure 3.12 (c) and (d) that the auto-correlation function of the warm sample appears to be less steep than the cool sample. This is just as we would

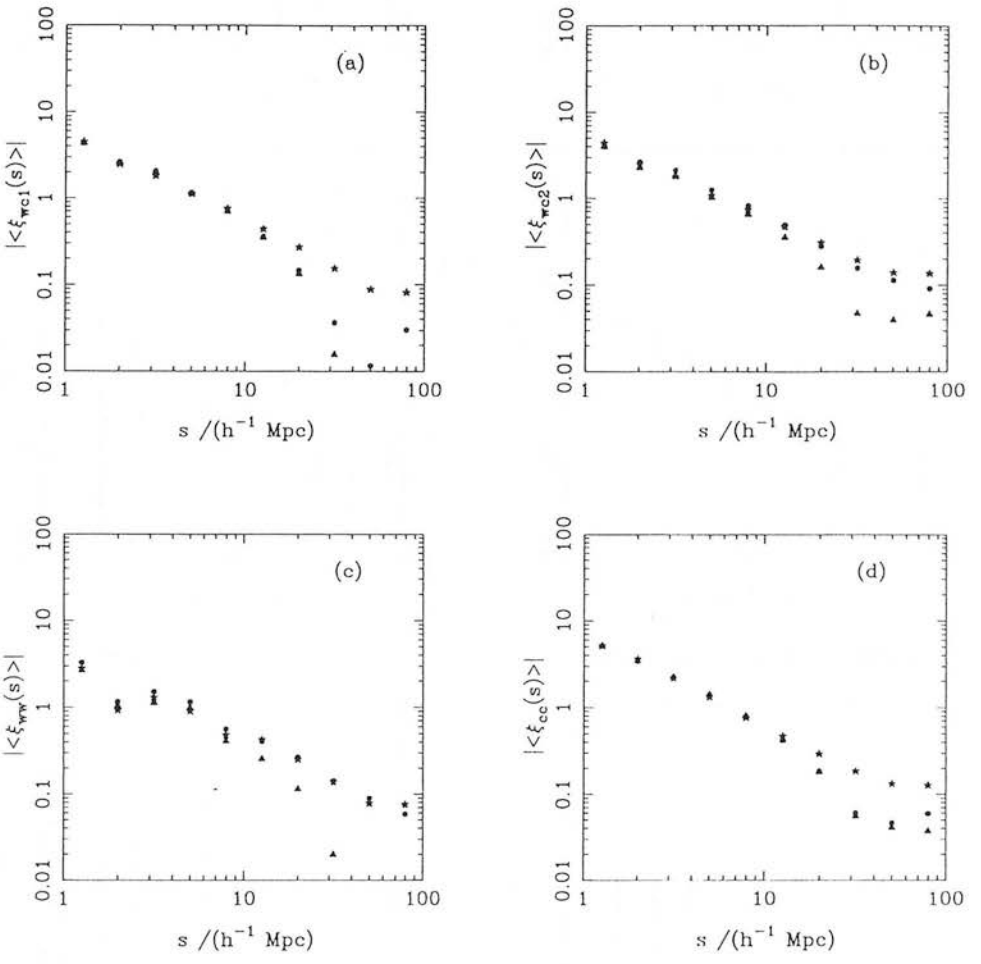


Figure 3.10: Results obtained using minimum variance weighting for samples w36.6 and c36.6 limited to 40 h^{-1} Mpc (stars), 100 h^{-1} Mpc (triangles) and 500 h^{-1} Mpc (circles). Filled symbols denote positive values of the correlation function and open symbols negative ones.

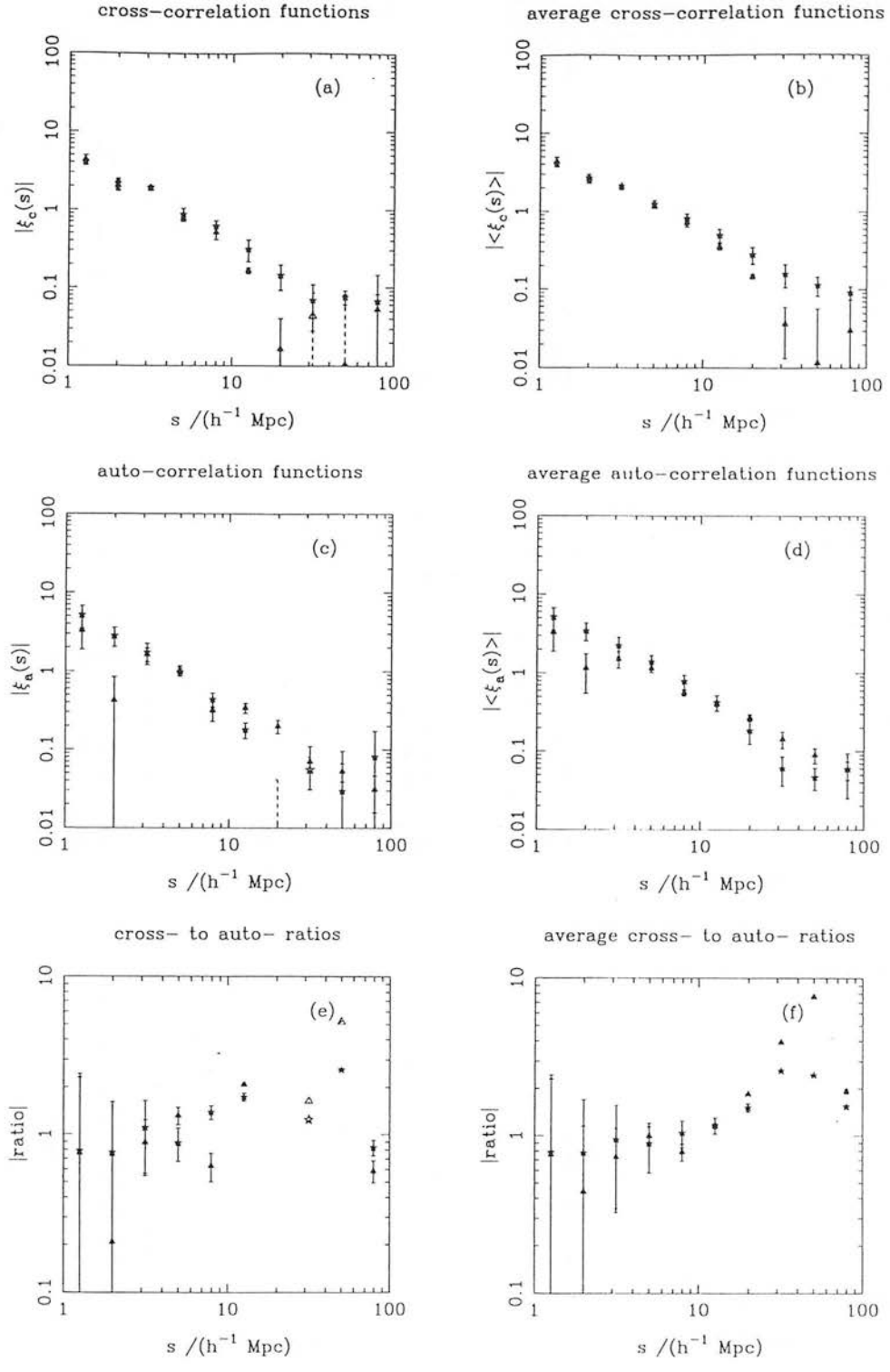


Figure 3.11: Results for w36.6 and c36.6 limited to 500 h^{-1} Mpc, using the minimum variance weighting scheme: in (a) and (b) the triangles and stars denote $\xi^{(1)}_{\rm wc}(s)$ and $\xi^{(2)}_{\rm wc}(s)$; in (c) and (d) the triangles [stars] denote $\xi_{\rm ww}(s)$ [$\xi_{\rm cc}(s)$]; in (e) and (f) the triangles and stars denote the ratios $\xi_{\rm ww}/\xi^{(1)}_{\rm wc}$ and $\xi^{(2)}_{\rm wc}/\xi_{\rm cc}$ respectively. Filled symbols denote positive values of the correlation function and open symbols negative ones.

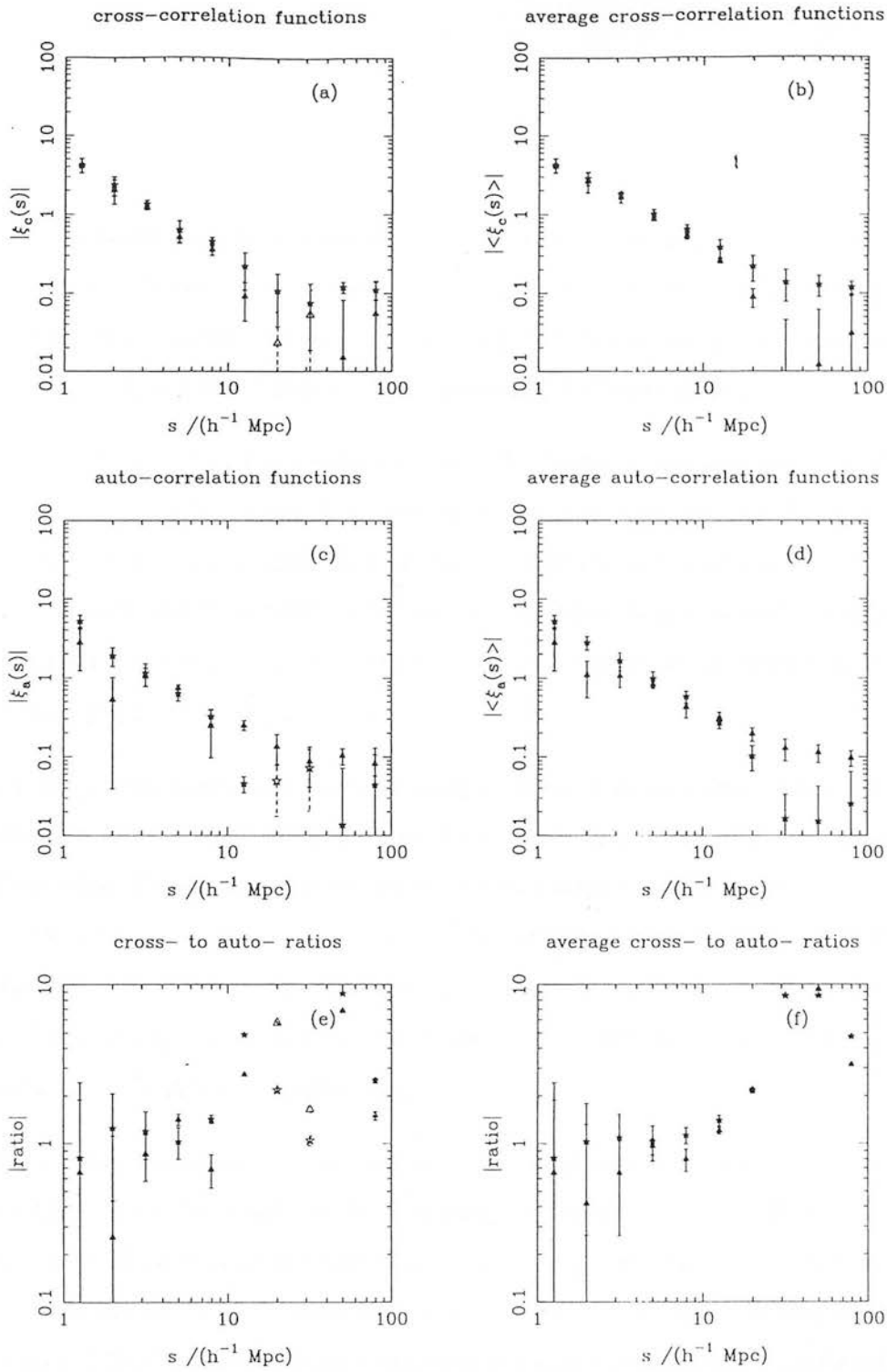


Figure 3.12: Results for w36.8 and c36.8 limited to 500 h^{-1} Mpc, using the minimum variance weighting scheme: in (a) and (b) the triangles and stars denote $\xi^{(1)}_{\rm wc}(s)$ and $\xi^{(2)}_{\rm wc}(s)$; in (c) and (d) the triangles [stars] denote $\xi_{\rm ww}(s)$ [$\xi_{\rm cc}(s)$]; in (e) and (f) the triangles and stars denote the ratios $\xi_{\rm ww}/\xi^{(1)}_{\rm wc}$ and $\xi^{(2)}_{\rm wc}/\xi_{\rm cc}$ respectively. Filled symbols denote positive values of the correlation function and open symbols negative ones.

expect if the warm sample is more strongly clustered in *real space* due to the warm galaxies being in denser regions since, as a result of their higher small-scale velocities, they would suffer a larger reduction in small-separation clustering upon translation to redshift space, while maintaining stronger clustering on larger scales.

Such an interpretation of the results may test the reader's credulity, so we shall summarise our discussion by saying that there is no unambiguous evidence for a difference between the redshift-space clustering strengths of warm and cool samples of QDOT galaxies, but that the data could be interpreted as supporting a hypothesised increase in the clustering strength of the warm population at radial velocities more than $\sim$ 10000 km s^{-1} from the Local Group.

We turn now to the correlation function analysis of our 2 Jy samples. The slopes of the 2 Jy selection functions are steeper above their knees than those of the corresponding QDOT samples. This means that we are unable to compare weighting the 2 Jy samples by the reciprocal of the selection function with using the minimum variance scheme: if we restricted ourselves to samples limited at the maximum distance out to which the reciprocal prescription could be applied then we would be left with an insufficient number of galaxies to perform a meaningful comparison.

We have no alternative but to use only the minimum variance weighting scheme. In Figure 3.13 we show the results for the 2 Jy samples limited to $500\,h^{-1}$ Mpc. They show no evidence at all of a more strongly clustered warm population. From Figure 3.3 we see that there are few 2 Jy galaxies beyond 20,000 km s^{-1}, so that a reader wishing to interpret our QDOT results as supporting the notion of a more strongly clustered warm population could achieve consistency with the 2 Jy results by arguing that the more strongly clustered population only appears beyond $200\,h^{-1}$ Mpc from us. We find that the J_3 computed from applying the minimum variance weighting prescription to our 2 Jy sample is consistent with the model J_3 used in its calculation, validating its use in the production of these results.

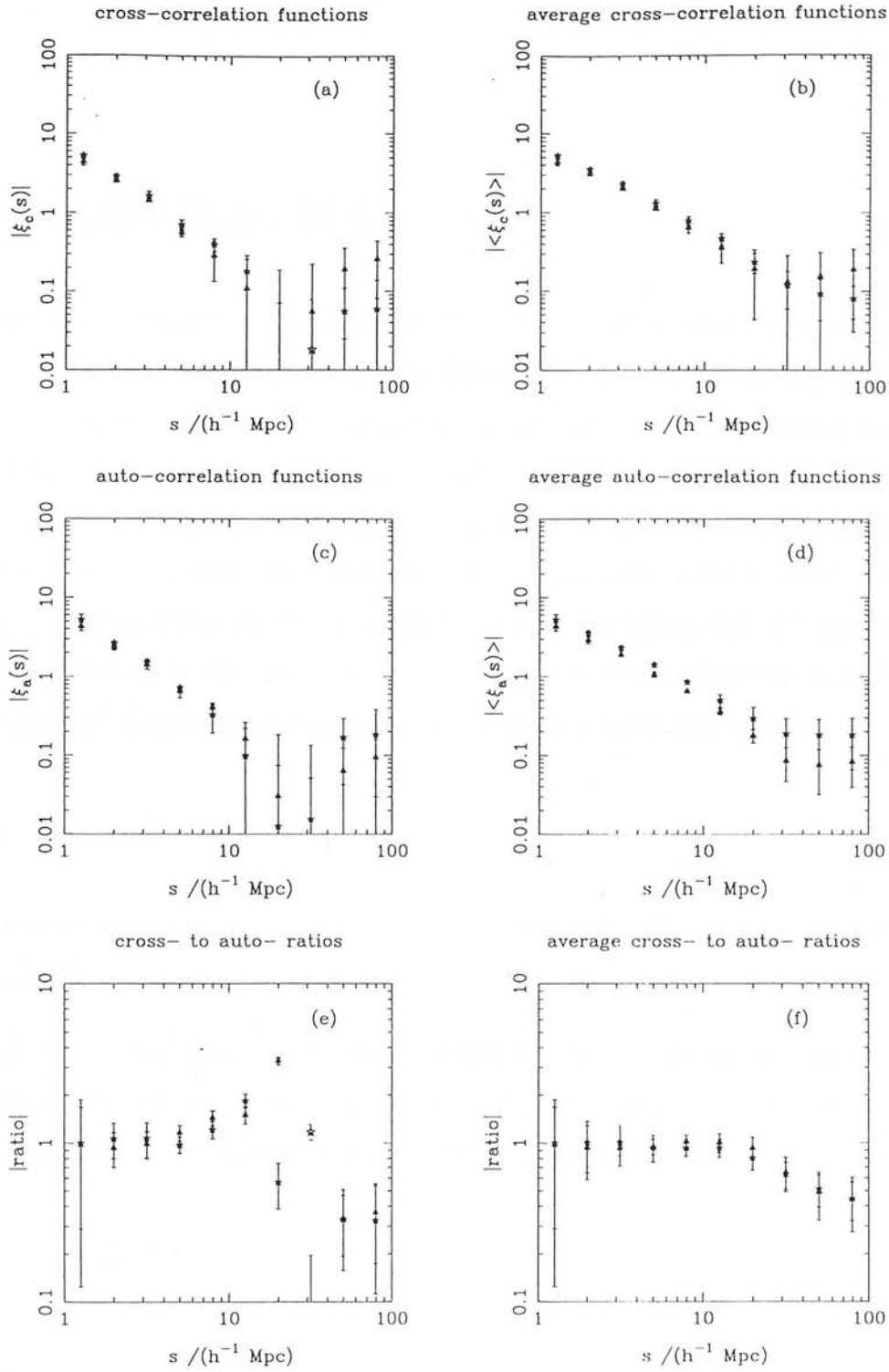


Figure 3.13: Results for w2j36 and c2j36 limited to 500 h^{-1} Mpc, using the minimum variance weighting scheme: in (a) and (b) the triangles and stars denote $\xi^{(1)}_{\rm wc}(s)$ and $\xi^{(2)}_{\rm wc}(s)$; in (c) and (d) the triangles [stars] denote $\xi_{\rm ww}(s)$ [$\xi_{\rm cc}(s)$]; in (e) and (f) the triangles and stars denote the ratios $\xi_{\rm ww}/\xi^{(1)}_{\rm wc}$ and $\xi^{(2)}_{\rm wc}/\xi_{\rm cc}$ respectively. Filled symbols denote positive values of the correlation function and open symbols negative ones.

3.4 COUNTS-IN-CELLS ANALYSIS

An alternative statistic for quantifying the clustering of a sample of galaxies is the variance in counts in cells (Peebles 1980; Efstathiou *et al.* 1990; Loveday *et al.* 1992). The principal advantage that counts-in-cells analysis has over analyses using correlation functions is that, by considering the variance between cells at the same redshift, no account need be taken of the radial decline in the mean density with redshift in a flux-limited sample. Our wish throughout this chapter has been to make a robust comparison of the clustering strengths of the warm and cool subsamples, and the counts-in-cells analysis presented in this section complements the correlation function analysis of the previous section, testing the sensitivity of its results to weighting schemes.

3.4.1 Method

The method of counts-in-cells analysis is ably described by Efstathiou *et al.* (1990) and we shall follow their treatment here.

Consider splitting the volume of space occupied by a flux-limited galaxy redshift sample up into radial shells of width l. These radial shells are then subdivided so as to form roughly cubical cells of volume $V = l^3$. We may define the following pair of quantities:

$$\bar{N} = \frac{1}{M}\sum_i N_i \tag{3.15}$$

and

$$S = \frac{1}{(M-1)}\sum_i (N_i - \bar{N})^2 - \bar{N}, \tag{3.16}$$

where the sums are performed over the M cells in a given radial shell and N_i is the galaxy count in the i^{th} cell in that shell. The expectation values of these are readily shown to be $\langle \bar{N} \rangle = nV$ and $\langle S \rangle = n^2V^2\sigma^2$, where n is the galaxy number density (assumed constant throughout the cell – *i.e.* the cell size is assumed to be smaller than the scale over

which the selection function varies appreciably) and the variance, $\sigma^2(l)$ is related to the autocorrelation function of the sample by

$$\sigma^2(l) = \frac{1}{V^2}\int_{V=l^3} \xi(r_{12})\, dV_1\, dV_2. \tag{3.17}$$

If the cells are independent then the variance in S is given by

$$\mathrm{Var}(S) = \langle S^2\rangle - \langle S\rangle^2 = \frac{\mu_4 - 2\mu_3 + \mu_2 - \mu_2^2}{M}, \tag{3.18}$$

in the limit of large M. In this limit S will approach a Gaussian distribution, by reason of the central limit theorem, with a variance given by equation 3.18, if the cells are independent: this, in turn, requires the cell size to be larger than the coherence length of the galaxy distribution. The quantities μ_2, μ_3 and μ_4 are, respectively, the second, third and fourth central moments of the cell counts. These will, of course, depend on the three- and four-point correlation functions, in general. If the fluctuations are Gaussian then all correlation functions of higher order than the two-point will vanish (see Chapter 1), so, with the assumption of Gaussian statistics, $\mathrm{Var}(S)$ may be shown to take the following simple form:

$$\mathrm{Var}(S) = \frac{2n^2V^2(1+\sigma^2) + 4n^3V^3\sigma^2 + 2n^4V^4\sigma^4}{M}, \tag{3.19}$$

which may be estimated using $\bar{N}$ and S.

The value of $\sigma^2(l)$ may be evaluated in each of the radial shells and these may be combined to yield a determination of $\sigma^2(l)$ for the whole sample which does not depend on the variation in mean galaxy number density with redshift. This is done by maximising the likelihood function

$$\mathcal{L}(\sigma^2) = \prod_i \frac{1}{[2\pi\mathrm{Var}(S_i)]^{1/2}} \exp\left[-\frac{(S_i - n_i^2V^2\sigma^2)^2}{2\mathrm{Var}(S_i)}\right], \tag{3.20}$$

where the product is taken over all the radial shells and the empirical estimates $\bar{N}_i$ are used in place of n_iV: this latter is justified because the errors on $\bar{N}_i$ are much smaller than those on S_i. It is clear from equation 3.20 that the maximum likelihood procedure

automatically assigns weights to shells dependent on the number of cells they contain, as $\mathrm{Var}(s_i) \propto 1/M_i$ from equation 3.18.

The variance on this maximum likelihood estimate of $\sigma^2(l)$ may be determined given the assumption of underlying Gaussian fluctuations, in which case

$$\mathrm{Var}(\sigma^2) = \left[\langle \frac{\partial^2 \ln(\mathcal{L})}{\partial \sigma^2} \rangle\right]^{-1}, \tag{3.21}$$

which, in the limit of large numbers of cells per shell takes the simple form

$$\mathrm{Var}(\sigma^2) \sim \left[\sum_i \frac{M_i}{2} \frac{\bar{N}_i^2}{\sigma^2 + (1 + \bar{N}_i \sigma^2)^2}\right]^{-1}, \tag{3.22}$$

where we take the empirically determined value of $\partial^2 \ln(\mathcal{L})/\partial\sigma^2$ as an estimate of its expectation value.

The foregoing analysis is complicated by the presence of a mask, since parts of some of the cells will lie in the regions of the sky that it excludes. It is necessary, therefore, to compute the unmasked volume of each cell and weight the observed counts according, as a result of which the estimator for S_i becomes

$$S_i = \frac{\sum_j (N_j - V_j \sum_k N_k / \sum_k V_k)^2 - [1 - \sum_k V_k^2 / (\sum_k V_k)^2] \sum_j N_j}{(\sum_j N_j / \sum_j V_j)^2 [\sum_k V_k^2 - 2 \sum_k V_k^3 / \sum_k V_k + (\sum_k V_k^2)^2 / (\sum_k V_k)^2]}, \tag{3.23}$$

where V_k is the unmasked volume of the k^{th} cell: this assumes that the principal effect of the mask is to alter the mean cell count, not the variance on the cell counts. Equation 3.23 differs slightly from the corresponding equation given by Efstathiou *et al.* (1990), due to the presence of a typographical error in that work (A. Taylor, private communication).

3.4.2 Results

A counts-in-cells analysis (conducted by A. Taylor, using software kindly provided by S. Oliver), using the procedure outlined in the previous subsection, was performed on the warm and cool subsamples defined in Section 3.2 and the results are presented below.

In Figures 3.14 and 3.15 (a) we show results from a counts-in-cells analysis performed on the full QDOT survey, to a depth of $300h^{-1}$Mpc. This was calculated using the corrected 1993 version of the QDOT catalogue, but Figure 3.14 shows that the presence of the $\sim$ 200 erroneous redshifts in the earlier version of the catalogue studied by Efstathiou *et al.* (1990) had a negligible effect on the results of their counts-in-cells analysis.

Informed by the results of the correlation function results presented in the previous section, we performed counts-in-cells analyses of warm and cool samples of QDOT galaxies to different limiting depths and selected above different 60 μm flux limits. Figure 3.14 shows that the values of $\sigma^2(l)$ for a given l vary greatly from radial shell to radial shell and each has a large uncertainty associated with it. The situation is even worse when we come to split QDOT into warm and cool samples and reduce the galaxy numbers by considering higher 60 μm flux limits. This leaves us very reliant on the ability of the maximum likelihood mechanism to compute the correct solution for $\sigma^2(l)$. We may try to facilitate this by excluding those shells where the counts are low, but we must inevitably contend with large uncertainties resulting from our small sample sizes.

From Figure 3.15 we see that the cell variances of the warm and cool samples are consistent within their errors for all three sample depths. A similar analysis of samples selected above 60 μm flux limits yields similar results: the $\sigma^2(l) - l$ relations for warm and cool subsamples selected above each flux limit and out to each depth are consistent with each other. The exception to this comes from the smallest cell size (l = 1000 km s^{-1}) in the 20000 and 30000 km s^{-1} samples selected above the 0.6 Jy flux limit, where the cool sample displays stronger clustering. Since this is not seen for the higher flux limit samples, nor seen in the 15000 km s^{-1} sample, despite the fact that there are very few cool galaxies beyond that distance, we must conclude that this result must be due to an erroneous solution being extracted by the maximum likelihood procedure from very noisy data.

From Figure 3.16 we see that a counts-in-cells analysis of warm and cool subsamples of 2 Jy galaxies gives a very similar picture to that from QDOT. The results from all three sample depths show that the cell count variances of the warm and cool 2 Jy galaxies are consistent, except in the 1000 km s^{-1} cells, where the maximum likelihood estimate of

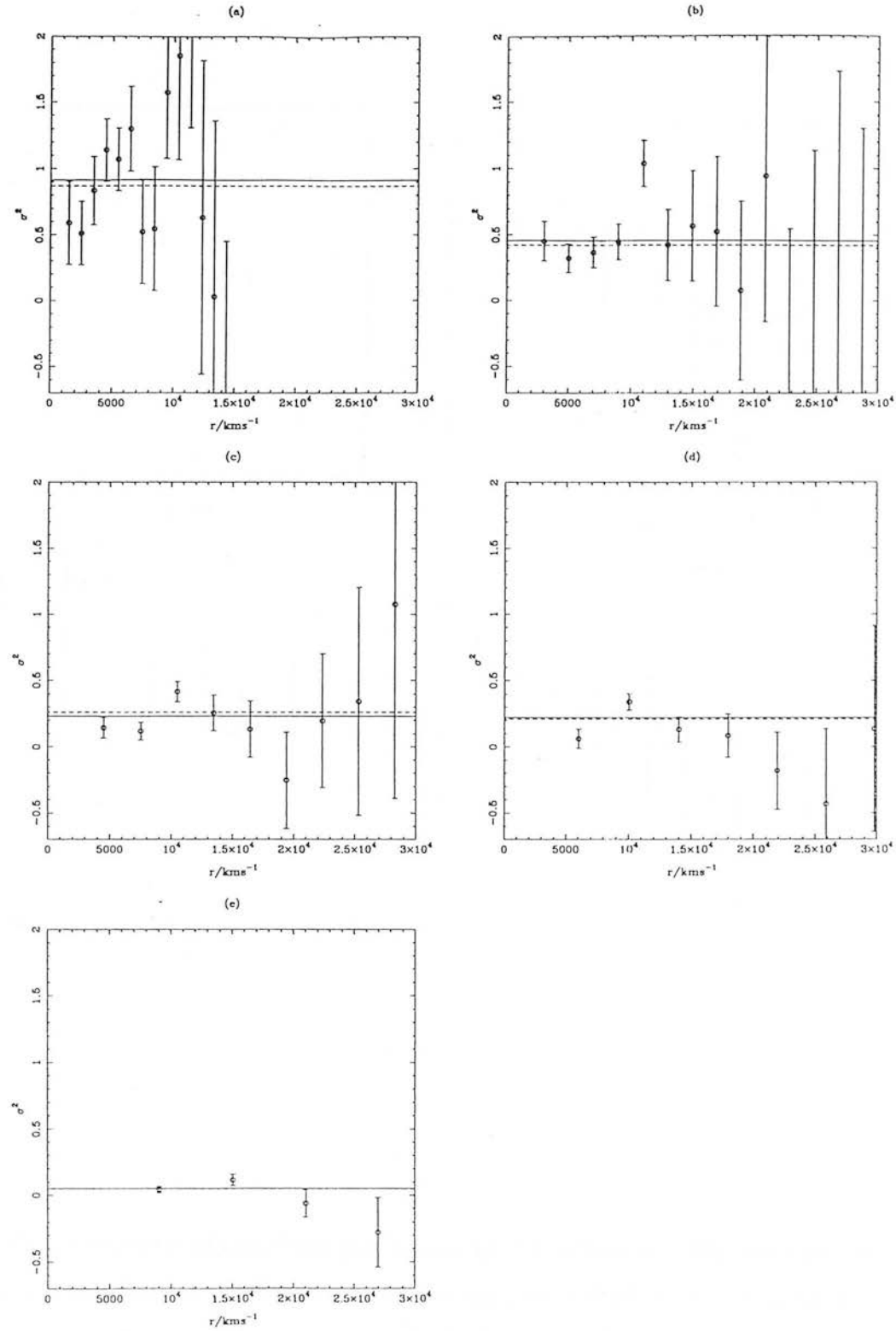


Figure 3.14: Results from a counts-in-cells analysis of the full QDOT survey, to a depth of 30000 km s^{-1}, for cells of size: (a) 1000 km s^{-1} (top left); (b) 2000 km s^{-1} (top right); (c) 3000 km s^{-1} (middle left); (d) 4000 km s^{-1} (middle right); and (e) 6000 km s^{-1} (bottom left). The errors are estimated using equation 3.19. In each plot the solid line denotes the maximum likelihood estimate of $\sigma^2(l)$ and the dashed line the estimate of the same quantity from Efstathiou *et al.* (1990).

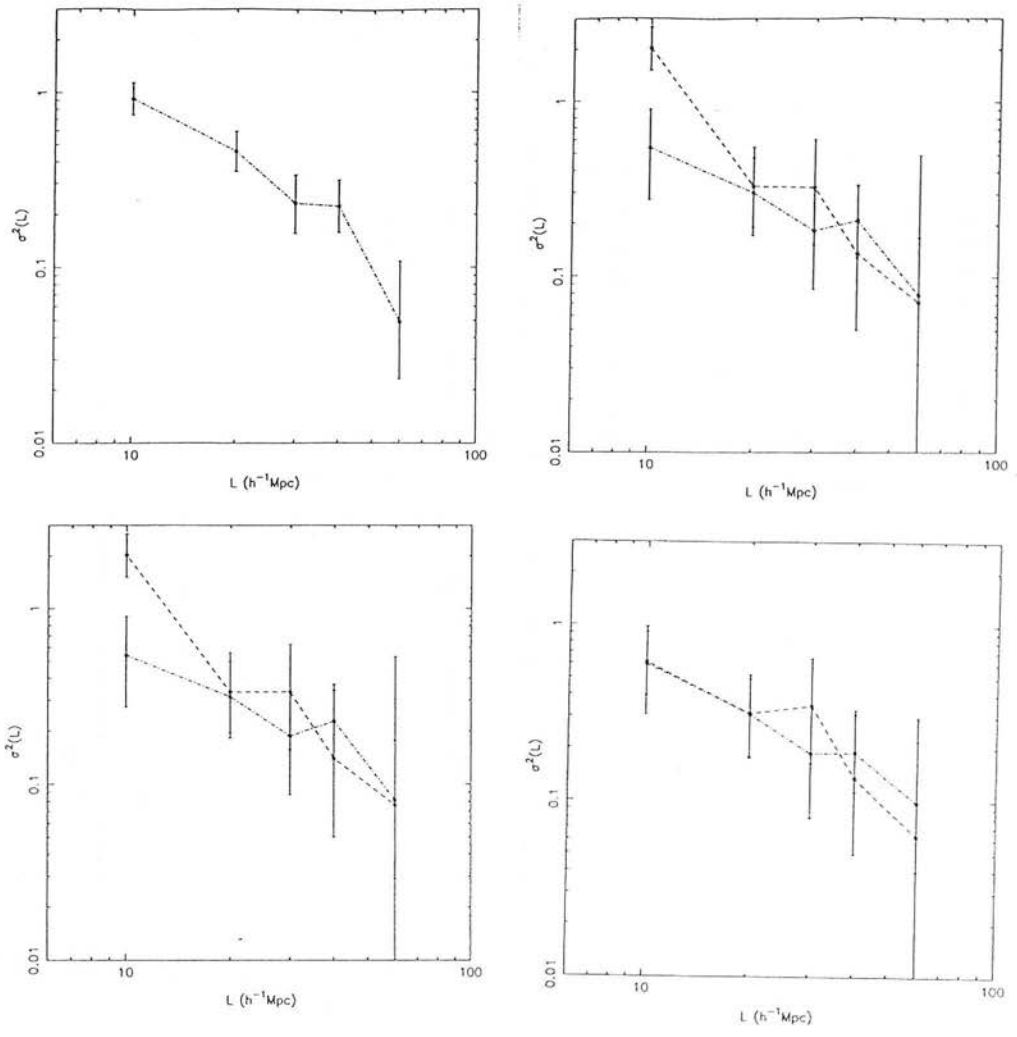


Figure 3.15: Estimated $\sigma^2(l)$ *vs.* l for: (a) the full QDOT survey to a limiting depth of 30000 km s^{-1} (top left); and for w36.6 and c36.6 samples limited to (b) 30000 km s^{-1} (top right), (c) 20000 km s^{-1} (bottom left) and (d) 15000 km s^{-1} (bottom right). In (b), (c) and (d) the dashed line denotes the results from the cool sample and the dot-dashed line those for the warm sample.

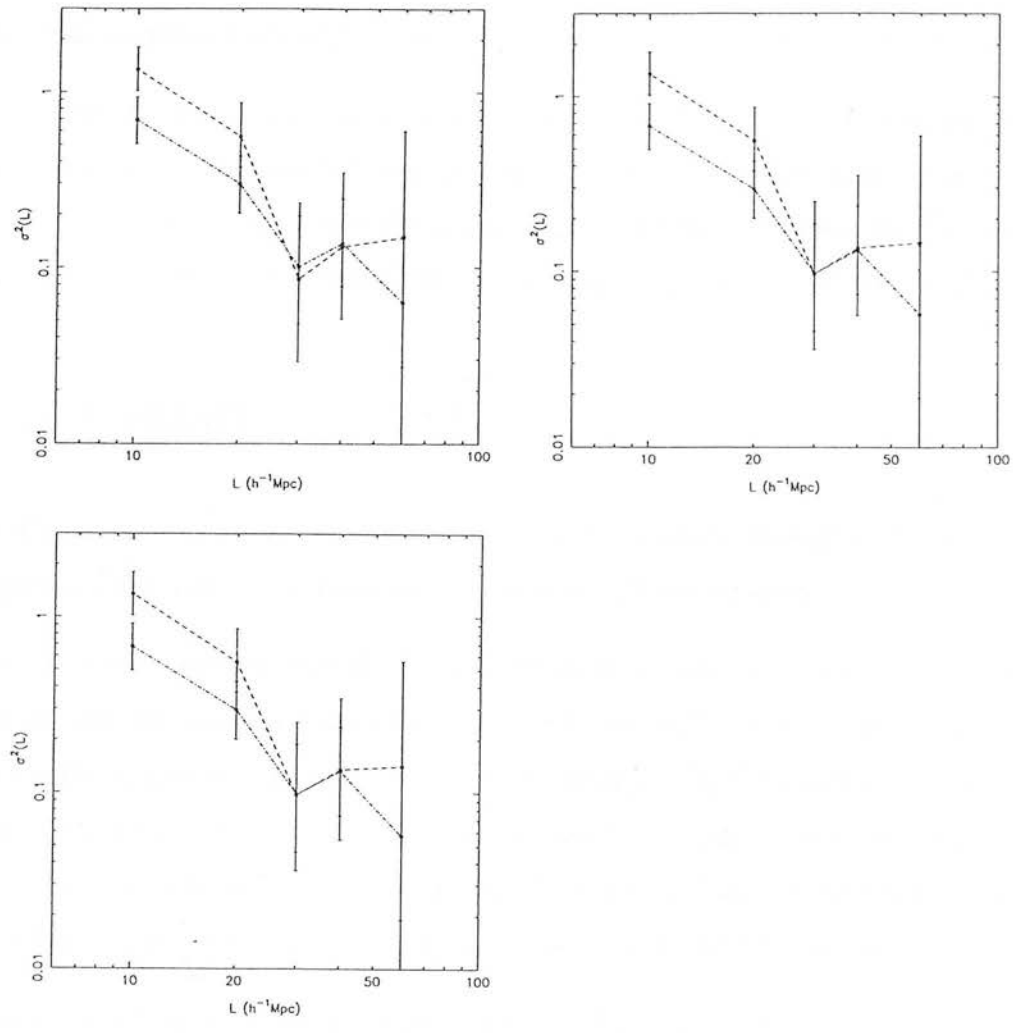


Figure 3.16: Estimated $\sigma^2(l)$ *vs.* l for warm and cool samples from the 2 Jy survey to depths of : (a) 15000 km s^{-1} (top left); (b) 20000 km s^{-1} (top right); (c) 30000 km s^{-1} (bottom). The dashed line denotes the results from the cool sample and the dot-dashed line those for the warm sample.

the cell count variance is slightly higher for the cool sample than for the warm one.

To summarise, we see that a counts-in-cells analysis of warm and cool subsamples of galaxies from the QDOT and 2 Jy surveys does not reveal any difference between the strength of their clustering in redshift space. This conclusion is borne out by results from samples selected above different 60 μm flux limits and out to different depths.

3.5 SUMMARY

In this Chapter we have studied the redshift-space clustering strengths of samples of *IRAS* galaxies selected on the basis of their far-infrared temperature.

Theoretical models predict that the far-infrared temperature of a galaxy (determined from its 60 μm/100 μm flux ratio) may be used to determine whether its far-infrared emission is dominated by: (a) a 'disc' component, produced by the scattering of photons from the ambient interstellar radiation field by dust in the galactic disc; or (b) a 'starburst' component, produced by emission from dust surrounding regions where massive stars are being born. A study of UGC galaxies observed by *IRAS* bears this out.

We have reviewed observational evidence linking starburst phenomena with interactions between galaxies and have argued that interactions are likely to take place preferentially in dense environments, where there are the most galaxies. This assertion led us to invoke the Kaiser mechanism to predict that subsamples of warm *IRAS* galaxies should be more strongly clustered than subsamples of cool ones. Such an effect would introduce a radial clustering strength gradient into flux-limited redshift surveys of *IRAS* galaxies, with serious consequences for cosmological studies using them as tracers of the underlying density field.

To look for such an effect we have constructed warm and cool subsamples of galaxies from the QDOT and 2 Jy catalogues, taking care not to have the samples biased by the variation between 100 μm upper limits assigned to galaxies in different portions of the sky in the absence of their confirmed detection in that band. We have computed the selection functions for the samples and used them in calculating their redshift-space

correlation functions.

We find no difference between the clustering strength of subsamples of warm and cool QDOT galaxies within 100 h^{-1} Mpc of the Local Group: this result is insensitive to the details of the weighting scheme used and to the 60 μm flux limit above which the galaxies are selected. We have found ambiguous evidence for a marginal enhancement in the clustering of the warm QDOT sample beyond 100 h^{-1} Mpc. We have seen no such effect in our 2 Jy samples, but there are few 2 Jy galaxies beyond 200 h^{-1} Mpc, so this may just be setting a limit on the distance beyond which the effect originates.

We have also studied the clustering of our samples using a counts-in-cells analysis, to test further whether our results could be artefacts of the different selection functions of the warm and cool populations. The results of our counts-in-cells analysis are very noisy, but they are consistent with there being no difference between the clustering of the warm and cool populations.

Having led the reader to suspect that there might be a difference in the clustering strengths of subsamples of *IRAS* galaxies selected on the basis of far-infrared temperature, it behoves us to account for our not having seen such a difference: this we shall do in Section 5.2 below.

Chapter 4

GALAXY CLUSTERING WITH EULERIAN BIAS

4.1 INTRODUCTION

In this Chapter we consider the use of galaxies as probes of the large-scale structure of the Universe. Our aim is to determine the set of galaxy clustering models which can provide a picture of large-scale structure consistent with that observed: for our purposes a galaxy clustering model is defined to be the combination of a linear density power spectrum and a physically well-motivated prescription relating the galaxy distribution to the underlying density field.

In pursuit of this aim we proceed as follows:

- we present a novel technique for studying a number of power spectra using results from a single numerical simulation
- we consider an Eulerian bias prescription motivated by N-body/hydrodynamic simulations of galaxy clustering: in doing this we eschew both *ad hoc* algorithms for grouping particles into "galaxies" and Lagrangian bias prescriptions which are likely to be poor at modelling the clustering of galaxy-sized halos

- we use the observed optical galaxy power spectrum in real space to select successful galaxy clustering models
- we construct *IRAS* galaxy distributions for our models through the use of the relation between the observed optical and *IRAS* galaxy distributions
- we compute the redshift-space power spectra for these models and compare them with observations
- we test our models further, by comparing with observations their predictions for the galaxy pairwise velocity dispersion.

This procedure provides us with firm constraints on the shape of the power spectrum of initial density fluctuations and the relation between the galaxy distribution and the underlying density field.

In Section 4.2 we review some of the previous work on this topic which has motivated this project and discuss some of the issues it has raised, concluding with a discussion of the simulations of Cen & Ostriker (1992, 1993), some of whose results we shall use in later sections. Section 4.3 describes the simulations that we ourselves use – these are the high-resolution P^3M CDM simulations run by Bertschinger & Gelb (Bertschinger & Gelb 1991, Gelb 1992, Gelb & Bertschinger 1993) – as well as detailing the novel approach we use in analysing them.

In Section 4.4 we review observational determinations of the power spectrum of optical galaxies and test galaxy clustering models on the basis of comparison with them. Those models which can account for the observed clustering of optical galaxies are then further tested by comparison with the observed power spectrum of *IRAS* galaxies in Section 4.6, using a relationship between the distributions of optical and *IRAS* galaxies deduced from observational data in Section 4.5: we use the redshift space *IRAS* power spectrum, since our work in Chapter 3 shows that, in redshift space at least, the observed clustering of *IRAS* galaxies is unaffected by any hypothesised relation between the emission temperature of an *IRAS* galaxy and the density of the environment in which it finds itself.

By testing galaxy clustering models against the *real space* optical galaxy power spectrum and the *IRAS* galaxy power spectrum in *redshift space* we are also implicitly studying the velocity fields they produce. We investigate these predicted velocity fields more explicitly in Section 4.7, where we compare the small-scale pairwise velocity dispersions of our successful galaxy clustering models with observational determinations of this quantity. Section 4.8 summarises the work presented in this Chapter, while a more general discussion of the issues it raises is postponed until Chapter 5.

4.2 SIMULATING GALAXY CLUSTERING

A great deal of effort has been devoted to numerical studies of galaxy clustering over the past decade or so. Most of these studies conform to a particular methodology. The starting point in this procedure is a model for the material constituents of the universe (usually CDM and/or HDM, plus a certain quantity of baryons) and a description of the primordial density fluctuations (usually adiabatic and obeying Gaussian statistics, with a scale-invariant power spectrum): the basic principle here is that these choices should be motivated by, and, to some extent, falsifiable by, particle physics.

Numerical integration of the perturbation equations produces the transfer function for the particular cosmogony, which yields the linear power spectrum of density fluctuations expected at early times (see Chapter 1). A realisation of this density field is then produced, for example by using the Zeldovich Approximation (see Chapter 2) to displace a (large) number of particles from positions on a three-dimensional grid in a manner consistent with the computed linear power spectrum. The evolution of this distribution of particles is then followed, using some suitable means to calculate the gravitational forces exerted on each particle (see Chapter 1 and Section 4.3 for more details). This process is continued until a time to be identified with the present (set, for example, by the value of σ_8), at which point the properties of the predicted galaxy distribution is deduced by some plausible means (to be discussed below) from the final distribution of particles and compared with observational data in a battery of tests, the results of which are used to judge the merits of the particular cosmogony.

The epitome of this approach is the series of papers (Efstathiou *et al.* 1985, Davis *et al.* 1985, Frenk *et al.* 1985, White *et al.* 1987a, White *et al.* 1987b, Frenk *et al.* 1988, Frenk *et al.* 1990) by the so-called *'Gang of Four'* (M. Davis, G. Efstathiou, C.S. Frenk, S.D.M. White), principally concerned with the development of the CDM model. The prime virtue of this approach is that it is well specified: its adherents claimed that its sole free parameter was the fluctuation amplitude. Another element of its appeal is the grandeur of its scope, unifying the microscopic world of particle physics with the vast scales of cosmology, reflecting the role that the interactions between fundamental particles in the early universe played in shaping the large-scale structure we observe today. The great successes of the CDM model developed by the *Gang of Four* have been reviewed elsewhere (*e.g.* Frenk 1991, Davis *et al.* 1992) and one must concur with Ostriker (1993) that there *"is no question about the historical utility of the CDM scenario. The existence of this model sharpened the questions to be asked of both observers and theoreticians."*

The dramatic observational results yielded by a number of large galaxy surveys over the past few years (*e.g.* Maddox *et al.* 1990a-c, Efstathiou *et al.* 1990, Loveday *et al.* 1992, Dalton *et al.* 1992) have not only overtaken the CDM model (at least in its classic form, as advanced by the *Gang of Four*) in the sense of indicating that its power spectrum is of the wrong shape (although see Bower *et al.* 1992 for an alternative view) but have also facilitated a new approach to studying large-scale structure. The Luddite theorist may respond to the great increase in the quantity and quality of the observations of large-scale structure by desperately trying to shore up a cherished model against the deluge of reality, but it is better to acknowledge that the study of large-scale structure is now observationally-driven and to adopt a more appropriate methodology. Such a methodology should be phenomenological in its outlook, relying on observational data (where possible) rather than theoretical prejudice to decide its course.

The determination of the amplitude of primordial fluctuations by *COBE* enables the new phenomenological method to be applied to the very N-body simulations that epitomise the philosophy of the 1980s. We saw in Chapter 1 that the transfer function of a pure CDM model features only one characteristic scale, set by the horizon scale at matter-

radiation equality and corresponding (Efstathiou 1990) to a wavelength

$$\lambda_{\rm eq} = \frac{ct_{\rm eq}a_0}{a(t_{\rm eq})} \simeq 10(\Omega_0 h)^{-1}\, h^{-1}{\rm Mpc}, \tag{4.1}$$

where the scaling with $(\Omega_0 h)^{-1}$ holds if the baryon density is sufficiently small that Silk damping is negligible. The action of gravitation imposes no further characteristic length scale on the density field during its evolution so, on scales larger than the spatial resolution of the simulation, all distances in the simulated density field continue to scale with $(\Omega_0 h)^{-1}$ at later times.

This may seem to be of limited utility, since inflation-inspired prejudice and the comparison of density and peculiar velocity fields by the *POTENT* method (see below for details) lead us to desire $\Omega_0 = 1$, while one of the early *Gang of Four* papers (Davis *et al.* 1985; DEFW) showed that for CDM to reproduce the observed strength of galaxy clustering on large scales the Hubble constant cannot greatly exceed 50 km s^{-1} Mpc^{-1} – *i.e.* $h = 0.5$. The great advantage of this freedom to scale simulation results arises if we regard equations 1.88 and 2.69 not as necessarily giving the transfer function for a CDM universe with particular values of Ω_0 and h, but, simply, as a convenient parametric form for the transfer function of a model that is scale-invariant on large scales and in which the power spectrum, $\Delta^2(k)$, flattens off at a scale determined by $\Omega_0 h$, which is now not to be regarded as the product of two cosmological quantities, but rather as a shape parameter – call it Γ^*.

The value of Γ^* labels members of a family of power spectra with transfer functions given by equations 1.88 and 2.69. Our phenomenological outlook leads us to regard all members of this family on an equal footing, as we have no pet particle physics model leading us to a particular value of Γ^*. It may seem, therefore, that we are left with the unattractive prospect of running a multitude of simulations, spanning a broad range of Γ^* values, and taking a long time to run. In fact, the very scaling that led us to consider the Γ^* parameter allows us to avoid such tedium.

Consider the set of different outputs from a numerical simulation of, say, a standard (*i.e.* $\Omega_0 h = 0.5$) CDM model with $\Omega_{\rm B} = 0$. At each output time, the simulation records the positions and velocities of each particle in units of, say, the length of the simulation

volume and the Hubble velocity associated with that distance, respectively. The size of the simulation box may be expressed as a certain multiple of λ_{eq} and, thus, the self-similarity of the simulation in these units may be invoked to scale the simulation results to yield those which would obtain in a CDM universe with a different value of Γ^* from that assumed by the simulators. Note that the amplitude of fluctuations in the simulation (expressed in terms of, say, σ_8 or ϵ: see Chapter 1 for definitions of these quantities) will be changed by this scaling, but in a way that is readily calculable.

In this way, we may identify the earlier epochs of our $\Omega_0 h = 0.5$ CDM simulation as the present epoch (defined, say, by the *COBE* ϵ) in cosmogonies with shape parameters $\Gamma^* < 0.5$. We may, therefore, study a number of cosmogonies using the results from a single $\Omega_0 h = 0.5$ CDM simulation although, clearly, we are restricted to the values of Γ^* which correspond to the output times of the simulation (for a given value of ϵ to define the present epoch) and we cannot use the very earliest outputs in the simulations, for which the particle distribution bears the mark of the grid of initial positions.

This procedure enables us to study a number of differently-shaped power spectra at the computational expense of running only one simulation. By adopting it we are also decoupling the question of the form of the initial power spectrum of density fluctuations from the question of what set of material constituents in the early universe produced it. This is in keeping with our phenomenological outlook: we let the observations of large-scale structure tell us the shape of the initial power spectrum and only thereafter do we speculate as to its origin.

The problem remains, however, that N-body codes simulate the dark matter distribution, not the distribution of the baryons which will go to produce the galaxies whose properties we wish to use to judge our cosmological models. Given the current lack of knowledge as to how galaxies form it is necessary to employ some plausible, but necessarily *ad hoc*, prescription to deduce the properties of the galaxy distribution from those of the dark matter distribution. A number of prescriptions have been used in previous work and we review some of them here.

One simple approach would be to follow the analogy with the study of cluster correlations presented in Chapter 2 and locate local maxima (above an appropriate threshold) in the

initial Gaussian density field (smoothed on an appropriate scale) and tag the particle nearest each maximum as a site of galaxy formation. These 'peak particles' may be followed through the course of the simulation and their distribution at the final output be compared with that of galaxies at redshift zero. Gelb & Bertschinger (1993) show that roughly half of the galaxy-sized halos they identify (using their DENMAX algorithm – see below) at redshift zero in their large, high-resolution CDM simulations contain no particles tagged in the initial density field following the prescription of Frenk *et al.* (1988): *i.e.* smoothing with Gaussians with filter radii of $R_s = 550$ and 800 kpc (comoving) and selecting peaks above thresholds of $\nu_s = 3.0$ and 2.5 respectively. It appears, therefore, that peak particles are poor tracers of galaxy-sized halos in the evolved density field and so should not be used (*e.g.* Park 1991) as a way of avoiding the requirement for good spatial resolution in the trade-off with maximising the simulation volume.

In simulations with sufficient resolution to have typically hundreds of particles per galaxy-sized halo, it is conventional to define halos using the so-called *'Friends-of-friends'* (FOF) algorithm. The FOF algorithm recursively links particles within a certain distance, b, called the linking length, typically taken to be $\sim 10\% - 20\%$ of the mean inter-particle separation. The justification (DEFW) for this is that such a prescription effectively corresponds to defining halos with overdensities above a certain threshold, the value of which is set by the linking length. The success of an FOF algorithm with a particular linking length varies with environment: a short linking length will miss some halos in underdense regions, while a longer one will fail to resolve halos in denser environments, joining together what the eye would regard as two distinct objects. This latter has a particularly serious consequence for the computation of small-scale velocity statistics and we shall discuss the possibility of its being the origin of 'velocity bias' in Section 4.7. To overcome this, some workers (*e.g.* Suto, Cen & Ostriker 1992; Suginohara & Suto 1992) have used an adaptive FOF algorithm whose linking length is set by the local density. The basic weakness of FOF remains, however: particles are linked on the basis of proximity, not physical association – *i.e.* there is no requirement that the halos selected by FOF are actually bound objects at all. The FOF algorithm has been adopted as something of an 'industry standard' by numerical simulators, with the unfortunate consequence that results obtained with it are all too often accepted as fact without a

proper investigation as to their sensitivity to the method of group-selection used. For example, the work of Suginohara & Suto (1992) and Gelb (1992) shows that the halo mass functions determined by the Press–Schechter (Press & Schechter 1974) formalism only match those found in N-body simulations if an FOF algorithm with $b \simeq 0.2$ is used and that discrepancies result from the use of equally plausible and/or higher resolution algorithms: this is serious, as it is this supposed agreement with N-body results that provides the pragmatic motivation for the widespread use of the conceptually-flawed Press-Schechter formalism. Perhaps more seriously, Suginohara & Suto (1992) show that the rotation curves of dark matter halos are very sensitive to the algorithm used to group particles into them and suggest that the production by the *Gang of Four* of flat rotation curves for the halos in their CDM simulations — one of the great triumphs of their CDM model — could simply be an artefact of their particular FOF algorithm and not reflect anything inherent to the CDM model at all.

These results must make one wary of using *ad hoc* algorithms, however plausible they may seem at first sight, and motivate one to adopt a more physical approach to finding groups of particles in N-body simulations. An algorithm developed in this spirit is DENMAX (Bertschinger & Gelb 1991, Gelb 1992, Gelb & Bertschinger 1993). In DENMAX, the density field is evaluated on a fine grid (typically 512^3) and particles are left to move in this fixed gravitational field to one of the local maxima. The particles are labelled to denote which maximum they end up in and all particles are returned to their true positions. Finally, the particles assigned to each maximum are analysed, to determine the largest set of them which constitute a bound object, by removing unbound particles. This "unbinding" process has the benefit of destroying small halos which arise solely as a result of Poisson fluctuations, as well as removing particles transiently in the proximity of a true, bound halo, with obvious consequences for the internal velocity dispersions of the halos.

DENMAX is well motivated physically, but it is very expensive computationally. One might also argue that any galaxy identification scheme should be implemented at an earlier time, corresponding to the epoch of galaxy formation. Such an approach is followed by Couchman & Carlberg (1992), who identify as "galaxy precursors" those particles lying within a minimum overdensity contour of 125 at $z = 3$. They then

construct "galaxies" at redshift zero by joining together those particles previously tagged as "galaxy precursors" which finish up in groups of a certain size that have overdensities greater than ~ 3000. This method is found to be rather inefficient – only $\sim 20\%$ of the "galaxy precursors" finish up in "galaxies" – but it selects a statistically similar halo population as the simulations of Carlberg, Couchman & Thomas (1990), which includes a dissipative component followed using the methods of Smooth Particle Hydrodynamics (SPH).

The prescriptions outlined above may vary in their sophistication, but none of them can overcome the fundamental objection that they do not analyse the baryon distribution, which is what is required for a realistic treatment of the properties of galaxies: for example, the high-resolution simulations of Bertschinger & Gelb, analysed using DENMAX, may provide an accurate description of the merging and evolution of dark matter halos but they can say nothing about the formation of galaxies through the cooling of baryons in the potential wells of the halos, nor about how such proto-galactic baryonic bodies respond to the interaction of their parent dark halos.

To study the baryon distribution directly requires the introduction of dissipative processes into numerical simulations. Advances in computing are now making such simulations possible, although the authors of the first reports on these mixed N-body/hydro codes (Katz, Hernquist & Weinberg 1992; Cen & Ostriker 1992, 1993) readily acknowledge that the subject is in its infancy and that their results should be treated with caution. The simulation by Cen & Ostriker (1992, 1993; hereafter COI, COII respectively) of a CDM universe incorporating a hydrodynamic treatment of the baryonic component and a *"heuristic but plausible"* [COI] prescription for galaxy formation produces particularly appealing results and we shall concentrate on their work in what follows.

At the heart of the Cen & Ostriker study (see COI, COII and references therein for further details) lies the simulation of an $80\,h^{-1}$ Mpc cube of 200^3 dark matter particles evolved using a PM code and a 200^3 cell Eulerian hydrodynamic code (based on the aerodynamics code of Jameson 1989) which computes (in a spatially-averaged fashion) the principal line and continuum atomic processes for an H/He plasma of primordial composition, through the modelling of the radiation field from 1 eV to 100 keV.

The following prescription is used by COI to model galaxy formation. At each time step the cells are analysed and flags raised marking those in which the baryonic component satisfies the following four criteria:

- dense: $\delta\rho_{\rm b}/\bar{\rho}_{\rm b} > \eta$
- contracting: $\nabla \cdot \boldsymbol{v} < 0$
- rapidly cooling: $t_{\rm cool} < t_{\rm dyn} \equiv \sqrt{\frac{3\pi}{32G\rho_{\rm t}}}$
- Jeans unstable: $m_{\rm b} > m_{\rm J} \equiv G^{-3/2}\rho_{\rm b}^{-1/2}c^3[1 + (\bar{\rho}_{\rm d}/\bar{\rho}_{\rm b})(\delta\rho_{\rm d}/\bar{\rho}_{\rm d})/(\delta\rho_{\rm b}/\bar{\rho}_{\rm b})]^{-3/2}$

where $\rho_{\rm b}, \rho_{\rm d}$ are, respectively, the baryonic and dark matter densities; $\rho_{\rm t} \equiv \rho_{\rm b} + \rho_{\rm d}$ is the total density; $m_{\rm b}, m_{\rm J}$ are the baryonic and Jeans masses of the cell; c is the sound speed; G is Newton's gravitational constant; and the cooling time, $t_{\rm cool}$, is defined (*e.g.* Rees & Ostriker 1977) to be $t_{\rm cool} = 3k_{\rm B}T/\Lambda(T)n$, where n is the number density of the gas and $\Lambda(T)$ is the cooling function. Cen & Ostriker use a density threshold of $\eta = 5.5$ and find that varying it affects the timing of galaxy formation events somewhat, but has little impact on the properties of the galaxy distribution at the present. The gas in cells flagged as satisfying the four criteria listed above is assumed to be heading towards inevitable collapse, with the resultant formation of a stellar system. This procedure is effected by a prescription inspired by the classic Eggen, Lynden-Bell & Sandage (1962, ELS) model for galaxy formation, assuming a simple relation between the galaxy formation time scale and the dynamical free-fall time. Cen & Ostriker compute the gas mass $\Delta m_{\rm b} = m_{\rm b}\Delta t/t_{\rm dyn}$ that would collapse in the next time step Δt and remove such an amount from the gas mass associated with the cell, replacing it by a particle of mass $m_* = \Delta m_{\rm b}$, placed at the centre of the cell and given the mean velocity of the gas in the cell. Each particle so created is given three labels: its mass (m_*), the redshift (z_*) of its formation and the free-fall time ($t_{\rm dyn}$) of its cell at redshift z_*. The distribution of these particles is followed in a dissipationless fashion thereafter, on an equal footing with the dark matter particles. At a later time, galaxies are formed by grouping these created particles using an adaptive FOF algorithm.

The conversion of the collapsed baryonic material into stars is (again following ELS) taken to be spread over a dynamical time, so that the mass $\Delta m_{\rm SF}$, converted into stars

in time step Δt following creation of a particle of mass m_* at time t_* is given by

$$\Delta m_{\rm SF} = m_* \left(\frac{\Delta t}{t_{\rm dyn}}\right) \left[\frac{(t-t_*)}{t_{\rm dyn}}\right] \exp\left[-(t-t_*)/t_{\rm dyn}\right]. \tag{4.2}$$

Star formation will result in the injection of energy into the interstellar medium from both supernovæ and the UV output from massive young stars. The supernova input rate is modelled as

$$\frac{\Delta E_{\rm SN}}{\Delta t} = \left(\frac{\Delta m_{\rm SF}}{\Delta t}\right) c^2 \epsilon_{\rm SN}, \tag{4.3}$$

where $\epsilon_{\rm SN} = 10^{-4.5}$ and this energy is added to the ISM locally, spread out using a Gaussian of width $\sigma = 1.2\,h^{-1}$ Mpc. The UV input rate is taken to be

$$\frac{\Delta E_\nu}{\Delta t} = \left(\frac{\Delta m_{\rm SF}}{\Delta t}\right) c^2 \epsilon_{\rm UV} g_\nu, \tag{4.4}$$

where $\epsilon_{\rm UV} = 10^{-4.0}$ and g_ν is the normalised spectrum of a young stellar association (Scalo 1986). This input is not added locally, but to the ambient radiation field: this must be considered as a weakness of the Cen & Ostriker method, since UV radiation, like the thermal input from supernovæ, would be expected to inhibit local galaxy formation (*e.g.* Bower *et al.* 1993).

In COII, Cen & Ostriker outline the results of this galaxy formation prescription. They find that galaxy formation peaks in the interval $z \simeq 3-2$ (with star formation peaking later, at $z \simeq 2-1$), but that there is a significant amount of galaxy formation as early as $z \simeq 8-5$ and that these oldest galaxies end up as the most massive systems by $z = 0$, due to a combination of higher masses at formation and addition of mass through merging. They investigate the variation of location (at $z = 0$) of galaxies as a function of age, finding that the youngest quartile of galaxies comprise a relatively uniform field population, the second quartile shows evidence for filamentary structure, the third quartile is found principally in filaments and the rich clusters at their intersections and the oldest quartile is found only in rich clusters. By looking at the distribution of the oldest galaxies at early times it is found that this density-age relationship is produced by galaxies having formed in the filaments and then drifting along them to the clusters at

their intersections. The observed correlation between galaxy mass and age leads Cen & Ostriker to identify the youngest 50% of their galaxies as spirals, the next 25% as S0s and the oldest 25% as ellipticals, following which they find a morphology-density relationship very reminiscent of that observed by Dressler (1980) and, consequently, a variation in the clustering strength with morphology like that observed by Davis & Geller (1976). One may question the identifications made by Cen & Ostriker between certain sorts of halo and certain classes of galaxy, but the picture they present is very appealing and certainly augurs well for the use of such N-body/hydro codes in the future.

Perhaps the most striking result in COI is the surprisingly tight relation between the density field at redshift zero (smoothed with a Gaussian of some filter radius) and the local galactic number density field (smoothed in the same way) which Cen & Ostriker deduce. In COII they give a simple fit to this relationship as being

$$\log_{10}\left(\frac{\rho_{\rm g}}{\langle\rho_{\rm g}\rangle}\right) = A + B\log_{10}\left(\frac{\rho_{\rm t}}{\langle\rho_{\rm t}\rangle}\right) + C\left[\log_{10}\left(\frac{\rho_{\rm t}}{\langle\rho_{\rm t}\rangle}\right)\right]^2, \tag{4.5}$$

where $\rho_{\rm g}, \rho_{\rm t}$ are, respectively, the galaxy density and total mass density fields at $z = 0$ and the values of A,B & C vary with the filter radius of the Gaussian used to smooth the fields. Cen & Ostriker note that this relationship does not reflect information about the process of galaxy formation so much as the combination of formation and subsequent movement: the most overdense regions may not be the best places for galaxies to form (as a results of their high virial gas temperatures, for example) but they will be the places to which many galaxies will drift, following their formation elsewhere.

The tightness of this $\rho_{\rm g} - \rho_{\rm t}$ relation motivates the use of an Eulerian bias scheme, where one deduces the present number density of galaxies at a certain point in space from its local mass density through a $\rho_{\rm g} - \rho_{\rm t}$ relation. Coles (1993) has recently considered Eulerian bias models in general terms and in this Chapter we perform a more detailed analysis, concentrating on the constraints placed on such models by the requirements of consistency with observations of galaxy clustering. Since the basic premise of this scenario is that the galaxy number density depends on the *total* mass density, we need only consider dissipationless simulations, so long as we are concerned with mass models where the dominant mass component is dissipationless.

4.3 DETAILS OF THE NUMERICAL SIMULATIONS

In this Chapter we use P^3M CDM simulations kindly provided by E. Bertschinger and J. Gelb: further details of the simulations are given in Gelb (1992), while scientific results deduced from them are presented by Bertschinger & Gelb (1991) and Gelb & Bertschinger (1993).

We use two sets of simulations, namely:

- "P3M": these are simulations of a box of comoving size 2560 km s^{-1} and contain 64^3 particles. Forces are evaluated using a standard Plummer (1911) potential with softening length $\epsilon = 40$ kpc and with the long-range component computed on a 128^3 grid. A Hubble constant of $H_0 = 50$ km s^{-1} Mpc^{-1} is assumed and a Holtzman (1989) CDM transfer function with $\Omega_0 = 1$ and 5% baryons is used. The mass of each particle is $3.52 \times 10^{10} M_\odot$. *N.B.* these simulations have the same physical parameters as the "CDM12" simulations discussed by Gelb (1992) and Gelb & Bertschinger (1993) but were run on different machines and had initial conditions determined by a different set of random numbers.

- "CDM16": these are simulations of a box of comoving size 5000 km s^{-1} and contain 144^3 particles. Forces are again computed using a Plummer potential, this time with a smoothing length of $\epsilon = 65$ kpc, and with the long-range, particle-mesh forces calculated on a 288^3 grid with (in a modification of the "adaptive P^3M" method of Couchman 1991) a finer (420^3) grid being used in dense regions. A zero-baryon CDM transfer function from Bardeen *et al.* (1986, BBKS) was used, again assuming $\Omega_0 = 1$ and $h = 0.5$. The mass of each particle in these simulations is $2.3 \times 10^{10} M_\odot$.

The Plummer (1911) model gives the force on a particle of mass m at a distance r from another particle of mass m as being $\boldsymbol{F}(\boldsymbol{r}) = Gm^2\boldsymbol{r}/(r^2 + \epsilon^2)^{3/2}$: a non-zero value of ϵ is needed to prevent two-body relaxation and the production of tightly-bound pairs of particles. Since this force is the sum of particle-particle and particle-mesh terms in a P^3M code, it follows that the potential used to compute the particle-particle forces

is not exactly of the Plummer form. In what follows we shall assume that there is a negligible difference between the fits obtained by Holtzman (1989) and by BBKS to the transfer function of a CDM model *of the same composition*: we shall, however, have to account for the 5% baryon fraction in the P3M simulations - Mann *et al.* (1993) give the scaling of the shape parameter, Γ^*, with baryon fraction as being well approximated by $\Gamma^* = \Omega h\, \mathrm{e}^{-2\Omega_{\mathrm{B}}}$ for values of Ω_{B} consistent with SBBN (see Chapter 1). *N.B.* our Γ^* parameter differs slightly from the shape parameter, Γ, introduced by Efstathiou, Bond & White (1992), which they define as being $\Gamma = \Omega_0 h$ despite the fact that they use a CDM transfer function with non-zero Ω_{B}.

We consider data from a number of output times, but wish to treat them, however, as being the present epoch (defined by a *COBE* fluctuation amplitude of $\epsilon = 2.9 \times 10^{-5}$) in simulations of $\Omega_0 = 1$ models with shape parameters differing from the original $\Gamma^* = 0.5\, \mathrm{e}^{-2\Omega_{\mathrm{B}}}$. Our choice of $\Omega_0 = 1$ is motivated by two considerations: (a) the desire to achieve consistency with the inflationary prediction of a spatially flat universe (consistent also with the *COBE* CBR anisotropy observations) without having to introduce a cosmological constant; and (b) the comparison performed by Dekel *et al.* (1993) of the local smoothed density field of *IRAS* 2 Jy galaxies and the local velocity field reconstructed using the *POTENT* (Bertschinger & Dekel 1989) procedure, which yields $\Omega_0^{0.6}/b_{\mathrm{I}} = 1.28^{+0.75}_{-0.59}$ (at the 95% confidence level), which favours $\Omega_0 \sim 1$ in the absence of a good physical reason why the linear bias factor, b_{I}, for *IRAS* galaxies should be significantly less than unity.

We must, therefore, determine the value of Γ^* corresponding to each value of σ_8: to avoid confusion we shall denote by σ_{SIM} the value of σ_8 in each data set's original guise as an $\Omega_0 h = 0.5$ CDM simulation. This may be done as follows, starting from the definition of the variance, $\sigma^2(R)$, in spheres of radius R:

$$\sigma^2(R) = \int_0^\infty \Delta^2(k) W^2(kR) \frac{\mathrm{d}k}{k}, \tag{4.6}$$

where $W(kR)$ is the window function for a spherical top hat: $W(x) = 3(\sin x - x\cos x)/x^3$. If we concern ourselves only with power spectra which are scale-invariant on large scales, then we know from Chapter 1 that the power spectrum, $\Delta^2(k)$, may be written in terms

of the transfer function, $T(k)$, as

$$\Delta^2(k) = \frac{4\epsilon^2}{9}\left(\frac{c}{H_0}\right)^4 k^4\, T^2(k), \tag{4.7}$$

where the normalisation constant, ϵ, may be determined from the *COBE* CBR anisotropy. In that case, equation 4.6 becomes

$$\sigma^2(R) = \frac{4\epsilon^2}{9}\left(\frac{c}{H_0}\right)^4 \int_0^\infty k^3\, T^2(k)\, W^2(kR)\, \mathrm{d}k. \tag{4.8}$$

Recalling that $T(k) \equiv T(k/\Gamma^*)$, we make a change of variables $x \equiv k/(\Gamma^*\, h^{-1}\ \mathrm{Mpc})$, whence equation 4.8 becomes

$$\sigma_R^2 = \frac{4\epsilon^2}{9}(3000)^4\, \Gamma^{*4} \int_0^\infty x^3 T^2(x)\, W^2(4R\Gamma^* x)\, \mathrm{d}x. \tag{4.9}$$

So, if we express R in terms of $(\Gamma^*)^{-1}h^{-1}$ Mpc (equivalent to expressing it in terms of λ_{eq}) the integral on the rhs of equation 4.9 remains constant under changes of Γ^*. Evaluating this integral then gives us

$$\Gamma^* = 0.455\, \sigma_{\mathrm{SIM}}^{1/2} \tag{4.10}$$

as the relation between the σ_8 value of an output in the original $\Gamma^* = 0.5$ simulation and the Γ^* value of the cosmogony that would yield the same numerical value for $\sigma(R = 4/\Gamma^*\, h^{-1}\ \mathrm{Mpc})$ in a *COBE*-normalised universe today.

Having determined the appropriate value of Γ^* for each output, we must then scale the simulation volumes accordingly. For example, the CDM16 simulations have comoving size 5000 km s^{-1}, corresponding to $25/\Gamma^*\, h^{-1}$ Mpc. So, the output with $\sigma_{\mathrm{SIM}} = 0.31$ becomes a simulation of size 100 h^{-1} Mpc of a $\Gamma^* = 0.25$ cosmogony. Not only are we, therefore, able to use the earlier outputs of a $\Gamma^* = 0.5$ simulation to represent the present in a number of cosmogonies with $\Gamma^* < 0.5$, but the comoving volumes of these models simulated are larger than those of the original simulation.

In Table 4.1 we list the simulation outputs studied, for each of which we give: the value of σ_{SIM}; the resultant value of Γ^*; and the original and final, scaled simulation sizes.

Simulation	Original Size [h^{-1} Mpc]	$\sigma_{\rm SIM}$	Γ^*	Scaled Size [h^{-1} Mpc]
P3M.11	25.6	0.20	0.18	64.0
P3M.12	25.6	0.30	0.23	50.0
P3M.13	25.6	0.40	0.26	44.3
P3M.14	25.6	0.45	0.28	41.1
P3M.15	25.6	0.50	0.29	39.7
P3M.16	25.6	0.55	0.31	37.2
P3M.17	25.6	0.60	0.32	36.0
P3M.18	25.6	0.65	0.33	34.9
P3M.19	25.6	0.68	0.34	33.9
CDM16.31	50.0	0.31	0.25	100.0
CDM16.39	50.0	0.39	0.28	89.3

Table 4.1: Parameters for the data sets used.

4.4 THE POWER SPECTRUM OF OPTICAL GALAXY CLUSTERING

The first stage of our procedure is to determine the set of galaxy clustering models which are consistent with the level of clustering observed in large optical galaxy catalogues: we do that in this Section.

4.4.1 Observational Data

Several determinations of the power spectrum of optical galaxy clustering have been published in recent years. Vogeley *et al.* (1992) analysed two volume-limited samples of galaxies from the CfA redshift survey (Huchra *et al.* 1992), which is now complete to Zwicky magnitude $m_{\rm B} = 15.5$ in eleven 6° slices. Loveday *et al.* (1992) determined the power spectrum of the Stromlo/APM redshift survey. This is a 1-in-20 sparse-sampled survey of bright ($b_{\rm J} < 17.15$) galaxies from the APM Galaxy Survey (Maddox *et al.* 1990a-c). Baugh & Efstathiou (1993; BE93) determined the *real space* power spectrum of the APM survey from its angular correlation function.

We shall summarise the work of BE93, as we shall principally be using their results here, since they provide the tightest available constraint on the power spectrum of optical galaxy clustering, due to the higher signal-to-noise that can be obtained from the use of a very large galaxy catalogue in projection rather than the three-dimensional distribution of a much smaller redshift sample drawn from it. BE93 use an iterative method (Lucy 1974) to invert an integral equation relating the angular correlation function to the three-dimensional real space power spectrum: this integral equation is the power spectrum analogue of Limber's equation (Limber 1953) and may be written in the form:

$$w(\varpi) \;=\; \int_0^\infty P(k)\, k\, g(k\varpi)\, \mathrm{d}k, \tag{4.11}$$

where $w(\varpi)$ is the angular correlation function for pairs of galaxies separated by angle $\theta = 2\sin^{-1}(2\varpi)$ on the sky, $P(k)$ is the three-dimensional real space power spectrum as a function of comoving wavenumber k and the kernel $g(k\varpi)$ is given by

$$g(k\varpi) \;=\; \frac{1}{2\pi}\frac{1}{(\mathcal{N}\Omega_{\rm s})^2}\int_0^\infty \frac{F}{(1+z)^\alpha}\left(\frac{\mathrm{d}N}{\mathrm{d}z}\right)^2\left(\frac{\mathrm{d}z}{\mathrm{d}x}\right) J_0(k\varpi x)\mathrm{d}z, \tag{4.12}$$

where the evolution of the power spectrum is modelled as

$$P(k,z) \;=\; \frac{P(k,0)}{(1+z)^\alpha}, \tag{4.13}$$

and $\mathcal{N}$ is the surface density of galaxies in the catalogue, which subtends a solid angle $\Omega_{\rm s}$. The quantities F and x are related to the metric by

$$\mathrm{d}s^2 \;=\; c^2\mathrm{d}t^2 \;-\; (\mathrm{d}x^2/F^2(x) \;-\; x^2\mathrm{d}\theta^2 \;+\; x^2\sin^2\theta\,\mathrm{d}\phi^2), \tag{4.14}$$

and $J_0(y)$ is the $0^{\rm th}$ order Bessel function of the first kind. A parametric form for the differential redshift distribution $\mathrm{d}N/\mathrm{d}z$ is chosen so as to be consistent with both (at bright magnitudes) the luminosity function of the Stromlo/APM redshift survey (Loveday *et al.* 1992) and (at faint magnitudes) the redshift distributions of the deep surveys of Broadhurst *et al.* (1988) and Colless *et al.* (1990, 1993).

Equation 4.11 may be inverted by Lucy's iterative method (Lucy 1974) to yield the power spectrum $P(k)$, given a parametric fit to $w(\varpi)$. We use the results of BE93 with $\alpha = 0$.

4.4.2 Method

Our aim in this Section is to determine candidate galaxy clustering models capable of accounting for the APM power spectrum determined by BE93. The data sets we use are the full 64^3 P3M particle distributions and 100,000–particle sparse samples from the CDM16 simulations: we found empirically that this size samples the distributions adequately for our purposes. From each data set we construct an optical galaxy distribution, by assigning a weight, $w_{\rm O}$, to each of its particles, according to the following prescription:

$$\log_{10}(w_{\rm O}) \;=\; c_1 \log_{10}\left(\frac{\rho}{\langle\rho\rangle}\right) \;+\; c_2 \left[\log_{10}\left(\frac{\rho}{\langle\rho\rangle}\right)\right]^2, \tag{4.15}$$

where ρ is the value in the original simulation (*i.e.* *not* in the sparse sample for the CDM16 data) of the local density field evaluated on a 128^3 grid and smoothed with a Gaussian of filter radius $1\,h^{-1}$ Mpc to approximate filtering on galaxy mass scale: $\langle\rho\rangle$ is the mean value of ρ on the grid. This prescription is intended to be the analogue of the Cen & Ostriker relation of equation 4.5, given that it is more direct for us to implement a prescription weighting particles in the data set than one relating smoothed density fields.

For each data set, we consider a grid of points on the (c_1, c_2) plane. For each of the optical galaxy clustering models specified by these (c_1, c_2) pairs, we assign the set of weights $\{w_{\rm O}\}$ to the particles in the data set and compute the power spectrum of this weighted particle distribution.

The first step in this procedure is to divide the simulation volume up into a number (we use 32^3) equal-volume cubic cells. A grid is formed by the set of points at the bottom left-hand corners of the cell and to each grid point is assigned the fraction of the total weight of the simulation to be found in that cell:

$$i.e.\;\; \frac{\sum_{\text{particles i in cell j}} w_{\rm O}(i,j)}{\sum_{\text{cells j}} \sum_{\text{particles i in cell j}} w_{\rm O}(i,j)}. \tag{4.16}$$

The derivation of the power spectrum of a distribution of weighted particles proceeds as follows (Peebles 1980). Consider dividing the volume containing the distribution of

particles into infinitesimal cells of volume $\{\delta V_i\}$, such that the occupancy, n_i, of each cell is either one or zero. The amplitude, $\delta_{\boldsymbol{k}}$, of the $\boldsymbol{k}^{\rm th}$ mode of the Fourier transform of this distribution is then given by

$$\delta_{\boldsymbol{k}} = \frac{\sum_i n_i w_i \, \mathrm{e}^{i\,\boldsymbol{k}\cdot\boldsymbol{r}_i}}{\sum_i w_i}, \tag{4.17}$$

where the sum is evaluated over the cells.

It follows that

$$(\sum_i w_i)^2 \, \langle \delta_{\boldsymbol{k}} \delta_{\boldsymbol{k}'} \rangle = \sum_i \langle n_i^2 w_i^2 \rangle \, \mathrm{e}^{i\,(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}_i} + \sum_{i\neq j} \langle n_i n_j w_i w_j \rangle \, \mathrm{e}^{i\,(\boldsymbol{k}\cdot\boldsymbol{r}_i - \boldsymbol{k}'\cdot\boldsymbol{r}_j)} \tag{4.18}$$

but $\langle n_i n_j w_i w_j \rangle = (\sum_i w_i)^2 \delta V_i \delta V_j [1 + \xi(r_{ij})]$, where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$, from the definition of the two-point correlation function. Setting $\boldsymbol{k}' = -\boldsymbol{k}$ in equation 4.18 then yields

$$\langle |\delta_{\boldsymbol{k}}|^2 \rangle = \frac{(\sum_i w_i^2)}{(\sum_i w_i)^2} + \frac{1}{V} \int \mathrm{d}^3\boldsymbol{r} \; [1 + \xi(\boldsymbol{r})] \; \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{4.19}$$

Since the power spectrum, $P(k)$, is related to the correlation function, $\xi(\boldsymbol{r})$, by (see Chapter 1)

$$P(k) = \frac{1}{V} \int \mathrm{d}^3\boldsymbol{r} \, \xi(\boldsymbol{r}) \, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}, \tag{4.20}$$

it follows that the power spectrum, $P(k)$, is related to the quantity $\langle |\delta_{\boldsymbol{k}}|^2 \rangle$ computed from the distribution of weighted particles by

$$P(k) = \langle |\delta_{\boldsymbol{k}}|^2 \rangle - \frac{(\sum_i w_i^2)}{(\sum_i w_i)^2} : \tag{4.21}$$

i.e. $(\sum_i w_i^2)/(\sum_i w_i)^2$ is the shot-noise term which must be subtracted from $\langle |\delta_{\boldsymbol{k}}|^2 \rangle$ to yield the estimate of the power spectrum of the weighted density field.

The power spectrum thus determined is taken to be that of the optical galaxy clustering model. Interpolating between these binned data yields estimates of the power spectrum

of the model at the wavenumbers corresponding to the centres of ten of the wavenumber bins used by BE93 in their determination of the APM power spectrum. For each candidate we compute the quantity χ^2, defined by

$$\chi^2 = \sum_{i=1}^{10} \left[\frac{y_{\rm obs}(k_i) - y_{\rm cand}(k_i)}{\sigma(k_i)} \right]^2 . \qquad (4.22)$$

In equation 4.22, $y_{\rm obs}(k_i) = \log_{10}[\Delta^2_{\rm obs}(k_i)]$, where $\Delta^2_{\rm obs}(k_i)$ is the BE93 estimate for the APM power spectrum in the bin centred on wavenumber k_i; $y_{\rm cand}(k_i) = \log_{10}[\Delta^2_{\rm cand}(k_i)]$, where $\Delta^2_{\rm cand}(k_i)$ is the interpolated estimate of the power spectrum of the candidate at wavenumber k_i; and $\sigma(k_i)$ is the estimated error on $y_{\rm obs}(k_i)$.

4.4.3 Results

From these results we determine the (c_1, c_2) pair that yields the lowest value of χ^2 (which we denote by $\chi^2_{\rm min}$) for each data set and produce a contour plot of $\Delta\chi^2 \equiv \chi^2 - \chi^2_{\rm min}$ on the (c_1, c_2) plane. In Figure 4.1 we show the resulting plot for the data set CDM16.31: the three contour levels marked are $\Delta\chi^2 = 2.3, 4.6$ and 6.0, which correspond, respectively, to the 68%, 90% and 95% confidence limits for the (c_1, c_2) pair. The interpretation of this plot is straightforward: for a given positive value of c_1 it is always possible to find a negative value of c_2 that will produce a weighted particle distribution with a power spectrum consistent with that of APM galaxies. The result of this trade-off between the linear and quadratic terms in equation 4.15 produces, for the most part, a weighted particle distribution in which the particle with the highest $\rho/\langle\rho\rangle$ does not have the highest $w_{\rm O}$ weight and the particle with the lowest $\rho/\langle\rho\rangle$ does not have the lowest $w_{\rm O}$ weight – *i.e.* the trade-off is effected by under-weighting the most overdense and over-weighting the most underdense regions of the simulations.

Even acknowledging our lack of understanding about the processes of galaxy formation, it is difficult to imagine how a physical process could yield the values $c_1 = 13.0, c_2 = -9.0$, for example. We must introduce some further condition of "reasonableness" to constrain the allowed region on the (c_1, c_2) plane and prevent what must be regarded as the unphysical excesses of the trade-off between the linear and quadratic terms. The choice

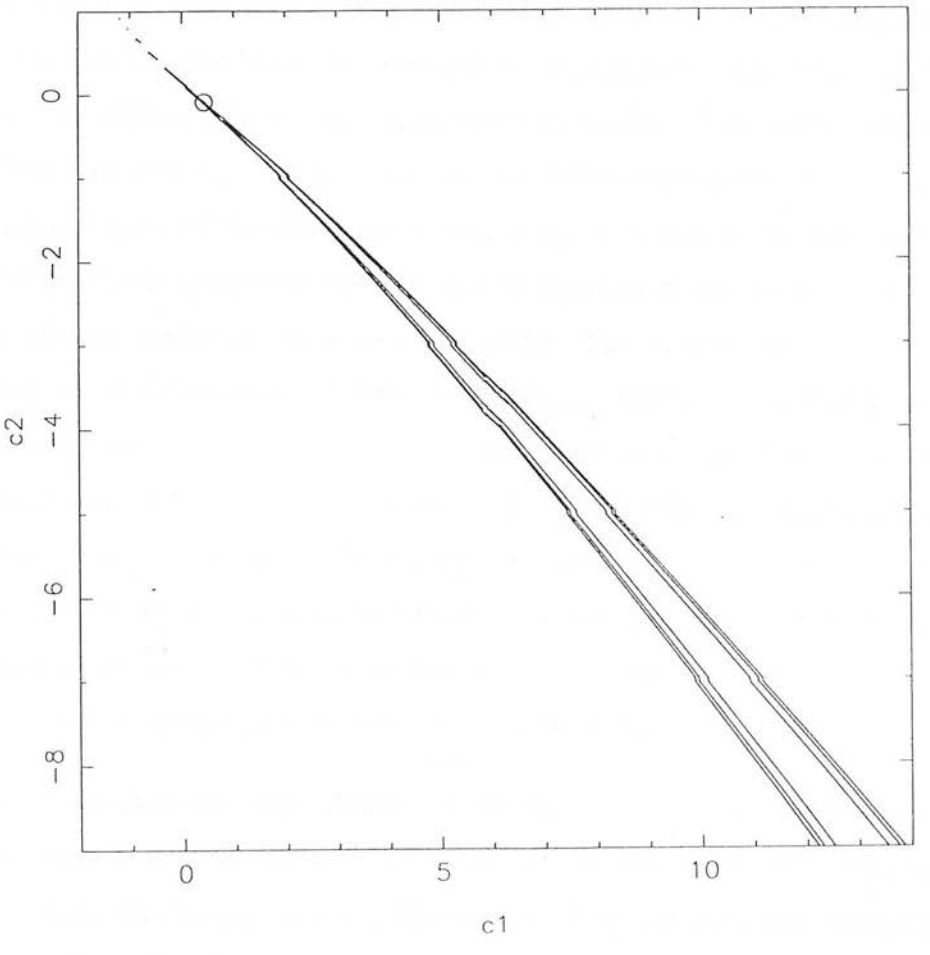


Figure 4.1: The χ^2 contour plot in the (c_1, c_2) plane for data set CDM16.31. The contour levels marked are $\Delta\chi^2 = 2.3$, 4.6 and 6.0 corresponding, respectively, to the 68%, 90% and 95% confidence limits for the (c_1, c_2) pair, while the dots mark the models allowed at the 95% level.

of such a constraint must be somewhat subjective, but we have no choice other than to select a plausible constraint and determine its consequences.

One obviously reasonable constraint would be to require that the weight w_O must be a monotonically-increasing function of the local density field ρ, in accordance with one's belief that galaxies are more likely to be found where the density is higher: the situation may well be more complicated for *IRAS* galaxies, due to gas-stripping by the intra-cluster medium and interaction-induced starbursting like that discussed in Chapter 3, but this should hold for optical galaxy samples like those under consideration here. From equation 4.15 we see that this condition corresponds to the inequality $c_1 + 2c_2 \log_{10}(\rho\,/\langle\rho\rangle) > 0$ for the range of $\log_{10}(\rho\,/\langle\rho\rangle)$ probed by the simulations. Imposing this constraint greatly reduces the allowed region of parameter space and, indeed, leaves no allowed models for data sets P3M.15-19. This may be readily explained by considering the transformation which the $w_O - (\rho/\langle\rho\rangle)$ relation has to effect to produce a weighted distribution with a power spectrum consistent with that of APM galaxies. For reasonable (by our definition above) values of the (c_1, c_2) pair, the weighting procedure has a greater effect on the amplitude of the power spectrum than on its slope: the slope of the mass power spectra for data sets P3M.15-19 are steeper than that of APM galaxies on the same scales, so it requires an unreasonable $w_O - (\rho\,/\langle\rho\rangle)$ relation to flatten the slope in the way required to produce accord with the APM power spectrum.

In Figure 4.2 we show the final allowed regions on the (c_1, c_2) plane: the dots denote the allowed models and the three contour lines are, as before, the 68%, 90% and 95% confidence limits for the (c_1, c_2) pair. The straight lines delineate the constraint $c_1 + 2c_2 \log_{10}(\rho\,/\langle\rho\rangle) = 0$ for the maximum and minimum values of $\log_{10}(\rho\,/\langle\rho\rangle)$ found in the simulation: the density range varies from $-1.076 \leq \log_{10}(\rho\,/\langle\rho\rangle) \leq 1.581$ for P3M.11 to $-1.569 \leq \log_{10}(\rho\,/\langle\rho\rangle) \leq 2.021$ for P3M.14 — the range increasing, as expected, with the amount of small-scale power in the density field (*i.e.* with increasing Γ^*).

Figure 4.3 shows contour plots of $\rho_O/\langle\rho_O\rangle$ against w_O for two models from Figure 4.2 (b). The left-hand panel shows an allowed model ($c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$), while the right-hand panel shows a model ($c_1 = 1.041, c_2 = -0.457$) which does not satisfy the $c_1 + 2c_2 \log_{10}(\rho\,/\langle\rho\rangle) > 0$ constraint. The behaviour shown in the right-hand panel is

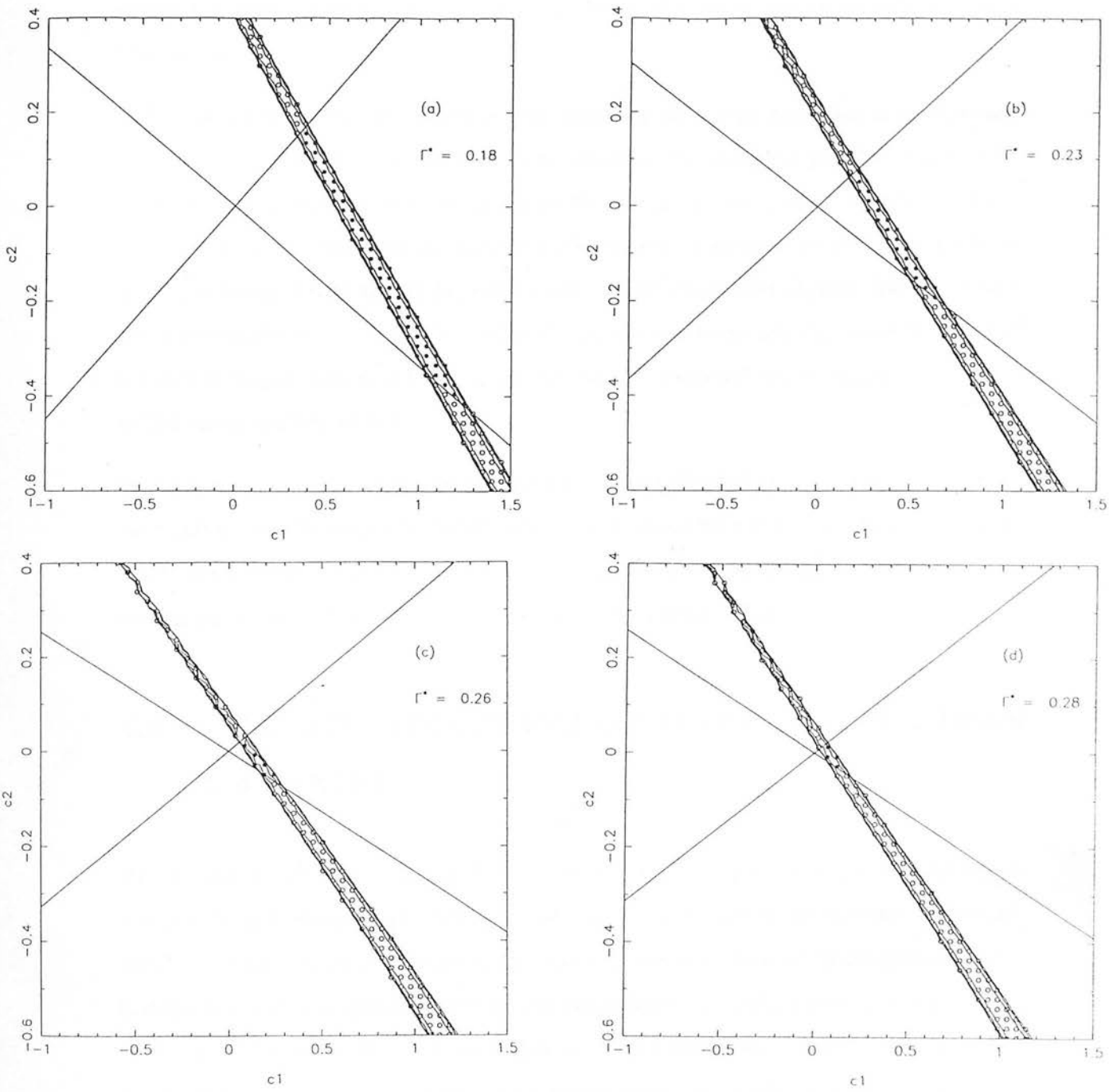


Figure 4.2: The allowed regions on the (c_1, c_2) plane for: (a) P3M.11; (b) P3M.12; (c) P3M.13; and (d) P3M.14. The circles denote the models with $\Delta\chi^2 < 6.0$: filled dots mark models which satisfy the constraint $\mathrm{d}\log_{10}(w_{\mathrm{O}})/\mathrm{d}\log_{10}(\rho/\langle\rho\rangle) > 0$, while empty circles mark those which do not.

generic to such models: the contours are concave, turning back on themselves to underweight the most overdense regions – as c_1 increases, this behaviour becomes increasingly pronounced.

In Figure 4.4 we show the optical power spectrum for a representative allowed model ($\Gamma^* = 0.23, c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$, as shown in the left-hand panel of Figure 4.3): the filled circles show the optical galaxy power spectrum predicted by the model, while the stars show the observational data from BE93. Other allowed models show a similar level of agreement with the APM power spectrum, although over slightly different scales: the wavelengths at which we determine the power spectrum are expressed in terms of the size of the simulation volume, so the interval in wavenumber over which we compute $\Delta^2(k)$ varies slightly with Γ^*.

We obtain very similar results from P3M.14 and and CDM16.39 throughout this procedure and the results from CDM16.31 follow closely those from P3M.13. We consider that ample evidence that the smaller volumes of the P3M simulations are perfectly adequate for our purposes and we shall consider only them in what follows.

4.5 THE DISTRIBUTIONS OF OPTICAL AND IRAS GALAXIES

Having found, for each value of Γ^* considered, the locus on the (c_1, c_2) plane where acceptable optical galaxy clustering models lie, we now wish to see whether these candidates can also account for the redshift-space power spectrum of *IRAS* galaxies. This is done in a way analogous to that for optical galaxies: we assign a set of weights $\{w_{\rm I}\}$ to the particles in the data set, such that we are to regard the power spectrum of this weighted distribution as being the *IRAS* power spectrum predicted by the galaxy clustering model.

We wish to deduce the set of *IRAS* weights $\{w_{\rm I}\}$ from the set of optical weights $\{w_{\rm O}\}$ through the use of a $w_{\rm O} - w_{\rm I}$ relation derived from observational comparisons of the distributions of optical and *IRAS* galaxies. A number of such studies (Babul & Postman 1990; Strauss *et al.* 1992; Oliver 1993; Hudson 1993) have been published in recent years,

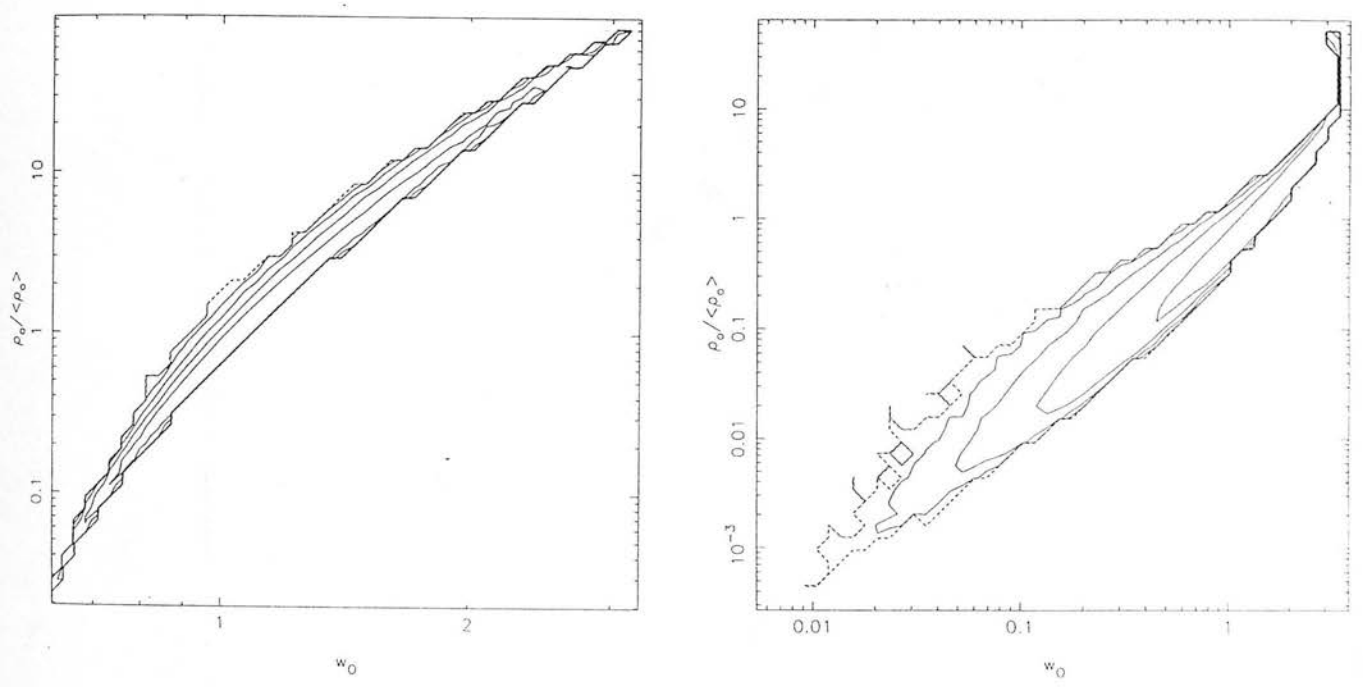


Figure 4.3: Contour plots of $\rho_O/\langle\rho_O\rangle$ against w_O for two P3M.12 models which give satisfactory agreement with the APM power spectrum. The left-hand panel shows the allowed model ($c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$), while the right-hand panel shows a model ($c_1 = 1.041, c_2 = -0.457$) which does not satisfy the $c_1 + 2c_2 \log_{10}(\rho/\langle\rho\rangle) > 0$ constraint. The contours are linear in surface density.

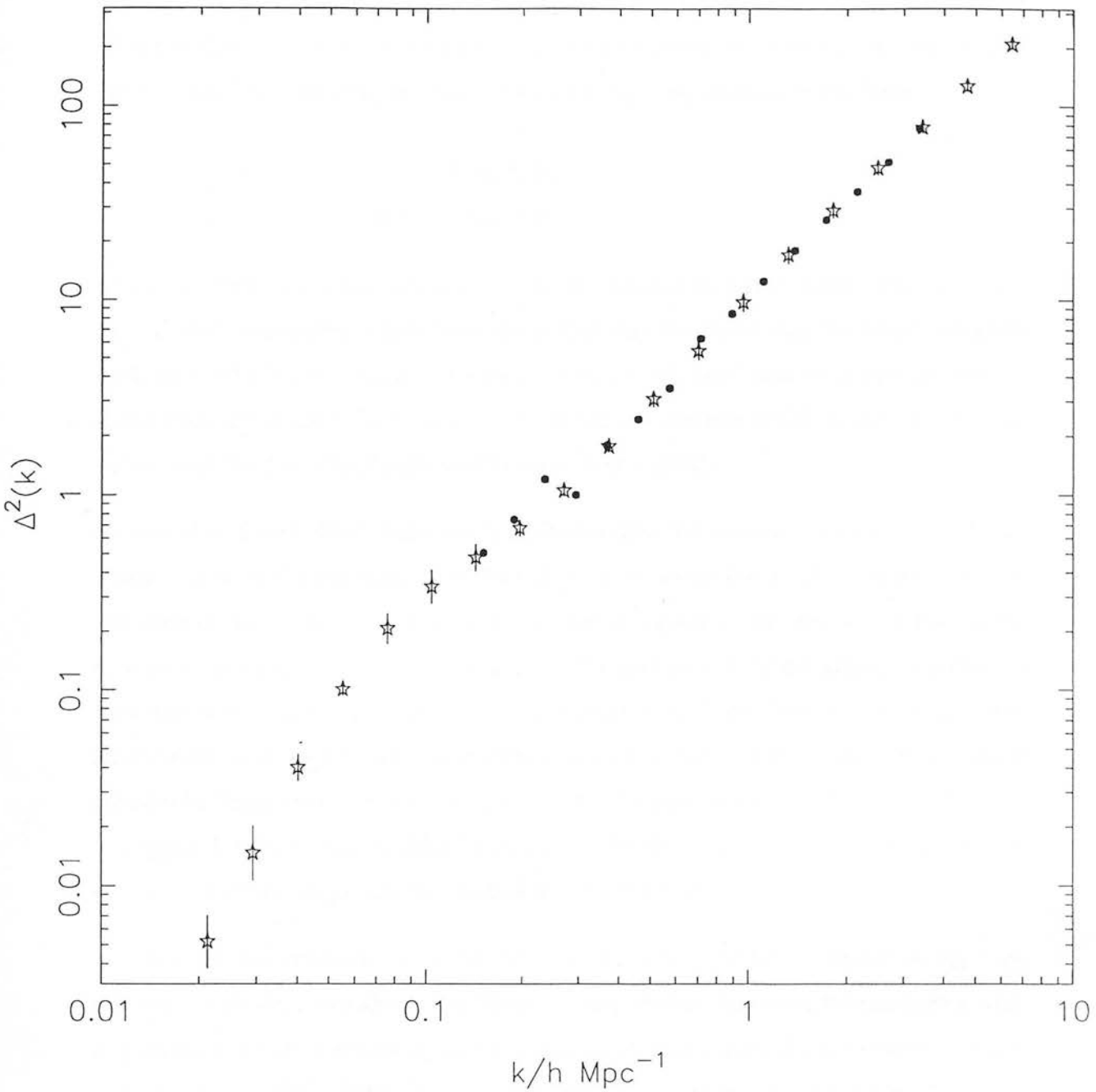


Figure 4.4: Comparison of the power spectrum of optical galaxy clustering in real space predicted by the model ($\Gamma^* = 0.23, c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$) and that computed from the APM catalogue by Baugh & Efstathiou (1993). The dots mark the model predictions and the stars the observational data with error bars derived from the scatter between the values of the power spectrum computed in four regions of the APM catalogue.

yielding a fairly clear picture of the two classes of galaxy following consistent distributions except in rich environments, which have a lower overdensities in *IRAS* galaxies than in optical galaxies. This suggests that we adopt a $w_{\rm O} - w_{\rm I}$ relation of the form

$$w_{\rm I} = \begin{cases} w_{\rm O} & \text{if } w_{\rm O} \leq w_{\rm C} \\ w_{\rm C} + (w_{\rm O} - w_{\rm C})^n & \text{if } w_{\rm O} \geq w_{\rm C} \end{cases}$$

where the parameters $w_{\rm C}$ and n are to be determined observationally. Most of these optical-*IRAS* comparisons have been quantified through smoothing the number density fields on $\sim 10\,h^{-1}$ Mpc scales, but we can extract the information we require on smaller scales from the study by Strauss *et al.* (1992) of five clusters found in the CfA redshift survey and the 2 Jy *IRAS* survey of Strauss *et al.* (1990).

Strauss *et al.* (1992) determine radial profiles for their five clusters (Fornax, Ursa Major, Virgo, Coma and Centaurus) and find that in all cases the profiles (defined as the variation of the average overdensity in a sphere of a given radius centred on the cluster core as a function of radius) delineated by the optical and *IRAS* galaxy populations have the same slope in redshift space. The clusters with the best statistics (*i.e.* the largest number of *IRAS* galaxies) are Virgo and Ursa Major. The overdensity in optical galaxies in Virgo averaged over spheres in redshift space out to a radius ~ 750 km s^{-1} is roughly 1.6 times that in *IRAS* galaxies, while the densities of the two species are consistent in Ursa Major out to a radius of ~ 1000 km s^{-1}.

We may use the information contained in these radial profiles to deduce a $w_{\rm O} - w_{\rm I}$ relation, as we shall now show. The Fourier-space window functions for smoothing with a spherical top hat of radius $R_{\rm s}$ and a Gaussian of filter length $R_{\rm g}$ are equal to second order for $R_{\rm s} = \sqrt{5}R_{\rm g}$. It follows, therefore, that we may identify with Ursa Major and Virgo Major particles in our simulations that have the same redshift-space overdensity in optical galaxies (evaluated using Gaussian smoothing with our fiducial filter radius of $1\,h^{-1}$ Mpc) as the clusters have averaged over spheres of radius $\sqrt{5}\,h^{-1}$ Mpc: it may be objected that not all particles in our simulations at those overdensities will be found in cluster cores, but recall that our basic premise is that the local galaxy density is determined solely by the *local* mass density. Since Virgo has an overdensity in optical galaxies of ~ 40 on this scale and Ursa Major one of ~ 10 we see that the $w_{\rm O}/w_{\rm I}$

ratio must change from unity to a value of 1.6 in the interval of w_O corresponding to $10 \lesssim [\rho_O / \langle \rho_O \rangle] \lesssim 40$, where ρ_O is the optical galaxy distribution in the simulation smoothed with our standard $1\,h^{-1}$ Mpc Gaussian. If we denote the optical weights of the particles to be identified with Virgo and Ursa Major as $w_{O,V}$ and $w_{O,UM}$ respectively, then the parameter n may be expressed in terms of w_C and $w_{O,V}$ for values of w_C in the interval $w_{O,UM} \leq w_C \leq w_{O,V}$. Unfortunately, the allowed (c_1, c_2) pairs do not, generally, yield particles with smoothed redshift space optical densities quite as high as that required to match Virgo. In such cases we take $w_{O,V}$ to equal the optical weight of the particle with the largest $\rho_O / \langle \rho_O \rangle$ value: typically "Virgo" becomes a particle with $\rho_O / \langle \rho_O \rangle \simeq 30 - 35$. It might be objected that, by so doing, we are compromising our $w_O - w_I$ relation, through weakening its link with observed reality but, in fact, we find a negligible change in the *IRAS* power spectrum if we vary the Virgo density from 40 to 25 for data sets that do have particles with $\rho_O / \langle \rho_O \rangle \geq 40$: a sufficiently small number of particles are given $w_O > w_I$ and the resultant $w_O - w_I$ is sufficiently small that the power spectrum is very insensitive to the details of the weighting-down of the *IRAS* galaxies in clusters.

To determine the $w_O - w_I$ relation to be applied to the set $\{w_O\}$ from each model we choose five equally-spaced values of w_C between $w_{O,UM}$ and $w_{O,V}$ and solve for n. This gives us, for each allowed (c_1, c_2) pair in each data set, five sets of *IRAS* weights, $\{w_I\}$, defining five predicted *IRAS* galaxy distributions, whose power spectra we compare with observations in the next Section.

4.6 THE POWER SPECTRUM OF IRAS GALAXY CLUSTERING

In this Section we test whether our galaxy clustering candidates can reproduce the clustering of *IRAS* galaxies, as well as optically-selected galaxies, by comparing the power spectra of our w_I–weighted particle distributions with the observed power spectrum of *IRAS* galaxy clustering.

4.6.1 Observational Data

Consistent results for the redshift-space power spectrum of *IRAS* galaxies have been published recently by Fisher *et al.* (1993) and Feldman, Kaiser & Peacock (1993) using samples drawn from the 1.2 Jy and QDOT redshift surveys respectively: we use both sets of data here, to maximise the range of wavenumbers over which we can make our comparison. Our results from Chapter 3 show that the redshift-space clustering of samples of QDOT galaxies is not significantly affected by any hypothesised relationship between the far-infrared temperature of an *IRAS* galaxy and the density of its environment, thus supporting our use of the redshift-space power spectrum rather than, for example, the Fourier transform of the *real space* correlation function determined by Saunders *et al.* (1992).

4.6.2 Method

Our method here mirrors that in Section 4.4.2: we simply determine the power spectrum of the $w_{\rm I}$–weighted particle distribution.

4.6.3 Results

In Figure 4.5 we compare the redshift-space *IRAS* power spectrum of the model ($\Gamma^* = 0.23, c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$) with observations. The empty circles show the QDOT power spectrum of Feldman *et al.* (1993), while the empty squares present the 1.2 Jy data of Fisher *et al.* (1993): Fisher *et al.* evaluate the 1.2 Jy power spectrum to $k \sim 10\,h$ Mpc^{-1}, but power is only detected significantly for $k \lesssim 0.2h$ Mpc^{-1}, so we restrict ourselves to that wavenumber interval.

Figure 4.5 shows predictions for five $(w_{\rm C}, n)$ pairs – namely (2.23,0.44), (2.43,0.38), (2.64,0.31), (2.84,0.22) and (3.04,0.12). The redshift-space *IRAS* power spectra predicted by other allowed models are very similar to that shown in Figure 4.5, of course, since they are produced by the same $w_{\rm O} - w_{\rm I}$ relation being applied to different $w_{\rm O}$–weighted particle distributions which have, by construction, very similar real-space power

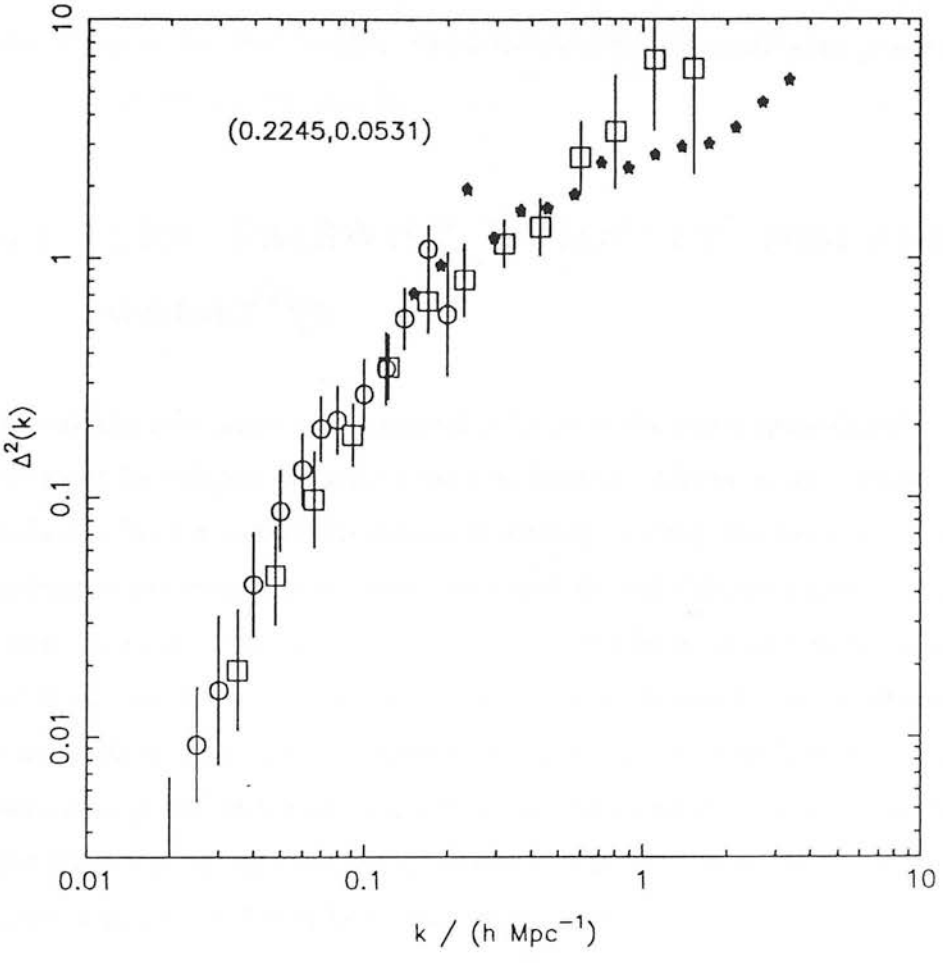


Figure 4.5: Comparison of the redshift-space power spectra of *IRAS* galaxy clustering predicted by the model ($\Gamma^* = 0.23, c_1 = 0.225, c_2 = 5.31 \times 10^{-2}$) and those computed by Feldman *et al.* (1993) (circles) and Fisher *et al.* (1993) (squares). The filled symbols show power spectra computed using five (w_C, n) pairs.

spectra. The only difference between *IRAS* power spectra computed from different models comes from the redshift-space distortions the models produce. The small size of this effect suggests that the velocity fields in the models must be very similar: we test this deduction in the next Section, where we compute the small-scale pairwise velocity dispersions in our allowed models.

4.7 THE PAIRWISE VELOCITY DISPERSION OF GALAXIES

As our aim is to construct a consistent picture of the phase-space distribution of galaxies, we must investigate the galaxy velocity field in addition to the spatial distribution of galaxies. We have done this implicitly already, through the use of the *real space* optical galaxy power spectrum in conjunction with the *redshift-space* power spectrum of *IRAS* galaxies: in this Section, however, we make our investigation more explicit, by considering the galaxy pairwise velocity dispersions predicted by our clustering models. The small-scale galaxy pairwise velocity dispersion has frequently been cited over the past decade as providing a very stringent constraint on small-scale power in the density field, but there are, as we shall see, good reasons, on both the observational and theoretical sides, why its use should be treated with caution.

4.7.1 Observational estimations of the pairwise velocity dispersion

The pairwise velocity dispersion of a redshift sample of galaxies is estimated from the effect that pairwise velocities have on $\xi(r_{\rm p}, \pi)$, which is the redshift-space correlation function as expressed as a function of projected separation, $r_{\rm p}$, and line-of-sight separation, π. The standard reference for this procedure is Davis & Peebles (1983, DP83) and we follow the general treatment they outline.

The correlation function $\xi(r_{\rm p}, \pi)$ is determined using the estimator

$$1 + \xi(r_{\rm p}, \pi) = \frac{DD(r_{\rm p}, \pi)}{DR(r_{\rm p}, \pi)}, \tag{4.23}$$

where $DD(r_p, \pi)$ is the number of galaxy pairs with separations (r_p, π) in the interval $(r_p - \Delta r_p/2, \pi - \Delta\pi/2) \rightarrow (r_p + \Delta r_p/2, \pi + \Delta\pi/2)$ and $DR(r_p, \pi)$ is the corresponding count of cross–pairs between the data sample and a random catalogue containing the same number of galaxies in the same volume of space and with the same radial selection function. Since the emphasis is on small-scale clustering, all pairs are given equal weight in equation 4.23 and only those pairs separated on the sky by angles, θ_{12}, of $\theta_{12} \leq 50^\circ$ are included in the analysis, to facilitate the small-angle approximation by which the physical separation between the pair is $r = (r_p^2 + \pi^2)^{1/2}$, with

$$\pi = \frac{v_1 - v_2}{H_0} \quad \text{and} \quad r_p = [(v_1 + v_2)/H_0] \tan(\theta_{12}/2), \tag{4.24}$$

where v_1, v_2 are the recessional velocities of the galaxies.

The correlation function $\xi(r_p, \pi)$ is a convolution of the *real space* correlation function $\xi(r)$ with the distribution function, $f(V, r)$, of the pairwise peculiar velocity:

$$1 + \xi(r_p, \pi) = H_0 \int_{-\infty}^{\infty} dy \left\{1 + \xi[(r_p^2 + y^2)^{1/2}]\right\} f(V, r). \tag{4.25}$$

The real space correlation function, $\xi(r)$, is deduced from the projected correlation function, $w(r_p)$, which is defined by

$$w(r_p) = \frac{1}{H_0} \int_{-v_L}^{v_L} d\pi \, \xi(r_p, \pi), \tag{4.26}$$

where a limit of $v_L = 2500$ km s^{-1} was used by DP83 and many subsequent workers. The spatial and projected correlation functions are related by

$$w(r_p) = 2 \int_0^{\infty} dy \, \xi \left[(r_p^2 + y^2)^{1/2}\right], \tag{4.27}$$

so, fitting a power law model, $w(r_p) = A r_p^{1-\gamma}$, to the projected correlation function data produces a power law model for the real space correlation function, $\xi(r_0/r)^\gamma$, with

$$r_0^\gamma = \frac{A(\gamma/2)}{(1/2)[(\gamma - 1)/2]}. \tag{4.28}$$

The velocity distribution function, $f(V, r)$, is assumed to vary only slowly with r and DP83 suggest the following form for it:

$$f(V) \;=\; C \,\exp\left(-2^{1/2}|V|/\sigma\right), \tag{4.29}$$

where

$$V \;=\; \pi \;-\; H_0 y \,\left\{1 - h\left[(r_{\rm p}^2 + y^2)^{1/2}\right]\right\}. \tag{4.30}$$

Many authors (*e.g.* DP83; Gelb, 1992; Mo, Jing & Borner 1993) have considered C as a free parameter whereas, clearly, its value is determined uniquely by a normalisation condition applied to the distribution function $f(V)$. DP83 justify this by arguing that this degree of freedom is necessary to account for the deviation of the observed $w(r_{\rm p})$ from a power law. This is a fair point, but one must be careful that, by introducing an extra degree of freedom, one may be hiding the failure of the model to give good agreement with the observations.

The term $H_0 y h(r)$ in equation 4.30 is to account for the mean streaming flow relative to Hubble expansion. DP83 use their earlier work on self-similar clustering using the hierarchy of BBGKY equations (Davis & Peebles 1977) to suggest that $h(r)$ takes the form

$$h(r) \;=\; F\,\left[1 + (r/r_0)^{\gamma}\right]^{-1}, \tag{4.31}$$

where r_0 is the (real space) correlation length of the galaxy sample. DP83 consider different values for the F parameter, to test the sensitivity of their results to the correction for streaming motions: their (Davis & Peebles 1977) similarity solution would suggest $F = 1$, while they note that the slowly collapsing clusters in the simulations of Efstathiou & Eastwood (1981) yield results consistent with $F \sim 1.5$.

Davis & Peebles obtained a value of $\sigma = 340 \pm 40$ km s^{-1} for the velocity dispersion at the fiducial separation of 1 h^{-1} Mpc from the CfA redshift survey using this method, while its application to the AAT redshift survey by Bean *et al.* (1983) yielded a slightly

lower figure of 250 ± 50 km s^{-1}. Both these values – particularly the DP83 result – have been used many times to constrain galaxy clustering models.

Recently, Mo *et al.* (1993) have performed a very thorough examination of the observational determination of the small-scale galaxy pairwise velocity dispersion. They investigate its sensitivity to sample composition (early– or late–type or *IRAS* galaxies, including or excluding galaxies in rich clusters) and analysis (with or without correcting for Virgocentric infall through the use of the $H_0 y h(r)$ in equation 4.30). Their results indicate that sampling effects are important: they find values for the dispersion in the projected separation bin $0.8 \leq r_p/(h^{-1}\ \mathrm{Mpc}) \leq 1.6$ that vary from 276 ± 17 km s^{-1} (for the CfA sample analysed by DP83) to 1861 ± 172 km s^{-1} (for the CfA2 sample, which includes many Coma galaxies). This is, of course, to be expected, given the morphology–density relation (*e.g.* Dressler 1980), but, more worryingly, the Virgocentric infall correction is seen to be quite significant, as is the selection function used for flux-limited samples. Mo *et al.* conclude that *"it is reasonable to doubt that the present data are a fair sample"* for the purposes of estimating the small-scale pairwise velocity dispersion and, thus, the stringent constraint implied by the Davis & Peebles (1983) estimate of 340 ± 40 km s^{-1} at a separation of $1\ h^{-1}$ Mpc must be treated with some suspicion.

4.7.2 Theoretical determinations of the pairwise velocity dispersion

Problems arise in the determination of the galaxy pairwise velocity dispersion from numerical simulations due to a variety of causes, further complicating its use as a test of cosmological models.

One of the principal uncertainties concerns the issue of so-called "velocity bias". A number of authors (*e.g.* Carlberg & Couchman 1989, Carlberg, Couchman & Thomas 1990, Couchman & Carlberg 1992) have observed that the pairwise velocity dispersion of the objects they identify as galaxies in their simulation is lower than that of the mass — *i.e.* galaxies are biased tracers of the velocity field in the same way that they are often assumed to be biased tracers of the density field. Carlberg *et al.* (1990) argue that this velocity bias has a physical origin, arising from dynamical friction: as galaxies move

through the sea of dark matter particles in a dark halo they experience a gravitational drag force due to dynamical friction (Chandrasekhar 1943), which results in the galaxy distribution becoming cooler and more centrally condensed than the distribution of dark matter in the halo. This picture is appealing and has been invoked to account for some dynamical evolution features seen in rich clusters (*e.g.* Kashlinsky 1986), but it is worrying that the magnitude of the velocity bias observed in simulations varies between authors: estimates of the galaxy velocity dispersion expressed as a fraction of that of the mass vary from 0.37 (Couchman & Carlberg 1992) to 0.8 (Cen & Ostriker 1992) to unity in the many simulations where no velocity bias is observed. There are, however, many other possible explanations. One possibility, to which we alluded earlier, is that velocity bias is an artefact of the algorithms used to group particles into halos or "galaxies" in dissipationless simulations. An FOF algorithm operating in a dense environment can link halos which the eye would regard as distinct, with the result that their relative velocity contributes to the internal velocity dispersion of the (wrongly-merged) halo rather than to the pairwise velocity dispersion of the halo distribution.

Gelb & Bertschinger (1993) note that the dispersion is also sensitive to whether the velocity of a halo is defined to be that of its most bound particle (as used by White *et al.* 1987, for example) or its centre-of-momentum velocity. The latter gives lower velocities to the halos, as it does not include the velocity of the most bound particle relative to the centre-of-momentum, which can be significant, given the internal velocity dispersions of massive halos. Gelb (1992) also notes that there can be a difference between the velocity dispersion measured directly in a simulation to that estimated using the Davis & Peebles observational approach: the dispersions estimated by the two methods differ typically at the $\sim 20\%$ level, although Gelb finds differences of up to 50% in extreme cases. Whilst it is probably best to view these differences as indicating failings in the observational approach of Davis & Peebles, it does suggest that caution should be exercised when comparing direct theoretical measures with observational data analysed using the Davis & Peebles treatment.

4.7.3 Method

We compute velocity dispersions from our simulation data using a procedure along the lines of the Davis & Peebles treatment. For computational speed we slightly modify the observational procedure, by fitting a power law model to the true real-space correlation function, rather than determining the model parameters r_0 and γ from $w(r_p)$: Gelb (1992) shows that the values of r_0 and γ derived from $w(r_p)$ do not differ greatly from those measured directly, so this should not compromise the claims of our procedure to being a fair representation of the Davis & Peebles observational method.

We compute the pairwise velocity dispersion of sparse samples of 10,000 particles drawn at random from our simulation data sets. For a given particle distribution (either unweighted to represent the mass, or w_O– or w_I–weighted to represent optical or *IRAS* galaxies respectively) we compute the real-space correlation function and fit it to the canonical power law form, $\xi(r) = (r_0/r)^\gamma$ in the interval $1 \leq r/(h^{-1}\,\mathrm{Mpc}) \leq 10$. We calculate the two-dimensional redshift-space correlation function, $\xi(r_p, \pi)$, in the same seven r_p and twenty π bins as DP83, using a 50,000–particle random catalogue: *N.B.* for computational convenience we employ the DD/RR correlation function estimator here, rather than the DD/DR estimator of equation 4.23.

We fit the DP83 model of equations 4.25, 4.29 and 4.30 to our model $\xi(r_p, \pi)$ data, finding, for each r_p bin, which of 25 equally-spaced values of σ in the range $300 \leq \sigma/(\mathrm{km\,s^{-1}}) \leq 800$ gives the least scatter. In common with other authors (*e.g.* Gelb 1992) we find that the DP83 model gives a particularly poor fit to the outermost r_p bin ($6.4 \leq r_p/(h^{-1}\,\mathrm{Mpc}) \leq 12.8$, which we exclude form our analysis. We also exclude the region $\pi > 750$ km s^{-1}, where $\xi(r_p, \pi)$ becomes very noisy, again in common with Gelb (1992). Finally, we fit the velocity dispersions from the various r_p bins to the power law form $\sigma(r_p) = \sigma_0(r_p/1\,h^{-1}\,\mathrm{Mpc})^\delta$, to obtain the values of σ_0 and δ.

4.7.4 Results

In Table 4.2 we present the results of our analysis for the unweighted particles in data sets P3M.11-14 – *i.e.* for the mass in models with $\Gamma^* = 0.18, 0.23, 0.26$ and 0.28. For

Simulation	r_0 [h^{-1} Mpc]	γ	σ_0 [km s^{-1}]	δ
P3M.11	1.73 ± 0.03	1.47 ± 0.01	460 ± 20	(-5.33 ±3.69) × 10^{-2}
P3M.12	2.54 ± 0.05	1.63 ± 0.02	485 ± 20	(-9.72 ±3.34) × 10^{-2}
P3M.13	3.07 ± 0.21	1.76 ± 0.05	538 ± 23	(-9.91 ±3.26) × 10^{-2}
P3M.14	3.37 ± 0.15	1.88 ± 0.03	557 ± 17	-0.11 ± 0.02

Table 4.2: Power law fit parameters: mass distribution.

Simulation	c_1	c_2	r_0 [h^{-1} Mpc]	γ	σ_0 [km s^{-1}	δ
P3M.11	0.43	0.11	3.95	1.83	500	-7.04 ×10^{-2}
P3M.12	0.23	5.3 ×10^{-2}	3.95	1.84	544	-9.94 ×10^{-2}
P3M.13	0.14	-8.2 ×10^{-3}	3.77	1.85	567	-9.95 ×10^{-2}
P3M.14	0.12	-2.9×10^{-2}	3.80	1.92	581	-0.12

Table 4.3: Power law fit parameters: w_O–weighted distribution.

each model we tabulate the parameters of the power law fits to the real-space correlation function and to $\sigma(r_p)$.

The results shown in Table 4.2 are unsurprising: as the small-scale power in the density field increases (*i.e.* as Γ^* increases), the correlation length of the mass distribution increases, as does its small-scale velocity dispersion. The striking feature of these results is that the values of σ_0 deduced from these power spectra are so much lower than that determined the value ($\sigma_0 \simeq 1300$ km s^{-1}) found by Couchman & Carlberg in their low-bias CDM simulation. This illustrates very nicely the problem with the shape of the CDM power spectrum: when normalised to *COBE* (as the Couchman & Carlberg simulation is, to a good approximation), $\Gamma^* \simeq 0.5$ CDM has too much small-scale power.

Having determined the small-scale velocity dispersions of the unweighted particle distributions in simulations P3M.11-14, we now turned to the w_O–weighted particle distributions. For each model, we consider the allowed (c_1, c_2) pair which gives the lowest χ^2 fit to the APM power spectrum and tabulate our results in Table 4.3: considerations of space dictate that we omit the uncertainties on these quantities, but they are similar to those of the corresponding quantities in Table 4.2.

We see from Table 4.3 that, as expected, if we give extra weight to the overdense regions

Simulation	c_1	c_2	w_C	n	r_0 [h^{-1} Mpc]	γ	σ_0 [km s^{-1}	δ
P3M.11	0.43	0.11	3.74	0.61	4.09	1.93	508	-5.94 $\times 10^{-2}$
P3M.12	0.23	5.3 $\times 10^{-2}$	2.09	0.24	3.85	1.79	535	-9.38 $\times 10^{-2}$
P3M.13	0.14	-8.2 $\times 10^{-3}$	1.46	0.25	3.66	1.82	554	-0.10
P3M.14	0.12	-2.9$\times 10^{-2}$	1.28	0.10	3.77	1.91	571	-0.12

Table 4.4: Power law fit parameters: w_I–weighted distribution.

(as we do in producing the w_O–weighted distribution), we increase the small-scale pairwise velocity dispersion. What is interesting is that we do not increase it by much and that the resultant pairwise "galaxy" dispersion of $\sim$ 550 km s^{-1} at the fiducial separation of 1 h^{-1} Mpc is perfectly consistent with the values deduced by Mo *et al.* (1993) in their recent re-analysis of the observational data.

Finally, we compute the velocity dispersion of the w_I-weighted particle distributions. We take the models from Table 4.3 and weight them according to our empirical $w_O - w_I$ relation using, for each model, the lowest of the five w_C values defined in Section 4.5: our results are given in Table 4.4, without errors which are, again, similar to those given for the corresponding quantities in Table 4.2.

From Table 4.4 we see the expected result that, since *IRAS* weights are lower than optical weights in our model, the small-scale pairwise velocity dispersion of the w_I–weighted distributions are lower than those of the parent w_O–weighted ones. The difference is not great, however, in accordance with what we would expect, given that the empirical $w_O - w_I$ relation only gives $w_O \neq w_I$ in the very richest environments.

4.8 SUMMARY

In this Chapter we have considered the clustering of optical and *IRAS* galaxies. This has been done with the aim of determining the set of galaxy clustering models – defined, for our purposes, to be combinations of linear density power spectra and prescriptions relating the galaxy distribution to the underlying density field – which are consistent with the observed large-scale structure of the Universe.

We have presented a novel method for studying a family of power spectra using the results of just one numerical simulation. We have implemented this procedure to investigate the family of CDM-like power spectra, studying nine different values of the shape parameter, Γ^*, which labels the members of the family.

Motivated by results from recent N–body/hydrodynamic simulations, we have investigated an Eulerian bias scheme, in which the local number density of galaxies at redshift zero is determined in a simple fashion from the *present* local density field. By adopting such a bias prescription we have avoided the need to use an *ad hoc* group-finding algorithm in our simulation data sets or to resort to biasing in Lagrangian space, which is likely to be a poor way to model the clustering of galaxies, given the nonlinearity of the cosmological density field on galactic scales.

We have used the observed power spectrum of optical galaxies to select those galaxy clustering models which are consistent with reality. We have found that an infinite set of models is selected if we give our weighting scheme free rein to trade off over– and under–dense regions against each other. With the addition of the constraint that the optical galaxy weight w_O should be a monotonically-increasing function of local density we find that no galaxy clustering model with $\Gamma^* > 0.29$ is able to account for the power spectrum of APM galaxy clustering.

From the observed relation between the redshift-space distributions of optical and *IRAS* galaxies we have constructed the *IRAS* galaxy distributions corresponding to our allowed galaxy clustering models. We have found the redshift-space power spectra of these distributions to be consistent with the clustering observed in the QDOT and 1.2 Jy *IRAS* redshift samples.

Through studying optical galaxy clustering in real space and *IRAS* galaxy clustering in redshift space we have implicitly been comparing the velocity fields in galaxy clustering models. We have made this study more explicit, by determining the small-scale pairwise velocity dispersions of sparse samples of our simulation data sets. We have found that our allowed galaxy clustering models yield pairwise velocity dispersions of 550 ± 50 km s^{-1} at the fiducial separation of $1\,h^{-1}$ Mpc, in good agreement with a recent analysis of observational determinations of this quantity.

We shall discuss the issues raised by this project in more detail in Chapter 5.

Chapter 5

DISCUSSION, CONCLUSIONS AND FUTURE WORK

καὶ τὸ οὖν σαφὲς 'ούτις 'ανὴρ 'ίδεν 'ουδέ τις 'έσται 'ειδὼς 'αμφὶ θεῶντε καὶ 'άσσα λέγω περὶ παντων. 'ει γὰρ καί τὰ μάλιστα τύχοι τετελεσμένον 'ειπών, 'αυτός 'όμως 'ουκ 'όιδε. δόκος δ' 'επὶ πᾶσι τέτυκται.

No man knows, or ever will know, the truth about the gods and about everything I speak of: for even if one chanced to say the complete truth, yet oneself knows it not; for seeming is wrought over all things.

Xenophanes

If the study of large-scale structure is to be cosmology rather than cosmography – or physics rather than stamp-collecting, to employ Rutherford's famous distinction – then we must understand how the distributions of astrophysical objects such as galaxies and clusters relate to the underlying density field: it is the study of this relation which is the theme of this thesis. We now review the salient points from earlier chapters, expand our previous discussions of the conclusions to be drawn from them and outline plans for future work.

5.1 THE SPATIAL CORRELATIONS OF CLUSTERS

As we have seen, we are currently enjoying a very active period of research into the spatial correlations of clusters of galaxies, on both observational and theoretical fronts.

The advent of machine-based galaxy catalogues like the EDSGC and the APM catalogue has facilitated the construction of cluster catalogues using objective selection criteria. These provide the theorist with reliable data with which to challenge cosmological models, as well as, one hopes, removing the necessity, and even the excuse, for theorists to try to reconstruct what George Abell *should* have seen, either by 'decontaminating' cluster samples or 'correcting' correlation functions computed from them. Advances are also being made using *ROSAT All-Sky Survey* data, but X-ray cluster selection with *ROSAT* has not proved to be as clean a procedure as some had hoped.

The continuing improvement in the capabilities of computers is now producing believable numerical simulations of cluster correlations, like those of Croft & Efstathiou and Cole described in Chapter 2. There remains, however, the need for those performing numerical simulations to demonstrate the degree of sensitivity of their results to the particular algorithms and parameters used in their codes. Analytical work is, by contrast, more transparent to the critical eye, as well as being untroubled by the dynamical scale of the problem and being much, much faster.

All these considerations motivate analytical studies of cluster correlations, such as that pursued here. The principal advantage of our method over those previously employed is its treatment of the mildly nonlinear evolution of the cosmological density field on cluster

scales and the resulting dynamical contribution to the clustering of galaxy clusters. The general degree of agreement between our results and those of Croft & Efstathiou gives us great confidence in our method. We must, however, understand the source of the differences between our predictions for the clustering strength of the richest clusters and those simulated numerically.

One possibility is that the correlations of the richest cluster samples are very sensitive to the prescription used to select them. Comparisons of Figures 2.13 and 2.14 do seem to show differences between the results from simulations run by Cole and by Croft & Efstathiou (which are comparable in all but their cluster selection procedures), although any such conclusion must be tempered by the uncertainty in the Cole data resulting from their coming from only two realisations. It is important that this hypothesised sensitivity to selection algorithms be investigated.

A second possibility is that the Zeldovich Approximation is unable to reproduce the degree of nonlinear evolution required to model accurately the cosmological density field. The results of our study of the mass correlation function in Section 2.6.2 lead one not to entertain that possibility too seriously, as the behaviour of $\xi_\rho(x)$ at small separations tends to suggest that, with an appropriate choice of filter radius, the Zeldovich Approximation may be used to evolve the smoothed density field accurately, without being compromised by shell-crossing: it may be tested, however, by directly comparing the mass correlation functions from N-body simulations and from our method.

Thirdly, it is possible that the difference results from the prescription we use to derive a cluster distribution from the evolved density field. Our approximation to the peak-background split neglects a term quadratic in the threshold ν_s (compare equations 2.25 and 2.55), which could become significant as ν_s increases with cluster richness. Evidence against this being the explanation for the difference between our results and those of Croft & Efstathiou comes from Figures 2.9, 2.12 & 2.13, which show that although the $\Gamma = 0.2$ model produces significantly larger values of ν_s, it is the $\Gamma = 0.5$ model for which there is the larger difference at high richness.

All three of these possibilities may be fruitfully addressed by further, more detailed comparisons between our analytical results and those from numerical simulations, along

the lines of those outlined in Section 2.7 and it is hoped to start such work shortly.

The most interesting possibility is the one that would remain after the elimination of the three listed above – that the Gaussian peaks prescription itself is inapplicable to the study of cluster correlations. As discussed in Section 2.5.1, the Kaiser effect is predicated on a model in which clusters form at the peaks in a Gaussian initial density field. A plausibility argument is invoked to suggest that clusters should form at density peaks, where there is the most mass, but this may not be a correct identification. Firstly, the density field may not be Gaussian: although work on the topology of large-scale structure (*e.g.* Gott *et al.* 1989 and references therein), the *COBE* DMR microwave background anisotropies (Smoot *et al.* 1992) and the distribution of long-wavelength power in the redshift-space power spectrum of the QDOT survey (Feldman, Kaiser & Peacock 1993) are all consistent with the hypothesis of a Gaussian initial density field. Secondly, perhaps other non-gravitational effects influence the locations where clusters form, although this is, perhaps, less likely for clusters than for galaxies, due to the larger scales in the density field appropriate to cluster formation and the finite range of feedback mechanisms like the injection of energy from supernovæ which could inhibit the collapse of density perturbations.

Turning now to the comparison between our results and the observational data presented in Section 2.8, we see that none of the models we study reproduces the strong trend of increasing correlation strength with richness that is required to account for the observed correlation lengths of all cluster samples, from APM $\mathcal{R} \geq 20$ clusters to Abell $R \geq 2$ clusters. In part this is just the result of how poor a description of the clustering of a real cluster sample the correlation length alone provides: from the $d - r_0$ relations shown in Figures 2.19 – 2.22 one would conclude that all four models are wildly inconsistent with the Abell $R \geq 2$ data, but the noise on the data presented in Figure 2.18 shows that the case is not so clear. We urge other workers in this field not to reduce the information contained in spatial correlation functions to a single number by quoting only a correlation length for their samples. There are formal, as well as pragmatic, reasons why power law fits to correlation functions are to be discouraged: values of the correlation function evaluated in different separation bins are themselves correlated, so it is incorrect to apply a simple χ^2-fitting procedure to obtain r_0 and γ, without accounting

for these correlations.

It is unfortunate that a number of the cluster samples yield negative values for the correlation function in single bins on large ($r \sim 40\,h^{-1}$ Mpc) scales, as this confuses the theoretical interpretation of these observational data. If, as seems likely, these negative values are simply the result of the binning of data from relatively small samples, then we may confidently rule out $\Gamma = 0.5$ CDM on the basis of its inability to reproduce the required strength of cluster correlations. On the other hand, if these negative values are real, then they lend strong support to this standard CDM model, since cosmogonies with relatively more large-scale power do not produce correlation functions with first zero-crossing points at such small separations. One must be very careful, however, when using the first zero-crossing point of the correlation function as a diagnostic tool since it is very sensitive to uncertainty in the mean cluster number density. This problem may be ameliorated by the construction of larger cluster samples. The APM cluster catalogue is still being expanded, while the forthcoming *Sloan Sky Survey* promises to yield large samples of even the richest clusters. Further samples of *ROSAT*-selected clusters should also provide additional observational determinations of the spatial correlations of clusters.

If we assume that the isolated negative data points in bins with separations of $r \sim 40h^{-1}$ Mpc result from the size of the data sets analysed, and do not reflect the correlations in the density field on those scales, then our results favour a spatially flat cosmogony with a linear power spectrum well approximated by the CDM transfer function with $\Gamma \simeq 0.2$. This is in accordance with the study of the angular correlations of APM galaxies by Efstathiou, Sutherland & Maddox (1990), verifying that galaxies and clusters do appear to delineate the same large-scale structure. Our results do not favour the ($\Gamma = 0.1$) model advanced by Bahcall & Cen (1992) which, when normalised correctly to *COBE*, predicts correlations for poorer clusters which greatly exceed those observed. We do not find theoretical support for the Bahcall $r_0 \propto d$ relation and doubt whether it will even remain a reasonable description of the observational data once the correlation function of the richer APM sample has been published.

There are other methods of studying the distributions of cluster environments beyond

those considered so far: for example, Peacock and collaborators (Peacock & Prestage 1988; Peacock & Nicholson 1991) have used radio galaxies as tracers of overdensities to great effect. Another probe of rich environments is the class of cD galaxies, as advocated by West & van den Bergh (1991) (see Section 2.4). The problem with their study is that the cD sample they use is drawn from the Abell catalogue and thus prone to the same projection effects and other oddities of selection as any other subsample of that catalogue. The advent of the *superCOSMOS* measuring machine at the Royal Observatory, Edinburgh presents for the first time the possibility of constructing samples of cD galaxies using objective selection criteria rather than serendipity and the author is leading a collaboration to investigate this possibility. The advantages that *superCOSMOS* has over the previous generation of plate-measuring machines (COSMOS and APM) for this task result from both its greater scanning speed and its greatly enhanced dynamical range. Most low redshift elliptical galaxies in the EDSGC have saturated cores when scanned by COSMOS, but *superCOSMOS*, with its far greater dynamical range, will be able to determine their surface brightness profiles very accurately and, through comparison with profiles of known cD galaxies obtained by CCD imaging, facilitate selection of cD galaxies from *superCOSMOS* scans of Schmidt plates. The scanning speed of *superCOSMOS* will also facilitate the construction of a relatively large catalogue, by enabling one to obtain good sky coverage in a reasonable length of time. It will be very instructive to compare, say, the clusters selected from the EDSGC by the methods of Lumsden *et al.* (see Section 2.3) with the cD galaxy distribution, to test the use of cD galaxies as tracers of dense environments.

With this exciting observational project and the continuing refinement of the method presented in Chapter 2 for predicting cluster correlations, the author intends to remain actively involved in this important and flourishing area of research.

5.2 THE CLUSTERING OF IRAS GALAXIES

In Chapter 3 we suggested that one might expect there to be a relationship between the far-infrared temperature of an *IRAS* galaxy and the density of the environment in which the galaxy finds itself. We argued that if this is so, and if warmer galaxies

are found preferentially in rich environments, then one would expect them to be more strongly clustered than cooler galaxies found preferentially in the field. We carefully constructed warm and cool subsamples of galaxies from the QDOT and 2 Jy surveys and studied their redshift-space correlations, using both correlation function and counts-in-cells methods. We found, contrary to the expectation engendered by the argument we had previously advanced, that there is no unambiguous evidence for a difference in the clustering strengths of warm and cool *IRAS* galaxies, at least in redshift space.

The caveat that our null result only applies to redshift space correlations deserves further comment. As we discussed briefly in Chapter 3, redshift-space effects could mask a difference in the real-space clustering of the two samples, by diluting the small-separation clustering of the species found in the richer environment. Certainly, it does appear from Figures 3.11 and 3.12 that the warm QDOT samples extending to 500 h^{-1} Mpc do show stronger large-scale and weaker small-scale redshift space clustering than the corresponding cool samples, in accordance with this possibility. No such effect is seen in the counts-in-cells analysis, but the uncertainties resulting from our small samples would hide all but the most extreme differences there. It is important, therefore, that we compare the *real space* correlations of our warm and cool subsamples too. Probably the best way to do this is to adopt the methods of Saunders *et al.* (1992) and so compute the spatial auto-correlation function of, say, the warm QDOT subsample via its projected cross-correlation function with the warm subset of QIGC galaxies. This method has the advantage of making use of the much larger number ($\sim$ 13000) of galaxies in the two-dimensional QIGC catalogue. We hope to undertake this project in the near future.

An alternative explanation for why we found no difference between the redshift space clustering strength of our warm and cool subsamples is, of course, that there is no difference between the strength of their clustering in real space. In that case we have to find the weak link in our chain of reasoning. One factor we have omitted to consider is that some cool *IRAS* galaxies will be ellipticals, which are concentrated in dense environments and, therefore, are more strongly clustered than spirals (Dressler 1980, Davis & Geller 1976). A quantitative assessment of the consequences of this omission must await the collection of morphological classifications of the galaxies in the QDOT and 2 Jy catalogues, but it is likely to be of minor importance, since ellipticals have very low

far-infrared luminosities, as a result of being dust-poor. Another possibility is that we are wrong to assume that the galactic interactions that may produce starburst phenomenonæ are more likely to occur in rich environments. Whilst it is, no doubt, true that galaxy-galaxy interactions, being two-body processes, take place more frequently in regions of high galaxy number density, it does not necessarily follow that starbursts will be found preferentially in rich environments. Numerical simulations of galactic interactions (see Barnes & Hernquist 1992 and references therein) reveal that an extended period of interaction between a pair of galaxies is likely to be required to trigger a starburst, since it is only prolonged interactions that have a strong effect on the subsequent evolution of the galaxies taking part in them. This consideration militates against the location of interaction-triggered starbursts in the densest environments since the pairwise velocities between the galaxies will be highest there and, hence, their interactions shortest-lived. The velocity dispersion within overdensities grows with time so, at some time in the past, the richest environments would have been the sites of starburst activity. It follows, therefore, that the clustering strength of warm *IRAS* galaxies should show a redshift dependence, tracing the redshift variation in the density of the typical environments in which starbursts are found. Such a trend could be seen in redshift surveys of sufficient depth but since, as we have shown, any such effect must originate at redshifts beyond that where many cool galaxies are seen, it would be impossible to distinguish this as being a temperature-dependent effect rather than a solely redshift-dependent one.

If this picture is correct, then starbursts are likely to be found in environments with a restricted range of densities at low redshift. This hypothesis may be tested observationally and the author is involved in a project to do just that. The first stage of the project is to obtain deep CCD images of the immediate environments of $\sim$ 100 *IRAS* galaxies selected from the QCCOD redshift survey (Oliver 1993). This sample is selected in a narrow band of 100 μm luminosity ($L_{100} \sim 10^{10}\, L_{\odot}$) and a narrow redshift interval around $z \simeq 0.03$. This sample is composed of fairly average galaxies – typical optical magnitudes are $m \simeq 14$, which is close to M_* at this redshift (Efstathiou *et al.* 1988) – so we are studying a more representative sample of galaxies than many previous projects (*e.g.* Lawrence *et al.* 1989, Leech *et al.* 1993) which have concentrated on ultra-luminous *IRAS* galaxies ($L_{60} \gtrsim 10^{11} L_{\odot}$). This sample can elucidate the relationship between star-

burst activity and environment on ~ 200 kpc scales and be used to calibrate the second portion of the project, which will study the same relationship on larger scales, using projected galaxy number surface densities measured from Schmidt plates.

Once again the author intends to pursue this line of investigation through a combination of further theoretical and observational work.

5.3 GALAXY CLUSTERING WITH EULERIAN BIAS

In Chapter 4 we considered the clustering of galaxies. In common with much of the previous work on this topic, our project was based on the analysis of the results of numerical simulations.

Our treatment differed from that of previous authors in two principal regards: (a) we did not set out to challenge a particular cosmogony with a battery of tests but, rather, we sought to use observational data to determine the shape of the mass power spectrum; and (b) we did not use an *ad hoc* algorithm to link particles into "galaxies", but rather constructed optical and *IRAS* galaxy distributions using weighting schemes motivated by the recent N–body/hydrodynamical simulations of Cen & Ostriker (1992) and the observed relation between the distributions of optical and *IRAS* galaxies.

We also showed how, with knowledge of the amplitude of the mass power spectrum (*e.g.* from *COBE*), it is possible to study a number of power spectra using the results from a single numerical simulation. This is a very powerful technique, which greatly increases the speed with which numerical simulations can cover the region of parameter space spanned by a family of power spectra.

To deduce a galaxy distribution from the density field, we employed an Eulerian bias scheme, in which the weights assigned to particles in the simulation are determined by the local mass density *at the present*. We showed how the constraint that the weight be a monotonically-increasing function of the local density implies that the power spectrum shape parameter, Γ^*, must be less than 0.29, to achieve consistency with the APM galaxy power spectrum.

At the heart of the project lies the Eulerian bias scheme, with its motivation from the Cen & Ostriker (1992) N–body/hydrodynamical simulations. The tightness of the mass density – galaxy density relationship they observe is, perhaps, surprising, given the picture of structure formation which emerges from their simulations, with galaxies forming in filaments and only thereafter drifting into clusters, but it is very important. It suggests a solution to one of the great problems of cosmology, by saying how the distributions of dark and luminous matter are related. The importance of knowing any such relationship means that it is essential to show whether or not the mass density – galaxy density relationship is generic or peculiar to CDM or to the particular simulation algorithms used by Cen & Ostriker: it is to be hoped that other authors will address this question with further simulations of different power spectra in the near future.

5.4 CONCLUDING REMARKS

This Thesis has been concerned with the use of galaxies and clusters as probes of the large-scale structure of the Universe. Central to this topic is the question of how the distributions of galaxies and clusters are related to the underlying density field. We have approached this question in several ways, using methods appropriate to the different properties of the galaxy and cluster populations. In all cases, however, we have turned to observational evidence in preference to theoretical prejudice to guide our course.

Cosmology has matured greatly over the course of the past few decades. This process has been particularly rapid in the past five years, when it has been characterised by the recognition that the quantity and quality of observational data describing large-scale structure have improved so greatly that theorists must adopt a phenomenological attitude, building up theoretical pictures from observations, rather than trying to make observations fit overarching theoretical notions. The next few years are sure to bring further observational advances, perhaps especially in galaxy surveys, the CBR and X-ray observations of clusters: these advances may tell us a great deal about large-scale structure, but only if theorists are flexible enough to let themselves be led by observations.

'ούτοι 'απ' 'αρχῆς πάντα θεὸι θνητοῖσ' 'υπέδειξαν, 'αλλὰ χρόνωι ζητοῦντες 'εφευρίσκουσιν 'άμεινον.

Yet the gods have not revealed all things to men from the beginning; but, by seeking, men find out better in time.

Xenophanes

MATHEMATICAL APPENDICES

"Innocent light-minded men, who think that astronomy can be learnt by looking at the stars without knowledge of mathematics will, in the next life, be birds."

Plato, *Timaeos*

A Derivation of $n_{\rm pk}(\nu_s)$

From BBKS (their equation A18), the differential number density of peaks in the range ν to $\nu + {\rm d}\nu$ is given by

$$\mathcal{N}_{\rm pk}(\nu){\rm d}\nu = \frac{1}{(2\pi)^2 R_*^3}\, {\rm e}^{-\nu^2/2}\, G(\gamma,\gamma\nu)\, {\rm d}\nu, \tag{A1}$$

where (BBKS equation A19) $G(\gamma,\gamma\nu)$ is defined by

$$G(\gamma,\gamma\nu) = \int_0^\infty \frac{{\rm d}x\; f(x)}{[2\pi(1-\gamma^2)]^{1/2}} \exp\left[\frac{-(x-\gamma\nu)^2}{2(1-\gamma^2)}\right]. \tag{A2}$$

In equation A2, ν is the height of the density perturbation field smoothed on scale $R_{\rm s}$ expressed in units of the rms perturbation, γ is the spectral parameter, $x = -\nabla^2\delta/\sigma_2$ and the function $f(x)$ has the closed form

$$\begin{aligned} f(x) &= \frac{(x^3-3x)}{2}\left\{{\rm erf}\left[\left(\frac{5}{2}\right)^{1/2} x\right] + {\rm erf}\left[\left(\frac{5}{2}\right)^{1/2}\frac{x}{2}\right]\right\} \\ &+ \left(\frac{2}{5\pi}\right)^{1/2}\left[\left(\frac{31x^2}{4}+\frac{8}{5}\right){\rm e}^{-5x^2/8} + \left(\frac{x^2}{2}-\frac{8}{5}\right){\rm e}^{-5x^2/2}\right], \end{aligned} \tag{A3}$$

from BBKS (their equation A15), who also give approximate fits to this exact form. It follows that the number density, $n_{\rm pk}(\nu_{\rm s})$, of peaks above a threshold $\nu_{\rm s}$ is given by

$$n_{\rm pk}(\nu_{\rm s}) = \int_{\nu_s}^\infty \mathcal{N}_{\rm pk}(\nu){\rm d}\nu = \frac{1}{(2\pi)^{5/2}(1-\gamma^2)^{1/2}R_*^3}\int_{x=0}^\infty {\rm d}x f(x) \exp\left[\frac{-x^2}{2(1-\gamma^2)}\right] I_{\rm A}, \tag{A4}$$

where

$$I_{\rm A} = \int_{\nu=\nu_s}^\infty {\rm d}\nu \exp\left[\frac{-\nu^2}{2(1-\gamma)^2} + \frac{\gamma\nu x}{(1-\gamma^2)}\right]. \tag{A5}$$

From Gradshteyn & Ryzhik (1980, § 3.222),

$$\int_u^\infty \exp\left(\frac{-x^2}{4\alpha} - \beta x\right){\rm d}x = \sqrt{\pi\alpha}\, e^{\alpha\beta^2}\left[1 - {\rm erf}\left(\beta\sqrt{\alpha} + \frac{u}{2\sqrt{\alpha}}\right)\right], \tag{A6}$$

and so equation A5 becomes

$$I_{\rm A} = \sqrt{\frac{\pi(1-\gamma^2)}{2}}\exp\left[\frac{x^2\gamma^2}{2(1-\gamma^2)}\right]{\rm erfc}\left[\frac{(\nu_{\rm s}-\gamma x)}{\sqrt{2(1-\gamma^2)}}\right]. \tag{A7}$$

Hence, from equations A4 and A7, it follows that the final form for the number density of peaks above the threshold is

$$n_{\rm pk}(\nu_{\rm s}) \;=\; \frac{1}{8\pi^2 R_*^3}\int_{x=0}^{\infty} {\rm d}x\; f(x)\,{\rm e}^{-x^2/2}\,{\rm erfc}\left[\frac{(\nu_{\rm s}-\gamma x)}{\sqrt{2(1-\gamma^2)}}\right]. \tag{A8}$$

B Derivation of $\langle\tilde{\nu_s}\rangle$

An expression for the n-point correlation function of peaks above a threshold is given in terms of a $2n$-dimensional integral by BBKS. This expression is valid for separations of more than a few times the filter scale, $R_{\rm s}$, used to select the peaks. In the limit $\nu_{\rm s}^2\psi(r) \ll 1$, which will obtain at large separations, this integral simplifies and the two-point function takes the form

$$\xi_{\rm pk-pk}(r) \;=\; \langle\tilde{\nu_{\rm s}}\rangle^2\,\psi(r), \tag{B1}$$

where $\psi(r) \equiv \xi(r)/\xi(0)$ is the normalized mass correlation function. The effective threshold, $\tilde{\nu}$, is defined by

$$\tilde{\nu} \;\equiv\; \frac{(\nu-\gamma x)}{(1-\gamma^2)}. \tag{B2}$$

The quantity $\langle\tilde{\nu_{\rm s}}\rangle$ is the average of this taken over all peaks above the threshold $\nu_{\rm s}$, and so is given by

$$\langle\tilde{\nu_{\rm s}}\rangle = \frac{\int_{\nu_s}^{\infty}\tilde{\nu_{\rm s}}\mathcal{N}_{pk}(\nu)\,{\rm d}\nu}{\int_{\nu_s}^{\infty}\mathcal{N}_{pk}(\nu)\,{\rm d}\nu} = \frac{1}{(2\pi)^{5/2}(1-\gamma^2)^{1/2}R_*^3\,n_{\rm pk}(\nu_{\rm s})}I_{\rm B}, \tag{B3}$$

where

$$I_{\rm B} \;=\; \int_{x=0}^{\infty}\int_{\nu=\nu_s}^{\infty} f(x)\left[\frac{(\nu-\gamma x)}{(1-\gamma^2)}\right]\,{\rm e}^{-S_0}\,{\rm d}x\,{\rm d}\nu, \tag{B4}$$

and

$$S_0 \;=\; \left[\frac{(\nu^2-2\gamma\nu x+x^2)}{2(1-\gamma^2)}\right]. \tag{B5}$$

If we make the substitution $y \equiv \nu - \gamma x$, then

$$S_0 = \left[\frac{y^2 + (1-\gamma^2)x^2}{(1-\gamma^2)}\right], \tag{B6}$$

and so the expression for $I_{\rm B}$ becomes

$$I_{\rm B} = \int_0^\infty f(x)\, {\rm e}^{-x^2/2} \exp\left[\frac{-(\nu_{\rm s} - \gamma x)^2}{2(1-\gamma^2)}\right] {\rm d}x. \tag{B7}$$

Finally, combination of equations A8, B3 and B7 yields the following form for $\langle \tilde{\nu_{\rm s}} \rangle$:

$$\langle \tilde{\nu_{\rm s}} \rangle = \left[\frac{2}{\pi(1-\gamma^2)}\right]^{1/2} \frac{\int_0^\infty f(x)\, {\rm e}^{-x^2/2} \exp\left[-(\nu_t - \gamma x)^2/2(1-\gamma^2)\right] {\rm d}x}{\int_0^\infty f(x)\, {\rm e}^{-x^2/2}\, {\rm erfc}\left[(\nu_t - \gamma x)/\sqrt{2(1-\gamma^2)}\right] {\rm d}x}. \tag{B8}$$

C Derivation of $\xi_{\rho {\rm Z}}(x)$

In this section we derive the form of the mass correlation function in the Zeldovich Approximation: this may not be the most elegant way to derive it, but it works. From equation 2.41 of Chapter 2 we see that the mass correlation function may be written as an integral over conditional probabilities, of the form

$$1 + \xi_\rho(x) = \int {\rm d}^3\boldsymbol{r}\, {\rm d}^3\boldsymbol{s}\, P(x\,|\,\boldsymbol{r},\boldsymbol{s})\, P(\boldsymbol{s}\,|\,\boldsymbol{r})\, P(\boldsymbol{r}),$$

where $P(x|\boldsymbol{r},\boldsymbol{s})$ is the conditional probability of two particles having an Eulerian separation x, given a Lagrangian separation $\boldsymbol{r}$ and a relative displacement $\boldsymbol{s}$; $P(\boldsymbol{s}|\boldsymbol{r})$ is the conditional probability of two particles having a relative displacement $\boldsymbol{s}$ given a Lagrangian separation $\boldsymbol{r}$; $P(\boldsymbol{r})$ is the probability of two particles having a Lagrangian separation $\boldsymbol{r}$; and the integral is taken over all $\boldsymbol{r}$ and $\boldsymbol{s}$.

We wish to calculate the covariance matrix $M_{ij} = \langle [s_i(\boldsymbol{y}+\boldsymbol{r}) - s_i(\boldsymbol{y})]\,[s_j(\boldsymbol{y}+\boldsymbol{r}) - s_j(\boldsymbol{y})] \rangle$, which enables us to determine $P(\boldsymbol{s}\,|\,\boldsymbol{r})$. Now,

$$\begin{aligned} & \langle [s_i(\boldsymbol{y}+\boldsymbol{r}) - s_i(\boldsymbol{r})]\,[s_j(\mathrm{y}+\mathrm{r}) - s_j\mathrm{r})] \rangle \\ = \; & \langle s_i(\boldsymbol{y}+\boldsymbol{r}) s_j(\boldsymbol{y}+\boldsymbol{r}) + s_i(\boldsymbol{r}) s_j(\boldsymbol{r}) - s_i(\boldsymbol{r}) s_j(\boldsymbol{y}+\boldsymbol{r}) - s_i(\boldsymbol{r}) s_j(\boldsymbol{y}+\boldsymbol{r}) \rangle \end{aligned} \tag{C1}$$

so we need to evaluate terms like $\langle s_i(\boldsymbol{y})\, s_j(\boldsymbol{y}+\boldsymbol{r})\rangle$. Recall (equation 2.39) that the components $\{\ \boldsymbol{s}_k\ \}$ and $\{\ \delta_k\ \}$ of the displacement and density fields are related by

$$(\boldsymbol{s}_k)_j \;=\; -i\,\frac{\delta_k}{k}\,(\hat{\boldsymbol{k}})_j, \tag{C2}$$

where $\hat{\boldsymbol{k}} \;=\; \boldsymbol{k}\,/\,|\,\boldsymbol{k}\,|$ and $(\boldsymbol{k})_j$ is the jth component of the vector $\boldsymbol{k}$, etc.

From equation C2 we can see that

$$\begin{aligned}\langle s_i(\boldsymbol{y})\, s_j(\boldsymbol{y}+\boldsymbol{r})\rangle \;&=\; \Big\langle \sum_{\boldsymbol{k}}\sum_{\boldsymbol{k}'} k^{-1}k'^{-1}\delta_k^*\,\delta_{k'}\,e^{i\,\boldsymbol{k}'\cdot\boldsymbol{r}}\,e^{i\,(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{y}}(\hat{\boldsymbol{k}})_i\,(\hat{\boldsymbol{k}}')_j\,\mathrm{e}^{-\frac{1}{2}(k^2+k'^2)R_f^2}\Big\rangle \\ &=\; \sum_{\boldsymbol{k}}\frac{|\,\delta_{\boldsymbol{k}}\,|^2}{k^2}\,e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(\hat{\boldsymbol{k}})_i\,(\hat{\boldsymbol{k}})_j\mathrm{e}^{-k^2R_f^2}\end{aligned}$$

since the cross-terms vanish, following the imposition of periodic boundary conditions.

We may evaluate this in a system of Cartesian coordinates, whose 3-axis lies along the direction of the vector $\hat{\boldsymbol{r}} = \boldsymbol{r}/\,|\,\boldsymbol{r}\,|$ and where ϕ is the angle measured from that axis, in which case

$$\begin{aligned}(\hat{\boldsymbol{k}})_j \;&=\; \sqrt{1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2}\,\cos\!\phi \quad \text{for } j=1 \\ &=\; \sqrt{1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2}\,\sin\!\phi \quad \text{for } j=2 \\ &=\; \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}} \quad \text{for } j=3\end{aligned}$$

and, thus,

$$\begin{aligned}\langle s_i(\boldsymbol{y})s_j(\boldsymbol{y}+\boldsymbol{r})\rangle \;&=\; 0 \quad \text{for } i\neq j \\ &=\; \int_0^\infty \Delta^2(k)\,e^{i\boldsymbol{k}\cdot\boldsymbol{r}}(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2\,\mathrm{e}^{-k^2R_f^2}\,\frac{\mathrm{d}k}{k^3} \quad \text{for } i=j=3 \\ &=\; \int_0^\infty \Delta^2(k)\,e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\,\frac{[1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2]}{2}\,\mathrm{e}^{-k^2R_f^2}\,\frac{\mathrm{d}k}{k^3} \quad \text{for } i=j\neq 3\end{aligned}$$

since $\langle\cos\!\phi\rangle = \langle\sin\!\phi\rangle = 0$ and $\langle\cos^2\!\phi\rangle = 1/2$.

In this coordinate system this may be expressed more concisely as

$$\langle s_i(\boldsymbol{y})s_j(\boldsymbol{y}+\boldsymbol{r})\rangle \;=\; \sigma_s^2\left[\psi_{\parallel}\,\hat{\boldsymbol{r}}_i\hat{\boldsymbol{r}}_j \;+\; \psi_{\perp}(\delta_{ij}-\hat{\boldsymbol{r}}_i\hat{\boldsymbol{r}}_j)\right] \tag{C3}$$

where $\sigma_s, \psi_{||}, \psi_{\perp}$ are as defined in equations 2.44, 2.55, 2.56 respectively.

Combining equations C1 and C3 we obtain

$$\langle [s_i(\boldsymbol{y}+\boldsymbol{r}) - s_i(\boldsymbol{r})]\,[s_j(\mathbf{y}+\mathbf{r}) - s_j\mathbf{r})]\rangle \;=\; 2\sigma_s^2(1-\psi_{\perp})\delta_{ij} + 2\sigma_s^2(\psi_{\perp}-\psi_{||})\hat{r}_i\hat{r}_j, \quad \text{(C4)}$$

so, writing A$\equiv\ 2\sigma_s^2(1-\psi_{\perp})$ and B$\equiv\ 2\sigma_s^2(\psi_{\perp}-\psi_{||})$ the determinant of the covariance matrix M_{ij} is given by

$$|| M ||= A^2(A+B) \quad \text{(C5)}$$

and its inverse is

$$M^{-1} = \frac{1}{|| M ||}\begin{pmatrix} A(A+B) & 0 & 0 \\ 0 & A(A+B) & 0 \\ 0 & 0 & A^2 \end{pmatrix}.$$

From equation C2 we see that the displacement field obeys Gaussian statistics if the linear density field is Gaussian (which it is by assumption) and, since, the difference of two Gaussian variables is itself Gaussian, it follows that $P(\boldsymbol{s} \mid \boldsymbol{r})$ will obey Gaussian statistics. $P(\boldsymbol{s} \mid \boldsymbol{r})$ then has the form

$$P(\boldsymbol{s} \mid \boldsymbol{r}) \;=\; \frac{1}{(2\pi)^{3/2}\ || M ||^{1/2}} \exp\left(\frac{-s_i M_{ij}^{-1} s_j}{2}\right)$$

$$\;=\; \frac{1}{(2\pi)^{3/2}\ || M ||^{1/2}}\ e^{-Q}$$

where

$$2Q \;=\; \frac{1}{|| M ||}(A^2s^2 + ABs^2 - ABs_is_j\hat{r}_i\hat{r}_j)$$

$$\;=\; \frac{A(A+B)s^2 - AB(\mathbf{s}\cdot\hat{\mathbf{r}})^2}{A^2(A+B)}$$

Combining this with

$$P(x \mid \ \boldsymbol{r}, \boldsymbol{s}) = \frac{\delta(x- \mid \boldsymbol{r}-\boldsymbol{s} \mid)}{4\pi x^2} \quad \text{(C6)}$$

and

$$P(\mathbf{r}) = 1 \tag{C7}$$

yields the following expression for the correlation function

$$\begin{aligned} 1 + \xi_Z(x) &= \int \mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \frac{1}{(2\pi)^{3/2}\,(2\sigma_s^2)^{3/2}\,(1-\psi_\perp)\,(1-\psi_{||})^{1/2}} \frac{\delta(x - \mid \boldsymbol{r} - \boldsymbol{s} \mid)}{4\pi x^2} \\ &\cdot \exp\left\{ - \left[\frac{(1-\psi_{||})s^2 - (\psi_\perp - \psi_{||})(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{4\sigma_s^2(1-\psi_\perp)(1-\psi_{||})} \right] \right\}. \end{aligned} \tag{C8}$$

The δ-function restricts the integration over $\boldsymbol{s}$ to the sphere at constant $\mid \boldsymbol{x} \mid$. Using the same coordinate system as before, we have

$$s^2 = r^2 + x^2 - 2\mathbf{x}\cdot\mathbf{r} = r^2 + x^2 - 2xr\cos\theta \tag{C9}$$

and

$$\boldsymbol{s}\cdot\hat{\boldsymbol{r}} = (\boldsymbol{r} - \boldsymbol{x})\cdot\hat{\boldsymbol{r}} = r - x\cos\theta \tag{C10}$$

and, thus, equation C8 becomes

$$\begin{aligned} 1 \;+\; \xi_Z(x) &= \int_0^\infty \frac{r^2\,\mathrm{d}r}{(2\pi)^{3/2}\,(2\sigma_s^2)^{3/2}\,(1-\psi_\perp)\,(1-\psi_{||})^{1/2}\,4\pi x^2} \\ &\cdot \int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} x^2\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\,\exp\left[\frac{(r - x\cos\theta)^2}{2D} - \frac{(r^2 + x^2 - 2xr\cos\theta)}{2A}\right] \end{aligned} \tag{C11}$$

where

$$\begin{aligned} D &\equiv \frac{A(A+B)}{B} \\ &= \frac{2\sigma_s^2(1-\psi_\perp)(1-\psi_{||})}{(\psi_\perp - \psi_{||})} \end{aligned}$$

Performing the trivial integration over ϕ and making the substitution $\mu \equiv \cos\theta$ we obtain

$$\begin{aligned} 1 + \xi_Z(\mathrm{x}) &= \int_0^\infty \frac{r^2\,\mathrm{d}r}{(2\pi)^{1/2}\,(2\sigma_s^2)^{3/2}\,(1-\psi_\perp)\,(1-\psi_{||})^{1/2}} \\ &\cdot \int_{\mu=-1}^{-1} \mathrm{d}\mu\,\exp\left[\frac{(x\mu)^2}{2D} + xr\mu E\right]\,\exp\left[\frac{-1}{2}\,\frac{(r^2+x^2)}{A} + \frac{r^2}{2D}\right] \end{aligned}$$

where $\mathrm{E} \equiv \frac{1}{A} - \frac{1}{D}$.

The integral over μ is facilitated by making the substitution

$$w = \frac{x\mu}{\sqrt{2D}} \tag{C12}$$

whence it becomes

$$I_\mu = \int_{w=\frac{-x}{\sqrt{2D}}}^{\frac{x}{\sqrt{2D}}} \mathrm{d}w \, \frac{\sqrt{2D}}{x} \exp\left[(w+a)^2 - a^2\right] \tag{C13}$$

where $a \equiv \mathrm{Er}\sqrt{D/2}$.

So, letting

$$\begin{cases} u_\pm & \equiv \; a \pm x/\sqrt{2D} \\ y & = \; w + a \\ F(u) & \equiv \; \int_0^u dy \; e^{y^2-u^2} \end{cases}$$

we may write

$$\begin{aligned} I_\mu &= \frac{(2D)^{1/2}}{x}\left[\int_{y=0}^{u_+} dy \; e^{y^2-u_+^2} \; e^{\left(\frac{x^2}{2D} + \frac{2ax}{\sqrt{2D}}\right)} \; - \; \int_{y=0}^{u_-} dy \; e^{y^2-u_-^2} \; e^{\left(\frac{x^2}{2D} - \frac{2ax}{\sqrt{2D}}\right)}\right] \\ &= \frac{\sqrt{2D}}{x}\left[F(u_+)\exp\left(\frac{x^2}{2D} + \frac{2ax}{\sqrt{2D}}\right) - F(u_-)\exp\left(\frac{x^2}{2D} - \frac{2ax}{\sqrt{2D}}\right)\right] \end{aligned} \tag{C14}$$

Thus, the integral equation for the correlation function becomes

$$\begin{aligned} 1 + \xi_{\mathrm{Z}}(x) &= \int_0^\infty \frac{r^2 \, \mathrm{d}r}{(2\pi)^{1/2}\,(2\sigma_s^2)^{3/2}\,(1-\psi_\perp)\,(1-\psi_{||})^{1/2}} \exp\left[-\frac{1}{2}\left(\frac{r^2+x^2}{A}\right) + \frac{r^2}{2D}\right] \\ &\quad \cdot \; \frac{\sqrt{2D}}{x}\left[F(u_+)\exp\left(\frac{x^2}{2D} + \frac{2ax}{\sqrt{2D}}\right) - F(u_-)\exp\left(\frac{x^2}{2D} - \frac{2ax}{\sqrt{2D}}\right)\right] \end{aligned} \tag{C15}$$

Consider the exponent in the $F(u_+)$ term:

$$= -\frac{1}{2}\left(\frac{r^2+x^2}{A}\right) + \frac{r^2}{2D} + \frac{x^2}{2D} + \frac{2ax}{\sqrt{2D}} = -\frac{(r-x)^2}{2}\left(\frac{1}{A}-\frac{1}{D}\right) \tag{C16}$$

and, similarly, the exponent in the $F(u_-)$ term $= -\frac{(r+x)^2}{2}\left(\frac{1}{A}-\frac{1}{D}\right)$.

Now,

$$\frac{1}{A}-\frac{1}{D} = \frac{1}{A}-\frac{B}{A(A+B)} = \frac{1}{A+B} = \frac{1}{2\sigma_s^2\,(1-\psi_{||})} \tag{C17}$$

and, recalling that

$$u_\pm = a \pm \frac{x}{\sqrt{2D}} \tag{C18}$$

we have

$$u_\pm = \sqrt{\frac{D}{2}}\,Er \pm \frac{x}{\sqrt{2D}} = \frac{1}{\sqrt{2D}}\,(DEr \pm x) \tag{C19}$$

but

$$\frac{1}{\sqrt{2D}} = \frac{(\psi_\perp - \psi_{||})^{1/2}}{2\,\sigma_s^2\,(1-\psi_\perp)^{1/2}\,(1-\psi_{||})^{1/2}} \tag{C20}$$

while

$$DE = D\left(\frac{1}{D}-\frac{1}{A}\right) = 1-\frac{D}{A} = \left(\frac{\psi_\perp - 1}{\psi_\perp - \psi_{||}}\right) \tag{C21}$$

So, finally, we have the following form for the correlation function

$$\begin{aligned} 1+\xi_Z(x) \;=\; & \frac{1}{2\sqrt{\pi}\sigma_s^2 x}\int_0^\infty \frac{r^2\mathrm{dr}}{(1-\psi_\perp)^{1/2}(\psi_\perp-\psi_{||})^{1/2}} \\ & \cdot\left\{F(u_+)\exp\left[-\frac{(r-x)^2}{4\sigma_s^2(1-\psi_{||})}\right] - F(u_-)\exp\left[-\frac{(r+x)^2}{4\sigma_s^2(1-\psi_{||})}\right]\right\} \end{aligned} \tag{C22}$$

where $F(u)$ is Dawson's integral,

$$F(u) \equiv \int_0^u e^{y^2-u^2} dy \tag{C23}$$

and where the arguments $u_\pm$ are given by

$$u_\pm \equiv \frac{(\psi_\perp - \psi_{||})^{1/2}}{2\sigma_s(1-\psi_{||})^{1/2}(1-\psi_\perp)^{1/2}} \left[r\frac{(1-\psi_\perp)}{(\psi_\perp - \psi_{||})} \pm x \right] \tag{C24}$$

N.B. This is slightly different from equation (2.1b) of Bond & Couchman, which contains a typographical error.

D Derivation of $\xi_{cZ}(x)$

In this section we present the derivation of the cluster correlation function given in equation 2.63 of Chapter 2. This derivation largely mirrors that of the previous section and, again, its utility exceeds its elegance.

Recall that the cluster correlation function may be written in terms of conditional probabilities as

$$1 + \xi_{cZ}(x) = \int \mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \mathrm{d}\nu_1\, \mathrm{d}\nu_2\, P(x\,|\,\boldsymbol{r},\boldsymbol{s})\, P(\boldsymbol{r})\, P(1c\,|\,\nu_1)\, P(2c\,|\,\nu_2)\, P(\boldsymbol{s},\nu_1,\nu_2\,|\,\boldsymbol{r}). \tag{D1}$$

If the fields ν_1, ν_2 are Gaussian, then so is the displacement field, $\boldsymbol{s}$,and, hence, $P(\boldsymbol{s},\nu_1,\nu_2 \mid \boldsymbol{r})$ will be a multivariate Gaussian, taking the form

$$P(\boldsymbol{s},\nu_1,\nu_2 \mid \mathbf{r}) = \frac{1}{(2\pi)^{5/2} \mid\mid M \mid\mid^{1/2}} \exp\left(-\frac{1}{2} y_i^T M_{ij}^{-1} y_j\right) \tag{D2}$$

where $\boldsymbol{y}^T = (\mathbf{s},\nu_1,\nu_2)$ and $M_{ij} = \langle y_i y_j \rangle$.

Since $P(\boldsymbol{r}) = 1$ and $P(x \mid \boldsymbol{r},\boldsymbol{s}) = \frac{1}{4\pi x^2}\delta(x - \mid \boldsymbol{r} - \boldsymbol{s} \mid)$, equation D1 takes the form

$$1 + \xi_{cZ} = \int \frac{\mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \mathrm{d}\nu_1\, \mathrm{d}\nu_2}{(2\pi)^{5/2} \mid\mid M \mid\mid^{1/2}} \frac{\delta(x - \mid \boldsymbol{r} - \boldsymbol{s} \mid)}{4\pi x^2}$$

$$\cdot \exp[-(E-1)^2\sigma_0^2] \exp[(E-1)\sigma_0(\nu_1+\nu_2)] \exp\left(-\frac{1}{2} y_i^T M_{ij}^{-1} y_j\right) \tag{D3}$$

Now,

$$\nu_1(\boldsymbol{r}) = \frac{1}{\sigma_0} \int \delta_{\boldsymbol{k}} \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{e}^{-\frac{1}{2}k^2 R_f^2} \mathrm{d}^3\boldsymbol{k} \tag{D4}$$

and, thus,

$$\langle \nu_1(\boldsymbol{r}_1)\nu_2(\boldsymbol{r}_1+\boldsymbol{r})\rangle = \int \nu_1^*(\boldsymbol{r}_1)\, \nu_2(\boldsymbol{r}_1+\boldsymbol{r})\, \mathrm{d}^3\boldsymbol{r}_1$$

$$= \frac{V^2}{\sigma_0^2(2\pi)^6} \int \mathrm{d}^3\boldsymbol{k}\, \mathrm{d}^3\boldsymbol{k}'\, \mathrm{d}^3\boldsymbol{r}\, \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{\eta}}\, \mathrm{e}^{-\frac{1}{2}k^2R_f^2}\, \delta_{\boldsymbol{k}}^*\, \mathrm{e}^{i\boldsymbol{k}'\cdot(\boldsymbol{\eta}+\boldsymbol{r})}\, \delta_{\boldsymbol{k}'}\, \mathrm{e}^{-\frac{1}{2}k'^2R_f^2}.$$

The imposition of periodic boundary conditions means that

$$\int \mathrm{e}^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{\eta}}\, \mathrm{d}^3\boldsymbol{r}_1 \longrightarrow \delta(\boldsymbol{k}'-\boldsymbol{k})\frac{(2\pi)^3}{V} \tag{D5}$$

and we obtain

$$\begin{aligned} \langle \nu_1\, \nu_2\rangle &= \frac{V}{\sigma_0^2(2\pi)^3} \int \mathrm{e}^{-k^2R_f^2}\, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, |\, \delta_{\boldsymbol{k}}|^2\, \mathrm{d}^3\boldsymbol{k} \\ &= \frac{1}{\sigma_0^2}\, \xi(\boldsymbol{r}) \equiv \psi \\ &= \langle \nu_2\nu_1\rangle, \end{aligned} \tag{D6}$$

where $\psi(r)$ is the scaled correlation function of the background density field. It follows that $\langle \nu_1\, \nu_1\rangle = \langle \nu_2\, \nu_2\rangle = 1$ since $\psi(0) = 1$, by definition.

From equation C2 and the definition of ν it follows that

$$\langle \nu(\boldsymbol{r}_1)\, s(\boldsymbol{r}_1+\boldsymbol{r})\rangle = \frac{-iV}{\sigma_0(2\pi)^3} \int |\, \delta_{\boldsymbol{k}}|^2\, \mathrm{e}^{-k^2R_f^2}\, \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}}\, \frac{\boldsymbol{k}}{k^2}\, \mathrm{d}^3\boldsymbol{k}. \tag{D7}$$

Evaluating this in the same coordinate system as before, we have

$$\boldsymbol{k} = k(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{D8}$$

and

$$\boldsymbol{s} = (s_1, s_2, s_3) \tag{D9}$$

then $\langle s_1 \nu \rangle = 0$ because $\langle \cos\phi \rangle = 0$ and, similarly, $\langle s_2 \nu \rangle = 0$ because $\langle \sin\phi \rangle = 0$, so the only non-zero term is $\langle s_3 \nu \rangle$:

$$\begin{aligned} \langle \nu(\boldsymbol{r}_1) s_3(\boldsymbol{r}_1 + \mathbf{r}) \rangle &= \frac{-iV}{\sigma_0 (2\pi)^3} \int |\, \delta_{\boldsymbol{k}} \,|^2 \; \mathrm{e}^{i\boldsymbol{k}\cdot\boldsymbol{r}} \mathrm{e}^{-k^2 R_f^2} \, \frac{k \cos\theta}{k^2} \, k^2 \, \mathrm{d}\cos\theta \, \mathrm{d}\phi \, \mathrm{d}k \\ &= \frac{-iV}{\sigma_0 (2\pi)^3} \int |\, \delta_{\boldsymbol{k}} \,|^2 \; k \, \mathrm{e}^{-k^2 R_f^2} \, \mathrm{d}k \cdot \int \mathrm{e}^{ikr\mu} \, \mu \, \mathrm{d}\mu \, \mathrm{d}\phi \end{aligned} \tag{D10}$$

where $\mu = cos\theta$. If we define $\sigma_s^2 \psi_{\mathrm{v}} \equiv \sigma_s^2 (2\psi_\perp + \psi_{||})/3$, then it follows from the definitions of $\psi_\perp$ and $\psi_{||}$ that

$$\sigma_s^2 \psi_{\mathrm{v}} = \frac{V}{3(2\pi)^3} \int |\, \delta_{\boldsymbol{k}} \,|^2 \; \mathrm{e}^{-k^2 R_f^2} \, \mathrm{d}k \cdot \int \mathrm{e}^{ikr\mu} \, \mathrm{d}\mu \, \mathrm{d}\phi \tag{D11}$$

and, therefore, that

$$\sigma_s^2 \, \frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r} \, \frac{iV}{3(2\pi)^3} \int |\, \delta_{\boldsymbol{k}} \,|^2 \, \mathrm{e}^{-k^2 R_f^2} \, k \, \mathrm{d}k \cdot \int \mu \mathrm{e}^{ikr\mu} \, \mathrm{d}\mu \mathrm{d}\phi = \frac{-\sigma_0}{3} \, \langle \nu s_3 \rangle \tag{D12}$$

i.e.

$$\langle \nu \; s_3 \rangle = -\frac{3\sigma_s^2}{\sigma_0} \left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r} \right). \tag{D13}$$

Thus, the covariance matrix for the Gaussian $P(\boldsymbol{s}, \nu_1, \nu_2 \mid \boldsymbol{r})$ is

$$M = \begin{pmatrix} A & 0 & 0 & 0 & 0 \\ 0 & A & 0 & 0 & 0 \\ 0 & 0 & A+B & C & C \\ 0 & 0 & C & 1 & D \\ 0 & 0 & C & D & 1 \end{pmatrix}$$

where

$$\begin{aligned} A &= 2\sigma_s^2 (1 - \psi_\perp) \\ B &= 2\sigma_s^2 (\psi_\perp - \psi_{||}) \\ C &= \frac{-3\sigma_s^2}{\sigma_0} \left(\frac{d\psi_{\mathrm{v}}}{\mathrm{d}r} \right) \\ D &= \psi \end{aligned}$$

and where we have used the results of the derivation of the equivalent expression from the previous section to give the matrix elements $\langle s_i\, s_j \rangle$.

This matrix is of block-diagonal form, so, clearly, we may consider separately the components for (s_1, s_2) and (s_3, ν_1, ν_2) since there will be no cross-terms in the resultant quadratic form linking these two sets of quantities. We consider first the covariant matrix for the vector $y^T = (s_3, \nu_1, \nu_2)$, which has the following form:

$$\tilde{M} = \begin{pmatrix} A+B & C & C \\ C & 1 & D \\ C & D & 1 \end{pmatrix}$$

and an inverse

$$\tilde{M}^{-1} = \frac{1}{\| \tilde{M} \|} \begin{pmatrix} 1-D^2 & C(D-1) & C(D-1) \\ C(D-1) & A+B-C^2 & C^2 - D(A+B) \\ C(D-1) & C^2 - D(A+B) & A+B-C^2 \end{pmatrix}$$

where

$$\| \tilde{M} \| = (1-D)[(A+B)(1+D) - 2C^2]. \tag{D14}$$

At this point it is convenient to make the following definitions:

$$\begin{cases} \alpha & \equiv & (1-D^2)/(2 \| \tilde{M} \|) \\ \beta & \equiv & [C(D-1)]/(2 \| \tilde{M} \|) \\ \gamma & \equiv & (A+B-C^2/(2 \| \tilde{M} \|) \\ \delta & \equiv & [C^2 - (A+B)D]/(2 \| \tilde{M} \|) \end{cases}$$

Then the quadratic form $-\frac{1}{2}(\tilde{y}^{\mathrm{T}} \cdot \tilde{\mathrm{M}}^{-1} \cdot \tilde{y})$ is given by

$$-\frac{1}{2}(\tilde{y}^{\mathrm{T}} \cdot \tilde{\mathrm{M}}^{-1} \cdot \tilde{y}) = -(s_3\ \nu_1\ \nu_2) \begin{pmatrix} \alpha & \beta & \beta \\ \beta & \gamma & \delta \\ \beta & \delta & \gamma \end{pmatrix} \begin{pmatrix} s_3 \\ \nu_1 \\ \nu_2 \end{pmatrix} \tag{D15}$$

$$= -\gamma\nu_1^2 - \nu_1\,[2(s_3\beta + \nu_2\delta)] - (\alpha s_3^2 + 2\beta s_3\nu_2 + \nu_2^2\gamma) \tag{D16}$$

Now, from equations D1 and D2, we see that we wish to evaluate the integral

$$I = \frac{1}{(2\pi)^{3/2}\ ||\ \tilde{M}\ ||^{-1/2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{(E-1)\sigma_0\nu_1}\ \mathrm{e}^{(E-1)\sigma_0\nu_2} \exp\left[-\frac{1}{2}\left(\tilde{y}^{\mathrm{T}} \cdot \tilde{\mathrm{M}}^{-1} \cdot \tilde{y}\right)\right] \mathrm{d}\nu_1\ \mathrm{d}\nu_2$$

$$= \frac{1}{(2\pi)^{3/2}\ ||\ \tilde{M}\ ||^{-1/2}} \int_{-\infty}^{\infty} \mathrm{e}^{\epsilon\nu_2}\ I_1\ \mathrm{d}\nu_2 \tag{D17}$$

where we define $\epsilon \equiv (E-1)\sigma_0$ and

$$I_1 \equiv \int_{-\infty}^{\infty} \mathrm{e}^{\epsilon\nu_1} \exp\left[-\frac{1}{2}\left(\tilde{y}^{\mathrm{T}} \cdot \tilde{\mathrm{M}}^{-1} \cdot \tilde{y}\right)\right] \mathrm{d}\nu_1 = \int_{-\infty}^{\infty} \mathrm{e}^{-Q_1}\ \mathrm{d}\nu_1 \tag{D18}$$

where Q_1 is given by

$$Q_1 = \gamma\nu_1^2 + \nu_1\,[2(s_3\beta + \nu_2\delta) - \epsilon] + (\alpha s_3^2 + 2\beta s_3\nu_2 + \nu_2^2\gamma) \tag{D19}$$

Now, using the standard result (Gradshteyn & Ryzhik 1980, § 3.323)

$$\int_{-\infty}^{\infty} \exp(-p^2x^2 \pm qx)\,\mathrm{d}x = \frac{\sqrt{\pi}}{p} \exp\left(\frac{q^2}{4p^2}\right) \tag{D20}$$

we find that

$$I_1 = e^{-(\alpha s_3^2 + 2\beta s_3\nu_2 + \nu_2^2\gamma)} \sqrt{\frac{\pi}{\gamma}} \exp\left\{\frac{[2(s_3\beta + \delta\nu_2) - \epsilon)^2]}{4\gamma}\right\}. \tag{D21}$$

Thus, from equations D14 - D18 it follows that

$$I = \frac{1}{\sqrt{8\pi^2\,\gamma\ ||\ \tilde{M}\ ||}} \int_{-\infty}^{\infty} \mathrm{e}^{\epsilon\nu_2}\ \mathrm{e}^{-(\alpha s_3^2 + 2\beta s_3\nu_2 + \nu_2^2\gamma)} \exp\left\{\frac{[2(\beta s_3 + \delta\nu_2) - \epsilon]^2}{4\gamma}\right\} \mathrm{d}\nu_2$$

$$= \frac{1}{\sqrt{8\pi^2\gamma\ ||\ \tilde{M}\ ||}} \int_{-\infty}^{\infty} \mathrm{e}^{-Q_2}\ \mathrm{d}\nu_2 \tag{D22}$$

where

$$Q_2 = \frac{\nu_2^2}{\gamma}(\gamma^2 - \delta^2) + \frac{\nu_2}{\gamma}[(2\beta s_3 - \epsilon)(\gamma - \delta)] + \frac{1}{4\gamma}[4s_3^2(\alpha\gamma - \beta^2) + 4s_3\beta\epsilon - \epsilon^2] \quad \text{(D23)}$$

So, using equation D20 we see that

$$I = \frac{1}{(2\pi)^{3/2} \mid\mid \tilde{M} \mid\mid^{1/2}} \sqrt{\frac{\pi}{\gamma}} \sqrt{\frac{\pi\gamma}{\gamma^2 - \delta^2}} \cdot \exp\left\{-\frac{[4s_3^2(\alpha\gamma - \beta^2) + 4s_3\beta\epsilon - \epsilon^2]}{4\gamma}\right\}$$

$$\cdot \exp\left\{\left[\frac{\{2\beta s_3 - \epsilon\}^2(\gamma - \delta)^2}{4\gamma^2}\right]\left(\frac{\gamma}{\gamma^2 - \delta^2}\right)\right\}$$

$$= \frac{1}{\sqrt{8\pi \mid\mid \tilde{M} \mid\mid (\gamma^2 - \delta^2)}} e^X \quad \text{(D24)}$$

where

$$X = \frac{(2\beta s_3 - \epsilon)^2(\gamma - \delta)^2}{4\gamma(\gamma^2 - \delta^2)} - \frac{[4s_3^2(\alpha\gamma - \beta^2) + 4s_3\beta\epsilon - \epsilon^2]}{4\gamma}$$

$$= \frac{-s_3^2}{\gamma(\gamma^2 - \delta^2)}\left[\beta^2(\gamma - \delta)^2 - (\alpha\gamma - \beta^2)(\gamma^2 - \delta^2)\right]$$

$$- \frac{\epsilon\beta s_3}{\gamma(\gamma^2 - \delta^2)}\left[(\gamma - \delta)^2 - (\gamma^2 - \delta^2)\right] + \frac{\epsilon^2}{4\gamma(\gamma^2 - \delta^2)}\left[(\gamma - \delta)^2 + (\gamma^2 - \delta^2)\right]$$

$$= s_3^2\left[\frac{2\beta^2}{(\gamma + \delta)} - \alpha\right] - \frac{2\epsilon\beta s_3}{(\gamma + \delta)} + \frac{\epsilon^2}{2(\gamma + \delta)}. \quad \text{(D25)}$$

Recalling the definitions of α, β, γ and δ we find

$$\gamma^2 - \delta^2 = \frac{1}{4 \mid\mid \tilde{M} \mid\mid^2}\{(A + B - C^2)^2 - [C^2 - (A + B)D]^2\}$$

$$= \frac{(A + B)(1 - D)}{4 \mid\mid \tilde{M} \mid\mid^2}[(A + B)(1 + D) - 2C^2] = \frac{A + B}{4 \mid\mid \tilde{M} \mid\mid} \quad \text{(D26)}$$

and so

$$\sqrt{8\pi \mid\mid \tilde{M} \mid\mid (\gamma^2 - \delta^2)} = \sqrt{2\pi(A + B)}. \quad \text{(D27)}$$

Similarly,

$$\gamma + \delta = \frac{(A + B)(1 - D)}{2 \mid\mid \tilde{M} \mid\mid}, \quad \text{(D28)}$$

so

$$\frac{2\beta^2}{(\gamma+\delta)} - \alpha = \frac{4 \mid\mid \tilde{M} \mid\mid}{(A+B)(1-D)} \cdot \frac{C^2(D-1)^2}{4 \mid\mid \tilde{M} \mid\mid^2} - \frac{(1-D^2)}{2 \mid\mid \tilde{M} \mid\mid}$$

$$= \frac{1}{2(A+B) \mid\mid \tilde{M} \mid\mid} \left[2C^2(1-D) - (A+B)(1-D^2)\right]$$

$$= \frac{-1}{2(A+B)} \tag{D29}$$

while

$$\frac{\beta}{\gamma+\delta} = \frac{C(D-1)}{2 \mid\mid \tilde{M} \mid\mid} \cdot \frac{2 \mid\mid \tilde{M} \mid\mid}{(A+B)(1-D)} = \frac{-C}{(A+B)} \tag{D30}$$

and

$$\frac{1}{(\gamma+\delta)} = \frac{2 \mid\mid \tilde{M} \mid\mid}{(A+B)(1-D)} = 2\left[\frac{(A+B)(1+D) - 2C^2}{(A+B)}\right]$$

$$= 2\left[(1+D) - \frac{2C^2}{(A+B)}\right] \tag{D31}$$

If we combine these with equations D21 and D22 we obtain

$$I = \frac{1}{\sqrt{2\pi(A+B)}} \exp\left[-\frac{s_3^2}{2(A+B)}\right] \exp\left[\frac{2(E-1)\sigma_0 C s_3}{(A+B)}\right]$$

$$\cdot \exp\left[(E-1)^2\sigma_0(1+D) - \frac{2(E-1)^2\sigma_0^2C^2}{(A+B)}\right]. \tag{D32}$$

Then

$$P(\boldsymbol{s}, \nu_1, \nu_2 \mid r) = \frac{I}{2\pi A} \exp\left[\frac{-(s_1^2+s_2^2)}{2A}\right] \tag{D33}$$

and so

$$1 + \xi_{cZ}(x) = \int \mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \frac{\delta(x - \mid \boldsymbol{r} - \boldsymbol{s} \mid)}{4\pi x^2} \exp[-(E-1)^2\sigma_0^2]\, P(\boldsymbol{s}, \nu_1, \nu_2 \mid r)$$

$$= \int \mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \frac{\delta(x - \mid \boldsymbol{r} - \boldsymbol{s} \mid)}{4\pi x^2} \frac{\exp[-(E-1)^2\sigma_0^2]}{\sqrt{(2\pi)^3A^2(A+B)}}$$

$$\cdot \exp\left[\frac{-s_3^2}{2(A+B)}\right] \exp\left[\frac{-(s_1^2+s_2^2)}{2A}\right] \exp\left[\frac{2(E-1)\sigma_0 C s_3}{(A+B)}\right]$$

$$\cdot \quad \exp\left[(E-1)^2\sigma_0^2(1+D) \; - \; \frac{2(E-1)^2\sigma_0^2C^2}{(A+B)}\right] \tag{D34}$$

$$= \int \frac{\mathrm{d}^3\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{s}}{\sqrt{(2\pi)^3A^2(A+B)}} \; \frac{\delta(x-\mid \boldsymbol{r}-\boldsymbol{s}\mid)}{4\pi x^2} \; \exp[(E-1)^2\xi_{\mathrm{L}}] \cdot \exp\left[\frac{-18(E-1)^2\sigma_s^4}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)^2\right]$$

$$\cdot \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} x^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \; \exp\left[\frac{(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{2A} - \frac{s^2}{2A} - \frac{(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{2(A+B)} - \frac{6(E-1)\sigma_s^2(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)\right] \tag{D35}$$

since, in our chosen coordinate system, $s_3 = \boldsymbol{s}\cdot\hat{\boldsymbol{r}}$ and $s^2 = s_1^2 + s_2^2 + s_3^2$. We have noted that the δ-function restricts the integration over $\boldsymbol{s}$ to the sphere at constant $\mid \boldsymbol{x} \mid$ and performed the resultant (trivial) integration over the azimuthal angle. Furthermore,

$$\begin{aligned} s^2 &= r^2 + x^2 - 2\boldsymbol{x}\cdot\boldsymbol{r} \\ &= r^2 + x^2 - 2xr\mu \end{aligned}$$

$$\begin{aligned} \boldsymbol{s}\cdot\hat{\boldsymbol{r}} &= (\boldsymbol{r}-\boldsymbol{x})\cdot\hat{\boldsymbol{r}} \\ &= r - x\mu \end{aligned} \tag{D36}$$

where $\mu = \cos\theta$, so the exponent in the integration over μ may be written as

$$\frac{(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{2A} - \frac{s^2}{2A} - \frac{(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{2(A+B)} - \frac{6(E-1)\sigma_s^2(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)$$

$$= \frac{(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{2}\left[\frac{1}{A} - \frac{1}{(A+B)}\right] - \frac{s^2}{2A} - \frac{6(E-1)\sigma_s^2(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)$$

$$= \frac{(r-x\mu)^2}{2E} - \frac{(r^2+x^2-2xr\mu)}{2A} - \frac{6(E-1)\sigma_s^2(r-x\mu)}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)$$

$$= \frac{(x\mu)^2}{2E} + \mu x\left[r\left(\frac{1}{A}-\frac{1}{E}\right) - \frac{6(E-1)\sigma_s^2}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)\right]$$

$$- \frac{(r^2+x^2)}{2A} + \frac{r^2}{2E} - \frac{6(E-1)\sigma_s^2 r}{(A+B)}\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right)$$

$$= \frac{(x\mu)^2}{2E} - \mu x F(r+\chi) - \frac{(r^2+x^2)}{2A} + \frac{r^2}{2E} + Fr\chi \tag{D37}$$

where

$$\chi = 6(E-1)\,\sigma_s^2\left(\frac{\mathrm{d}\psi_{\mathrm{v}}}{\mathrm{d}r}\right) \tag{D38}$$

$$E = \frac{A(A+B)}{B} \tag{D39}$$

so

$$\frac{1}{E} - \frac{1}{A} = \frac{B}{A(A+B)} = \frac{-1}{(A+B)} = F. \tag{D40}$$

Now, if we let

$$\tilde{w} = \frac{x\mu}{\sqrt{2E}} \tag{D41}$$

and

$$\tilde{a} = -F(r+\chi)\sqrt{E/2} \tag{D42}$$

and then we may write the integral equation for the correlation function in the form

$$1 + \xi_{cZ} = \int_0^\infty \frac{r^2\,dr\,\exp[(E-1)^2\xi_L]\,2\pi x^2}{\sqrt{(2\pi)^3 A^2(A+B)}} \cdot \exp\left[\frac{-(r^2+x^2)}{2A} + \frac{r^2}{2E} + Fr\chi - \frac{F\chi^2}{2}\right] \cdot I_\mu \tag{D43}$$

where

$$I_\mu = \int_{-1}^{1} d\mu\,\exp[(\tilde{w}+\tilde{a})^2 - \tilde{a}^2]$$

$$= \int_{-x/\sqrt{2E}}^{x/\sqrt{2E}} d\tilde{w}\,\exp[(\tilde{w}+\tilde{a})^2 - \tilde{a}^2] \cdot \frac{\sqrt{2E}}{x} \tag{D44}$$

If we now define

$$\begin{cases} \tilde{y} = \tilde{w} + \tilde{a} \\ \tilde{u}_\pm = \tilde{a} \pm x/\sqrt{2E} \end{cases}$$

then

$$I_\mu = \int_{\tilde{u}_-}^{\tilde{u}_+} d\tilde{y} \exp(\tilde{y}^2 - \tilde{a}^2) \cdot \frac{\sqrt{2E}}{x}$$

$$= \frac{\sqrt{2E}}{x} \cdot \left[\int_0^{\tilde{u}_+} d\tilde{y} \exp(\tilde{y}^2 - \tilde{a}^2) - \int_0^{\tilde{u}_-} d\tilde{y} \exp(\tilde{y}^2 - \tilde{a}^2) \right]$$

$$= \frac{\sqrt{2E}}{x} \left\{ \left[\int_0^{\tilde{u}_+} d\tilde{y} \exp(\tilde{y}^2 - \tilde{u}_+^2) \right] \exp\left(\frac{x^2}{2E} + \frac{2\tilde{a}x}{\sqrt{2E}} \right) \right.$$

$$\left. - \left[\int_0^{\tilde{u}_-} d\tilde{y} \exp(\tilde{y}^2 - \tilde{u}_-^2) \right] \exp\left(\frac{x^2}{2E} - \frac{2\tilde{a}x}{\sqrt{2E}} \right) \right\} \tag{D45}$$

and, thus,

$$1 + \xi_{cZ} = \int_0^\infty \frac{r^2 \, dr \exp[(E-1)^2 \xi_L]}{\sqrt{(2\pi)^3 A^2 (A+B)}} 2\pi x^2 \frac{\sqrt{2E}}{x} \exp\left[\frac{-(r^2+x^2)}{2A} + \frac{r^2}{2E} + Fr\chi + \frac{\chi^2 F}{2} \right]$$

$$\cdot \left[F(\tilde{u}_+) \exp\left(\frac{x^2}{2E} + \frac{2\tilde{a}x}{\sqrt{2E}} \right) - F(\tilde{u}_-) \exp\left(\frac{x^2}{2E} - \frac{2\tilde{a}x}{\sqrt{2E}} \right) \right] \tag{D46}$$

Consider the exponent in the $F(\tilde{u}_+)$ term:

$$- \frac{(r^2+x^2)}{2A} + \frac{r^2}{2E} + Fr\chi - \frac{F\chi^2}{2} + \frac{x^2}{2E} - \frac{2xF(r+\chi)}{\sqrt{2E}} \sqrt{\frac{E}{2}}$$

$$= \frac{F}{2}[x - (r+\chi)]^2 = \frac{F}{2}(x - \tilde{r})^2 \tag{D47}$$

where $\tilde{r} = r + \chi$ and, similarly, the exponent in the $F(\tilde{u}_-)$ term is $(x + \tilde{r})^2 (F/2)$. Equation D46 now takes the form

$$1 + \xi_{cZ} = \int_0^\infty dr \, r^2 \exp[(E-1)^2 \xi_L] \, x \sqrt{\frac{2E}{(2\pi)^3 A^2 (A+B)}}$$

$$\cdot \left\{ F(\tilde{u}_+) \exp\left[\frac{F}{2}(x+\tilde{r})^2 \right] - F(\tilde{u}_-) \exp\left[\frac{F}{2}(x-\tilde{r})^2 \right] \right\}, \tag{D48}$$

but,

$$\frac{2\pi\sqrt{2E}}{\sqrt{(2\pi)^3 A^2 (A+B)}} = \frac{1}{\sqrt{\pi}} \sqrt{\frac{E}{A^2(A+B)}} = \frac{1}{\sqrt{AB\pi}} \tag{D49}$$

so that, making the final substitutions into equation D48 we obtain the final expression for the correlation function as

$$1 + \xi_{cZ}(x) = \frac{1}{2\sqrt{\pi} x \sigma_s^2} \int_0^\infty \frac{r^2 \, dr \exp[(E-1)^2 \xi_L]}{(1 - \psi_\perp)^{1/2} (\psi_\perp - \psi_\parallel)^{1/2}}$$

$$\cdot \quad \left\{ F(\tilde{u}_+)\exp\left[\frac{-(x-\tilde{r})^2}{4\sigma_s^2(1-\psi_{\|})}\right] \; - \; F(\tilde{u}_-)\exp\left[\frac{-(x+\tilde{r})^2}{4\sigma_s^2(1-\psi_{\|})}\right]\right\} \qquad \text{(D50)}$$

where

$$\begin{cases} \tilde{r} & = & r + 6(E-1)\sigma_s^2\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right) \\ \\ \tilde{u}_{\pm} & = & \frac{(\psi_\perp - \psi_{\|})^{1/2}}{2\sigma_s(1-\psi_{\|})^{1/2}(1-\psi_\perp)^{1/2}}\left[\tilde{r}\frac{(1-\psi_\perp)}{(\psi_\perp-\psi_{\|})} \pm x\right] \end{cases}$$

and where ξ_{L} is the correlation function of the linear density field.

REFERENCES

Aaronson M., Bothun G., Mould J., Huchra J., Schommer R., Cornell M.E., 1986, ApJ, 302, 536

Abell G.O., 1958, ApJS, 3, 211

Abell G.O., 1961, AJ, 66, 607

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

Adams W.S., 1941, ApJ, 93, 11

Adler R.J., 1981, The Geometry of Random Fields. Wiley, Chichester

Albrecht A., Steinhardt P.J., 1982, Phys. Rev. Lett., 48, 1220

Alcock C., Fahri H., 1985, Phys. Rev., D32, 1273

Alimi J.-M., Valla-Gabaud D., Blanchard A., 1988, A&A, 206, L11

Alpher R.A., Herman R., 1948, Nat, 162, 774

Arp H.C., Burbidge G., Hoyle F., Narlikar J.V., Wickramasinghe N.C., 1990, Nat, 346, 807

Ashman K.M., 1992, PASP, 104, 1109

Athanassoula, E., Bosma A., Papaioannou S., 1987, A&A, 179, 23

Babul A., Katz N., 1993, ApJ, 406, L51

Babul A., Postman M., 1990, ApJ, 359, 280

Bahcall J.N., 1984, ApJ, 276, 169

Bahcall J.N., Flynn C., Gould A., 1992, ApJ, 389, 234

Bahcall N.A., 1979, ApJ, 232, 689

Bahcall N.A., 1988, ARA&A, 26, 631

Bahcall N.A., Cen R.Y., 1992, ApJ, 398, L81

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

Bahcall N.A., Soneira R.M., 1984, ApJ, 277, 27

Bahcall N.A., Soneira R.M., Burgett W.S., 1986, ApJ, 311, 15

Bahcall N.A., West M.J., 1992, ApJ, 392, 419

Bardeen J.M., 1980, Phys. Rev., D22, 1882

Bardeen J.M., Bond J.R., Efstathiou G., 1987, ApJ, 321, 28

Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., 1986, ApJ, 304, 15

Barnes J., Dekel A., Efstathiou G., Frenk C.S., 1983, ApJ, 295, 368

Barnes J.E., Hernquist L., 1992, ARA&A, 30, 705

Batuski D.J., Bahcall N.A., Olowin R.P., Burns J.O., 1989, ApJ, 341, 599

Batuski D.J., Melott A.L., Burns J.O., 1987, ApJ, 322, 48

Batuski D.J., Melott A.L., Scherrer R.J., Bertschinger E., 1991, ApJ, 367, 393

Baugh C.M., Efstathiou G., 1993, MNRAS, in press

Bean A.J., Efstathiou G., Ellis R.S., Peterson B.A., Shanks T., 1983, MNRAS, 205, 605

Beichman C.A., Neugebauer G., Habing H.J., Clegg P.E., Chester T.J., 1988, eds, *IRAS* Catalogs and Atlases: Explanatory Supplement. US Govt. Printing Office, Washington, DC

Bertschinger E., 1985, ApJS, 58, 39

Bertschinger E., Dekel A., 1989, ApJ, 336, L5

Bertschinger E., Gelb J.M., 1991, Comput. Phys., 5, 164

Bienaymé O., Robin A.C., Crézé M., 1987, A&A, 180, 94

Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ Press, Princeton, NJ

Blau S.K., Guth A.H., 1987, in S.W. Hawking, W. Israel, eds, 300 Years of Gravitation. Cambridge Univ Press, Cambridge

Bogart R.S., Wagoner R.V., 1973, ApJ, 181, 609

Bond J.R., Couchman H.M.P., 1987, preprint

Bond J.R., Couchman H.M.P., 1988, in Coley A., Dyer C.C., Tupper B.O.J., eds, Proc. Second Canadian Conference on General Relativity & Relativistic Astrophysics. World Scientific, Singapore

Bond J.R., Efstathiou G., 1984, ApJ, 285, L45

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

Bondi H., Gold T., 1948, MNRAS, 108, 252

Bonnor W.B., 1956, MNRAS, 117, 104

Bosma A., 1978, Ph.D. thesis, Univ of Groningen

Bothun G.D., Lonsdale C.J., Rice W.L., 1989, ApJ, 341, 129

Bower R.G., Coles P., Frenk C.S., White S.D.M., 1993, ApJ, 405, 403

Briel U.G., Henry J.P., Bohringer H., A&A, 259, L31

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

Burstein D., Heiles C., 1982, ApJ, 87, 1165

Bushouse H.A., Lamb S.A., Werner M.W., 1988, ApJ, 335, 74

Canizares C.R., 1982, ApJ, 263, 50

Carlberg R.G., Couchman H.M.P., 19889, ApJ, 340, 47

Carlberg R.G., Couchman H.M.P., Thomas P.A., 1990, ApJ, 352, L29

Carroll S.M., Press W.H., Turner E.L., 1992, ARA&A, 30, 499

Castertano S., van Albada T.S., 1990, in D. Lynden-Bell, G. Gilmore, eds, Baryonic Dark Matter. Kluwer, Dordrecht

Casertano S., van Gorkom J.H., AJ, 101, 1231

Cen R.Y., Ostriker J.P., 1992, ApJ, 399, L13

Cen R.Y., Ostriker J.P., 1993, Princeton Observatory Preprint, POP-493

Chandrasekhar S., 1943, ApJ, 97, 255

Coleman S., 1977, Phys. Rev., D15, 2929

Coleman S., Weinberg E.J., 1973, Phys. Rev., D7, 1888

Coles P., 1986, MNRAS, 222, 9P

Coles P., 1989, MNRAS, 238, 319

Coles P., 1990, MNRAS, 243, 171

Coles P., 1993, MNRAS, 262, 1065

Coles P., Jones B., 1991, MNRAS, 248, 1

Coles P., Melott A.L., Shandarin S.F., 1993, MNRAS, 260, 765

Colless M.M., Ellis R.S., Taylor K., Hook R.N., 1990, MNRAS, 244, 405

Colless M.M., Ellis R.S., Broadhurst T.J., Taylor K., Peterson B.A., 1993, preprint

Collins C.B., Hawking S.W., 1973, ApJ, 180, 317

Couchman H.M.P., ApJ, 368, L23

Couchman H.M.P., Carlberg R.G., 1992, ApJ, 389, 453

Cowsik R., M^{c}Lelland J., 1973, ApJ, 180, 7

Crézé M., Robin A.C., Bienaymé O., 1989, A&A, 211,1

Crittenden R., Bond J.R., Davis R.L., Efstathiou G., Steinhardt P.J., 1993, preprint

Croft R.A.C., Efstathiou G., 1993, preprint

Dalton G.B., 1992, Ph.D. thesis, Univ of Oxford

Dalton G.B., Efstathiou G., Maddox S.J., Sutherland W.J., 1992, ApJ, 390, L1

Davis M., Efstathiou G., Frenk C.S., White S.D.M., 1985, ApJ, 292, 371

Davis M., Efstathiou G., Frenk C.S., White S.D.M., 1992, Nat, 356, 489

Davis M., Geller M.J., 1976, ApJ, 208, 13

Davis M., Huchra J.P., 1982, ApJ, 254, 437

Davis M., Peebles P.J.E., 1977, ApJS, 34, 425

Davis M., Peebles P.J.E., 1983, ApJ, 267, 465

Davis M., Summers F.J., Schlegel D., 1992, Nat, 359, 393

Dekel A., Bertschinger E., Yahil A., Strauss M.A., Davis M., Huchra J., 1993, preprint

Dekel A., Blumenthal G.R., Primack J.R., Olivier S., ApJ, 338, L5

Dicke R.H., Peebles P.J.E., Roll P.G., Wilkinson D.T., 1965, ApJ, 142, 414

Dicke R.H., Peebles P.J.E., 1979, in S.W. Hawking, W. Israel, eds, General Relativity: an Einstein Centenary Survey. Cambridge Univ Press, Cambridge

Doroshkevich A.G., Novikov I., 1964, Dokl. Akad. Nauk. SSSR, 154, 809

Dressler A., 1980, ApJ, 326, 351

Efstathiou G., 1988, in A. Lawrence, ed, Comets to Cosmology, Proc. Third *IRAS* Conference. Springer-Verlag, Berlin

Efstathiou G., 1990, in J.A. Peacock, A.F. Heavens, A.T. Davis, eds, Physics of the Early Universe. SUSSP, Edinburgh

Efstathiou G., Bond J.R., White S.D.M., 1992a, MNRAS, 258, 1P

Efstathiou G., Dalton G.B., Sutherland W.J., Maddox S.J., 1992b, MNRAS, 257, 125

Efstathiou G., Davis M., Frenk C.S., White S.D.M., 1985, ApJS, 57, 241

Efstathiou G., Eastwood J.W., 1981, MNRAS, 194, 503

Efstathiou G., Ellis R.S., Peterson B.A., 1988, MNRAS, 232, 431

Efstathiou G., Kaiser N., Saunders W., Lawrence A., Rowan-Robinson M., Ellis R.S., Frenk C.S., 1990, MNRAS, 247, 10P

Efstathiou G., Silk J., 1983, Fund. Cosmic Phys., 9, 1

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

Eggen O.J., Lynden-Bell D., Sandage A.R., 1962, ApJ, 136, 748

Ellis G.F.R., 1988, in Coley A., Dyer C.C., Tupper B.O.J., eds, Proc. Second Canadian Conference on General Relativity & Relativistic Astrophysics. World Scientific, Singapore

Feldman H.A., Kaiser N., Peacock J.A., 1993, ApJ, accepted for publication

Fesenko B.I., 1979a, Sov. Astron., 23, 524

Fesenko B.I., 1979b, Sov. Astron., 23, 657

Fetisova T.S., 1982, Sov. Astron., 25, 647

Fisher K.B., Davis M., Strauss M.A., Yahil A., Huchra J.P., 1993, ApJ, 402, 42

Fisher K.B., Strauss M.A., Davis M., Yahil A., Huchra J.P., 1992, ApJ, 389, 188

Frenk C.S., 1991, Physica Scripta, T36, 70

Frenk C.S., White S.D.M., Davis M., Efstathiou G., 1988, ApJ, 327, 507

Frenk C.S., White S.D.M., Efstathiou G., Davis M., 1985, Nat, 317, 595

Frenk C.S., White S.D.M., Efstathiou G., Davis M., 1990, ApJ, 351, 10

Gelb J.M., 1992, PhD thesis, MIT

Gelb J.M., Bertschinger E., 1993, preprint

Gelb J.M., Gradwohl B.-A., Frieman J.A., 1993, ApJ, 403, L5

Gott J.R. et al., 1989, ApJ, 340, 625

Gradshteyn I.S., Ryzhik I.M., 1980, Tables of Integrals, Series and Products. Academic Press, New York, NY

Gundersen J.O. et al., 1993, ApJ, 413, L1

Gunn J.E., Peterson B.A., 1965, ApJ, 142, 1633

Gurbatov S.N., Saichev A.I., Shandarin S.F., 1989, MNRAS, 236, 385

Guth A.H., 1981, Phys. Rev., D23, 347

Guth A.H., Tye S.-H., 1980, Phys. Rev. Lett., 44, 631: *Erratum*; Phys. Rev. Lett., 44, 963

Guth A.H., Weinberg E.J., 1981, Phys. Rev., D23, 876

Harrison E.R., 1970, Phys. Rev., D1, 2726

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

Hawking S.W., Ellis G.F.R., 1973, The Large Scale Structure of Space-Time. Cambridge Univ Press, Cambridge

Hawking S.W., Penrose R., 1970, Proc. Roy. Soc. Lond., A314, 529

Helou G., Walker D.W., 1988, *IRAS* Small Scale Structure Catalog. US Govt. Printing Office, Washington, DC

Hewett P.C., 1982, MNRAS, 201, 867

Heydon-Dumbleton N.H., Collins C.A., MacGillivray H.T., 1989, MNRAS, 238, 379

Hoessel J.G., Gunn J.E., Thaun T.X., 1980, ApJ, 241, 486

Holtzman J.A., 1989, ApJS, 71, 1

Holtzman J.A., Primack J.R., 1993, ApJ, 405, 428

Hoyle F., 1948, MNRAS, 108, 372

Huchra J.P. et al., 1992, in preparation

Huchra J., Brodie J., 1987, AJ, 93, 779

Huchra, J., Henry J.P., Postman M., Geller M.J., 1990, ApJ, 365, 66

Hudson M.J., 1993, MNRAS, 265, 43

Hughes J.P., 1989, ApJ, 337, 21

Humason M.L., Mayall N.O., Sandage A.R., 1956, AJ, 61, 97

IRAS Point Source Catalog, 1985, US Govt. Printing Office, Washington, DC

Islam J.N., 1992, An Introduction to Mathematical Cosmology. Cambridge Univ Press, Cambridge

Jacoby G.H. et al., 1992, PASP, 104, 599

Jameson A., 1989, Science, 245, 361

Jeans J.H., 1928, Astronomy and Cosmology. Cambridge Univ Press, Cambridge

Jensen L.G., Szalay A.S., 1986, ApJ, 305, L5

Jing Y.-P., Plionis M., Valdarnini R., 1992, ApJ, 389, 499

Kaiser N., 1984, ApJ, 284, L9

Kaiser N., 1986, MNRAS, 219, 785

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., Davis M., 1985, ApJ, 297, 365

Kaiser N., Malaney R.A., Stockman G.D., 1993, preprint

Kalnajs A., 1983, in E. Athanassoula, ed, Internal Kinematics and Dynamics of Galaxies, IAU Symposium No. 100. Reidel, Dordrecht

Kashlinsky A., 1986, ApJ, 306, 374

Katz N., Hernquist L., Weinberg D.H., 1992, ApJ, 399, L109

Kauffmann G., White S.D.M., 1992, MNRAS, 258, 511

Kent S.M., Gunn J.E., 1982, AJ, 87, 945

Klypin A.A., Kopylov A.I. 1983, Sov. Astron. Lett., 9, 41

Knapp G.R., Helou G., Stark A.A., 1987, ApJ, 94, 54

Kolb E.W., Turner M.S., 1990, The Early Universe. Addison-Wesley, Redwood City, CA

Kuhn J.R., Miller R.H., 1989, ApJ, 341, L41

Kuijken K., Gilmore G., 1989a, MNRAS, 239, 571

Kuijken K., Gilmore G., 1989b, MNRAS, 239, 605

Kuijken K., Gilmore G., 1989c, MNRAS, 239, 651

Kulessa A.S., Lynden-Bell D., 1992, MNRAS, 255, 105

Lachièze-Rey M., 1993a, ApJ, 407, 1

Lachièze-Rey M., 1993b, ApJ, 408, 403

Lahav O., Edge A.C.,-Fabian A.C., Putney A., MNRAS, 238, 881

Larson R.B., Tinsley B.H., 1978, ApJ, 219, 46

Lawrence A., Rowan-Robinson M., Crawford J., Parry I., Xia X.-Y., Ellis R.S., Frenk C.S., Saunders W., Efstathiou G., Kaiser N., 1993, in preparation

Lawrence A., Rowan-Robinson M., Leech K.J., Jones D.H.P., Wall J.V., 1989, MNRAS, 240, 329

Leech K.J., Rowan-Robinson M., Lawrence A., Hughes J.D., 1993, MNRAS, in press

Leir A.A., van den Bergh S., 1977, ApJS, 34, 381

Lilje P.B., Efstathiou G., 1988, MNRAS, 231, 635

Limber D.N., 1953, ApJ, 117, 134

Lin C.C., Mestel L., Shu F., 1965, ApJ, 142, 1431

Linde A.D., 1982, Phys. Rev. Lett., 108B, 389

Ling E.N., Frenk C.S., Barrow J.D., 1986, MNRAS, 223, 21P

Little B., Tremaine S., 1987, ApJ, 320, 493

Loveday J., Efstathiou G., Peterson B.A., Maddox S.J., 1992, ApJ, 400, L43

Low F.J. et al., 1984, ApJ, 278, L19

Lucey J.R., 1983, MNRAS, 204, 33

Lucy L.B., 1974, AJ, 79, 745

Lumsden S.L., Heavens A.F., Peacock J.A., 1989, MNRAS, 238, 293

Lumsden S.L., Nichol R.C., Collins C.A., Guzzo L., 1992, MNRAS, 258, 1

Lyubimov V.A., Novikov E.G., Nozik V.Z., Tretyakov E.F., Kozek V.S., 1980, Phys. Lett., 394, 266

M^{c}Kellar A., 1941, Publ. Dominion Astrophys. Obs., 7, 251

Maddox S.J., Efstathiou G., Sutherland W., 1990a, MNRAS, 246, 433

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

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

Madsen J., 1992, Phys. Rev. Lett., 69, 571

Mann R.G., Heavens A.F., Peacock J.A., 1993, MNRAS, 263, 798

Marx G., Szalay A.S., 1972, in Neutrino '72. Technoinform, Budapest

Mather J.C., 1993, *COBE* preprint 93-10

Mather J.C. et al., 1990, ApJ, 354, L37

Meinhold P. et al., 1993, ApJ, 409, L1

Merritt D., 1987, ApJ, 313, 121

Mészáros P., 1974, A&A, 37, 225

Mo H.J., Jing Y.P., Borner G., 1992, ApJ, 392, 452

Mo H.J., JIng Y.P., Borner G., 1993, MNRAS, 264, 825

Mo H.J., Peacock J.A., Xia X.-Y., 1993, MNRAS, 260, 121

More J.G., Heavens A.F., Peacock J.A., 1986, MNRAS, 220, 189

Moshir M. et al., 1989, Explanatory Supplement to the *IRAS* Faint Source Survey. Jet Propulsion Lab, Pasadena

Mould J.R., Oke J.B., de Zeeuw P.T., Nemec J.M., 1990, AJ, 99, 1823

Neyman J., Scott E.L., 1952, ApJ, 116, 144

Neyman J., Scott E.L., Shane C.D., 1953, ApJ, 117, 92

Neyman J., Scott E.L., Shane C.D., 1954, ApJS, 1, 269

Nichol R.C., 1992, Ph.D. thesis, Univ of Edinburgh

Nichol R.C., Collins C.A., Guzzo L., Lumsden S.L., 1992, MNRAS, 255, 21P

Nilson P., 1973, *Acta Upsaliensis, Nova Regiae Societatis, Scientiarium Upsaliensis*, Ser V:A; Vol. 1

Nilsson J.S., Gustafsson B., Skagerstam B.-S., 1991, Physica Scripta, T36

Nusser A., Dekel A., 1993, ApJ, 405, 437

Ohm E.A., Bell Syst. Tech. J., 40, 1065

Olive K.A., 1990, Phys. Reports, 190, 307

Olivier S., Blumenthal G.R., Dekel A., Primack J.R., Stanhill D., 1990, ApJ, 356, 1

Olivier S., Primack J.R., Blumenthal G.R., Dekel A., ApJ, 1993, 408, 17

Oliver S.J., 1993, Ph.D. thesis, Univ of London

Oort J.H., 1932, Bull. Astron. Inst. Netherlands, 6, 249

Oort J.H., 1960, Bull. Astron. Inst. Netherlands, 15, 45

Ostriker J.P., 1993, ARA&A, 31, 689

Ostriker J.P., Peebles P.J.E., 1973, ApJ, 186, 467

Otto S., Politzer H.D., Wise M.B., 1986, Phys. Rev. Lett., 56, 1878

Park C., 1991, MNRAS, 251, 167

Peacock J.A., 1991, MNRAS, 253, 1P

Peacock J.A., Heavens A.F., 1985, MNRAS, 217, 805

Peacock J.A., Nicholson D., 1991, MNRAS, 253, 307

Peacock J.A., West M.J., 1992, MNRAS, 259, 494

Peccei R.D., Quinn H.R., 1977, Phys. Rev. Lett., 38, 1440

Peebles P.J.E., 1980, The Large-Scale Structure of the Universe. Princeton Univ Press, Princeton, NJ

Peebles P.J.E., 1993, Principles of Physical Cosmology. Princeton Univ Press, Princeton, NJ

Peebles P.J.E., Schramm D.N., Turber E.L., Kron R.G., 1991, Nat, 352, 769

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

Persic M., Salucci P., 1988, MNRAS, 234, 131

Petrosian V., Salpeter E.E., Szekeres P., 1967, ApJ, 147, 1222

Plionis M., Barrow J.D., Frenk C.S., 1991, MNRAS, 249, 662

Plionis M., Valdarnini R., 1991, MNRAS, 249, 46

Plionis M., Valdarnini R., Coles P., 1992, MNRAS, 258, 114

Plionis M., Valdarnini R., Jing Y.-P., 1992, ApJ, 398, 12

Plummer H.C., 1911, MNRAS, 71, 460

Politzer H.D., Wise M.B., 1984, ApJ, 285, L1

Postman M., Geller M.J., Huchra J.P., 1986, AJ, 91, 1267

Postman M., Huchra J.P., Geller M.J., 1992, ApJ, 384, 404

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

Prestage R.M., Peacock J.A., 1988, MNRAS, 230, 131

Puche D., Carignan C., 1988, AJ, 95, 1025

Puche D., Carignan C., 1991, ApJ, 378, 487

Rainey G.W., 1977, Ph.D. thesis, Univ of California

Rees M.J., 1987, in S.W. Hawking. W. Israel, eds, 300 Years of Gravitation. Cambridge Univ Press, Cambridge

Rees M.J., Ostriker J.P., 1977, MNRAS, 179, 541

Rees M.J., Sciama D.W., 1968, Nat, 217, 511

Rice W.L., Lonsdale C.J., Soifer B.T., Neugebauer G., Kopan E.L., Lloyd L.A., de Jong T., Habing H., 1988, ApJS, 68, 91

Romer A.K., Collins C.A., Ebeling H., MacGillivray H., Cruddace R.G., Böhringer H., 1993, preprint

Roll P.G., Wilkinson D.T., 1966, Phys. Rev. Lett., 16, 405

Rood H.J., 1976, ApJ, 207, 16

Rowan-Robinson M., 1985, The Cosmological Distance Ladder. Freeman, New York, NY

Rowan-Robinson M., Crawford J., 1989, MNRAS, 238, 523

Rowan-Robinson M., Helou G., Walker D., 1987, MNRAS, 227, 589

Rowan-Robinson M., Saunders W., Lawrence A., Leech K., 1991, MNRAS, 253, 485

Rubin V.C., Ford W.K., Thonnard N., ApJ, 238, 471

Sachs R.K., Wolfe A.M., 1967, ApJ, 147, 73

Salucci P., Ashman K.M., Persic M., 1991, ApJ, 379, 89

Sancisi R., van Albada T.S., 1990, in J. Kormendy, G.R. Knapp, eds, Dark Matter in the Universe, IAU Symposium No. 117. Reidel, Dordrecht

Sato K., Teresawa T., 1991, Physica Scripta, T36, 60

Saunders W., Rowan-Robinson M., Lawrence A., Kaiser N., Ellis R.S., Frenk C.S., 1990, MNRAS, 242, 318

Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS, 258, 134

Scalo J.M., 1986, Fund. Cosmic Phys., 11, 1

Scaramella R., Zamorani G., Vettolani G., Chincarini G., 1991, AJ, 101, 342

Scott E.L., 1957, AJ, 62, 248

Scott E.L., Shane C.D., Swanson M.D., 1954, ApJ, 119, 91

Seldner M., Peebles P.J.E., 1977, ApJ, 215, 703

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

Shandarin S.F., Zeldovich Ya.B., 1989, Rev. Mod. Phys., 61, 185

Shane C.D., Wirtanen C.A., 1967, Pub. Lick Obs., 22

Shectman S.A., 1985, ApJS, 57, 77

Silk J., 1968, ApJ, 151, 459

Silk J., Vittorio N., 1987, ApJ, 317, 564

Smith M.S., Kawano L.H., Malaney R.A., 1993, ApJS, 85, 219

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

Soifer B.T., Houck J.R., Neugebauer G., 1987, ARA&A, 25, 187

Soltan A., 1988, MNRAS, 231, 309

Stewart G.C., Canizares C.R., Fabian A.C., Nulsen P.E., 1984, ApJ, 278, 536

Strauss M.A., Davis M., Yahil A., Huchra J.P., 1990, ApJ, 361, 49

Strauss M.A., Davis M., Yahil A., Huchra J.P., 1992, ApJ, 385, 421

Struble M.F., Rood H.J., 1987, ApJS, 63, 555

Struble M.F., Rood H.J., 1991, ApJ, 374, 395

Suginohara T., Suto Y., 1992, ApJ, 396, 395

Sulentic J.W., Keel W.C., Telesco C.M., 1990, Paired and Interacting Galaxies. NASA Conf. Publ. 3098

Sunyaev R.A., Zeldovich Ya.B., 1972, Comments Ap. Space Sci., 4, 173

Surace J.A., Mazzarella J.M., Soifer B.T., Werhle A.E., 1993, AJ, 105, 864

Sutherland W., 1988, MNRAS, 234, 159

Sutherland W., Efstathiou G., 1991, MNRAS, 248, 159

Suto Y., Cen R.Y., Ostriker J.P., 1992, ApJ, 395, 1

Suto Y., Suginohara T., 1991, ApJ, 370, L15

Szalay A.S., 1988, ApJ, 333, 21

Szalay A.S., Hollósi J., Tóth G., 1989, ApJ, 339, L5

Taylor A.N., Rowan-Robinson M., 1992, Nat, 359, 396

The L.S., White S.D.M., 1986, AJ, 92, 1248

The L.S., White S.D.M., 1988, AJ, 95, 15

Tisdall E.K.M., 1993, Ph.D. thesis, Univ of Edinburgh

Toomre A., 1981, in S.M. Fall, D. Lynden-Bell, eds, The Structure and Evolution of Normal Galaxies. Cambridge Univ Press, Cambridge

Tremaine S., Gunn J.E., 1979, Phys. Rev. Lett., 42, 407

Trimble V., 1987, ARA&A, 25, 425

van den Bergh S., 1992, PASP, 104, 680

van Dalen A., Schaefer R.K., 1992, ApJ, 398, 33

Vogeley M.S., Park C., Geller M.J., Huchra J.P., 1992, ApJ, 391, L5

Xu C., Sulentic J.W., 1991, ApJ, 374, 407

Wagoner R.V., Fowler W.A., Hoyle F., 1967, ApJ, 148, 3

Weinberg S., 1972, Gravitation and Cosmology. Wiley, New York

West M.J., van den Bergh S., 1991, ApJ, 373, 1

White S.D.M., Davis M., Efstathiou G., Frenk C.S., 1987b, Nat, 330, 451

White S.D.M., Efstathiou G., Frenk C.S., 1992, MNRAS, 262, 1023

White S.D.M., Frenk C.S., Davis M., Efstathiou G., 1987a, ApJ, 313, 505

White S.D.M., Huchra J., Latham D., Davis M., 1983, MNRAS, 203, 701

White S.D.M., Navarro J.F., Evrard A.E., Frenk C.S., 1993, preprint

Williams B.G., Heavens A.F., Peacock J.A., Shandarin S.F., 1991, MNRAS, 250, 458

Witten E., 1984, Phys. Rev., D30, 272

Zaritsky D., Olszewski E.W., Schommer R.A., Peterson R., Aaronson M., 1989, ApJ, 345, 759

Zeldovich Ya.B., 1965, Adv. Astron. Ap., 3, 241

Zeldovich Ya.B., 1970, A&A, 5,84

Zeldovich Ya.B., 1972, MNRAS, 160, 1P

Zwicky F., 1933, Helv. Phys. Acta, 6, 110

Zwicky F., Herzog E., Wild P., Karpowicz, Kowal C.T., 1961-68, Catalog of Galaxies and Clusters of Galaxies. California Institute of Technology, Pasadena, CA

PUBLISHED PAPER

Mann R.G., Heavens A.F., Peacock J.A., 1993, MNRAS, 263,798.

The richness dependence of cluster correlations

Mon. Not. R. Astron. Soc. **263**, 798–816 (1993)

The richness dependence of cluster correlations

R.G. Mann,[1] A.F. Heavens[1] and J.A. Peacock[2]
[1]*Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[2]*Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 1993 March 18. Received 1993 March 10; in original form 1992 November 9

ABSTRACT

We report on a study of the richness dependence of the spatial correlations of clusters of galaxies. We employ the method devised by Bond & Couchman, which combines the theory of the statistics of peaks in Gaussian random fields with the evolution of the cosmological density field by the Zeldovich Approximation, to calculate analytically both the statistical and dynamical contributions to the clustering.

We compute the cluster correlation function for a variety of popular cosmological models and compare our results with data from four recent cluster samples. We find no model able to account for all of the observations, although, conversely, the observational data are of insufficient quality to rule out firmly any of the models. The model that fares best is one which has been advocated as accounting for the angular correlation function of APM galaxies: that the same power spectrum gives the best fit to both the galaxy and cluster correlation data may be taken as support for the standard picture, due to Kaiser, in which objects form at the sites of peaks in an initial cosmological density field which obeyed Gaussian statistics.

No model is able to reproduce the correlation length of a sample of Abell $R \geq 2$ clusters. This result may just indicate the inadequacy of the correlation length when taken alone as a diagnostic statistic, or that this cluster sample is seriously corrupted by projection effects. We consider, however, alternative explanations, including the possibility that non-Gaussian initial conditions are required and that the identification of peaks in the linear density field as sites of nascent clusters may break down for the highest peaks, corresponding to the richest clusters.

Key words: galaxies: clustering – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

The large-scale structure of the Universe is delineated by the distribution of galaxies and of clusters of galaxies. These distributions may be quantitatively described using the hierarchy of n-point spatial correlation functions (e.g. Peebles 1980). The two-point function, $\xi(r)$, has become the dominant statistic for describing large-scale structure and for challenging cosmological theories through comparison with observation, despite the fact that, in many ways, the use of its Fourier transform, the power spectrum, is preferable (e.g. Peacock 1991; Peacock & West 1992, hereafter PW92). It is observed that the two-point functions for many classes of objects, from galaxies (Davis & Peebles 1983) to rich clusters (Bahcall & Soneira 1983, BS83), are well approximated by the same power-law form, $\xi(r) = (r_0/r)^\gamma$ with $\gamma \simeq 1.8$, with different values of the correlation length, r_0, over a wide range of separations, r, where the correlation functions are relatively well determined. The correlation length of rich clusters has, however, been a matter of great controversy throughout the past decade. The most widely used cluster catalogue has been that of Abell (1958), with a southern extension by Abell, Corwin & Olowin (1989, ACO). From a redshift survey of 104 Abell clusters with richness $R \geq 1$, BS83 obtained a correlation length of $r_0 = 25\ h^{-1}$ Mpc (where, as usual, h denotes the Hubble parameter in units of 100 km s^{-1} Mpc^{-1}). Many workers have deemed this number to be too large to be readily explained by many cosmological theories, in particular by the standard Cold Dark Matter (CDM) model of White et al. (1987). The importance of any such conclusion has prompted detailed study of the BS83 result.

The clusters in the Abell/ACO catalogue were selected by scanning photographic plates by eye, so their selection was dependent on many unquantifiable physiological and psychological factors. A number of authors (e.g. Lucey 1983; Sutherland 1988; Dekel et al. 1989; Sutherland & Efstathiou 1991; Efstathiou et al. 1992b) have questioned the reliability of this method of

cluster selection and, hence, of the Abell/ACO catalogue as the basis for statistical analyses of the spatial distribution of clusters. They detect gross projection effects, which Sutherland (1988) defined as 'angular correlations that are not due to genuine clustering in redshift space' and which manifest themselves in a redshift-space clustering pattern with a higher degree of anisotropy than can be accounted for by reasonable cluster peculiar velocities. Sutherland (1988) presented a method for correction for projection effects, and its application to a sample of Abell clusters from the redshift survey of Struble & Rood (1987) resulted in a much reduced correlation length of $r_0 = 14\ h^{-1}$ Mpc. Sutherland & Efstathiou (1991) and Efstathiou et al. (1992b) have applied the same procedure to further samples of Abell clusters, with similar results.

It is clearly preferable to start with a reliable sample, rather than to have to correct an initially unreliable one, especially if there may be some doubt as to the validity of the correction procedures. The advent of plate-measuring machines has facilitated the creation of cluster catalogues whose selection criteria should be much more objective than those of the Abell/ACO catalogue, and which were chosen to circumvent some of the problems associated with it. Recently, Nichol et al. (1992) have reported a correlation length of $r_0 = 16.4 \pm 4.0\ h^{-1}$ Mpc for the clusters of the Edinburgh/Milan Redshift Survey (EMRS), which are drawn from the Edinburgh/Durham Cluster Catalogue (EDCC) of Lumsden et al. (1992), while a sample of APM clusters (Dalton et al. 1992) yields $r_0 = 12.9 \pm 1.4\ h^{-1}$ Mpc. Bahcall & West (1992) have claimed that these latest results are consistent with the larger correlation length of BS83, given that the number densities of the EMRS and APM samples are higher than that of Abell $R \geq 1$ clusters and provided that there is a strong dependence of correlation length on number density, as previously claimed (Bahcall & Burgett 1986; Bahcall 1988) to be followed by objects ranging from individual galaxies to superclusters.

In this paper we investigate whether the required strength of this richness dependence can be reproduced theoretically. Our method is an analytic one, that of Bond & Couchman (1987, 1988; hereafter collectively BC), which uses the Zeldovich Approximation to perform the evolution of the cosmological density field. We show how its use facilitates the computation of predictions of the cluster correlation function for cosmological models in a much shorter time than is possible through N-body simulation. The advantages of analytic methods over numerical simulation extend, however, beyond the simple considerations of enhanced speed and the much more efficient coverage of parameter space that they facilitate. Studies of cluster correlations through N-body simulation face difficulties resulting from the dynamic range inherent in the problem: the numerical simulations should ideally have sufficient spatial resolution to be capable of producing plausibly galaxy-like objects, while the low number density of clusters means that a very large simulation volume is required to generate a sample of clusters large enough to facilitate the computation of the correlation function with reasonable statistical accuracy. These conflicting requirements place tight constraints on those running N-body simulations, given the capabilities of present-day computers. For example, Bahcall & Cen (1992) are able to simulate a cube of side $400\ h^{-1}$ Mpc, which gives them a sample of several hundred clusters, but only through the use of a PM code, the limited spatial resolution of which ($\sim 0.8\ h^{-1}$ Mpc) makes it far from ideal for addressing the important related problem of the cluster mass function, which they define in terms of the mass contained within a sphere of radius $1.5\ h^{-1}$ Mpc. Other numerical modellers (e.g. Dalton et al. 1992) have been forced to make assumptions about the relation between the cluster distribution and the Mpc-scale mass distribution. Given this necessity, it is clearly preferable to implement such assumptions analytically, which is the approach we adopt here.

We compute the cluster correlation functions for a number of models and for the number densities appropriate to observed cluster samples, comparing our results with those observed. We also calculate the relationship between the correlation length, r_0, and the characteristic separation, d, of a sample of clusters, which is related to its number density, N_c, by $d \equiv N_c^{-1/3}$. In Section 2 we review some of the features of the theory of the statistics of peaks in Gaussian random fields that relate to the study of large-scale structure. In particular, we look at the use of the peak–peak correlation function as an approximation to the cluster correlation function, and we show how a simple collapse model may be used to relate a population of clusters to a set of peaks in a particular Gaussian random field. The Zeldovich Approximation is the subject of Section 3, and the models that we have chosen to consider are described in Section 4. In Section 5 we compare our results, both for the cluster correlation function and for the $d - r_0$ relation, with observation. A discussion of our work and the conclusions that we draw from it are presented in Section 6.

2 GAUSSIAN RANDOM FIELDS AND LARGE-SCALE STRUCTURE

Most theoretical studies of the formation of large-scale structure through gravitational instability assume that the primordial density perturbation field obeyed Gaussian statistics. There are both theoretical and observational reasons for supposing that this might be the case. Observational support comes from work by Gott and collaborators (Gott, Melott & Dickinson 1986; Hamilton, Gott & Weinberg 1986; Weinberg, Gott & Melott 1987; Melott, Weinberg & Gott 1988; Gott et al. 1989) and others (e.g. Coles & Plionis 1991; Moore et al. 1992) on the topology of large-scale structure, as well as from the analysis of the QDOT redshift survey of *IRAS* galaxies presented by Saunders et al. (1991) (see also Juszkiewicz & Bouchet 1992) and, most recently, from the results of the *COBE* DMR experiments presented by Smoot et al. (1992). On the theoretical side, the assumption of Gaussian statistics is particularly well motivated in inflationary models, which predict Gaussian statistics as a generic feature, resulting from the random nature of the quantum processes that generate the primordial density perturbations (Bardeen, Steinhardt & Turner 1983, and references therein).

2.1 Peaks in Gaussian random fields

The simplifying assumption of Gaussian statistics facilitates an analytic study of the properties of the cosmological density field, which is complementary and, in some ways, preferable to that pursued through N-body simulations. Much work has been done to elaborate a scenario where non-linear structures form preferentially in thresholded regions (Kaiser 1984; Jensen & Szalay 1986) or at peaks [Peacock & Heavens 1985; Bardeen et al. 1986 (BBKS); Coles 1986; Couchman 1987; Lumsden, Heavens & Peacock 1989 (LHP)] in a Gaussian primordial density field, following the insight of Kaiser (1984) that such a model could explain the enhancement of the correlations of Abell clusters to a strength well above that of galaxies. BBKS present the most comprehensive treatment of the statistics of peaks in Gaussian random fields, and we shall generally use their terminology in what follows.

The basic assumption of BBKS's theory is that the cosmological density perturbation field obeyed Gaussian statistics at early times and that, after smoothing, those of its peaks that lie above some threshold may be identified as sites of nascent objects. It follows that a particular class of objects is defined by the pair R_s and ν_s, which are, respectively, the smoothing scale of the (conventionally Gaussian) filter which, when applied to the density field, identifies the correct population of peaks (mass of object $\propto R_s^3$) and the threshold in this smoothed density field above which peaks are to be identified with nascent objects. Williams et al. (1991) have studied the relative merits of various forms of filter function in a one-dimensional implementation of peaks theory and they have found that, indeed, the use of a Gaussian filter gives the best agreement with results of simulations based on the adhesion model, which is exact in one dimension: we consider only Gaussian filters in this paper.

BBKS derive a formal expression for the n-point correlation function of peaks in a three-dimensional Gaussian random field that satisfy the threshold condition. The joint probability function of n peaks is a function of $10n$ variables and the n-point correlation function may be reduced, using suitable approximations, only as far as a $2n$-dimensional integral with a very complicated integrand, so that calculation of even the two-point function is problematic. Approximations to the peak–peak correlation function are considered by BBKS, as well as by LHP (and references therein).

If these methods are to be used to model the correlation function of galaxies or clusters as observed today, then one must assume that the clustering in the present-day distribution is due to that in the distribution of the sites of nascent objects, resulting from their being rare events in a Gaussian process, and that there has been negligible movement of the objects from the sites of their formation as a result of their mutual gravitational attraction, since this will change the correlation function from its initial peak–peak form. BBKS provide a simplified approximate prescription for modification of the peak–peak correlation function to account for evolution while the density field remains linear, and this has been used recently by Holtzman & Primack (1993) to study several CDM variants. The deficiencies of linear theory, in comparison with a fuller dynamical treatment using the Zeldovich Approximation to perform the evolution of the density field into the non-linear regime, are discussed by Coles, Melott & Shandarin (1993). Such a treatment is required to facilitate the accurate computation of the cluster–cluster correlation function and BC have shown how this may be done, as will be described in Section 3.

The static peak–peak correlation function should remain a good approximation to the true correlation function at very large separations, however, where the effects of dynamical evolution are negligible. In the limit of arbitrarily large separations, the peak–peak correlation function is simply a multiple of the autocorrelation function of the linear density field, and is shown by BBKS and LHP to take the form

$$\xi_{\mathrm{pk-pk}}(r) \;\rightarrow\; \langle \tilde{\nu}_s \rangle^2 \, \psi(r) \;\text{ as }\; r \rightarrow \infty, \tag{1}$$

where $\psi(r) \equiv \xi(r)/\xi(0)$ is the scaled autocorrelation function of the linear density field, smoothed on scale R_s, and where $\langle \tilde{\nu}_s \rangle$ is the value of an effective threshold quantity, $\tilde{\nu}_s$, averaged over all peaks above the threshold ν_s, and is given by

$$\langle \tilde{\nu}_s \rangle = \left[\frac{2}{\pi(1-\gamma^2)} \right]^{1/2} \frac{\int_0^\infty f(x) \mathrm{e}^{-x^2/2} \exp\left[-(\nu_s - \gamma x)^2 / 2(1-\gamma^2) \right] \mathrm{d}x}{\int_0^\infty f(x) \mathrm{e}^{-x^2/2} \mathrm{erfc}\left[(\nu_s - \gamma x)/\sqrt{2(1-\gamma^2)} \right] \mathrm{d}x}. \tag{2}$$

A derivation of this closed form for $\langle \tilde{\nu}_s \rangle$ is given in Appendix B, where the function $f(x)$ is defined. The spectral parameter, γ, indicates the range of wavenumbers over which there is power in the power spectrum: a low value of γ indicates that there is power on a wide range of scales, while, as $\gamma \rightarrow 1$, the power spectrum tends towards a δ-function. This spectral parameter is defined in terms of moments, σ_j, of the power spectrum of the density field, by $\gamma = \sigma_1^2/\sigma_2 \sigma_0$. We define the moments by

$$\sigma_j^2 \;=\; \int \Delta^2(k) \, k^{2j} \, \frac{\mathrm{d}k}{k}, \tag{3}$$

and employ the dimensionless form of the power spectrum given by $\Delta^2(k)$ [where $\Delta^2(k) \propto k^3 P(k)$], which is equal to the contribution to the variance of the density perturbation field per unit interval of $\ln k$, and which is related to the autocorrelation function of the density field by

$$\Delta^2(k) \;=\; \frac{2k^3}{\pi} \int_0^\infty \xi(r) \, \frac{\sin kr}{kr} \, r^2 \mathrm{d}r. \tag{4}$$

2.2 Defining a population of objects

In Section 3 we shall use the asymptotic form of the correlation function given in equation (1) to normalize the prescription used by BC to relate the cluster distribution to the underlying density field. To do this we require a procedure for calculation of the values of R_s and ν_s appropriate to the set of peaks which we identify with a population of clusters with number density N_c : we assume that the number density of clusters is conserved even though the individual cluster positions change over time. There is no unique way of making this identification. We employ a prescription based on the following simple collapse model.

Collapse models commonly identify the formation of an object with the time when the rms density contrast of the extrapolated local *linear* density field, smoothed on the appropriate mass scale, exceeds some critical value δ_c, which, in an Einstein–de Sitter universe, equals 1.7 for a spherical perturbation. Lower values ($1.7 \gtrsim \delta_c \gtrsim 1$) will be appropriate for a triaxial proto-object (More, Heavens & Peacock 1986), which is the generic case for peaks in a Gaussian random field (Peacock & Heavens 1985; BBKS). It is found (e.g. White et al. 1992) that the value of δ_c is very insensitive to Ω_0 for a flat universe, so we shall consider the range $1.7 \geq \delta_c \geq 1$ for all models in what follows. An independent constraint on the value of δ_c to be used is that this procedure should produce sensible cluster mass estimates: this point is discussed further in Section 5, with reference to the comparison of our results with those from numerical simulations.

It is clear that the first objects of a given mass to form in this picture will be those corresponding to the highest peaks in the initial density field and that, as time progresses, the objects of that mass which collapse will correspond to increasingly low peaks in the initial density field. A lower bound to the threshold ν_s may be obtained, therefore, by equating it with the peak height corresponding to peaks collapsing today. This would clearly be an underestimate of the threshold for objects like galaxies which formed some time ago, but for clusters, which are the most massive objects to have collapsed by the present and which are observed to be dynamically young, it is reasonable to equate the threshold with the height of peaks collapsing today, and that is what we shall do here. This provides the first constraint on the values of R_s and ν_s, which may be written as

$$\nu_s \sigma_0(R_s) = \delta_c. \tag{5}$$

A second constraint comes from matching the number density of clusters, N_c, to $n_{pk}(\nu_s)$, which is the number density of peaks in the initial density field, smoothed on scale R_s, which lie above the threshold ν_s, and which is given by

$$n_{pk}(\nu_s) = \frac{1}{8\pi^2 R_*^3} \int_0^\infty \mathrm{d}x\, f(x) \mathrm{e}^{-x^2/2} \mathrm{erfc}\left[\frac{(\nu_s - \gamma x)}{\sqrt{2(1-\gamma^2)}}\right], \tag{6}$$

where $R_* = \sqrt{3}\,\sigma_1/\sigma_2$: a derivation of this formula is presented in Appendix A.

From equations (5) and (6) we obtain a unique solution for R_s and ν_s for a particular cluster number density and a given power spectrum: this prescription is very similar to that used by BBKS and Bardeen, Bond & Efstathiou (1987). We discuss the application of this procedure to power spectra of practical interest in Section 4.

3 CLUSTERING IN THE ZELDOVICH APPROXIMATION

The statistical peak–peak correlation function of BBKS and LHP is a poor approximation to the true cluster correlation function on scales where the dynamical evolution of the cluster population since formation is not negligible. A method must be found to combine the dynamical and statistical components to the clustering, and BC have shown how this may be achieved, within an analytic framework, through the use of the Zeldovich Approximation (Zeldovich 1970; Shandarin & Zeldovich 1989) to perform the dynamical evolution of the density field into the mildly non-linear regime that is of relevance to the cluster distribution. We present a full description of this method here, since that given by BC is somewhat terse and contains a number of typographical errors.

3.1 The Zeldovich Approximation

Consider a three-dimensional space populated by a dissipationless fluid. In proper coordinates, the (Eulerian) position, $\boldsymbol{y}(\boldsymbol{r},t)$, of a particle at a time t is related to its initial (Lagrangian) position, $\boldsymbol{r}$, by

$$\boldsymbol{y}(\boldsymbol{r},t) = a(t)\boldsymbol{r} - \boldsymbol{D}(\boldsymbol{r},t), \tag{7}$$

where $a(t)$ is the cosmological scale factor, $\boldsymbol{D}(\boldsymbol{r},t)$ is the (negative of the) displacement field and the minus sign is a matter of convention. The Zeldovich Approximation consists of the assumption that the displacement field may be expressed as the product of spatial and temporal parts, of the form

$$\boldsymbol{D}(\boldsymbol{r},t) = b(t)\boldsymbol{d}(\boldsymbol{r}), \tag{8}$$

so that, in the Zeldovich Approximation, equation (7) becomes

$$\boldsymbol{x}(\boldsymbol{r},t) = \boldsymbol{r} - \frac{b(t)}{a(t)}\boldsymbol{d}(\boldsymbol{r}), \tag{9}$$

where we define comoving coordinates, $\boldsymbol{x}(\boldsymbol{r},t) = \boldsymbol{y}(\boldsymbol{r},t)/a(t)$. The function $b(t)$ depends on the value of the cosmological density parameter, Ω, and determines the evolution of density perturbations: an expression for the growing-mode solution for $b(t)$ is given in the appendix to Zeldovich (1970). Our concern here is with the distribution of matter at a specific epoch, the present, rather than with following the evolution of the density field with time, so it is convenient to define the variable

$$\boldsymbol{s}(\boldsymbol{r}) \equiv \frac{b(t_0)}{a(t_0)}\boldsymbol{d}(\boldsymbol{r}), \tag{10}$$

where t_0 denotes the present epoch, from which it follows that

$$\boldsymbol{x} = \boldsymbol{r} - \boldsymbol{s}. \tag{11}$$

Given knowledge of the density perturbation field, equation (11) is all that is required to determine the distribution of particles, since the displacement field, $\boldsymbol{s}$, is related to the density perturbation field by

$$\delta \equiv \frac{\rho - \bar{\rho}}{\bar{\rho}} = \nabla_{\boldsymbol{r}} \cdot \boldsymbol{s}. \tag{12}$$

This result, which holds for small perturbations at early times, may be derived through a consideration of mass conservation. At later times it may be taken as defining the extrapolation of the linear density field obtained using the Zeldovich Approximation, which is what we shall henceforth denote by the symbol δ, and it may be expressed in terms of the Fourier components, $\{\boldsymbol{s}_k\}, \{\delta_k\}$, of the two fields as

$$(\boldsymbol{s}_k)_j = -\mathrm{i}\frac{\delta_k}{k}(\hat{\boldsymbol{k}})_j, \tag{13}$$

where $\hat{\boldsymbol{k}} = \boldsymbol{k} / |\boldsymbol{k}|$ and $(\boldsymbol{k})_j$ is the jth component of the vector $\boldsymbol{k}$, etc.

The Zeldovich Approximation is exact in one dimension until caustics form – i.e. $\rho \to \infty$ locally. After this time it gives an increasingly poor description of the density field, as it leads to shell crossing, with particles passing through each other rather than coalescing, resulting in an underestimation of the density of collapsed objects. This may be remedied by introducing a viscosity term into the dynamical equations, to produce 'sticky' particles: this is the adhesion model, for a discussion of which the reader is referred to Williams et al. (1991) and references therein (particularly Gurbatov, Saichev & Shandarin 1989). The advantage afforded by the introduction of 'sticky' particles is outweighed, in the present context, by the impossibility of incorporating the adhesion model into a straightforward analytic framework like that outlined here for the Zeldovich Approximation.

An analytically tractable solution to the shell-crossing problem in the Zeldovich Approximation is to apply a filter to the initial density field, to remove the small-scale power which causes shell crossing. It has been shown (Efstathiou & Silk 1983; Efstathiou 1990; Coles et al. 1993) that, with the inclusion of such a filter mechanism, the Zeldovich Approximation can provide a remarkably accurate picture of the evolution of the density field well into the non-linear regime, as judged by comparison with the results of N-body simulations using the same initial conditions. Coles et al. (1993) favour the simple truncation of the power spectrum at a particular wavenumber to remove small-scale power, while BC prefer smoothing with a Gaussian filter. We adopt the latter approach here and discuss the choice of smoothing scale for the Gaussian filter in Section 3.2.

In addition to this direct illustration of the efficacy of the Zeldovich Approximation in this application, through comparison with N-body simulations, BC and Coles (1990) have shown that the Zeldovich Approximation reproduces the quadratic non-linear terms in the full dynamical power spectrum in the mildly non-linear regime appropriate to analysis through second-order perturbation theory.

Given that the Zeldovich Approximation is rooted firmly in linear theory, and one-dimensional linear theory at that, one might be surprised at how well it can describe the mildly non-linear evolution of the cosmological density field, but this can be readily explained through a consideration of the properties of the collapsing density perturbations involved. It can be shown that proto-objects in Gaussian density fields are generically triaxial (Peacock & Heavens 1985; BBKS) and that gravitation acts to accentuate that deviation from spherical symmetry (Lin, Mestel & Shu 1965), leading to collapse along one axis first (Zeldovich 1970), with subsequent particle motion being preferentially directed parallel to this axis. While the details of this procedure depend on the power spectrum of the density field, it will be generally true that gravitational collapse in structure formation is well approximated by a one-dimensional process, and hence that the Zeldovich Approximation will give a good description of the evolution of the density field well into the non-linear regime, for a wide range of cosmological models (Coles et al. 1993).

3.2 Mass correlations in the Zeldovich Approximation

The first application of the Zeldovich Approximation to be considered is in the determination of the correlation function of a density field with known linear power spectrum. We require knowledge of the mass correlation function as part of the procedure we adopt for determining the smoothing length of the filter that we must apply to the density field to prevent shell crossing. It

is also instructive to consider this first part of the BC formalism (the evolution of the density field) before the addition of the further complication of the prescription to relate the cluster distribution to the density field. The mass correlation function, $\xi_\rho(x)$, is related to the probability, $P(x\,|1,2)$, of finding masses at points 1,2 a distance x apart by

$$1 + \xi_\rho(x) = P(x\,|\,1,2). \tag{14}$$

This may be written as

$$1 + \xi_\rho(x) = \int \mathrm{d}^3\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{s}\,P(x\,|\boldsymbol{r},\boldsymbol{s})\,P(\boldsymbol{s}\,|\boldsymbol{r})\,P(\boldsymbol{r}), \tag{15}$$

where $P(x|\boldsymbol{r},\boldsymbol{s})$ is the conditional probability that two particles have an Eulerian separation x, given that their Lagrangian separation is $\boldsymbol{r}$ and that their relative displacement through subsequent motion is $\boldsymbol{s}$; $P(\boldsymbol{s}|\boldsymbol{r})$ is the conditional probability that two particles have a relative displacement $\boldsymbol{s}$ given a Lagrangian separation $\boldsymbol{r}$; $P(\boldsymbol{r})$ is the probability of there being two particles with Lagrangian separation $\boldsymbol{r}$; and the integral is taken over all $\boldsymbol{r}$ and $\boldsymbol{s}$.

This method would calculate the correlation function for a distribution of particles that had been perturbed in any way that produces a Gaussian displacement field: the Zeldovich Approximation is required only for relation of the displacement field to the initial density field. From equation (11) we see that the Zeldovich Approximation results in the conditional probability $P(x\,|\boldsymbol{r},\boldsymbol{s})$ taking the simple form $P(x|\boldsymbol{r},\boldsymbol{s}) = \delta_{\mathrm{D}}(x-|\boldsymbol{r}-\boldsymbol{s}|)/(4\pi x^2)$, where δ_{D} denotes the Dirac δ-function.

In Appendix C we sketch the derivation of the mass correlation function in the Zeldovich Approximation from this point through its reduction to a one-dimensional integral, which reads

$$1+\xi_{\rho Z}(x) = \frac{1}{2\sqrt{\pi}\sigma_s^2 x}\int_0^\infty \frac{r^2\mathrm{d}r}{(1-\psi_\perp)^{1/2}(\psi_\perp-\psi_\parallel)^{1/2}}\left\{F(u_+)\exp\left[-\frac{(r-x)^2}{4\sigma_s^2(1-\psi_\parallel)}\right] - F(u_-)\exp\left[-\frac{(r+x)^2}{4\sigma_s^2(1-\psi_\parallel)}\right]\right\}. \tag{16}$$

In equation (16), $F(u)$ is Dawson's integral,

$$F(u) \equiv \int_0^u \mathrm{e}^{y^2-u^2}\mathrm{d}y, \tag{17}$$

and the arguments $u_\pm$ are given by

$$u_\pm \equiv \frac{(\psi_\perp-\psi_\parallel)^{1/2}}{2\sigma_s(1-\psi_\parallel)^{1/2}(1-\psi_\perp)^{1/2}}\left[r\,\frac{(1-\psi_\perp)}{(\psi_\perp-\psi_\parallel)} \pm x\right], \tag{18}$$

which differs slightly from the equivalent equation in BC, which contains typographical errors. The quantity σ_s is the one-dimensional rms particle displacement and is given in terms of the density field by

$$\sigma_s^2 = \frac{1}{3}\int_0^\infty \Delta^2(k)\,\mathrm{e}^{-k^2R_f^2}\,\frac{\mathrm{d}k}{k^3}, \tag{19}$$

and $\psi_\parallel$ and $\psi_\perp$ are defined by

$$\sigma_s^2\psi_\parallel(r) = \int_0^\infty \Delta^2(k)\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2\,\mathrm{e}^{-k^2R_f^2}\,\frac{\mathrm{d}k}{k^3}, \tag{20}$$

and

$$\sigma_s^2\psi_\perp(r) = \int_0^\infty \Delta^2(k)\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left[\frac{1-(\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}})^2}{2}\right]\mathrm{e}^{-k^2R_f^2}\,\frac{\mathrm{d}k}{k^3}. \tag{21}$$

Equations (19)–(21) contain a Gaussian smoothing term ($\mathrm{e}^{-k^2R_f^2}$) to filter small-scale power from the power spectrum $\Delta^2(k)$. The choice of the filter radius, R_f, for use with a particular power spectrum is motivated by several considerations: if R_f is too small, then shell crossing will not be suppressed; if R_f is too large, it will filter out real power at short wavelengths and produce a two-point function that underestimates the strength of correlations on small scales. This behaviour can clearly be seen in Fig. 1, which shows the mass correlation function for a CDM model in the Zeldovich Approximation, with the linear power spectrum smoothed with Gaussian filters with different smoothing lengths, R_f. The largest smoothing length of $R_f = 5\ h^{-1}$ Mpc filters out real small-scale power: as the smoothing scale is reduced, the small-scale correlations increase, until a maximum is produced by a filter radius of $3\ h^{-1}$ Mpc. As R_f is reduced further, it is clear that shell crossing is occurring to a large extent since, with filter radii of $R_f \leq 1\ h^{-1}$ Mpc, the correlations are reduced at small separations, due to the smearing out of structure as shells of particles pass through each other. From this it is clear that for this particular power spectrum a filter length of $R_f = 3\ h^{-1}$ Mpc is appropriate. Using this method, we may determine the correct filter radius to be used for each of the models we consider.

When calculating the cluster correlation function, however, consideration of the motion of the peaks in the density field leads us to modify this procedure and to impose the constraint that R_f cannot be smaller than the filter length, R_s, used to define the peaks. This follows from seeing that if the peaks are defined by smoothing the density field on some scale R_s then the motion of those peaks can only be due to modes with wavelengths longer than R_s and so, whenever the procedure outlined in the preceding paragraph gives $R_f < R_s$, we set R_f equal to R_s.

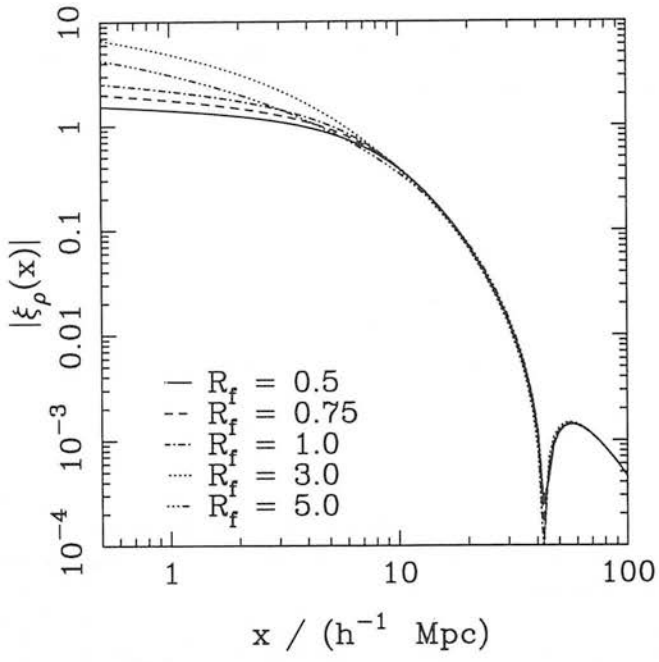


Figure 1. The mass correlation function for an $\Omega_0 = 1$, Γ=0.5 model with *COBE* amplitude, calculated for a number of values of the filter length, R_f, as indicated. (R_f in units of h^{-1} Mpc).

3.3 Cluster correlations in the Zeldovich Approximation

In order to calculate the cluster correlation function, it is necessary to supplement our mechanism for evolution of the density field with a prescription for relation of the cluster distribution to the underlying density field. This involves, of course, a rather subjective choice, given the current level of understanding as to how galaxies and clusters form. Given such a prescription, the cluster correlation function may be written in terms of conditional probabilities in a form analogous to, and derived from, equation (15):

$$1 + \xi_c(x) = \int \mathrm{d}^3\boldsymbol{r}\, \mathrm{d}^3\boldsymbol{s}\, \mathrm{d}\nu_1\, \mathrm{d}\nu_2\, P(x\,|\,\boldsymbol{r},\boldsymbol{s})\, P(\boldsymbol{r})\, P(1c\,|\,\nu_1) P(2c\,|\,\nu_2) P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)\, P(\nu_1,\nu_2\,|\,\boldsymbol{r}). \tag{22}$$

In this equation, $P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)$ is the conditional probability that two points, where the smoothed density field takes values ν_1,ν_2 (in units of the rms perturbation, σ_0), should have a relative displacement $\boldsymbol{s}$, given that their initial separation was $\boldsymbol{r}$. The term $P(\nu_1,\nu_2\,|\,\boldsymbol{r})$ is the conditional probability that the smoothed density field should take values ν_1,ν_2 at points with initial separation $\boldsymbol{r}$. From the definition of conditional probabilities it follows that

$$P(\boldsymbol{s}\,|\,\boldsymbol{r},\nu_1,\nu_2)\, P(\nu_1,\nu_2\,|\,\boldsymbol{r}) \;=\; P(\boldsymbol{s},\nu_1,\nu_2\,|\,\boldsymbol{r}), \tag{23}$$

which has a Gaussian distribution. The term $P(c\,|\,\nu)$, the probability of there being a cluster at a point where the density field (smoothed on a scale larger than that used to define the peaks) takes a value ν, encapsulates our statistical prescription for cluster formation. As discussed above, BBKS and others have studied models where objects form in thresholded regions or at peaks in the linear density field. Some authors have argued, however, that the physics of structure formation is more likely to produce a smooth non-linear relation between the mass density and the number density of objects, rather than a sharp threshold (e.g. Kaiser & Davis 1985; BBKS; Szalay 1988; Coles & Jones 1991). In particular, Kaiser & Davis (1985) suggest an exponential relationship between the number density of objects and the density field. This gives the correct form for the statistical contribution to the correlation function at large separations (see below) and is very convenient computationally, so we follow BC in employing it here, giving the form of $P(c\,|\,\nu)$ as

$$P(c\,|\,\nu) \;=\; \exp[(E-1)\sigma_0\nu]\, \exp[-(E-1)^2\sigma_0^2/2], \tag{24}$$

where the second exponential term normalizes $P(c\,|\,\nu)$, so that it conserves mean density, and where the choice of the value of the enhancement factor, E, is considered below. The background smoothing scale used here is arbitrary, except for the constraint that it cannot be smaller than R_s, so we take it to be equal to R_f, the filter scale used to suppress shell crossing in the Zeldovich Approximation, which satisfies this constraint by definition.

The derivation of the final expression for the cluster correlation function from this point is sketched in Appendix D. It proceeds similarly to that for the mass correlation function and yields the following result for $\xi_{cZ}(x)$, the cluster correlation function in the Zeldovich Approximation:

$$1+\xi_{cZ}(x)=\frac{1}{2\sqrt{\pi}x\sigma_s^2}\int_o^\infty \frac{r^2\,\mathrm{d}r\exp[(E-1)^2\xi(r)]}{(1-\psi_\perp)^{1/2}(\psi_\perp-\psi_\parallel)^{1/2}}\left\{F(\tilde{u}_+)\exp\left[\frac{-(x-\tilde{r})^2}{4\sigma_s^2(1-\psi_\parallel)}\right]-F(\tilde{u}_-)\exp\left[\frac{-(x+\tilde{r})^2}{4\sigma_s^2(1-\psi_\parallel)}\right]\right\}, \tag{25}$$

where

$$\begin{cases}\tilde{r} &= r+6(E-1)\sigma_s^2\left(\mathrm{d}\psi_v/\mathrm{d}r\right)\\ \tilde{u}_\pm &= \frac{(\psi_\perp-\psi_\parallel)^{1/2}}{2\sigma_s(1-\psi_\parallel)^{1/2}(1-\psi_\perp)^{1/2}}\left[\tilde{r}\frac{(1-\psi_\perp)}{(\psi_\perp-\psi_\parallel)}\pm x\right]\end{cases}$$

and where $\psi_v(r)\equiv(2\psi_\perp+\psi_\parallel)/3$ and $\xi(r)$ is the correlation function of the smoothed linear density field. This also differs from the equivalent expression given by BC, due to the presence of typographical errors in that work.

The appropriate value of the enhancement factor, E, may be determined readily by matching the large-separation asymptotic limit of the statistical contribution to ξ_{cZ} to that of the statistical peak–peak correlation function. The one-dimensional rms particle displacement, σ_s, is a measure of the degree of dynamical evolution of the density field, so that the statistical contribution to ξ_{cZ} comes from considering the limit $\sigma_s\to 0$. It may readily be shown that

$$1+\xi_{cZ}(x)\to\exp\left[(E-1)^2\xi(x)\right]\quad\text{as}\quad\sigma_s\to 0, \tag{26}$$

in which case the large-separation asymptotic form of the statistical contribution to $\xi_{cZ}(x)$ is given by

$$\xi_{cZ}(x)\to(E-1)^2\xi(x)\quad\text{as}\quad x\to\infty,\ \sigma_s\to 0. \tag{27}$$

Matching of this with the result given in equation (1) yields the correct value of E to be used. From the results of LHP it is also possible to verify that the form of $P(c\mid\nu)$ given in equation (24) provides a good approximation to $\xi_{\mathrm{pk-pk}}(r)$ down to $r\sim R_s$.

4 MODELS AND NORMALIZATION

We consider models in which the universe is spatially flat and in which the initial density fluctuations are Gaussian and have scale-invariant power spectra. Such models are consistent with the predictions of inflation and with the results of the *COBE* DMR experiment (Smoot et al. 1992), which we use to constrain the amplitude of the fluctuations at the present epoch.

The power spectrum of density fluctuations is given by

$$\Delta^2(k)=\frac{4\epsilon^2}{9}\left(\frac{ck}{H_0}\right)^4 T^2(k), \tag{28}$$

where $T(k)$ is the transfer function and the dimensionless normalization parameter, ϵ, is that defined by Peacock (1991).

We follow Efstathiou, Bond & White (1992a, EBW) in adopting the following parametric form for the transfer function:

$$T(k)=\left\{1+\left[a_1k+(a_2k)^{3/2}+(a_3k)^2\right]^{a_4}\right\}^{-1/a_4}, \tag{29}$$

where $a_1=(6.4/\Gamma)\,h^{-1}$ Mpc, $a_2=(3.0/\Gamma)\,h^{-1}$ Mpc, $a_3=(1.7/\Gamma)\,h^{-1}$ Mpc and $a_4=1.13$. The origin of this parametric form is in fitting the linear power spectrum of CDM models (Bond & Efstathiou 1984), but, with suitable choices of the parameter Γ, it can fit the power spectra of a wide range of cosmological models to a reasonable accuracy. In retrospect, the definition of Γ chosen by EBW is slightly unfortunate. In pure CDM models, where Ω_B (the fraction of the critical density in baryons) is zero, $T(k)$ is a function of $k/\Omega h^2$ only, so one might interpret Γ as Ωh. For models where $\Omega_B\neq 0$, the appropriate scaling is found empirically to be roughly with $k/\Omega h^2\mathrm{e}^{-2\Omega_B}$. The EBW transfer function for CDM is a fit to a model with $\Omega_B=0.03$, which means that Γ differs from Ωh by about 6 per cent, which can lead to significant differences between a transfer function of the EBW form with a given value of Γ and, say, a BBKS CDM transfer function (for negligible baryonic content) which has Ωh equal to the same numerical value.

EBW also show how the *COBE* observations may be used to fix the amplitude of the power spectrum through the quantity Q_{rms}/T_0, where Q_{rms} is the rms quadrupole of the fluctuations in the background radiation and T_0 is its mean temperature, found by Mather et al. (1990) to be $T_0=2.735\pm0.06$ K. This analysis assumes that the temperature anisotropies seen by *COBE* were (excepting the monopole and dipole, of course) produced by fluctuations in the gravitational potential on the surface of last scattering and that the standard treatment of this effect (Sachs & Wolfe 1967) is applicable. In using it, we neglect the possibility that a significant portion of the *COBE* signal may be due to gravitational waves (e.g. Liddle & Lyth 1992; Lidsey & Coles 1992): the magnitude of the gravitational wave component is dependent on the form of the potential employed in the inflationary model that is used to generate the primordial fluctuations and it is thought to be $\lesssim 10$ per cent for most standard inflationary models. If this assumption is false, then lower values of ϵ are appropriate and the importance of the dynamical component of the clustering is reduced relative to the statistical part. We also assume that the Universe is spatially flat.

The background radiation temperature fluctuations on the sky may be expanded in spherical harmonics, $Y_l^m(\theta,\phi)$:

$$\frac{\Delta T}{T} = \sum_{l,m} a_l^m Y_l^m(\theta, \phi), \tag{30}$$

where the angular power spectrum, $C_l = \langle | a_l^m |^2 \rangle$, is related to the power spectrum of density perturbations at the present epoch. For a scale-invariant spectrum of the form

$$\Delta^2(k) = \frac{4\epsilon^2}{9} \left(\frac{ck}{H_0} \right)^4, \tag{31}$$

in a spatially flat universe, this relationship is given (by EBW) as

$$C_l = \frac{2\pi}{9} \frac{\Omega_0^{1.54}}{l(l+1)} \epsilon^2. \tag{32}$$

Smoot et al. (1992) fit such a formula to the *COBE* data for multipoles higher than the quadrupole and express the results in terms of the value of $Q_{\rm rms}$, which is related to C_2 by

$$Q_{\rm rms} = \left(\frac{5C_2}{4\pi} \right)^{1/2} T_0, \tag{33}$$

and so using equations (36) and (37) this gives, for a scale–invariant spectrum,

$$\epsilon^2 = \frac{108}{5} \left(\frac{Q_{\rm rms}}{T_0} \right)^2 \Omega_0^{-1.54}. \tag{34}$$

Fitting to the *COBE* data, subject to the constraint of a scale-invariant spectrum, Smoot et al. (1992) obtain a value of $Q_{\rm rms} = 16.7 \pm 4\ \mu$K. This does not include, however, the effect of the 'cosmic variance', an added uncertainty due to the power spectrum of the observed sky not being equal to that of the mean of the ensemble of all possible skies in the Universe. Inclusion of the cosmic variance requires another fit to the higher order multipoles, subject, for our purposes, to the constraint of a scale-invariant spectrum, which will modify the $Q_{\rm rms}$ value and increase its uncertainty slightly. We take the fluctuation amplitude indicated by *COBE* to be $Q_{\rm rms} = 17 \pm 5\ \mu$K, which, from equation (34), yields $\epsilon = 2.9 \pm 0.9 \times 10^{-5}\ \Omega_0^{-0.77}$. This defines the interval of fluctuation amplitudes that we shall consider in this work.

We consider the following four spatially flat models:

CDM: $\Omega_0 = 1$, $\lambda_0 = 0$, $\Gamma = 0.5$ (Model 1),
low-density CDM: $\Omega_0 = 0.2$, $\lambda_0 = 0.8$, $\Gamma = 0.2$ (Model 2),
CDM + extra large-scale power: $\Omega_0 = 1$, $\lambda_0 = 0$, $\Gamma = 0.2$ (Model 3),
Bahcall & Cen model: $\Omega_0 = 0.2$, $\lambda_0 = 0.8$, $\Gamma = 0.1$ (Model 4),

where $\lambda_0 = \Lambda/3H_0^2$. Model 2 is that advanced by Efstathiou, Sutherland & Maddox (1990) to account for the strong angular correlations seen in the APM galaxy catalogue: cluster–cluster correlations in this model were considered in linear theory by Bardeen et al. (1987). From EBW we see that one of the possible cosmogonies that could correspond to Model 3 is an $\Omega_0 = 1$, h=0.5, CDM-dominated universe in which one species of massive neutrino contributes $\Omega_\nu \simeq 0.3$ (van Dalen & Schaefer 1992): such a model is elaborated by Davis, Summers & Schlegel (1992) and Taylor & Rowan-Robinson (1992). Models 2 and 3 have the same transfer function, so any differences between them can only result from the difference in the fluctuation amplitude that they require to match the *COBE* observations: one result of the higher amplitude of Model 2 is that the importance of the dynamical component of the clustering, relative to the statistical component, is greater than for Model 3. Model 4 is the transfer function which Bahcall & Cen (1992) claim is consistent with all cluster observations. The amplitude used by Bahcall & Cen is excluded by *COBE* , and we consider only the range of amplitudes that *COBE* allows. Bahcall & Cen (1992) claim that their cluster correlation function results are unchanged if this model is open or has a non-zero cosmological constant to make it spatially flat: we consider only the spatially flat model here.

In Fig. 2 we show the variation of R_s with characteristic separation, d, that our procedure from Section 2.2 produces for the four models listed above. The differences in the value of R_s for a given separation, d, between the different models may be readily understood from a consideration of the relative amounts of small-scale power in their respective power spectra. In all cases the filter radii correspond to reasonable mass estimates for, say, Abell $R \geq 1$ clusters, although we stress that the 'cloud-in-cloud' problem precludes the results shown in Fig. 2 from being expressed as a cluster mass function.

5 RESULTS

In this section we present the results of our computations of the cluster correlation function and of the $d - r_0$ relationship, and compare them with observations. Before that, however, we shall consider the sensitivity of the results to variation in the parameters of the model and compare our results with those obtained by Bahcall & Cen (1992) by numerical simulation.

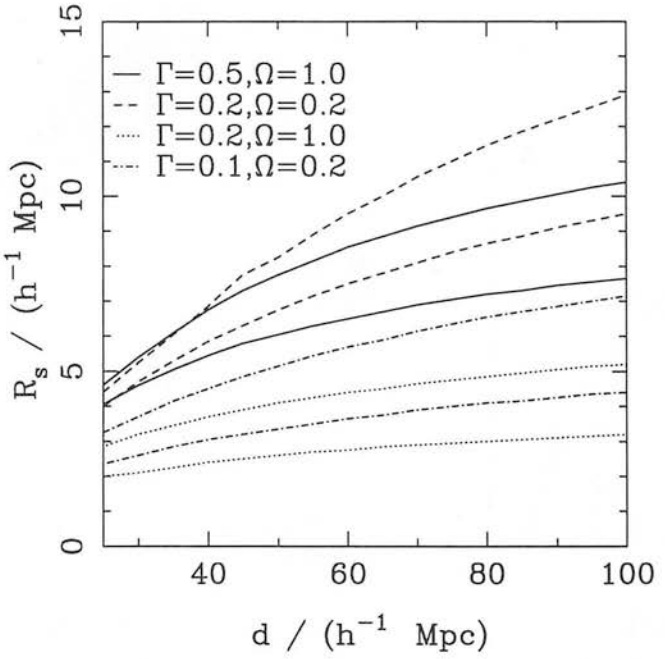


Figure 2. The variation of R_s with characteristic separation, d, for the four models we consider, normalized to *COBE*. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively. In all cases the upper line of the pair for each model is for δ_c=1.0 and the lower is for δ_c=1.7.

5.1 Sensitivity to parameters

In our formalism, the cluster correlation function for a given transfer function and cluster number density is a function of three parameters: the background smoothing scale, R_f; the linear density contrast at collapse, δ_c; and the amplitude of primordial fluctuations, ϵ. In the preceding sections we have discussed constraints on all these quantities, and now we consider the sensitivity of our results to variations in their values.

In Section 3 we discussed a procedure for determining the correct value of the smoothing scale, R_f, through the maximization of the small-scale correlations of the density field. This is subject to a constraint that $R_f \geq R_s$, to ensure that excessive small-scale cluster correlations are not produced by power spectra with little small-scale power, resulting from our biasing prescription, $P(c \mid \nu)$. The effect of varying the value of R_f is, as one would naively expect, largely confined to separations of less than a few filter radii, where one must always be suspicious that the filter function will be dominant. For example, upon changing the value of the smoothing length from $R_f = 0.5\ h^{-1}$ Mpc to $2.6\ h^{-1}$ Mpc, the correlation function for an $\Omega = 1$, $\Gamma = 0.2$ model only changes by more than 10 per cent for $r \leq 6\ h^{-1}$ Mpc. Our results are, therefore, quite robust to variations in the selected value of R_f.

The sensitivity to the value of δ_c is illustrated in Fig. 4, which is discussed in more detail in the next subsection. We find that there is a degree of sensitivity to the value of δ_c for all four models that we consider, but it is not too serious: the difference in the correlation function for $\delta_c = 1.0$ and $\delta_c = 1.7$, for a given power spectrum and cluster number density, is only very rarely found to be larger than the error bar on the observational point with which it is to be compared.

The sensitivity of our results to the value of the fluctuation amplitude, ϵ, is shown in Fig. 3. This figure shows the correlation function for clusters of EMRS number density, calculated for the four models with the extreme allowed values of ϵ. It is clear that Model 3 is less sensitive to the fluctuation amplitude than are Models 1, 2 and 4. This is because Model 3 has less small-scale power and so its cluster correlations are dominated by the statistical contribution, rather than by the dynamical contribution, which means that its cluster correlation function will be less sensitive to the value of ϵ, which has a strong influence on the dynamical contribution to clustering, but not on the statistical one. The results in Fig. 3 should be borne in mind when looking at the plots shown in Fig. 5, which, for the sake of clarity, show the correlation function only for the *COBE* normalization and do not indicate the effects of uncertainty in the value of ϵ.

5.2 Comparison with the results of numerical simulations

In Fig. 4 we show a comparison between our results and those calculated through numerical simulation by Bahcall & Cen (1992). We show two of their models, CDM and a $\Gamma = 0.1$ model, for two different cluster number densities. These cluster correlation functions are calculated by wholly independent methods: our formalism uses a prescription based on peak selection in the linear density field, while Bahcall & Cen use an adaptive 'friends-of-friends' algorithm to link particles in the non-linear density field at

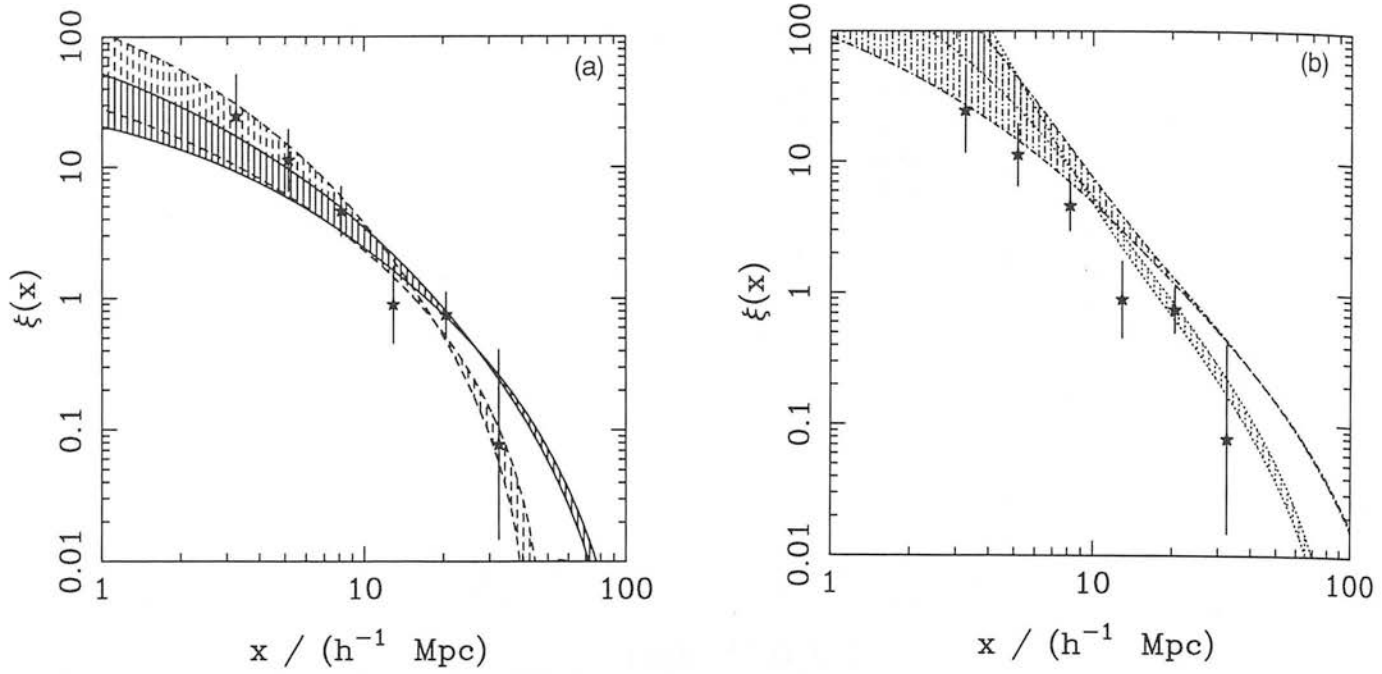


Figure 3. The effect on the correlation function of varying the amplitude of primordial fluctuations. The number density considered is that of the EMRS clusters of Nichol et al. (1992), from where the data points are taken. In (a) the dashed and solid lines give results for Models 1 and 2 respectively, while in (b) the dotted and dash-dotted lines are for Models 3 and 4 respectively: all models are for $\delta_c = 1.0$. In each case the hatched area indicates the region between the curves for the correlation function for the maximum and minimum values of the fluctuation amplitude ϵ.

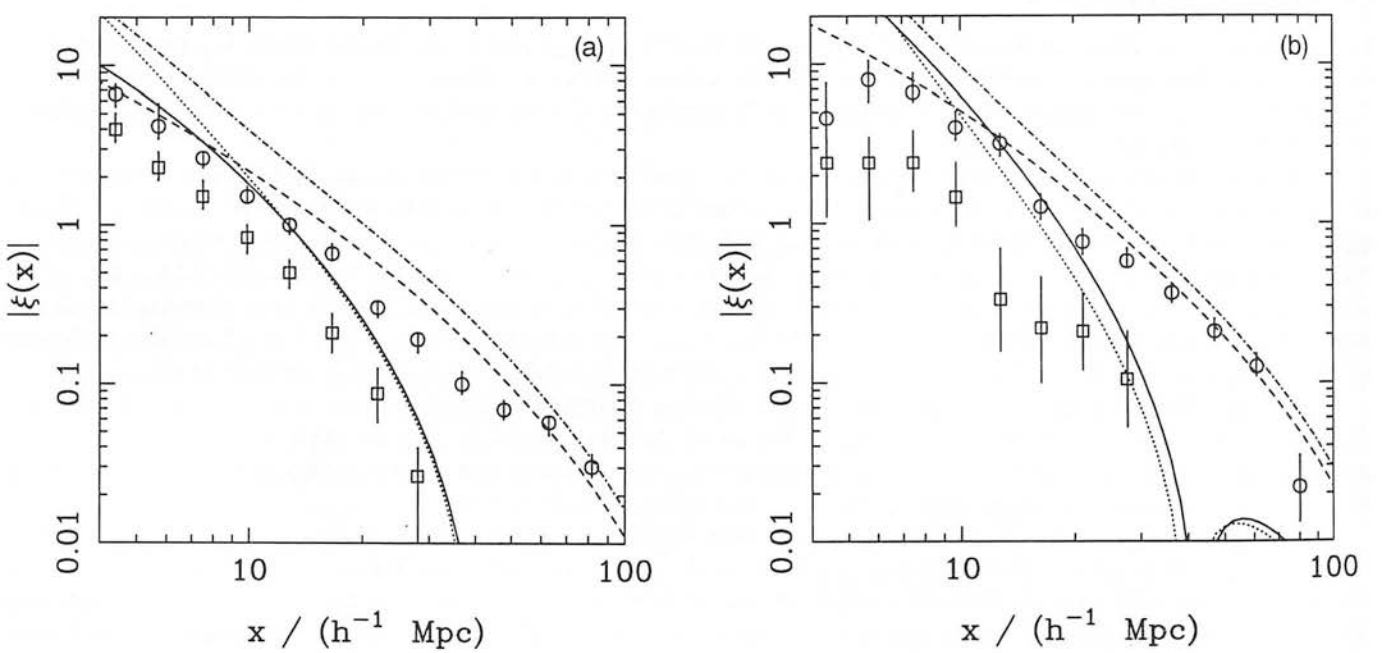


Figure 4. Comparison with numerical results from Bahcall & Cen (1992), for clusters with the number density of (a) APM clusters and (b) Abell $R \geq 1$ clusters. The solid (dotted) line shows the correlation function for an $\Omega_0 = 1, \Gamma = 0.5$ model with $\delta_c = 1.0$ (δ_c=1.7), while the dashed (dash-dotted) line shows that for an $\Omega_0 = 0.2, \Gamma = 0.1$ model with δ_c=1.0 ($\delta_c = 1.7$). The squares and circles are the numerical results of Bahcall & Cen for the $\Gamma = 0.5$ and $\Gamma = 0.1$ models respectively.

the end of their simulation to form clusters. We reproduce the results for the $\Gamma = 0.1$ model very well at both number densities using δ_c=1, but neither value of δ_c can reproduce the low level of clustering that Bahcall & Cen find in their $\Gamma = 0.5$ model: to do that we require $\delta_c \sim 5$, which produces clusters whose masses are too low. In view of the problems mentioned earlier concerning the use of low-resolution N-body codes in simulating clusters, we do not regard this discrepancy too seriously and await the advent of more sophisticated numerical simulations with which to test our analytic theory.

5.3 Comparison with observations

We have computed the cluster correlation function for the four models for four cluster number densities, those for the EMRS and APM automated surveys and the Abell $R \geq 1$ and $R \geq 2$ samples of PW92: we note that PW92 urge caution in the use of their $R \geq 2$ sample, due to its small size. We do not consider Abell cluster samples including $R = 0$ clusters, since it has been amply demonstrated (Sutherland 1988; Sutherland & Efstathiou 1991; Efstathiou et al. 1992b; PW92) that such samples are corrupted by projection effects, which lead to spurious clustering. Sutherland (1988) also claimed that the $R \geq 1$ sample of BS83 is similarly corrupted, but a recent analysis of a much larger sample of $R \geq 1$ clusters by PW92 found no significant projection effects. In what follows, we shall assume the truth of the hypothesis of PW92, that the correlations of Abell $R \geq 1$ clusters are genuine, although we note that the results of PW92 are consistent with a low level of corruption by projection effects, albeit one much less serious than that found in samples of $R \geq 0$ clusters.

We choose not to make comparisons of our results with observations to which the Sutherland method for correction for redshift anisotropy has been applied. Such a procedure can only be reliably and profitably applied when it can be unambiguously shown that the clustering at large radial separation that it removes is spurious, and this is possible only for samples like the EMRS where there are many measured redshifts per cluster. Efstathiou et al. (1992b) estimate that the removal of genuine cluster pairs with large radial separations can introduce spurious contributions of ~ 0.01 into the correlation function of a cluster sample that is free from projection biases. This is unfortunate, since it is at low ξ_{cc} ($\xi_{cc} \leq 0.1$) that discrimination between theoretical models is easiest, and so this effect could seriously impair the comparison of theory and observation. Whilst it is true that such biases could have no significant impact on the correlation function at $\xi_{cc} \sim 1$, we shall see below that the correlation length taken alone is a poor way of characterizing the spatial distribution of a sample of clusters.

We show our results for the cluster correlation function in Fig. 5. The observational points have 1σ Poisson error bars. Where no correlation is detected with significance at the 1σ level we plot a 90 per cent confidence upper limit, marked with an arrow. We compare theoretical predictions of configuration-space correlation functions with observed results measured in redshift space. Redshift-space corrections are considered by Kaiser (1987), who finds that the (isotropic) real-space correlation function, $\xi_r(x)$, is related to the direction-averaged redshift-space correlation function, $\xi_{(s)}(x)$, by $\xi_{(s)}(x) = [1 + (2\Omega_0^{0.6}/3E) + (\Omega_0^{1.2}/5E^2)]\xi_r(x)$. For the models that we consider, this factor is a small fraction of the uncertainty in the *COBE* amplitude and much less than the errors in the observations with which we shall compare our results, so we neglect this correction. Fig. 5(a) shows results for APM $\mathcal{R} \geq 20$ clusters, with number density $N_c = 2.4 \times 10^{-5} h^3\,\mathrm{Mpc}^{-3}$. The observational data are taken from Efstathiou et al. (1992b) and include clusters from a slightly wider redshift range than that considered by Dalton et al. (1992). In Fig. 5(b) we give results for EMRS clusters, from Nichol et al. (1992), which have a number density of $N_c = 1.0 \times 10^{-5} h^3\,\mathrm{Mpc}^{-3}$. Figs 5(c) and 5(d) show correlation functions for the Abell $R \geq 1$ and $R \geq 2$ samples of PW92, which have number densities of 7.2×10^{-6} and $1.7 \times 10^{-6} h^3\,\mathrm{Mpc}^{-3}$ respectively.

None of the models can account for the correlation functions of all four cluster samples, although, conversely, the uncertainty in the *COBE* amplitude and the size of the error bars on the observations mean that it is difficult to exclude any of the models with a high degree of confidence on the basis of the observations as they stand. CDM with $\Gamma = 0.5$ appears incapable of reproducing the strength of clustering observed for rich Abell clusters, although it gives the best fit to the APM $\mathcal{R} \geq 20$ clusters and is the only model in which the first zero-crossing point of the correlation function is at a separation of $\sim 40\ h^{-1}$ Mpc, as indicated by the APM and EMRS (Nichol 1992) data. One must be careful, however, in using the first zero-crossing point as a diagnostic tool, since the observational data become very noisy at low values of ξ_{cc} and are also prone to significant systematic errors as a result of uncertainty in the cluster number density. Model 2 gives the best overall agreement with the correlation functions of the four cluster samples and, indeed, its only real inconsistency with the data comes at large scales for the APM number density, where the reliability of the observations is in some doubt. The presence of a cosmological constant affects the epoch and rate at which structure forms, but has little effect on the final phase-space distribution of the particles (Kauffmann & White 1992, and references therein), so that the results of Model 2 should be very similar to those of an open model ($\Omega_0 = \Gamma = 0.2, \lambda_0 = 0$) with the same fluctuation amplitude, and, indeed, Bahcall & Cen (1992) report such a finding on the basis of their numerical simulations of models with and without a cosmological constant. The problem, however, in making this identification between open and flat models with the same Γ lies with the amplitude: the standard analyses of the Sachs–Wolfe effect are for a spatially flat universe, so it is not clear how to determine the fluctuation amplitude for open models from *COBE*.

In Fig. 6 we show results for the $d - r_0$ relationship. In addition to the four cluster samples discussed above, we also include observational data from the APM $\mathcal{R} \geq 35$ sample of Efstathiou et al. (1992b) and the Abell $R \geq 1$ samples of BS83 and Postman, Huchra & Geller (1992, PHG). We determine the value of r_0 by interpolation in log r space from a correlation function evaluated at a hundred separations in the range $0.5 \leq r/h^{-1}\mathrm{Mpc} \leq 100$. If, instead, a power-law fit is made to the (positive portion of the) correlation function in this interval, the resultant values of r_0 differ, typically, by $\sim 1\ h^{-1}$ Mpc or less from those obtained by interpolation, and the slope of the best-fitting power law is seen to increase with d.

Two points are clear from Figs 5 and 6. First, no model is able to produce the strength of the trend of correlation length increasing with richness that is required to account for all of the observations from APM $\mathcal{R} \geq 20$ clusters to Abell $R \geq 2$. Secondly, the $d - r_0$ relation, taken on its own, may be a misleading statistic: from Fig. 6 one would conclude that all of the models are

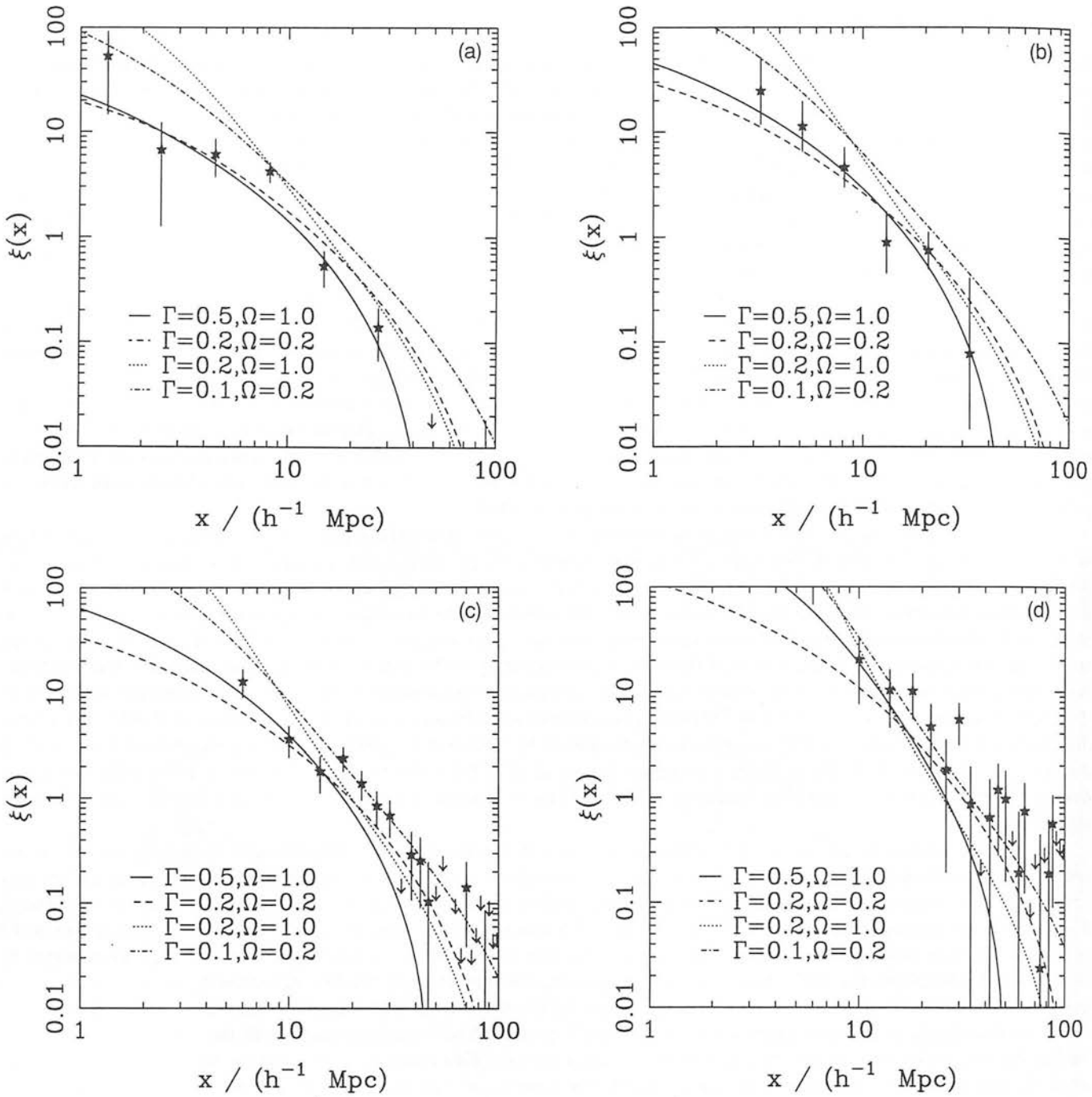


Figure 5. The cluster correlation function for (a) APM $\mathscr{R} \geq 20$ clusters, (b) EMRS clusters, (c) Abell $R \geq 1$ clusters, and (d) Abell $R \geq 2$ clusters. The solid, dashed, dotted and dash-dotted lines give the correlation function for Models 1,2,3 and 4 respectively with *COBE* normalization and $\delta_c = 1.0$.

wildly inconsistent with the Abell $R \geq 2$ data, although the noise on the data given in Fig. 5(d) shows that the situation is not so clear and, in addition, the Poisson error quoted for the correlation length may be something of an underestimate in this case.

The variation between models in the sensitivity of the $d - r_0$ relation to the value of ϵ can be readily explained by the same argument as given in Section 5.1 above: the sensitivity to ϵ is greater for models with more small-scale power, since these are the ones where the amplitude-sensitive dynamical contribution to the clustering is enhanced relative to the statistical component produced by the initial peak distribution.

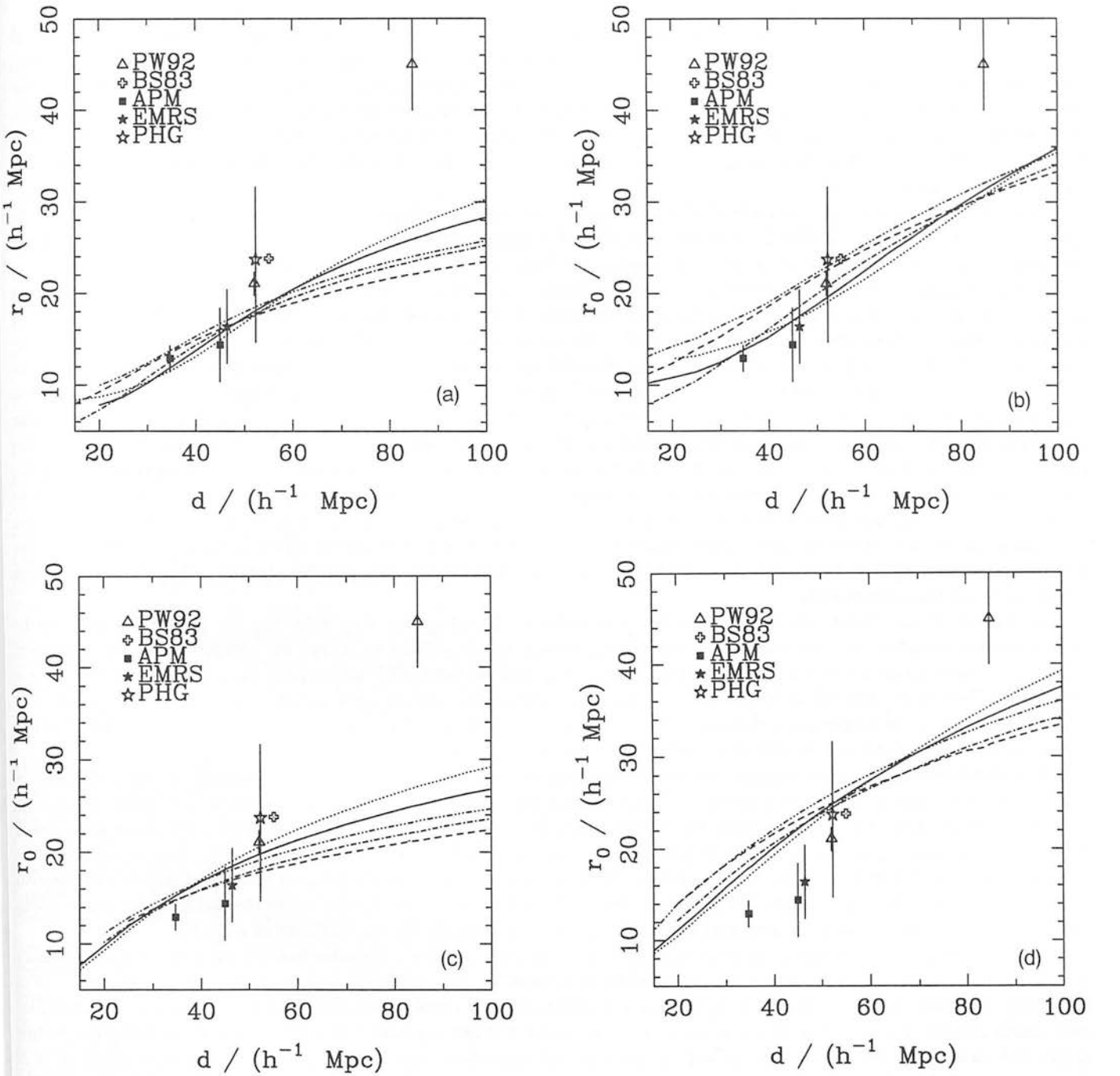


Figure 6. The $d - r_0$ relation as a function of ϵ and δ_c for (a) Model 1, (b) Model 2, (c) Model 3 and (d) Model 4 (see Section 4). The observational data points are from the sources indicated in the figure and are discussed in Section 5. In (a) and (b) the curves are for (i) $\epsilon = 3.8 \times 10^{-5}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 2.9 \times 10^{-5}$, $\delta_c = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 2.0 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line). In (c) and (d) the curves are for (i) $\epsilon = 1.3 \times 10^{-4}$, $\delta_c = 1.0$ (dotted line) and $\delta_c = 1.7$ (dot-dot-dot-dashed line), (ii) $\epsilon = 1.0 \times 10^{-4}$, $\delta = 1.0$ (solid line) and $\delta_c = 1.7$ (dashed line) and (iii) $\epsilon = 6.9 \times 10^{-5}$, $\delta_c = 1.0$ (dot-dashed line).

6 DISCUSSION AND CONCLUSIONS

In this paper we have reported on the first analytic calculation of the cluster correlation function to take account of the non-linear evolution of the cosmological density field. We employ the Zeldovich Approximation, which has been shown to evolve the cosmological density field very accurately well into the non-linear regime, along with a physically well-motivated prescription to relate the evolved density field to the distribution of clusters.

We have applied our method to a range of cosmological models with Gaussian primordial fluctuations in a spatially flat universe. CDM models seem incapable of reproducing the strength of the correlations observed in samples of the richest clusters, but these are drawn from the Abell/ACO catalogue and so may be regarded with a degree of suspicion. The models that give the

best agreement with the cluster correlation data are those with low Γ, as required to fit the galaxy correlation data. This lends a degree of support to the hypothesis of Gaussian primordial fluctuations, since, through the Kaiser mechanism (Kaiser 1984), Gaussian models would be expected to produce a picture such as this: a power spectrum capable of accounting for both galaxy and cluster correlations, with a richness dependence qualitatively like that which is observed, with r_0 increasing with d initially, but with the $d - r_0$ relation flattening off eventually at very high richness. This flattening of the $d - r_0$ relation at high richness follows from BBKS, who show that $\xi_{\mathrm{pk-pk}} \propto \nu_s^2$ and that the number density of peaks above the threshold, $n_{\mathrm{pk}}(\nu_s)$, varies as $n_{\mathrm{pk}}(\nu_s) \propto \nu_s^2 \mathrm{e}^{-\nu_s^2/2}$ for large ν_s.

In all of the models we have considered, this saturation in the value of r_0 happens well below the correlation length of the Abell $R \geq 2$ cluster sample of PW92. There are a number of possible explanations for this, the simplest of which is that the correlation length of the $R \geq 2$ sample is incorrect. Another possibility is that the primordial fluctuations are not quite Gaussian, in which case one might expect any deviations from Gaussian predictions to be most apparent for the richest clusters, corresponding to the highest peaks. It is also possible that the deviation between the observed correlation length of the PW92 Abell $R \geq 2$ sample and those predicted by our models is due to the breakdown of the identification of peaks in the linear density field as sites of nascent clusters for the very highest peaks. This identification of peaks in the linear and non-linear density fields is known to be very poor for low peaks, but is seen to improve with increasing peak height, so one might expect it to be closest for the highest peaks. This may be so, but the number density of peaks with $\nu_s \gg 1$ is so low that even a tiny error in ν_s can mean that linear and non-linear peaks are completely different sets of objects. The Voronoi cellular model may have an insight to offer here. This model produces a cluster correlation in which r_0 scales proportional to d (Williams 1992). This happens because the Voronoi model produces its richest clusters around the periphery of the largest voids, since mass drains from the proto-voids to the proto-clusters. A change in the cluster richness does not alter the typical pair separation very much – just the distance to the next 'clump' of clusters. This may appear rather different from the picture of rich cluster formation in Gaussian models, but the idea of clusters gaining mass from voids is, none the less, generic in such models. It is possible, therefore, that an $r_0 \propto d$ scaling can emerge from Gaussian models.

Analysis of further *COBE* data will reduce the uncertainty in the value of ϵ, thus tightening the predictions that can be made using the formalism we have outlined in this paper, although this is unlikely to change the conclusions to any significant degree, since the uncertainties in the predictions for ξ_{cc} due to the error on the *COBE* quadrupole are no larger than those in the observations with which they are to be compared. Better observational data are required, therefore, to improve the effectiveness of the cluster correlation function as a diagnostic tool in the study of large-scale structure, and it is hoped that samples of X-ray clusters selected by *ROSAT* will provide an important new observational input.

A more fundamental problem remains, however, concerning operational definitions of what constitutes a cluster of galaxies. Theoretical work tends to be expressed in terms of mass thresholds and peak heights, whereas observational selection criteria are far more complex than this. This is particularly true for optically selected clusters, where the observer's definition of a cluster in terms of an excess in the surface density of galaxies projected on the sky is not trivial to relate to the theorist's image of an overdense clump of matter that collapses and virializes. There is, therefore, always the problem of whether observers and theorists are comparing like with like, although this is perhaps less of a problem for X-ray clusters, where the picture of localized X-ray emission from gas confined within a deep potential well is nearer to a simplistic theoretical image of a cluster.

Despite these caveats, the spatial distribution of clusters of galaxies remains a powerful tool for the study of the large-scale structure of the Universe, and we believe that the analytic framework we have described here provides an important new way of addressing this topic. Our results show that, while there are models that can readily reproduce the correlation functions determined from recent automated cluster catalogues, we have found no model that can reproduce the strong trend of increasing correlation length with richness that is required to account for the observed correlation lengths of all cluster samples, from APM $\mathscr{R} \geq 20$ clusters to Abell $R \geq 2$ clusters. The model that comes closest overall to reproducing the observed cluster correlations is that advocated on the basis of its ability to reproduce the strong angular correlations of APM galaxies: this may be taken as support for the Kaiser model, in which objects form at peaks in a Gaussian initial density field. Better observational data are required before the cluster correlation function can realize its full potential as a cosmological diagnostic, but we believe that our results place the theoretical side of its study on a firmer footing.

ACKNOWLEDGMENTS

RGM acknowledges receipt of an SERC studentship and would like to thank Gavin Dalton for patiently answering a number of queries, Bob Nichol for many stimulating discussions and Andy Taylor for a very thorough reading of a draft of this paper. We thank the anonymous referee for helpful suggestions which clarified a couple of points in this paper.

REFERENCES

Abell G.O., 1958, ApJS, 3, 211
Abell G.O., Corwin H.G., Olowin R.P., 1989, ApJS, 70, 1 (ACO)
Bahcall N.A., 1988, ARA&A, 26, 631
Bahcall N.A., Burgett W.S., 1986, ApJ, 300, L35
Bahcall N.A., Cen R.Y., 1992, ApJ, 398, L81
Bahcall N.A., Soneira R.M., 1983, ApJ, 270, 20 (BS83)
Bahcall N.A., West M.J., 1992, ApJ, 392, 419
Bardeen J.M., Steinhardt P.J., Turner M.S., 1983, Phys. Rev. D, 28, 679
Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., 1986, ApJ, 304, 15 (BBKS)
Bardeen J.M., Bond J.R., Efstathiou G., 1987, ApJ, 321, 28
Bond J.R., Couchman H.M.P., 1987, preprint (BC)
Bond J.R., Couchman H.M.P., 1988, in Coley A., Dyer C.C., Tupper B.O.J., eds, Proc. Second Canadian Conference on General Relativity & Relativistic Astrophysics. World Scientific, Singapore (BC)
Bond J.R., Efstathiou G., 1984, ApJ, 285, L45
Coles P., 1986, MNRAS, 222, 9P
Coles P., 1990, MNRAS, 243, 171
Coles P., Jones B.J.T., 1991, MNRAS, 248, 1
Coles P., Plionis M., 1991, MNRAS, 250, 75
Coles P., Melott A.L., Shandarin S.F., 1993, MNRAS, 260, 765
Couchman H.M.P., 1987, MNRAS, 225, 777
Dalton G.B., Efstathiou G., Maddox S.J., Sutherland W.J., 1992, ApJ, 390, L1
Davis M., Peebles P.J.E., 1983, ApJ, 267, 465
Davis M., Summers F.J., Schlegel D., 1992, Nat, 359, 393
Dekel A., Blumenthal G.R., Primack J.R., Olivier S., 1989, ApJ, 338, L5
Efstathiou G., 1990, in Peacock J.A., Heavens A.F., Davis A.T., eds, Proc. 36th Scot. Univ. Summer School in Physics, Physics of the Early Universe. SUSSP, Edinburgh
Efstathiou G., Silk J., 1983, Fundam. Cosmic Phys., 9, 1
Efstathiou G., Sutherland W.J., Maddox S.J., 1990, Nat, 348, 705
Efstathiou G., Bond J.R., White S.D.M., 1992a, MNRAS, 258, 1P (EBW)
Efstathiou G., Dalton G.B., Sutherland W.J., Maddox S.J., 1992b, MNRAS, 257, 125
Gott J.R., Melott A.L., Dickinson M., 1986, ApJ, 306, 341
Gott J.R. et al., 1989, ApJ, 340, 625
Gradshteyn I.S., Ryzhik I.M., 1980, Tables of Integrals, Series and Products. Academic Press, New York, NY
Gurbatov S.N., Saichev A.I., Shandarin S.F., 1989, MNRAS, 236, 385
Hamilton A.J.S., Gott J.R., Weinberg D.H., 1986, ApJ, 309, 1
Holtzman J.A., Primack J.R., 1993, ApJ, 405, 428
Jensen L.G., Szalay A.S., 1986, ApJ, 305, L5
Juszkiewicz R., Bouchet F.R., 1992, preprint
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1987, MNRAS, 227, 1
Kaiser N., Davis M., 1985, ApJ, 297, 365
Kauffmann G., White S.D.M., 1992, MNRAS, 258, 511
Liddle A.R., Lyth D., 1992, preprint
Lidsey J.E., Coles P., 1992, MNRAS, 258, 57P
Lin C.C., Mestel L., Shu F.H., 1965, ApJ, 142, 1431
Lucey J.R., 1983, MNRAS, 204, 33
Lumsden S.L., Heavens A.F., Peacock J.A., 1989, MNRAS, 238, 293 (LHP)
Lumsden S.L., Nichol R.C., Collins C.A., Guzzo L., 1992, MNRAS, 258, 1
Mather J.C. et al., 1990, ApJ, 354, L37
Melott A.L., Weinberg D.H., Gott J.R., 1988, ApJ, 328, 50
Moore B. et al., 1992, MNRAS, 256, 477
More J.G., Heavens A.F., Peacock J.A., 1986, MNRAS, 220, 189
Nichol R.C., 1992, Ph.D. thesis, Univ. of Edinburgh
Nichol R.C., Collins C.A., Guzzo L., Lumsden S.L., 1992, MNRAS, 255, 21P
Peacock J.A., 1991, MNRAS, 253, 1P

Peacock J.A., Heavens A.F., 1985, MNRAS, 217, 805
Peacock J.A., West M.J., 1992, MNRAS, 259, 494 (PW92)
Peebles P.J.E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Postman M., Huchra J.P., Geller M.J., 1992, ApJ, 384, 404 (PHG)
Sachs R.K., Wolfe A.M., 1967, ApJ, 147, 73
Saunders W. et al., 1991, Nat, 349, 32
Shandarin S.F., Zeldovich Ya.B., 1989, Rev. Mod. Phys., 61, 185
Smoot G.F. et al., 1992, ApJ, 396, L1
Struble M.F., Rood H.J., 1987, ApJS, 63, 543
Sutherland W., 1988, MNRAS, 234, 159
Sutherland W., Efstathiou G., 1991, MNRAS, 248, 159
Szalay A.S., 1988, ApJ, 333, 21
Taylor A.N., Rowan-Robinson M., 1992, Nat, 359, 396
van Dalen A., Schaefer R.K., 1992, ApJ, 398, 33
Weinberg D.H., Gott J.R., Melott A.L., 1987, ApJ, 321, 2
White S.D.M., Frenk C.S., Davis M., Efstathiou G., 1987, ApJ, 313, 505
White S.D.M., Efstathiou G., Frenk C.S., 1992, preprint
Williams B.G., 1992, Ph.D. thesis, Univ. of Edinburgh
Williams B.G., Heavens A.F., Peacock J.A., Shandarin S.F., 1991, MNRAS, 250, 458
Zeldovich Ya.B., 1970, A&A, 5, 84

APPENDIX A: DERIVATION OF $n_{\rm pk}(\nu_{\rm s})$

From BBKS (their equation A18), the differential number density of peaks in the range ν to $\nu + {\rm d}\nu$ is given by

$$\mathcal{N}_{\rm pk}(\nu){\rm d}\nu = \frac{1}{(2\pi)^2 R_*^3}\,{\rm e}^{-\nu^2/2}\,G(\gamma,\gamma\nu)\,{\rm d}\nu, \tag{A1}$$

where (BBKS equation A19) $G(\gamma,\gamma\nu)$ is defined by

$$G(\gamma,\gamma\nu) = \int_0^\infty \frac{{\rm d}x\,f(x)}{[2\pi(1-\gamma^2)]^{1/2}}\exp\left[\frac{-(x-\gamma\nu)^2}{2(1-\gamma^2)}\right]. \tag{A2}$$

In equation (A2), ν is the height of the density perturbation field smoothed on scale $R_{\rm s}$ expressed in units of the rms perturbation, γ is the spectral parameter, $x = -\nabla^2\delta/\sigma_2$ and the function $f(x)$ has the closed form

$$f(x) = \frac{(x^3-3x)}{2}\left\{{\rm erf}\left[\left(\frac{5}{2}\right)^{1/2}x\right] + {\rm erf}\left[\left(\frac{5}{2}\right)^{1/2}\frac{x}{2}\right]\right\} + \left(\frac{2}{5\pi}\right)^{1/2}\left[\left(\frac{31x^2}{4}+\frac{8}{5}\right){\rm e}^{-5x^2/8} + \left(\frac{x^2}{2}-\frac{8}{5}\right){\rm e}^{-5x^2/2}\right], \tag{A3}$$

from BBKS (their equation A15), who also give approximate fits to this exact form.

It follows that the number density, $n_{\rm pk}(\nu_{\rm s})$, of peaks above a threshold $\nu_{\rm s}$ is given by

$$n_{\rm pk}(\nu_{\rm s}) = \int_{\nu_{\rm s}}^\infty \mathcal{N}_{\rm pk}(\nu){\rm d}\nu = \frac{1}{(2\pi)^{5/2}(1-\gamma^2)^{1/2}R_*^3}\int_{x=0}^\infty {\rm d}x\,f(x)\exp\left[\frac{-x^2}{2(1-\gamma^2)}\right] I_{\rm A}, \tag{A4}$$

where

$$I_{\rm A} = \int_{\nu=\nu_{\rm s}}^\infty {\rm d}\nu\exp\left[\frac{-\nu^2}{2(1-\gamma)^2} + \frac{\gamma\nu x}{(1-\gamma^2)}\right]. \tag{A5}$$

From Gradshteyn & Ryzhik (1980, section 3.322),

$$\int_u^\infty \exp\left(\frac{-x^2}{4\alpha}-\beta x\right){\rm d}x = \sqrt{\pi\alpha}\,e^{\alpha\beta^2}\left[1-{\rm erf}\left(\beta\sqrt{\alpha}+\frac{u}{2\sqrt{\alpha}}\right)\right], \tag{A6}$$

and so equation (A5) becomes

$$I_{\rm A} = \sqrt{\frac{\pi(1-\gamma^2)}{2}}\exp\left[\frac{x^2\gamma^2}{2(1-\gamma^2)}\right]{\rm erfc}\left[\frac{(\nu_{\rm s}-\gamma x)}{\sqrt{2(1-\gamma^2)}}\right]. \tag{A7}$$

Hence, from equations (A4) and (A7), it follows that the final form for the number density of peaks above the threshold is

$$n_{\rm pk}(\nu_{\rm s}) = \frac{1}{8\pi^2 R_*^3}\int_{x=0}^\infty {\rm d}x\,f(x)\,{\rm e}^{-x^2/2}\,{\rm erfc}\left[\frac{(\nu_{\rm s}-\gamma x)}{\sqrt{2(1-\gamma^2)}}\right]. \tag{A8}$$

APPENDIX B: DERIVATION OF $\langle \tilde{\nu}_s \rangle$

An expression for the n-point correlation function of peaks above a threshold is given in terms of a $2n$-dimensional integral by BBKS. This expression is valid for separations of more than a few times the filter scale, R_s, used to select the peaks. In the limit $\nu_s^2 \psi(r) \ll 1$, which will obtain at large separations, this integral simplifies and the two-point function takes the form

$$\xi_{\rm pk-pk}(r) = \langle \tilde{\nu}_s \rangle^2 \, \psi(r), \tag{B1}$$

where $\psi(r) \equiv \xi(r)/\xi(0)$ is the normalized mass correlation function. The effective threshold, $\tilde{\nu}$, is defined by

$$\tilde{\nu} \equiv \frac{(\nu - \gamma x)}{(1-\gamma^2)}. \tag{B2}$$

The quantity $\langle \tilde{\nu}_s \rangle$ is the average of this taken over all peaks above the threshold ν_s, and so is given by

$$\langle \tilde{\nu}_s \rangle = \frac{\int_{\nu_s}^{\infty} \tilde{\nu}_s \mathcal{N}_{\rm pk}(\nu)\, {\rm d}\nu}{\int_{\nu_s}^{\infty} \mathcal{N}_{\rm pk}(\nu)\, {\rm d}\nu} = \frac{1}{(2\pi)^{5/2}(1-\gamma^2)^{1/2} R_*^3 \, n_{\rm pk}(\nu_s)} I_{\rm B}, \tag{B3}$$

where

$$I_{\rm B} = \int_{x=0}^{\infty} \int_{\nu=\nu_s}^{\infty} f(x) \left[\frac{(\nu - \gamma x)}{(1-\gamma^2)}\right] {\rm e}^{-S_0} \, {\rm d}x \, {\rm d}\nu, \tag{B4}$$

and

$$S_0 = \left[\frac{(\nu^2 - 2\gamma\nu x + x^2)}{2(1-\gamma^2)}\right]. \tag{B5}$$

If we make the substitution $y \equiv \nu - \gamma x$, then

$$S_0 = \left[\frac{y^2 + (1-\gamma^2)x^2}{(1-\gamma^2)}\right], \tag{B6}$$

and so the expression for $I_{\rm B}$ becomes

$$I_{\rm B} = \int_0^{\infty} f(x) \, {\rm e}^{-x^2/2} \exp\left[\frac{-(\nu_s - \gamma x)^2}{2(1-\gamma^2)}\right] {\rm d}x. \tag{B7}$$

Finally, combination of equations (A8), (B3) and (B7) yields the following form for $\langle \tilde{\nu}_s \rangle$:

$$\langle \tilde{\nu}_s \rangle = \left[\frac{2}{\pi(1-\gamma^2)}\right]^{1/2} \frac{\int_0^{\infty} f(x) \, {\rm e}^{-x^2/2} \exp\left[-(\nu_t - \gamma x)^2/2(1-\gamma^2)\right] {\rm d}x}{\int_0^{\infty} f(x) \, {\rm e}^{-x^2/2} \, {\rm erfc}\left[(\nu_t - \gamma x)/\sqrt{2(1-\gamma^2)}\right] {\rm d}x}. \tag{B8}$$

APPENDIX C: DERIVATION OF $\xi_{\rho Z}(x)$

From equation (15) of Section 3 we see that the mass correlation function may be written as an integral over conditional probabilities, of the form

$$1 + \xi_\rho(x) = \int {\rm d}^3\boldsymbol{r} \, {\rm d}^3\boldsymbol{s} \, P(x \,|\, \boldsymbol{r}, \boldsymbol{s}) \, P(\boldsymbol{s} \,|\, \boldsymbol{r}) \, P(\boldsymbol{r}),$$

where $P(\boldsymbol{r}) = 1$ and $P(x|\boldsymbol{r},\boldsymbol{s}) = \delta(x - |\,\boldsymbol{r} - \boldsymbol{s}\,|)/(4\pi x^2)$.
The probability $P(\boldsymbol{s}|\boldsymbol{r})$ has a Gaussian distribution:

$$P(\boldsymbol{s} \mid \boldsymbol{r}) = \frac{1}{(2\pi)^{3/2} \, \| \, \mathbf{M} \, \|^{1/2}} \exp\left(\frac{-s_i M_{ij}^{-1} s_j}{2}\right), \tag{C1}$$

where the covariance matrix of the distribution is $M_{ij} = \langle s_i(\boldsymbol{y}) \, s_j(\boldsymbol{y} + \boldsymbol{r}) \rangle$, where $\{s_i\}$ are components of the relative displacement field. This may be readily evaluated in a Cartesian coordinate system whose 3-axis lies along the direction of the vector $\hat{\boldsymbol{r}} = \boldsymbol{r}/\,|\,\boldsymbol{r}\,|$, using the relationship between the Fourier components of the displacement and density fields (equation 13) to obtain

$$M_{ij} = 2\sigma_s^2 \left[(1 - \psi_\perp)\delta_{ij} + (\psi_\perp - \psi_\parallel)\hat{r}_i \hat{r}_j\right], \tag{C2}$$

with σ_s, $\psi_\perp$ and $\psi_\parallel$ as defined in equations (19) – (21).
It follows that equation (C1) may be written as

$$P(\boldsymbol{s} \mid \boldsymbol{r}) = \frac{1}{(2\pi)^{3/2}(2\sigma_s)^{3/2}(1-\psi_\perp)(1-\psi_\parallel)^{1/2}} \exp\left\{-\left[\frac{(1-\psi_\parallel)s^2 - (\psi_\perp - \psi_\parallel)(\boldsymbol{s}\cdot\hat{\boldsymbol{r}})^2}{4\sigma_s^2(1-\psi_\perp)(1-\psi_\parallel)}\right]\right\}. \tag{C3}$$

Since $P(x|r,s)$ takes the form of a δ-function, the integration over s is restricted to the sphere at constant $|\,x\,|$. In our coordinate system this means that $s^2 = r^2 + x^2 - 2xr\cos\theta$ and $s\cdot\hat{r} = r - x\cos\theta$. These may be substituted into the expression for $P(s|r)$ and the resultant integrations over θ, ϕ performed, the latter one trivially. This leaves a one-dimensional integration (over r), which is straightforward to convert into the form given in equation (16) of Section 3.

APPENDIX D: DERIVATION OF $\xi_{cZ}(x)$

Equations (22) and (23) in Section 3 gives the following expression for the cluster correlation function:

$$1 + \xi_{cZ}(x) = \int \mathrm{d}^3\boldsymbol{r}\,\mathrm{d}^3\boldsymbol{s}\,\mathrm{d}v_1\,\mathrm{d}v_2\,P(x|\boldsymbol{r},\boldsymbol{s})\,P(\boldsymbol{r})\,P(1c|v_1)\,P(2c|v_2)\,P(\boldsymbol{s},v_1,v_2|\boldsymbol{r}). \tag{D1}$$

The term $P(\boldsymbol{s},v_1,v_2|\boldsymbol{r})$ is given by

$$P(\boldsymbol{s},v_1,v_2|\boldsymbol{r}) = \frac{1}{(2\pi)^{5/2}\,||\,\tilde{\mathbf{M}}\,||^{1/2}}\exp\left(-\frac{1}{2}\tilde{y}_i^{\mathrm{T}}\tilde{M}_{ij}^{-1}\tilde{y}_j\right), \tag{D2}$$

where $\tilde{\boldsymbol{y}}^{\mathrm{T}} = (\boldsymbol{s},v_1,v_2)$ and $\tilde{M}_{ij} = \langle\tilde{y}_i(\boldsymbol{z})\tilde{y}_j(\boldsymbol{z}+\boldsymbol{r})\rangle$. Using the same coordinate system as in Appendix C, the only non-vanishing matrix elements of the form $\langle vs_i\rangle$ are those involving the 3-component of the displacement field, which read

$$\langle vs_3\rangle = -\frac{3\sigma_s^2}{\sigma_0}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right), \tag{D3}$$

where $\psi_v \equiv (2\psi_\perp + \psi_\parallel)/3$ and σ_0 is the rms perturbation in the background density field. The matrix elements $\langle s_is_j\rangle$ are as given above in equation (C2), while the remaining elements of $\tilde{\mathbf{M}}$ are given by

$$\langle v_1v_2\rangle = \langle v_2v_1\rangle = \frac{\xi(r)}{\sigma_0^2} = \psi(r) \tag{D4}$$

and

$$\langle v_1v_1\rangle = \langle v_2v_2\rangle = \psi(0) = 1, \tag{D5}$$

where $\psi(r)$ is the normalized correlation function of the background density field.

The covariance matrix, $\tilde{\mathbf{M}}$, is of block-diagonal form, so there are no terms linking the sets $\{s_1,s_2\}$ and $\{s_3,v_1,v_2\}$ in the exponentiated quadratic form in the probability distribution $P(\boldsymbol{s},v_1,v_2|\boldsymbol{r})$. So, to perform the integrations over v_1,v_2 in equation (D2), we need only consider the reduced covariance matrix $\hat{\mathbf{M}}$ for the vector $\hat{\boldsymbol{y}}^{\mathrm{T}} = (s_3,v_1,v_2)$. It follows that the integral over v_1,v_2 takes the form

$$I_v = \frac{1}{(2\pi)^{3/2}\,||\,\tilde{\mathbf{M}}\,||^{-1/2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{(E-1)\sigma_0(v_1+v_2)}\exp\left(-\frac{1}{2}\tilde{y}_i^{\mathrm{T}}\tilde{M}_{ij}^{-1}\tilde{y}_j\right)\mathrm{d}v_1\,\mathrm{d}v_2. \tag{D6}$$

Both integrations may be perfomed using the standard result (Gradshteyn & Ryzhik 1980, section 3.323)

$$\int_{-\infty}^{\infty}\exp(-p^2x^2 \pm qx)\,\mathrm{d}x = \frac{\sqrt{\pi}}{p}\exp\left(\frac{q^2}{4p^2}\right), \tag{D7}$$

to obtain

$$I_v = \frac{1}{\sqrt{4\pi\sigma_s(1-\psi_\parallel)}}\exp\left[\frac{-s_3^2}{4\sigma_s^2(1-\psi_\parallel)}\right]\exp\left[\frac{-3(E-1)s_3}{(1-\psi_\parallel)}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right)\right]$$
$$\times\ \exp\left[(E-1)^2\sigma_0(1+\psi) - \frac{9\sigma_s^2(E-1)^2}{(1-\psi_\parallel)}\left(\frac{\mathrm{d}\psi_v}{\mathrm{d}r}\right)^2\right]. \tag{D8}$$

From this point, the derivation proceeds as in Appendix C to yield, after much tedious rearrangement, the expression for ξ_{cZ} given in equation (25) of Section 3.

This paper has been produced using the Blackwell Scientific Publications LaTeX style file.

FINIS OPERIS